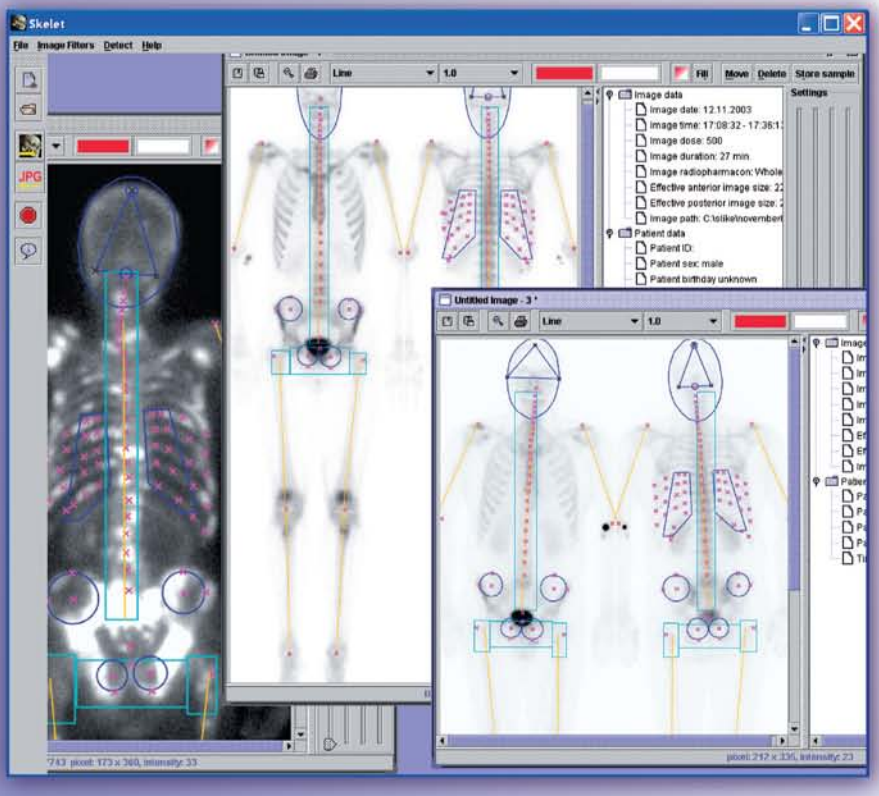
COMPUTATIONAL INTELLIGENCE IN MEDICAL IMAGING

TECHNIQUES AND APPLICATIONS



EDITED BY

GERALD SCHAEFER

ABOUL ELLA HASSANIEN

JIANMIN JIANG

CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

Computational Intelligence in Medical Imaging

Techniques and Applications

Computational Intelligence in Medical Imaging

Techniques and Applications

Edited by
Gerald Schaefer
Aboul Ella Hassanien
Jianmin Jiang

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-6059-1 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Computational intelligence in medical imaging techniques and applications / editors, Gerald Schaefer, Aboul Ella Hassanien, and Jianmin Jiang.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-4200-6059-1 (alk. paper)
1. Diagnostic imaging--Data processing. 2. Computational intelligence. I. Schaefer, Gerald. II. Hassanien, Aboul Ella. III. Jiang, J., Ph. D. IV. Title.
[DNLM: 1. Diagnosis, Computer-Assisted--methods. 2. Diagnosis, Computer-Assisted--trends. WB 141 C738 2008]

RC78.7.D53C656 2008
616.07'540285--dc22 2008040413

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

Medical imaging is an indispensible tool for many branches of medicine. It enables and facilitates the capture, transmission, and analysis of medical images and aids in medical diagnoses. The use of medical imaging is still on the rise with new imaging modalities being developed and continuous improvements being made to devices' capabilities. Recently, computational intelligence techniques have been employed in various applications of medical imaging and have been shown to be advantageous compared to classical approaches, particularly when classical solutions are difficult or impossible to formulate and analyze. In this book, we present some of the latest trends and developments in the field of computational intelligence in medical imaging.

The first three chapters present the current state of the art of various areas of computational intelligence applied to medical imaging. Chapter 1 details neural networks, Chapter 2 reviews evolutionary optimization techniques, and Chapter 3 covers in detail rough sets and their applications in medical image processing.

Chapter 4 explains how neural networks and support vector machines can be utilized to classify wound images and arrive at decisions that are comparable to or even more consistent than those of clinical practitioners. Neural networks are also explored in Chapter 5 in the context of accurately extracting the boundaries of skin lesions, a crucial stage for the identification of melanoma. Chapter 6 discusses tabu search, an intelligent optimization technique, for feature selection and classification in the context of prostate cancer analysis.

In Chapter 7, the authors demonstrate how image processing techniques based on intuitionistic fuzzy sets can successfully handle the inherent uncertainties present in mammographic images. Fuzzy logic is also employed in Chapter 8, where fuzzy set–based clustering techniques for medical image segmentation are discussed.

A comprehensive system for handling and utilizing biomedical image databases is described in Chapter 9: The features extracted from medical images are encoded within a Bayesian probabilistic framework that enables learning from previously retrieved relevant images. Chapter 10 explores how machine learning techniques are used to develop a statistical parts-based appearance model that can be used to encapsulate the natural intersubject anatomical variance in medical images.

In Chapter 11, a multistage image segmentation algorithm based on reinforcement learning is introduced and successfully applied to the problem of prostate segmentation in transrectal ultrasound images. Chapter 12 presents a machine learning approach for automatic segmentation and diagnosis of bone scintigraphy. Chapter 13 employs a set of intelligent agents that communicate via a blackboard architecture to provide accurate and efficient 3-D medical image segmentation.

Chapter 14 explains how Monte Carlo simulations are employed to perform reconstruction of SPECT and PET tomographic images. Chapter 15 discusses the use of artificial life concepts to develop intelligent, deformable models that segment and analyze structures in medical images.

Obviously, a book of 15 chapters is nowhere near sufficient to encompass all the exciting research that is being conducted in utilizing computational intelligence techniques in the context of medical imaging. Nevertheless, we believe the chapters that were selected from among almost 40 proposals and rigorously reviewed by three experts present a good snapshot of the field. This work will prove useful not only in documenting recent advances but also in stimulating further research in this area.

Gerald Schaefer, Aboul Ella Hassanien, Jianmin Jiang

We are grateful to the following reviewers:

Rocio Alba-Flores
Lucia Ballerini
Hueseyin Cakmak
Marina Chukalina
Sergio Damas
Sabine Dippel
Christian Dold
Emanuele Frontoni
Peter Goebel
Yu-Le Huang
Jiri Jan
Jean-Claude Klein
Syoji Kobashi
Francesco Masulli
Kalana Mendis
Roberto Morales
Henning Mueller
Mike Nachtegael
Tomoharu Nakashima
Dmitry Nikolaev
Lars Nolle
Peter Plassmann
Stefan Schulte
George Sergiadis
Stephen Smith
Roman Starosolski
Kenji Suzuki
Gui Yun Tian
Michal Zavisek
Primo Zingaretti
Reyer Zwiggelaar

Editors

Gerald Schaefer obtained his BSc in computing from the University of Derby and his PhD in computer vision from the University of East Anglia. He worked as a research associate (1997–1999) at the Colour & Imaging Institute, University of Derby, as a senior research fellow at the School of Information Systems, University of East Anglia (2000–2001), and as a senior lecturer in computing at the School of Computing and Informatics at Nottingham Trent University (2001–2006). In September 2006, he joined the School of Engineering and Applied Science at Aston University. His research interests include color image analysis, image retrieval, medical imaging, and computational intelligence. He is the author of more than 150 scientific publications in these areas.

Aboul Ella Hassanien is an associate professor in the Computer and Information Technology Department at Cairo University. He works in a multidisciplinary environment involving computational intelligence, information security, medical image processing, data mining, and visualization applied to various real-world problems. He received his PhD in computer science from Tokyo Institute of Technology, Japan. He serves on the editorial board of several reputed international journals, has guest edited many special issues on various topics, and is involved in the organization of several conferences. His research interests include rough set theory, wavelet theory, medical image analysis, fuzzy image processing, information security, and multimedia data mining.

Jianmin Jiang is a full professor of digital media at the School of Informatics, University of Bradford, United Kingdom. He received his BSc from Shandong Mining Institute, China in 1982; his MSc from China University of Mining and Technology in 1984; and his PhD from the University of Nottingham, United Kingdom in 1994. His research interests include image/video processing in compressed domains, medical imaging, machine learning and AI applications in digital media processing, retrieval, and analysis. He has published more than 200 refereed research papers and is the author of one European patent (EP01306129) filed by British Telecom Research Lab. He is a chartered engineer, a fellow of IEE, a fellow of RSA, a member of EPSRC College, an EU FP-6/7 proposal evaluator, and a consulting professor at the Chinese Academy of Sciences and Southwest University, China.

Contributors

Ajith Abraham
Center for Quantifiable Quality of Service in Communication Systems
Norwegian University of Science and Technology
Trondheim, Norway

Tal Arbel
Centre for Intelligent Machines
McGill University
Montreal, Quebec, Canada

Brahima Belem
Department of Computing and Mathematical Sciences
University of Glamorgan
Pontypridd, Wales, United Kingdom

Ahmed Bouridane
Institute of Electronics, Communications and Information Technology (ECIT)
Queen's University Belfast
Belfast, Ireland

Edward Chaum
University of Tennessee Health Science Center
Hamilton Eye Institute
Memphis, Tennessee

Jens Gregor
Department of Computer Science
University of Tennessee
Knoxville, Tennessee

Ghassan Hamarneh
School of Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada

Aboul Ella Hassanien
Information Technology Department
Cairo University
Giza, Egypt

Adrian A. Hopgood
Faculty of Computing Sciences and Engineering
De Montfort University
Leicester, United Kingdom

Jianmin Jiang
School of Informatics
University of Bradford
Bradford, West Yorkshire, United Kingdom

Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

Thomas P. Karnowski
Image Science and Machine Vision Group
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Igor Kononenko
Faculty of Computer and Information Science
University of Ljubljana
Ljubljana, Slovenia

Ignace Lemahieu
University of Ghent – DOZA
Ghent, Belgium

Tim McInerney
Department of Computer Science
Ryerson University
Toronto, Ontario, Canada

Chris McIntosh
School of Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada

Lars Nolle
School of Science and Technology
Nottingham Trent University
Nottingham, United Kingdom

Y. H. Peng
School of Informatics
University of Bradford
Bradford, West Yorkshire, United Kingdom

James F. Peters
Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada

Peter Plassmann
Department of Computing and Mathematical Sciences
University of Glamorgan
Pontypridd, Wales, United Kingdom

Jeffery R. Price
Image Science and Machine Vision Group
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Maher I. Rajab
Computer Engineering Department
College of Computer and Information Systems
Umm Al-Qura University
Mecca, Saudi Arabia

Mohammad Ali Roula
Department of Electronics and Computer Systems Engineering
University of Glamorgan
Pontypridd, Wales, United Kingdom

Farhang Sahba
Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada

Luka Šajn
Faculty of Computer and Information Science
University of Ljubljana
Ljubljana, Slovenia

Magdy M. A. Salama
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, Canada

Gerald Schaefer
School of Engineering and Applied Science
University of Birmingham
Birmingham, United Kingdom

George D. Sergiadis
Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki
Thessaloniki, Greece

Steven Staelens
IBITECH-MEDISIP
Ghent University – IBBT
Ghent, Belgium

Muhammad Atif Tahir
Faculty of CEMS
University of the West of England
Bristol, United Kingdom

Roger J. Tait
School of Computing and Informatics
Nottingham Trent University
Nottingham, United Kingdom

Demetri Terzopoulos
Department of Computer Science
University of California
Los Angeles, California

Hamid R. Tizhoosh
Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada

Kenneth W. Tobin
Image Science and Machine Vision Group
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Matthew Toews
Centre for Intelligent Machines
McGill University
Montreal, Quebec, Canada

Ioannis K. Vlachos
Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki
Thessaloniki, Greece

Jonathan Wall
Amyloid and Preclinical and Diagnostic Molecular Imaging Laboratory
University of Tennessee Graduate School of Medicine
Knoxville, Tennessee

Z. Q. Wu
School of Informatics
University of Bradford
Bradford, West Yorkshire, United Kingdom

Huiyu Zhou
Department of Electronic and Computer Engineering
School of Engineering and Design
Brunel University
Uxbridge, Middlesex, United Kingdom

Chapter 1

Computational Intelligence on Medical Imaging with Artificial Neural Networks

Z. Q. Wu, Jianmin Jiang, and Y. H. Peng

Contents

Neural networks have been widely reported in the research community of medical imaging. In this chapter, we provide a focused literature survey on neural network development in computer-aided diagnosis (CAD), medical image segmentation and edge detection toward visual content analysis, and medical image registration for its preprocessing and postprocessing. From among all these techniques and algorithms, we select a few representative ones to provide inspiring examples to illustrate (a) how a known neural network with fixed structure and training procedure can be applied to resolve a medical imaging problem; (b) how medical images can be analyzed, processed, and characterized by neural networks; and (c) how neural networks can be expanded further to resolve problems relevant to medical imaging. In the concluding section, a comparison of all neural networks is included to provide a global view on computational intelligence with neural networks in medical imaging.

1.1 Introduction

An artificial neural network (ANN) is an information processing system that is inspired by the way biological nervous systems store and process information like human brains. It contains a large number of highly interconnected processing neurons working together in a distributed manner to learn from the input information, to coordinate internal processing, and to optimize its final output. In the past decades, neural networks have been successfully applied to a wide range of areas, including computer science, engineering, theoretical modeling, and information systems. Medical imaging is another fruitful area for neural networks to play crucial roles in resolving problems and providing solutions. Numerous algorithms have been reported in the literature applying neural networks to medical image analysis, and we provide a focused survey on computational intelligence with neural networks in terms of (a) CAD with specific coverage of image analysis in cancer screening, (b) segmentation and edge detection for medical image content analysis, (c) medical image registration, and (d) other applications covering medical image compression, providing a global view on the variety of neural network applications and their potential for further research and developments.

Neural network applications in CAD represent the mainstream of computational intelligence in medical imaging. Their penetration and involvement are comprehensive for almost all medical problems because (a) neural networks can adaptively learn from input information and upgrade themselves in accordance with the variety and change of input content; (b) neural networks can optimize the relationship between the inputs and outputs via distributed computing, training, and processing, leading to reliable solutions desired by specifications; (c) medical diagnosis relies on visual inspection, and medical imaging provides the most important tool for facilitating such inspection and visualization.

Medical image segmentation and edge detection remains a common problem fundamental to all medical imaging applications. Any content analysis and regional inspection requires segmentation of featured areas, which can be implemented via edge detection and other techniques. Conventional approaches are typified by a range of well-researched algorithms, including watershed, region-growing, snake modeling, and contour detection. In comparison, neural network approaches exploit the learning capability and training mechanism to classify medical images into content-consistent regions to complete segmentations as well as edge detections.

Another fundamental technique for medical imaging is registration, which plays important roles in many areas of medical applications. Typical examples include wound care, disease prediction, and health care surveillance and monitoring. Neural networks can be designed to provide alternative solutions via competitive learning, self-organizing, and clustering to process input features and find the best possible alignment between different images or data sets.

The remainder of this chapter provides useful insights for neural network applications in medical imaging and computational intelligence. We explain the basics of neural networks to enable beginners to understand the structure, connections, and neuron functionalities. Then we present detailed descriptions of neural network applications in CAD, image segmentation and edge detection, image registration, and other areas.

1.2 Neural Network Basics

To enable understanding of neural network fundamentals, to facilitate possible repetition of those neural networks introduced and successfully applied in medical imaging, and to inspire further development of neural networks, we cover essential basics in this section about neural networks to pave the way for the rest of the chapter in surveying neural networks. We start from a theoretical model of one single neuron and then introduce a range of different types of neural networks to reveal their structure, training mechanism, operation, and functions.

The basic structure of a neuron can be theoretically modeled as shown in Figure 1.1.

Figure 1.1 shows the model of a single neuron, where $X\{x_i, i = 1, 2, \ldots, n\}$ represents the inputs to the neuron and Y represents the output. Each input is multiplied by its weight w_i, a bias b is associated with each neuron, and their sum goes through a transfer function f. As a result, the relationship between input and output can be described as follows.

$$Y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{1.1}$$

A range of transfer functions have been developed to process the weighted and biased inputs. Four of the basic transfer functions widely adopted for medical image processing are illustrated in Figure 1.2.

Via selection of transfer function and connection of neurons, various neural networks can be constructed to be trained for producing the specified outputs. Major neural networks commonly used for medical image processing are

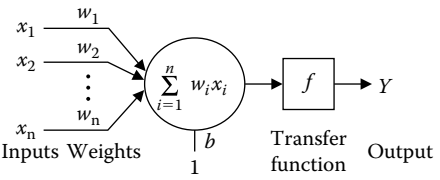


FIGURE 1.1: The model of a neuron.

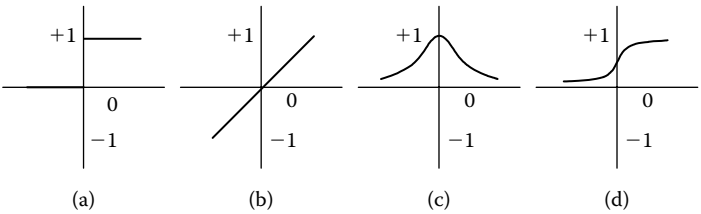


FIGURE 1.2: Four widely adopted transfer functions: (a) hardlimit, (b) linear, (c) RBF, and (d) sigmoid.

classified as feedforward neural network, feedback network, and self-organizing map. The learning paradigms for the neural networks in medical image processing generally include supervised networks and unsupervised networks. In supervised training, the training data set consists of many pairs in the source and target patterns. The network processes the source inputs and compares the resulting outputs against the target outputs, and adjusts its weights to improve the correct rate of the resulting outputs. In unsupervised networks, the training data set does not include any target information.

A general feedforward network [1] often consists of multiple layers, typically including one input layer, a number of hidden layers, and an output layer. In the feedforward neural networks, the neurons in each layer are only fully interconnected with the neurons in the next layer, which means signals or information being processed travel along a single direction.

A back-propagation (BP) network [2] is a supervised feedforward neural network, and it is a simple stochastic gradient descent method to minimize the total squared error of the output computed by the neural network. Its errors propagate backwards from the output neurons to the inner neurons. The processes of adjusting the set of weights between the layers and recalculating the output continue until a stopping criterion is satisfied.

The radial basis function (RBF) [3] network is a three-layer, supervised feedforward network that uses a nonlinear transfer function (normally the Gaussian) for the hidden neurons and a linear transfer function for the output neurons. The Gaussian is applied to the net input to produce a radial function of the distance between each pattern vector and each hidden unit weight vector.

The feedback (or recurrent) neural network [4] can have signals traveling in both directions by introducing loops. Their state is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium must be found. They are powerful but can get extremely complicated.

The Hopfield network [4] is a typical feedback, and its inspiration is to store certain patterns in a manner similar to the way the human brain stores memories. The Hopfield network has no special input or output neurons, but all neurons are both input and output, and all of them connect to all others in both directions. After receiving the input simultaneously by all the neurons,

they output to each other, and the process does not stop until a stable state is reached. In the Hopfield network, it is simple to set up the weights between neurons in order to set up a desired set of patterns as stable class patterns. The Hopfield network is an unsupervised learning network and thus does not require a formal training phase.

Quite different from feedforward and feedback networks, the Kohonen neural network (a self-organizing map, SOM) [5] learns to classify input vectors according to how they are grouped in the input space. In the network, a set of artificial neurons learns to map points in an input space to the coordinates in an output space. Each neuron stores a weight vector (an array of weights), each of which corresponds to one of the inputs in the data. When presented with a new input pattern, the neuron whose weight is closest in Euclidian space to the new input pattern is allowed to adjust its weight so that it gets closer to the input pattern. The Kohonen neural network uses a competitive learning algorithm to train itself in an unsupervised manner.

In Kohonen neural networks, each neuron is fed by input vector (data point) $x \in R^n$ through a weight vector $w \in R^n$. Each time a data point is input to the network, only the neuron j whose weight vector most resembles the input vector is selected to fire, according to the following rule:

$$j = \arg\min(\|x - w\|^2) \tag{1.2}$$

The firing or winning neuron j and its neighboring neurons i have their weight vectors w modified according to the following rule:

$$w_i(t+1) = w_i(t) + h_{ij}(\|r_i - r_j\|, t) \cdot (x(t) - w_i(t)) \tag{1.3}$$

where $h_{ij}(||r_i - r_j||, t)$ is a kernel defined on the neural network space as a function of the distance $||r_i - r_j||$ between the firing neuron j and its neighboring neurons i, and the time t defines the number of iterations. Its neighboring neurons modify their weight vectors so they also resemble the input signal, but less strongly, depending on their distance from the winner.

The remainder of the chapter provides detailed descriptions of computational intelligence in medical imaging with neural networks. Their recent applications are classified into four categories: CAD, image segmentation, registration, and other applications. Each section gives more details on an application in one of these categories and provides overviews of the other relevant applications. A comparison of neural networks is presented in Section 1.7.

1.3 Computer-Aided Diagnosis (CAD) with Neural Networks

Neural networks have been incorporated into many CAD systems, most of which distinguish cancerous signs from normal tissues. Generally, these

systems enhance the images first and then extract interesting regions from the images. The values of many features are calculated based on the extracted regions and are forwarded to neural works that make decisions in terms of learning, training, and optimizations. Among all applications, early diagnosis of breast cancers and lung cancers represents the most typical examples in the developed CAD systems.

Ge and others [6] developed a CAD system to identify microcalcification clusters automatically on full-field digital mammograms. The main procedures of the CAD system included six stages: preprocessing, image enhancement, segmentation of microcalcification candidates, false positive (FP) reduction for individual microcalcifications, regional clustering, and FP reduction for clustered microcalcifications.

To reduce FP individual microcalcifications, a convolution neural network (CNN) was employed to analyze 16×16 regions of interest centered at the candidate derived from segmentations. The CNN was designed to simulate the vision of vertebrate animals and could be considered a simplified vision machine designed to perform the classification of the regions into two output types: disease and nondisease. Their CNN contained an input layer with 14 neurons, two hidden layers with 10 neurons each, and one output layer. The convolution kernel sizes of the first group of filters between the input and the first hidden layer were designed as 5×5, and those of the second group of filters between the first and second hidden layers were 7×7. The images in each layer were convolved with convolution kernels to obtain the pixel values to be transferred to the following layer. The logistic sigmoid function was chosen as the transfer function for both the hidden neurons and output neurons. An illustration of the neural network structure and its internal connections between the input layer, hidden layer, and output layers is given in Figure 1.3.

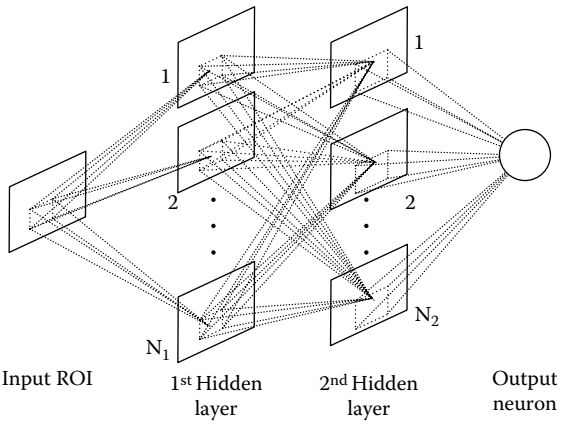


FIGURE 1.3: Schematic diagram of a CNN.

The convolution kernels are arranged in a way to emphasize a number of image characteristics rather than those less correlated values derived from feature spaces of input. These characteristics include (a) the horizontal versus vertical information, (b) local versus nonlocal information, and (c) image processing (filtering) versus signal propagation [7].

The CNN was trained using a backpropagation learning rule with the sum-of-squares error (SSE) function, which allowed a probabilistic interpretation of the CNN output, that is, the probability of correctly classifying the input sample as a true microcalcification region of interest (ROI).

At the stage of FP reduction for clustered microcalcifications, morphological features (such as the size, mean density, eccentricity, moment ratio, axis ratio features, and number of microcalcifications in a cluster) and features derived from the CNN outputs (such as the minimum, maximum, and mean of the CNN output values) were extracted from each cluster. For each cluster, 25 features (21 morphological plus 4 CNN features) were extracted. A linear discriminating analysis (LDA) classifier was then applied to differentiate clustered microcalcifications from FPs. The stepwise LDA feature selection involved the selection of three parameters for selection.

In the study by Ge and colleagues, a set of 96 images was split into a training set and a validation set, each with 48 images. An appropriate set of parameters was selected by searching in the parameter space for the combination of three parameters of the LDA that could achieve the highest classification accuracy with a relatively small number of features in the validation set. Then the three parameters of LDA were applied to select a final set of features and the LDA coefficients by using the entire set of 96 training images, which contained 96 true positive (TP) and over 500 FP clusters. The trained classifier was applied to a test subset to reduce the FPs in the CAD system [6].

To develop a computerized scheme for the detection of clustered microcalcifications in mammograms, Nagel and others [8] examined three methods of feature analysis: rule based (the method currently used), an ANN, and a combined method. The ANN method used a three-layer error-backpropagation network with five input units corresponding to the radiographic features of each microcalcification and one output unit corresponding to the likelihood of being a microcalcification. The reported work revealed that two hidden units were insufficient for good performance of the ANN, and it was necessary to have at least three hidden units to achieve adequate performance. However, the performance was not improved any further when the number of hidden units was increased over three. Therefore, the finalized ANN had five inputs, three hidden units, and one output unit. It was reported that such a combined method performed better than any method alone.

Papadopoulossa, Fotiadisb, and Likasb [9] presented a hybrid intelligent system for the identification of microcalcification clusters in digital mammograms, which could be summarized in three steps: (a) preprocessing and segmentation, (b) ROI specification, and (c) feature extraction and classification. In the classification schema, 22 features were automatically computed

that referred either to individual microcalcifications or to groups of them. The reduction of FP cases was performed using an intelligent system containing two subsystems: a rule-based system and a neural network-based system. The rule construction procedure consisted of the feature identification step as well as the selection of the particular threshold value for each feature. Before using the neural network, the reduction in the number of features was achieved through principal component analysis (PCA), which transforms each 22-dimensional feature vector into a 9-dimensional feature vector as the input to the neural network. The neural network used for ROI characterization was a feedforward neural network with sigmoid hidden neuron (multilayer perceptron, MLP).

Christoyiani, Dermatas, and Kokkinakis [10] presented a method for fast detection of circumscribed mass in mammograms employing an RBF neural network (RBFNN). In the method, each neuron output was a nonlinear transformation of a distance measure of the neuron weights and its input vector. The nonlinear operator of the RBFNN hidden layer was implemented using a Cauchy-like probability density function. The implementation of RBFNN could be achieved by using supervised or unsupervised learning algorithms for an accurate estimation of the hidden layer weights. The k-means unsupervised algorithm was adopted to estimate the hidden-layer weights from a set of training data containing statistical features from both circumscribed lesions and normal tissue. After the initial training and the estimation of the hidden-layer weights, the weights in the output layer were computed by using Wincer-filter theory, or minimizing the mean square error (MSE) between the actual and the desired filter output.

Patrocinio and others [11] demonstrated that only several features, such as irregularity, number of microcalcifications in a cluster, and cluster area, were needed as the inputs of a neural network to separate images into two distinct classes: suspicious and probably benign. Setiono [12] developed an algorithm by pruning a feedforward neural network, which produced high accuracy rates for breast cancer diagnosis with a small number of connections. The algorithm extracted rules from a pruned network by considering only a finite number of hidden-unit activation values. Connections in the network were allowed only between input units and hidden units and between hidden units and output units. The algorithm found and eliminated as many unnecessary network connections as possible during the training process. The accuracy of the extracted rules from the pruned network is almost as high as the accuracy of the original network.

The abovementioned applications cover different aspects of applying neural networks, such as the number of neurons in the hidden layer, the reduction of features in classifications, and the reduction of connections for better efficiency. Similar improvements could be made in applying ANN to other practical utilizations rather than just in identifying microcalcification clusters.

ANN also plays an important role in detecting the cancerous signs in lungs. Xu and colleagues [13] developed an improved CAD scheme for the automated

detection of lung nodules in digital chest images to assist radiologists who may miss up to 30% of the actually positive cases in their daily practice. In the CAD scheme, nodule candidates were selected initially by multiple gray-level thresholds of the difference image (subtraction of a signal-enhanced image and a signal-suppressed image) and then classified into six groups. A large number of FPs were eliminated by adaptive rule-based tests and an ANN.

Zhou and others [14] proposed an automatic pathological diagnosis procedure called neural ensemble-based detection that utilized an ANN ensemble to identify lung cancer cells in the specimen images of needle biopsies obtained from the bodies of the patients to be diagnosed. An ANN ensemble formed a learning paradigm while several ANNs were jointly used to solve a problem. The ensemble was built on a two-level ensemble architecture, and the predictions of those individual networks were combined by plurality voting.

Keserci and Yoshida [15] developed a CAD scheme for automated detection of lung nodules in digital chest radiographs based on a combination of morphological features and the wavelet snake. In their scheme, an ANN was used to efficiently reduce FPs by using the combined features. The scheme was applied to a publicly available database of digital chest images for pulmonary nodules. Qian and others [16] trained a computer-aided cytologic diagnosis (CACD) system to recognize expression of the cancer biomarkers histone H2AX in lung cancer cells and then tested the accuracy of this system to distinguish resected lung cancer from preneoplastic and normal tissues. The major characteristics of CACD algorithms were to adapt detection parameters according to cellular image contents. Coppini and colleagues [17] described a neural network–based system for the computer-aided detection of lung nodules in chest radiograms. The approach was based on multiscale processing and feedforward neural networks that allowed an efficient use of a priori knowledge about the shape of nodules and the background structure.

Apart from the applications in breast cancer and lung cancer, ANN has been adopted in many other analyses and diagnosis. Mohamed and others [18] compared bone mineral density (BMD) values for healthy persons and identified those with conditions known to be associated with BMD obtained from dual X-ray absorptiometry (DXA). An ANN was designed to quantitatively estimate site-specific BMD values in comparison with reference values obtained by DXA. Anthropometric measurements (i.e., sex, age, weight, height, body mass index, waist-to-hip ratio, and the sum of four skinfold thicknesses) were fed to an ANN as input variables. The estimates based on four input variables were generated as output and were generally identical to the reference values among all studied groups.

Scott [19] tried determining whether a computer-based scan analysis could assist clinical interpretation in this diagnostically difficult population. An ANN was created using only objective image-derived inputs to diagnose the presence of pulmonary embolism. The ANN predictions performed comparably to clinical scan interpretations and angiography results.

In all the applications mentioned above, the roles of ANNs have a common principle in the sense that most of them are applied to reduce FP detections in both mammograms and chest images via examining the features extracted from the suspicious regions. As a matter of fact, ANN is not limited to academic research but also plays important roles in commercially available diagnosis systems, such as ImageChecker for mammograms.

1.4 Medical Image Segmentation and Edge Detection with Neural Networks

Medical image segmentation is a process for dividing a given image into meaningful regions with homogeneous properties. Image segmentation is an indispensable process in outlining boundaries of organs and tumors and in the visualization of human tissues during clinical analysis. Therefore, segmentation of medical images is very important for clinical research, diagnosis, and applications, leading to requirement of robust, reliable, and adaptive segmentation techniques.

Kobashi and others [20] proposed an automated method to segment the blood vessels from three-dimensional (3-D) time-of-flight magnetic resonance angiogram (MRA) volume data. The method consisted of three steps: removal of the background, volume quantization, and classification of primitives by using an artificial neural network.

After volume quantization by using a watershed segmentation algorithm, the primitives in the MRA image stand out. To further improve the result of segmentation, the obtained primitives had to be separated into the blood vessel class and the fat class. Three features and a three-layered, feedforward neural network were adopted for the classification. Compared with the fat, the blood vessel is like a tube—long and narrow. Two features, vascularity and narrowness, were introduced to measure such properties. Because the histogram of blood vessels is quite different from that of the fat in shapes, the third feature, histogram consistency, was added for further improvement of the segmentation.

The feedforward neural network is composed of three layers: an input layer, a hidden layer, and an output layer. The structure of the described neural network is illustrated in Figure 1.4.

As seen, three input units were included at the input layer, which was decided by the number of features extracted from medical images. The number of neurons in the output layer was one to produce two classes. The number of neurons in the hidden layer was usually decided by experiments. Generally, a range of different numbers were tried in the hidden layer, and the number that achieved the best training results was selected.

In the proposed method, the ANN classified each primitive, which was a clump of voxels, by evaluating the intensity and the 3-D shape. In their

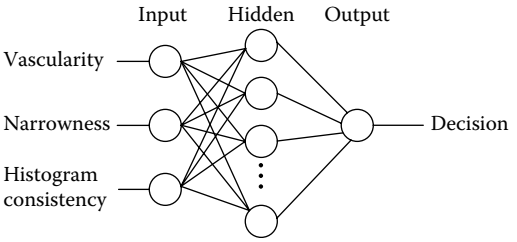


FIGURE 1.4: Three-layer feedforward neural network.

experiments, the ANN was trained using 60 teaching data sets derived from an MRA data set. Each primitive was classified into the blood vessel (indicated by the value of 1) or the fat (indicated by the value of 0), and the values of the three features were calculated. All these values were fed into the feedforward ANN for training the weights of the neurons. Seven new MRA data, whose primitives were unclassified, were fed into the trained neural network for testing. The segmentation performance was measured by the value of accuracy, as defined in Equation 1.4, and the rate achieved by the reported algorithm is 80.8% [20].

$$Accuracy = \frac{Number\ of\ correctly\ classified\ primitives}{Total\ number\ of\ primitives} \times 100\% \tag{1.4}$$

Apart from the work proposed by Kobashi and colleagues in ANN-based segmentation, there are many applications for the images generated by computed tomography (CT) and magnetic resonance imaging (MRI). Middleton and Damper [21] combined use of a neural network (an MLP, a type of feedforward neural network) and active contour model ("snake") to segment structures in magnetic resonance (MR) images. The highlights of the reported work can be summarized by the following two steps:

1. The perceptron was trained to produce a binary classification of each pixel as either a boundary or a nonboundary;

2. The resulting binary (edge-point) image formed the external energy function for a snake model, which was applied to link the candidate boundary points into a continuous and closed contour.

Lin [22] applied the Hopfield neural network (a feedback neural network) with penalized fuzzy c-means (FCM) technique to medical image segmentation. In the algorithm, the pixels with their first- and second-order moments constructed from their n nearest neighbors as a training vector were mapped to a two-dimensional (2-D) Hopfield neural network for the purpose of classifying the image into suitable regions.

Lin and colleagues [23] generalized the Kohonen competitive learning (KCL) algorithm with fuzzy and fuzzy-soft types called fuzzy KCL (FKCL)

and fuzzy-soft KCL (FSKCL). These KCL algorithms fused the competitive learning with soft competition and FCM membership functions. These generalized KCLs were applied to MRI and MRA ophthalmological segmentations. It was found that these KCL-based MRI segmentation techniques were useful in reducing medical image noise effects using a learning mechanism. The FSKCL algorithm was recommended for use in MR image segmentation as an aid to small lesion diagnosis.

Dokur and Olmez [24] proposed a quantizer neural network (QNN) for the segmentation of MR and CT images. QNN was a novel neural network structure and was trained by genetic algorithms. It was comparatively examined with an MLP and a Kohonen network for the segmentation of MR and CT head images. They reported that QNN achieved the best classification performance with fewer neurons after a short training time.

Stalidis and others [25] presented an integrated model-based processing scheme for cardiac MRI, which was embedded in an interactive computing environment suitable for quantitative cardiac analysis. The scheme provided a set of functions for the extraction, modeling, and visualization of cardiac shape and deformation. In the scheme, a learning segmentation process incorporating a generating–shrinking neural network was combined with a spatiotemporal parametric model through functional basis decomposition.

Chang and Ching [26] developed an approach for medical image segmentation using a fuzzy Hopfield neural network based on both global and local gray-level information. The membership function simulated with neuron outputs was determined using a fuzzy set, and the synaptic connection weights between the neurons were predetermined and fixed in order to improve the efficiency of the neural network.

Shen and others [27] proposed a segmentation technique based on an extension to the traditional FCM clustering algorithm. In their work, a neighborhood attraction, which was dependent on the relative location and features of neighboring pixels, was shown to improve the segmentation performance, and the degree of attraction was optimized by a neural-network model. Simulated and real brain MR images with different noise levels were segmented to demonstrate the superiority of the technique compared to other FCM-based methods.

Chang and Chung [28] designed a two-layer Hopfield neural network called the competitive Hopfield edge-finding neural network (CHEFNN) to detect the edges of CT and MRI images. To effectively remove the effect of tiny details or noises and the drawback of disconnected fractions, the CHEFNN extended the one-layer 2-D Hopfield network at the original image plane to a two-layer 3-D Hopfield network with edge detection to be implemented on its third dimension. Under the extended 3-D architecture, the network was capable of incorporating a pixel's contextual information into a pixel-labeling procedure. In addition, they [29] discovered that high-level contextual information could not be incorporated into the segmentation procedure in techniques

using traditional Hopfield neural networks and thus proposed the contextual constraint-based Hopfield neural cube (CCBHNC) for image segmentation. The CCBHNC adopted a 3-D architecture with pixel classification implemented on its third dimension. Recently, still for the edge detection, Chang [30] presented a specially designed Hopfield neural network called the contextual Hopfield neural network (CHNN). The CHNN mapped the 2-D Hopfield network at the original image plane. With direct mapping, the network was capable of incorporating pixels' contextual information into an edge-detecting procedure. As a result, the CHNN could effectively remove the influence of tiny details and noise.

Most of these applications were developed based on CT or MRI images but the neural networks adopted are in quite different ways. ANN can reduce the influence of noise in the image and hence make the segmentation more robust. Further, ANN can classify different tissues and then combine them according to segmentation requirements, which is beyond the power of traditional segmentation.

1.5 Medical Image Registration with Neural Networks

Image registration is the process of transforming the different sets of data into one coordinate system. Registration is necessary in order to be able to compare or integrate the images from different measurements, which may be taken at different points in time from the same modality or obtained from the different modalities such as CT, MR, angiography, and ultrasound. Medical imaging registration often involves elastic (or nonrigid) registration to cope with elastic deformations of the body parts imaged. Nonrigid registration of medical images can also be used to register a patient's data to an anatomical atlas. Medical image registration is the preprocessing needed for many medical imaging applications with strong relevance to the result of segmentation and edge detection.

Generally, image registration algorithms can be classified into two groups: area-based methods and feature-based methods. For area-based image registration methods, the algorithm looks at the structure of the image via correlation metrics, Fourier properties, and other means of structural analysis. Most feature-based methods fine-tune their mapping to the correlation of image features: lines, curves, points, line intersections, boundaries, and so on.

To measure the volume change of lung tumor, Matsopoulos and colleagues [31] proposed an automatic, 3-D, nonrigid registration scheme that applied SOM to thoracic CT data of patients for establishing correspondence between the feature points. The practical implementation of this scheme could provide estimations of lung tumor volumes during radiotherapy treatment planning. In the algorithm, the automatic correspondence of the interpolant points

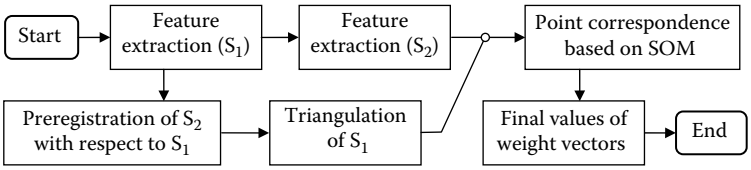


FIGURE 1.5: The elastic registration scheme.

was based on the initialization of the Kohonen neural network model able to identify 500 corresponding pairs of points approximately in the two CT sets S_1 and S_2. An overview of the described algorithm is illustrated in Figure 1.5.

In the algorithm, two sets of points were defined: S_2 is the set of points for vertebrae, ribs, and blades segmented from the reference data; and S_1 is the set of points for the same anatomical structures from the second data set, called float data. Preregistration took place between these sets of points, and triangulation of S_1 was performed. The preregistration process was applied in three dimensions and was applied in order to realign the two data sets in all coordinates. After preregistration, two steps were performed to obtain the interpolant points:

1. Triangulating S_1 and producing a wire frame based on the topology of S_1; the triangulation was based on Feitzke's work [32] and was performed by defining an SOM with the following characteristics:

   a. A grid of neurons with 20 rows by 100 columns (20×100) was chosen for the specific implementation.

   b. The initial weighting vectors of the neurons of the grid were set equal to the coordinates of a set of points extracted from an enclosing surface, typically a cylindrical surface.

   c. The input to the neural network consisted of the Cartesian coordinates of the set of points to be triangulated.

   After the process of adaptation of the neural network, the weighting vectors of the neurons had values identical to the appropriate points of S_1. A wire frame consisting of one node for each neuron could be constructed, with Cartesian coordinates of each node equal to the weight vector of the corresponding neuron. The wire frame was triangulated according to the connectivity of the neurons.

2. Establishing an SOM in terms of the topology of S_1 and training the SOM by using S_2; the search for corresponding points was based on replicating the topology of the set S_1 on the input layer of an SOM model. In the SOM model, one neuron was allocated to each node of the wire frame and the connections between the neurons were identical to the connections of the wire frame. No connection between two neurons was

accepted when the two corresponding nodes were not directly connected in the float set. The initial weight vector of the neurons was the Cartesian coordinates of the corresponding wire frame nodes in the 3-D space.

The training of the network was realized by providing the network with the coordinates of randomly selected points sampled from the reference set S_2. The neuron with weight vector closest to signal was selected to fire. The firing neuron adjusted its weight vector, and its neighboring neurons modified their weight vectors as well but less strongly. The neighboring neurons were restricted to a window of 3×3 neurons during the network training.

The convergence of the SOM network during the triangulation of S_1 set of points leads to a triangulated subset of points (S_1'). Each node of subset S_1' corresponded to a neuron of the SOM network (20×100 neurons), whose initial weighting vector (wx_0, wy_0, wz_0) in S_1 was set to the initial Cartesian coordinates of this node. In S_1, this node was moved to new coordinates and equal to the final weighting vector (wx_1, wy_1, wz_1). The new position always coincided with a point in S_2.

Although SOM lateral interactions between neurons generated a one-to-one point correspondence, more than one point from S_1' might correspond to one point in S_2. However, most such point mismatches are avoided by using a distance threshold criterion to exclude corresponding points exceeding a distance of more than five voxels. With the help of this process, excessive deformation of the final warped image was also prohibited. Therefore, the total number of successful corresponding points was cut down to approximately 500 pairs of points for all patient data [31].

SOM also has been used in many other applications. Shang, Lv, and Yi [33] developed an automatic method to register CT and MR brain images by using first principal directions of feature images. In the method, a PCA neural network was used to calculate the first principal directions from feature images, and then the registration was realized by aligning feature images' first principal directions and centroids.

Coppini, Diciotti, and Valli [34] presented a general approach to the problem of image matching that exploits a multiscale representation of local image structure. In the approach, a given pair of images to be matched were named target and stimulus, respectively, and were transformed by Gabor wavelets. Correspondence was calculated by exploiting the learning procedure of a neural network derived from Kohonen's SOM. The SOM neurons coincided with the pixels of the target image, and their weights were pointers to those in the stimulus images. The standard SOM rule was modified to account for image features.

Fatemizadeh, Lucas, and Soltanian-Zadeh [35] proposed a method for automatic landmark extraction from MR brain images. In the method, landmark was extracted by modifying growing neural gas (GNG), which was a neural network–based cluster-seeking algorithm. Using the modified GNG (a splitting–merging SOM), corresponding dominant points of contours

extracted from two corresponding images are found. The contours were the boundaries of the regions generated by segmenting the MR brain image.

Di Bona and Salvetti [36] developed the volume-matcher 3-D project, an approach for a data-driven comparison and registration of 3-D images. The approach was based on a neural network model derived from self-organizing maps and extended to match a full 3-D data set of a source volume with the 3-D data set of a target volume.

These applications suggest that SOM is a promising algorithm for elastic registration, which is probably due to its clustering characteristics.

1.6 Other Applications with Neural Networks

In addition to those mentioned previously, ANN has been applied to other relevant areas such as medical image compression, enhancement, and restoration. In image compression [37,38], medical images such as mammograms are usually quite large in size and are stored in databases inside hospitals, which causes some difficulties in image transfer over the Internet or intranet. Some researchers applied ANN to existing compression algorithms to select interesting regions transmission for transmission or reduce the errors during the quantization in compression [40–43,47].

Panagiotidis and others [39] proposed a neural network architecture to perform lossy compression on medical images. To achieve higher compression ratio while retaining the significant (from a medical viewpoint) image content, the neural architecture adaptively selected ROI in the images.

Karlik [40] presented a combined technique for image compression based on the hierarchical finite state vector quantization and neural networks. The algorithm performed nonlinear restoration of diffraction-limited images concurrently with quantization. The neural network was trained on image pairs consisting of a lossless compression algorithm named hierarchical vector quantization.

Meyer-Bäse and colleagues [41] developed a method based on topology-preserving neural networks to implement vector quantization for medical image compression. The method could be applied to larger image blocks and represented better probability distribution estimation methods. A "neural-gas" network for vector quantization converged quickly and reached a distortion error lower than that from Kohonen's feature map. The influence of the neural compression method on the phantom features and the mammograms was not visually perceptible up to a high compression rate.

Jaiswal and Gaikwad [42] trained a resilient backpropagation neural network to encode and decode the input data so that the resulting difference between input and output images was minimized. Lo, Li, and Freedman [43] developed a neural network–based framework to search for an optimal wavelet kernel that could be used for a specific image processing task. In the algorithm,

a linear convolution neural network was applied to seek a wavelet that minimized errors and maximized compression efficiency for an image or a defined image pattern such as microcalcifications in mammograms and bone in CT head images.

To enhance original images, ANN has been used to suppress unwanted signals such as noise and tissues affecting cancerous signs. Suzuki and others [44] proposed an analysis method that makes clear the characteristics of the trained nonlinear filter, which is based on multilayer neural networks, and developed an approximate filter that achieves very similar results but was computational cost-efficient.

To detect lung nodules overlapped with ribs or clavicles in chest radiographs, Suzuki and colleagues [45] developed an image-processing technique for suppressing the contrast of ribs and clavicles in chest radiographs by means of a multiresolution massive training artificial neural network (MTANN). The structure of this neural network is illustrated in Figure 1.6, in which "bone" images are obtained by use of a dual-energy subtraction technique [46] as the teaching images to facilitate the neural network training. After that, the multiresolution MTANN was able to provide "bone-image-like" images that were similar to the teaching bone images. By subtracting the bone-image-like images from the corresponding chest radiographs, they were able to produce "soft-tissue-image-like" images where ribs and clavicles were substantially suppressed.

The MTANN consists of a linear-output multilayer ANN model, which was capable of operating on image data directly. The linear-output multilayer ANN model employed a linear function as the transfer function in the output layer because the characteristics of an ANN were improved significantly with a linear function when applied to the continuous mapping of values in image processing [47]. The inputs of the MTANN are the pixel values in a size-fixed subimage and can be written as

$$\vec{I}_{x,y} = \{I_1, I_2, \ldots, I_N\} \tag{1.5}$$

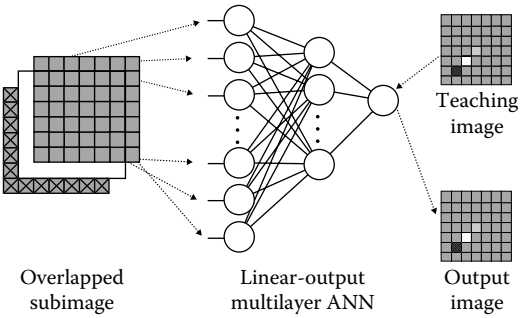


FIGURE 1.6: Architecture of MTANN.

where N is the number of inputs (i.e., the number of pixels inside a subimage). The output of the nth neuron in the hidden layer is represented by

$$O_n = f_h\left\{\sum_{m=1}^{N} w_{mn} \cdot I_m - b_n\right\} \tag{1.6}$$

where w_{mn} is a weight between the mth unit in the input layer and the nth neuron in the hidden layer, f_h is a sigmoid function, and b_n is an offset of the nth unit in the hidden layer. The output of the neuron in the output layer is represented by

$$f(x,y) = f_o\left\{\sum_{m=1}^{N_h} w_m^o \cdot O_m^H - b_o\right\} \tag{1.7}$$

where w_m^o is a weight between the mth neuron in the hidden layer and the neuron in the output layer, b_o is an offset of the neuron in the output layer, and f_o is a linear function.

To train MTANN, a dual-energy subtraction technique [48] was used to obtain the teaching image T (i.e., "bone" images) for suppression of ribs in chest radiographs. Input chest radiographs were divided pixel by pixel into a large number of overlapping subimages. Each subimage $I(x,y)$ corresponds to a pixel $T(x,y)$ in the teaching image, and the MTANN was trained with massive subimage pairs as defined in Equation 1.8:

$$\{I(x,y), T(x,y) \,|x,y \in R_T\} = \left\{\left(\vec{I}_1, T_1\right), (\vec{I}_2, T_2), \ldots, (\vec{I}_{N_T}, T_{N_T})\right\} \tag{1.8}$$

where R_T is a training region corresponding to the collection of the centers of subimages, and N_T is the number of pixels in R_T. After training, the MTANN is expected to produce images similar to the teaching images (i.e., bone-image-like images).

Since ribs in chest radiographs included various spatial–frequency components and it was difficult in practice to train the MTANN with a large subimage, multiresolution decomposition/composition techniques were employed in the algorithm. Three MTANNs for different resolutions were trained independently with the corresponding resolution images: a low-resolution MTANN was used for low-frequency components of ribs, a medium-resolution MTANN was used for medium-frequency components, and a high-resolution MTANN was used for high-frequency components. After training, the MTANNs produced a complete high-resolution image based on the images with different resolution [45].

Hainc and Kukal [49] found the ANN could also be employed as a kind of a sophisticated nonlinear filter on a local pixel neighborhood (3×3), since linear system sensitivity to impulse (isolated) noise was not good.

Chen, Chiueh, and Chen [50] introduced an ANN architecture for reducing the acoustic noise level in MRI processes. The proposed ANN consisted of

two cascaded time-delay ANNs. The ANN was employed as the predictor of a feedback active noise control (ANC) system for reducing acoustic noises. Preliminary results showed that with the proposed ANC system installed, acoustic MR noises were greatly attenuated, while verbal communication during MRI sessions was not affected.

Apart from compression and enhancement, ANN has been applied to medical image processing for other purposes. Wu [51] developed a new method to extract the patient information number field automatically from the film-scanned image using a multilayer cluster neural network. Cerveri and others [52] presented a hierarchical RBF network to correct geometric distortions in X-ray image intensifiers, which reduced the accuracy of image-guided procedures and quantitative image reconstructions.

Hsu and Tseng [53] established a method to predict and create a profile of bone defect surfaces by a well-trained 3-D orthogonal neural network. To train the neural network to team the scattering characteristic, the coordinates of the skeletal positions around the boundary of bone defects were input into the network. After the neural network had been well trained, the mathematic model of the bone defect surface was generated, and the pixel positions were obtained. The 3-D orthogonal neural network avoided local minima and converges rapidly.

It is difficult to generalize all these applications of ANN into several unified models. However, it might be possible to analyze the general pattern of applying ANNs. In Section 1.7, a comparison is made by studying the applications described in all previous sections.

1.7 Conclusions

As described in the previous five sections, applications of neural networks are classified into four major categories. These applications seem quite different from one another and cover many aspects of medical image processing. To summarize all the neural networks successfully applied to medical imaging, we highlight the comparisons of their application patterns, structures, operations, and training design in Table 1.1. Because there is no theory to indicate what is the best neural network structure for medical image processing and pattern recognition, the information such as type of network, type of input, number of inputs, neurons in hidden layers, and neurons in output is listed to help with searching and designing similar neural networks for future applications. Although these applications may come from different areas, such as CAD and segmentation, and inputs for neural networks are various, the essential purpose of applying these neural networks lies in their classifications, providing inspiring summary for existing modes of neural network applications and thus leading to further developments. Since the data sets for these applications are quite different, it is not possible to compare their results and the performance

TABLE 1.1: Comparative summary of feedforward neural network applications in medical imaging.

Application	Type of Network	Purpose	Type of Input	Number of Inputs	Neurons in Hidden Layers	Neurons in Output	Train/Test/Validation
[6]	CNN*(BP*)	Detect FP*	Intensity of pixels	256	14/10	1	268ROI*/267ROI
[8]	BP	Reduce FP	Value of features	5	5	1	1448 clusters/leave-one-out
[9]	MLP*	Reduce FP	Value of features	9	20/10	1	Unknown
[10]	RBFNN*	Classify tissues	Value of features	4	5	2	44 regions/54 images
[13]	BP	Detect FP	Value of features	11	9	1	100 images/100 images/Jackknife [54]
[15]	BP	Detect FP	Value of features	10	5	1	397ROI/397 ROI/Jackknife
[17]	Feedforward	Classify boundary	Coordinate/magnitude	3	30/10	1	100 images/147 images &
		Classify region	Coordinate/intensity	3	50	1	65-image CV
[19]	BP	Predict tissue	Value of features	8	5	1	262/leave-one-
				7	3	1	out/Jackknife
[20]	BP	Classify tissues	Value of features	3	10	1	60 primitives/983 primitives
[27]	MLP	Classify boundary	Intensity of pixels	49	30	1	1200 patterns/400 slices
[25]	BP	Classify tissues	Statistical indexes	3	Unknown	3	Small number, improved by interaction
[45]	BP	Remove noise	Intensity of pixels	25	20	1	Unknown
[46]	MTANN*(BP)	Classify tissues	Intensity of pixels	81	20	1	5000 regions/118 images
[55]	BP	Classify MC	Value of features	14	13	1	100 ROI/leave-one-out
[56]	MLBNN*(BP)	Classify MC	Vectors from SOM	5	25/14	7	32 cases/64 cases
[57]	BP	Classify tissues	Vectors from SOM	3	7	7	Unknown/80 images
[58]	BP	Detect edge	Intensity of pixels	121	20	1	24 images/fourfold CV

BP, backpropagation; CNN, convolution neural network; CV, cross-validation; FP, false positive MC or regions; MC, microcalcification cluster; MLBNN, multilayered BP neural network; MLP, multilayer perceptron; RBFNN, radial basis function neural network; ROI, region of interest; SOM, self-organizing map; MTANN, massive training artificial neural network.

of these algorithms. Some applications are ignored in the list because the details about their neural networks are limited. The total number of neurons needed in the hidden layers somewhat depends on the total number of training samples.

In contrast to feedforward neural network, the applications of feedback neural networks for medical image processing have been quite limited in the past decade, and most of them are in the area of image segmentation and are primarily based on Hopfield neural networks. The similarities among these applications are quite limited, but all of them need to minimize an energy function during convergence of the network. The energy function must be designed individually, which might affect its application in medical imaging. Because the Hopfield neural network is unsupervised, it may not work for CAD like the feedforward neural network, which requires a priori knowledge in classifications.

Although the applications of Kohonen's SOM are not as many as those of feedforward neural networks, its clustering and unsupervised properties make it very suitable for image registration. SOM converges to a solution that approximates its input data by adapting to prototype vectors. During this process, the relation of its neighborhood neurons is also taken into account, leading to preservation of topology and mapping of training sets. For the applications of image registration, the input vectors of the neurons in SOM usually contain the spatial coordinate and intensity of pixels. For applications in image compression, SOM is used as a topology-preserving feature map to generate vector quantization for code words. Sometimes, SOM produces the segmentation results for feedforward neural networks due to its unsupervised clustering property.

In summary, the applications of ANN in medical image processing have to be analyzed individually, although many successful models have been reported in the literature. ANN has been applied to medical images to deal with the issues that cannot be addressed by traditional image processing algorithms or by other classification techniques. By introducing ANNs, algorithms developed for medical image processing and analysis often become more intelligent than conventional techniques. While this chapter provided a focused survey on a range of neural networks and their applications to medical imaging, the main purpose here is to inspire further research and development of new applications and new concepts in exploiting neural networks.

Acknowledgment

The authors wish to acknowledge the financial support from the British Cancer Research Trust, through which this survey was carried out as part of a funded research project.

References

[1] Ripley, B. D. 1996. Feed-forward neural networks. In *Pattern Recognition and Neural Networks*, 143–80. Cambridge: Cambridge University Press.

[2] Wasserman, P. D. 1993. Backpropagation and beyond. In *Advanced Methods in Neural Computing*, 97–118. New York: Van Nostrand Reinhold.

[3] Wasserman, P. D. 1993. Radial basis function networks. In *Advanced Methods in Neural Computing*, 147–176. New York: Van Nostrand Reinhold.

[4] Carling, A. 1992. Recurrent networks. In *Introducing Neural Networks*, 189–94. Wilmslow: Sigma.

[5] Kohonen, T. 1990. The self-organizing map. *Proceedings of the IEEE* 78: 1464–80.

[6] Ge, J., B. Sahiner, L. M. Hadjiiski, H. P. Chan, J. Wei, M. A. Helvie, and C. Zhou. 2006. Computer-aided detection of clusters of microcalcifications on full field digital mammograms. *Medical Physics* 33: 2975–88.

[7] Lo, S. C. B., H. P. Chan, J. S. Lin, H. Li, M. T. Freedman, and S. K. Mun. 1995. Artificial convolution neural network for medical image pattern recognition. *Neural Networks* 8: 1201–14.

[8] Nagel, R. H., R. M. Nishikawa, J. Papaioannou, and K. Doi. 1998. Analysis of methods for reducing false positives in the automated detection of clustered microcalcifications in mammograms. *Medical Physics* 25: 1502–6.

[9] Papadopoulossa, A., D. I. Fotiadisb, and A. Likasb. 2002. An automatic microcalcification detection system based on a hybrid neural network classifier. *Artificial Intelligence in Medicine* 25: 149–67.

[10] Christoyiani, I., E. Dermatas, and G. Kokkinakis. 2000. Fast detection of masses in computer-aided mammography. *IEEE Signal Processing Magazine* 17: 54–64.

[11] Patrocinio, A. C., H. Schiabel, R. H. Benatti, C. E. Goes, and F. L. S. Nunes. 2000. Investigation of clustered microcalcification features for an automated classifier as part of a mammography CAD scheme. *Proceedings of the 22nd Annual EMBS International Conference* 1203–5.

[12] Setiono, R. 1996. Extracting rules from pruned neural networks for breast cancer diagnosis. *Artificial Intelligence in Medicine* 8: 37–51.

[13] Xu, X. W., K. Doi, T. Kobayashi, H. MacMahon, and M. L. Giger. 1997. Development of an improved CAD scheme for automated detection of lung nodules in digital chest images. *Medical Physics* 24: 1395–1403.

[14] Zhou, Z. H., Y. Jiang, Y. B. Yang, and S. F. Chen. 2002. Lung cancer cell identification based on artificial neural network ensembles. *Artificial Intelligence in Medicine* 24: 25–36.

[15] Keserci, B., and H. Yoshida. 2002. Computerized detection of pulmonary nodules in chest radiographs based on morphological features and wavelet snake model. *Medical Image Analysis* 6: 431–47.

[16] Qian, W., T. Zhukov, D. S. Song, and M. S. Tockman. 2007. Computerized analysis of cellular features and biomarkers for cytologic diagnosis of early lung cancer. *Analytical and Quantitative Cytology and Histology* 29: 103–11.

[17] Coppini, G., S. Diciotti, M. Falchini, N. Villari, and G. Valli. 2003. Neural networks for computer-aided diagnosis: Detection of lung nodules in chest radiograms. *IEEE Transactions on Information Technology in Biomedicine* 7: 344–57.

[18] Mohamed, E. I., C. Maiolo, R. Linder, S. J. Poppl, and A. De Lorenzo. 2003. Artificial neural network analysis: A novel application for predicting site-specific bone mineral density. *Acta Diabetologica* 40: S19–22.

[19] Scott, J. A. 2004. The lung scan and the abnormal chest X-ray: difficult diagnoses. *Nuclear Medicine Communications* 25: 1137–41.

[20] Kobashi, S., N. Kamiura, Y. Hata, and F. Miyawaki. 2001. Volume-quantization-based neural network approach to 3D MR angiography image segmentation. *Image and Vision Computing* 19: 185–93.

[21] Middleton, I., and R. I. Damper. 2004. Segmentation of magnetic resonance images using a combination of neural networks and active contour models. *Medical Engineering & Physics* 26: 71–86.

[22] Lin, J. S. 2000. Segmentation of medical images through a penalized fuzzy Hopfield network with moments preservation. *Journal of the Chinese Institute of Engineers* 23: 633–43.

[23] Lin, K. C. R., M. S. Yang, H. C. Liu, J. F. Lirng, and P. N. Wang. 2003. Generalized Kohonen's competitive learning algorithms for ophthalmological MR image segmentation. *Magnetic Resonance Imaging* 21: 863–70.

[24] Dokur, Z., and T. Olmez, 2003. Segmentation of MR and CT images by using a quantiser neural network. *Neural Computing & Applications* 11: 168–77.

[25] Stalidis, G., N. Maglaveras, S. N. Efstratiadis, A. S. Dimitriadis, and C. Pappas. 2002. Model-based processing scheme for quantitative 4-D cardiac MRI analysis. *IEEE Transactions on Information Technology in Biomedicine* 6: 59–72.

[26] Chang, C. L., and Y. T. Ching. 2002. Fuzzy Hopfield neural network with fixed weight for medical image segmentation. *Optical Engineering* 41: 351–58.

[27] Shen, S., W. Sandham, M. Granat, and A. Sterr. 2005. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine* 9: 459–67.

[28] Chang, C. Y., and P. C. Chung. 2000. Two-layer competitive based Hopfield neural network for medical image edge detection. *Optical Engineering* 39: 695–703.

[29] Chang, C. Y., and P. C. Chung. 2001. Medical image segmentation using a contextual-constraint-based Hopfield neural cube. *Image and Vision Computing* 19: 669–78.

[30] Chang, C. Y. 2006. Contextual-based Hopfield neural network for medical image edge detection. *Optical Engineering* 45: Art. No. 037006.

[31] Matsopoulosa, G. K., N. A. Mouravlianskya, P. A. Asvestasa, K. K. Delibasisa, and V. Koulouliasb. 2005. Thoracic non-rigid registration combining self-organizing maps and radial basis functions. *Medical Image Analysis* 9: 237–54.

[32] Fritzke, B. 1994. A growing neural gas network learns topologies. In G. Tesauro, D. S. Touretzky, and T. L. Leen, eds., *Advances in Neural Information Processing Systems*, 7th ed., 625–32. Cambridge, MA: MIT Press.

[33] Shang, L. F., J. C. Lv, and Z. Yi. 2006. Rigid medical image registration using PCA neural network. *Neurocomputing* 69: 1717–22.

[34] Coppini, G., S. Diciotti, and G. Valli. 2004. Matching of medical images by self-organizing neural networks. *Pattern Recognition Letters* 25: 341–52.

[35] Fatemizadeh, E., C. Lucas, and H. Soltanian-Zadeh. 2003. Automatic landmark extraction from image data using modified growing neural gas network. *IEEE Transactions on Information Technology in Biomedicine* 7: 77–85.

[36] Di Bona, S., and O. Salvetti. 2002. Neural method for three-dimensional image matching. *Journal of Electronic Imaging* 11: 497–506.

[37] Bovik, A. C. 2005. *Handbook of Image and Video Processing.* London: Academic Press.

[38] Jiang, J. 1999. Image compression with neural networks: A survey. *Signal Processing—Image Communication* 14: 737–60.

[39] Panagiotidis, N. G., D. Kalogeras, S. D. Kollias, and A. Stafylopatis. 1996. Neural network-assisted effective lossy compression of medical images. *Proceedings of the IEEE* 84: 1474–87.

[40] Karlik, B. 2006. Medical image compression by using vector quantization neural network (VQNN). *Neural Network World* 16: 341–48.

[41] Meyer-Bäse, A., K. Jancke, A. Wismüller, S. Foo, and T. Martinetz. 2005. Medical image compression using topology-preserving neural networks. *Engineering Applications of Artificial Intelligence* 18: 383–92.

[42] Jaiswal, R. R., and A. N. Gaikwad. 2006. Neural network assisted effective lossy compression of medical images. *IETE Technical Review* 23: 119–26.

[43] Lo, S. C. B., H. Li, and M. T. Freedman. 2003. Optimization of wavelet decomposition for image compression and feature preservation. *IEEE Transactions on Medical Imaging* 22: 1141–51.

[44] Suzuki, K., I. Horiba, and N. Sugie. 2002. Efficient approximation of neural filters for removing quantum noise from images. *IEEE Transactions on Signal Processing* 50: 1787–99.

[45] Suzuki, K., H. Abe, H. MacMahon, and K. Doi. 2006. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Transactions on Medical Imaging* 25: 406–16.

[46] Suzuki, K., I. Horiba, K. Ikegaya, and M. Nanki. 1995. Recognition of coronary arterial stenosis using neural network on DSA system. *Systems and Computers in Japan* 26: 66–74.

[47] Suzuki, K. 2003. Neural edge enhancer for supervised edge enhancement from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25: 1582–96.

[48] Ishigaki, T., S. Sakuma, Y. Horikawa, M. Ikeda, and H. Yamaguchi. 1986. One-shot dual energy subtraction imaging. *Radiology* 161: 271–73.

[49] Hainc, L., and J. Kukal. 2006. Role of robust processing in ANN de-noising of 2D image. *Neural Network World* 16: 163–76.

[50] Chen, G. K., T. D. Chiueh, and J. H. Chen. 1999. Active cancellation system of acoustic noise in MR imaging. *IEEE Transactions on Biomedical Engineering* 46: 186–91.

[51] Wu, H. H. P. 2004. Patient information extraction in digitized X-ray imagery. *Image and Vision Computing* 22: 215–26.

[52] Cerveri, P., C. Forlani, N. A. Borghese, and G. Ferrigno. 2002. Distortion correction for x-ray image intensifiers: Local unwarping polynomials and RBF neural networks. *Medical Physics* 29: 1759–71.

[53] Hsu, J. H., and C. S. Tseng. 2001. Application of three-dimensional orthogonal neural network to craniomaxillary reconstruction. *Computerized Medical Imaging and Graphics* 25: 477–82.

[54] Dorfman, D. D., K. S. Berbaum, and C. E. Metz. 1992. Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method. *Investigative Radiology* 27: 723–31.

[55] Kallergi, M. 2004. Computer-aided diagnosis of mammographic microcalcification clusters. *Medical Physics* 31: 314–26.

[56] Glass, J. O., W. E. Reddick, O. Goloubeva, V. Yo, and R. G. Steen. 2000. Hybrid artificial neural network segmentation of precise and accurate inversion recovery (PAIR) images from normal human brain. *Magnetic Resonance Imaging* 18: 1245–53.

[57] Reddick, W. E., R. K. Mulhern, T. D. Elkin, J. O. Glass, T. E. Merchant, and J. W. Langston. 1998. A hybrid neural network analysis of subtle brain volume differences in children surviving brain tumors. *Magnetic Resonance Imaging* 16: 413–21.

[58] Suzuki, K., I. Horiba, N. Sugie, and M. Nanki. 2004. Extraction of left ventricular contours from left ventriculograms by means of a neural edge detector. *IEEE Transactions on Medical Imaging* 23: 330–39.

Chapter 2

Evolutionary Computing and Its Use in Medical Imaging

Lars Nolle and Gerald Schaefer

Contents

Many scientific problems can be formulated as optimization problems. Among the many classes of algorithms for solving such problems, one interesting, biologically inspired group is that of evolutionary optimization techniques. In this chapter, we provide an overview of such techniques, in particular of genetic algorithms and genetic programming and its related subtasks of selection, crossover, mutation, and coding. We then explore some applications of genetic techniques in the context of medical imaging.

2.1 Evolutionary Computing

Many scientific problems can be viewed as search or optimization problems, where an optimum input parameter vector for a given system has to be found in order to maximize or to minimize the system response to that input vector. Often, auxiliary information about the system, like its transfer function and derivatives, is not known, and the measures might be incomplete and distorted by noise. This makes such problems difficult to solve by traditional mathematical methods. Evolutionary optimization algorithms, which are based on biological principles borrowed from nature, can offer a solution. These algorithms work on a population of candidate solutions, which are iteratively improved so that an optimal solution evolves over time.

This chapter discusses the general problem of search and optimization before it introduces the system's view, followed by a definition of search space and fitness landscape. It then explains the process of optimization and the concept of optimization loops. It introduces biologically inspired evolutionary optimization algorithms: genetic algorithms and genetic programming. Finally, it provides an overview of same sample applications of evolutionary approaches in medical imaging.

2.1.1 Systems

Every process or object can be seen as a system. Fenton and Hill (1993) define a system as "an assembly of components, connected together in an organised way, and separated from its environment by a boundary. This organised assembly has an observable purpose which is characterised in terms of how it transforms input from the environment into output to the environment." By definition, a system has exactly one input channel x and exactly one output channel y (see Figure 2.1). All interactions with the environment must be made through these interfaces.

Both input and output can be vectors or scalars. The input is called the *independent variable* or *parameter*, because its value(s) can be chosen freely, and it results in the output y, the so-called *dependent variable*. If the present state of the system does not depend on previous states but only on the current input, the system is said to be a *steady-state system*, and the output of the system can be described as a function of the input $y = f(x)$.

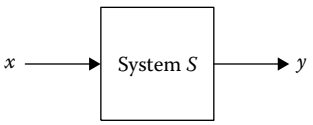


FIGURE 2.1: Generic system.

2.1.2 Objective function

In order to rate the quality of a candidate solution x, it is necessary to transform the system response to x into an appropriate measure, called the *objective* or *fitness*. If the system has only one output variable, the system output y equals the fitness. If y has more than one component, the output variables of the system have to be combined into a single value, computed by the *objective function* or *fitness function*. In general, there are four approaches to judge the system output: aggregation, the changing objectives method, the use of niche techniques, and Pareto-based methods (Fonseca and Fleming, 1995). The most often used method is aggregation. In its simplest case, the fitness function $F(x)$ equals the weighted sum of the components $y_i = c_i \cdot F_i(x)$ of y, where c_i is the weight for component i:

$$F(x) = c_0 + c_1 \cdot F_1(x) + \cdots + c_n \cdot F_n(x) \tag{2.1}$$

2.1.3 Search space and fitness landscape

If all the possible candidate solutions are collected in an ordered way, this collection is called the search space and sometimes the input space. For an optimization problem of dimension n, that is, a system with n independent parameters, the search space also has dimension n. Adding the dimension *fitness* or *costs* to the search space results in the $(n + 1)$ dimensional fitness landscape (Wright, 1931); see Figure 2.2.

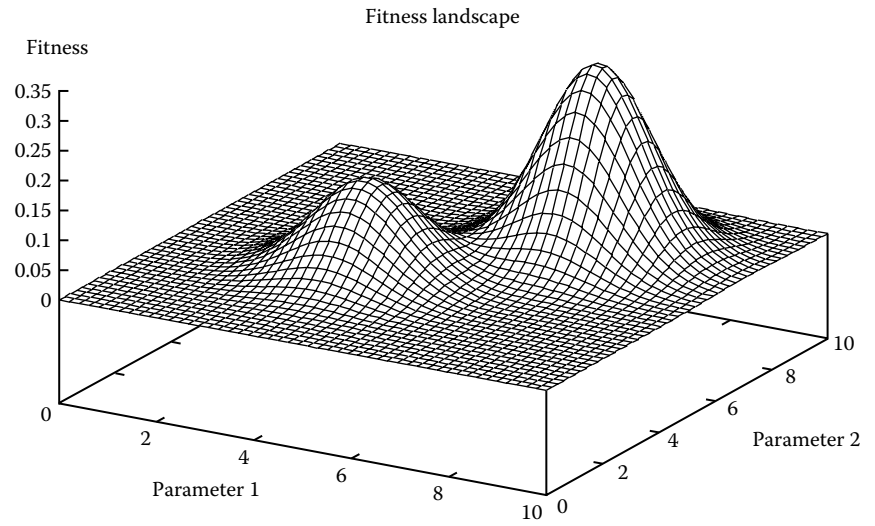


FIGURE 2.2: Example of a fitness landscape for a system with two input parameters.

2.1.4 Optimization

Optimization (Schwefel, 1995) is the process of selecting the best candidate solution from a range of possibilities (i.e., from the search space). In other words, a system S that has to be optimized in terms of a quality output value y is brought into a new state that has a better quality output value y than the previous state. This is done by changing the independent input parameters x. The error function describes the difference between the predefined objective $y_{desired}$ and system response $f(x)$ to the input x.

$$Error(x) = y_{desired} - f(x) \tag{2.2}$$

Usually, the aim is to find the vector x' that leads to a minimal error for the system S, that is, the minimal departure from the optimal output value:

$$Error(x') = 0 \tag{2.3}$$

Often, a predefined target value is not known. In this case, one tries to gain a fitness value that is as high as possible in the case of maximization or as low as possible in the case of minimization. Ideally, one would evaluate all possible candidates and choose the best one. This is known as exhaustive search. However, often it is not feasible to consider all possible solutions, for example, if the search space is too large and the evaluation of a single candidate is too expensive. In such cases, only a subset of the solutions can be evaluated.

Optimization problems can be either function optimization problems or combinatorial problems. The first class of problems can be divided into continuous optimization and discrete optimization problems. In continuous function optimization, the independent variables are real numbers, whereas for discrete function optimization, the independent variables can be chosen only from a predefined set of allowed and somehow ordered numbers, such as $\{10, 20, 30, 40\}$.

In combinatorial optimization problems, the optimum sequence or combination of a fixed set of input values has to be found. Here, the input values are symbols and might not be connected or ordered, for example {apple, orange, strawberry}. An example of a combinatorial optimization problem is the classical traveling salesman problem (TSP), where a sales agent needs to visit a predefined set of cities and return to base. The problem here is to find an optimal route that connects all cities while having the shortest travel distance by choosing the order in which the cities are visited.

2.1.5 Optimization loop

Mathematical or calculus-based methods use known functional relationships among variables and objectives to calculate the optimum of the given system. Therefore, an exact mathematical model of the process must exist. Edelbaum (1962) introduced the differentiation of calculus-based methods in direct methods and indirect methods.

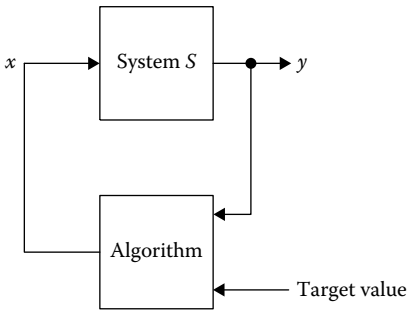


FIGURE 2.3: Closed optimization loop consisting of a system and an optimization algorithm.

Direct methods solve the optimization problem by iterative calculation and derivation of the error function and moving in a direction to the maximum slope gradient. Indirect methods solve the optimization problem in one step—without testing—by solving a set of equations (usually nonlinear). These equations result from setting the derivative of the error function equal to zero.

Both classes of methods are local in scope: they tend to find only local optima. Therefore, they are not robust. They depend on the existence of derivatives. Real problem functions tend to be perturbed by noise and are not smooth (i.e., derivations may not exist for all points of functions). This class of problem cannot be solved by mathematical methods.

If the functional relations among input variables and objectives are not known, one can experiment on the real system (or a model of this system) in order to find the optimum. Access to the independent variables must exist for the whole multidimensional search space—the collection of all possible candidate solutions. Also, a possibility of measuring the independent variable and the objective must be given. The optimization process is iterative; that is, it has to be done in a closed optimization loop (Figure 2.3).

Experimental optimization methods can therefore be seen as a search for the optimum by traversing over the fitness landscape.

2.2 Genetic Algorithms

As Darwin's theory of natural selection articulates, nature is very effective at optimization (e.g., enabling life forms to survive in a unfriendly and changing environment by means of simple trial and error). Genetic algorithms (GAs) simulate this evolutionary mechanism by using heredity and mutation. They were first introduced by Holland (1975), who also provided a theoretical framework for genetic algorithms, the schemata theorem (Goldberg, 1989).

For genetic algorithms, the independent input parameters of a system S (Figure 2.4) are coded into a binary string, the *genotype* of an individual (Figure 2.5).

The individual represented by genotype is called a *phenotype*. This phenotype has a certain quality or fitness to survive, which can be determined by presenting the phenotype to the system S and measuring the system response.

The search is undertaken not only by one individual but by a population of n genotypes, the *genepool* (Figure 2.6). Therefore, the search space is tested at n points in parallel. All the individuals of the genepool at a time t_n are called a *generation*.

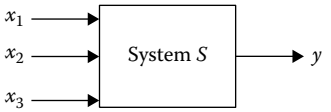


FIGURE 2.4: System to be optimized.

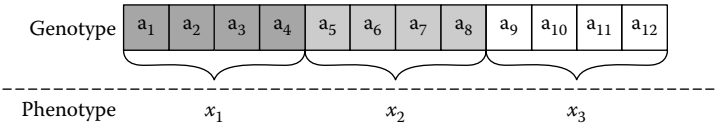


FIGURE 2.5: Binary string representing one input pattern of the system.

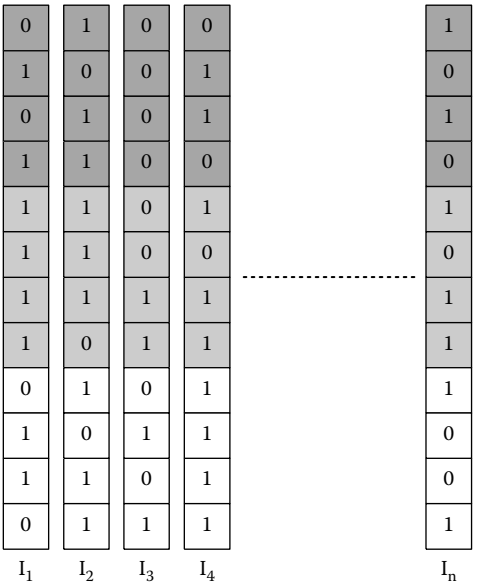


FIGURE 2.6: Genepool consisting of individuals $I_1 \ldots I_n$.

A new generation for time t_{n+1} is generated by selecting N individuals from the current population for breeding. They are copied into the genepool of the next generation, and their genetic information is then recombined, using the *crossover* operator (see Section 2.2.2), with a predefined crossover probability p_c. The resulting offspring is then copied into the new genepool, and mutation is applied to the offspring. Figure 2.7 shows the flowchart of a simple genetic algorithm.

The search is carried out until at least one individual has a better fitness than the defined minimum fitness or a maximum number of generations has been reached.

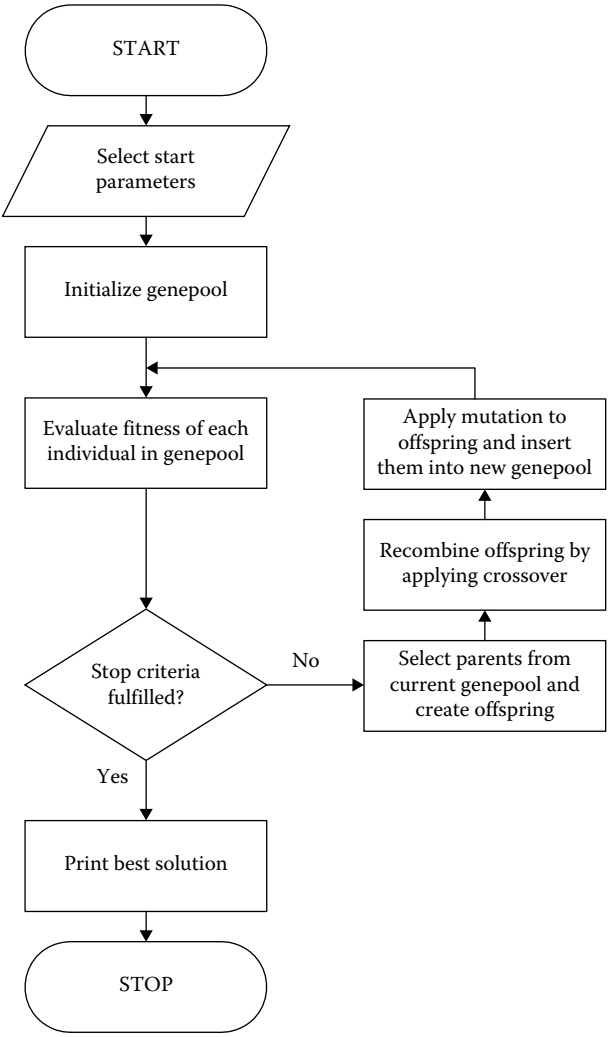


FIGURE 2.7: Flowchart of basic GA algorithm.

2.2.1 Selection

In general, there are three approaches to choose individuals from the current generation for reproduction: tournament selection, fitness proportional selection, and rank-based selection. In tournament selection, two or more individuals are randomly selected from the current generation of N genotypes to compete with each other. The individual with the highest fitness of this set is the winner and is selected for generating offspring. The process is repeated N times in order to create the new population. Using tournament selection, the least fit individual can never be selected.

In fitness proportional selection, the chance of an individual to be selected is related to its fitness value. The most commonly used method of this type is roulette wheel selection in which proportions of an imaginary roulette wheel are distributed in proportion to the relative fitness of an individual. Figure 2.8 shows an example for $N = 3$. In this example, the fitness of individual 3 is approximately four times higher than the fitness of individual 1, which means its chance of selection is four times greater than that of individual 1. For a population of N individuals, the wheel is spun N times, and the individual under the pointer is selected. In fitness proportional selection, all individuals have a chance of selection, but high-fitness individuals are more likely to be selected because they occupy a larger portion of the wheel.

However, there is the statistical chance that the actual selected distribution might not reflect the expected distribution based on the fitness values. If the selection is too strong, it can lead to premature convergence: the population would converge before it found the region of the search space that contains the global optimum. In other words, the exploitation would start before the search space is fully explored. On the other hand, if the selection is too weak, it can lead to stalled evolution, which means the search is reduced to randomly walking through search space.

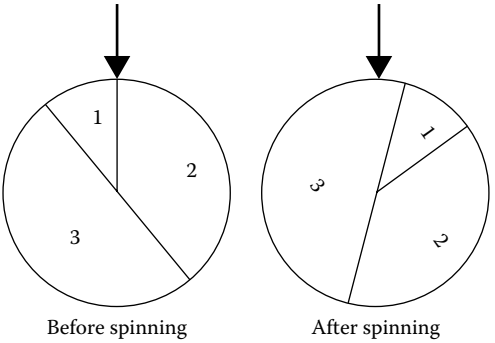


FIGURE 2.8: Roulette wheel selection.

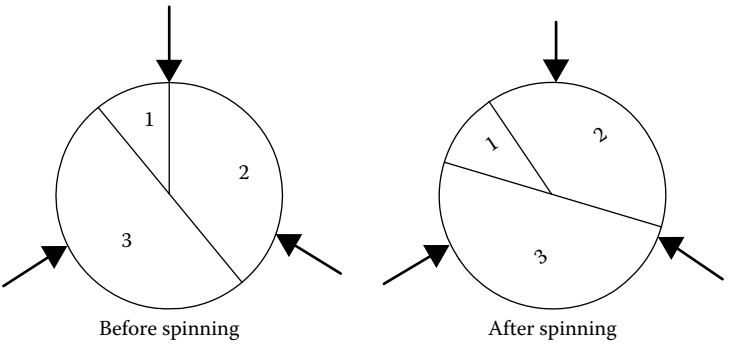


FIGURE 2.9: SUS selection.

These effects are overcome using stochastic universal selection (SUS). Here, the same roulette wheel is used, but instead of using a single pointer, N equally spaced pointers are used for a population of N individuals, and the wheel is spun only once (Figure 2.9).

Instead of using the fitness of an individual for selection, a selective s value can be used, which is based on the rank position of an individual in the population (Equation 2.4).

$$s_i = Min + (Max - Min)\frac{rank_i - 1}{N - 1} \tag{2.4}$$

where

Min: minimum fitness within a generation
Max: maximum fitness within a generation
$rank_i$: rank of individual i within the population in a generation
N: number of individuals within population

So, instead of using the raw fitness to determine the proportion for an individual, the rank of the individual within the generation is used.

Sometimes the m fittest individuals in a generation are cloned into the next generation in order to preserve their genetic material. This is known as elitism.

2.2.2 Crossover

The most important operator in terms of robustness of the algorithm is the crossover operator. Figure 2.10 shows the one-point crossover operator, which combines the information of two parents. They are aligned and then both cut at a randomly chosen crossover point, and the tails are swapped successively.

Instead of a single crossover point, two or more random crossover points can be used for recombining the genetic information of the parents.

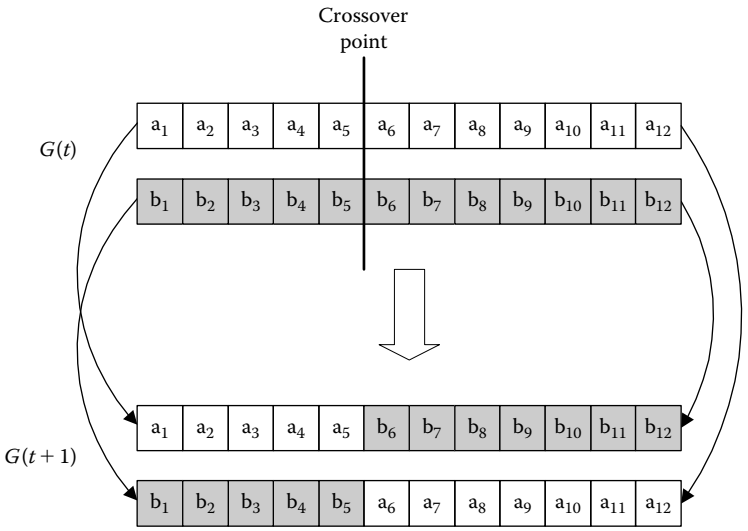


FIGURE 2.10: Crossover operator.

Another form of crossover is called uniform crossover (Syswerda, 1989). Here, every component of a parent individual X is randomly passed on either to offspring A or offspring B. If X passes on its component to A, the position in B is filled using the component from parent Y, and vice versa.

2.2.3 Mutation

After the genetic information of the parents is recombined using crossover, mutation is applied to every individual of the new generation. Every bit of the offspring is inverted (mutated) with probability p_m. The mutation operator is important for restoring lost information and producing a better effectiveness of the genetic algorithm.

2.2.4 Discussion

The advantages of genetic algorithms are that they use payoff (objective function) information, not derivatives or other auxiliary knowledge; that is, they are black-box optimization methods. Genetic algorithms tend to converge toward the global optimum rather than getting stuck in a local optimum, and therefore they are very robust. On the other hand, it is not always straightforward to find the right GA parameters for a particular optimization problem, such as a suitable genepool size or mutation probability. Also, the efficiency of genetic algorithms relies heavily on the right coding of the input parameters (i.e., the chosen mapping function from phenotype to genotype), and they tend to fail if the inputs of the system are heavily correlated.

2.2.5 Schemata theorem

Holland provided a theoretical foundation of genetic algorithms—a theoretical proof of convergence—which he called the schemata theorem. A schema is a template for binary strings, but built from a three-letter alphabet containing the symbols *, 0, and 1. The * symbol is the "don't care" symbol, which stands for either 0 or 1. Figure 2.11 shows an example of a schema for chromosomes consisting of 12 bits, of which 3 are set to the don't care symbol and the remaining 9 bits are set to fixed values.

The distance between the first and the last fixed bit is called the defined length of the schema, and the number of fixed bits is called the order of the schema. Figure 2.12 shows an example of a schema H and the different instances it represents.

A binary string s is an instance of a schema H if it fits into the template. Therefore, any binary string of length l does not represent just one candidate solution; it is simultaneously an instance of 2^1 schemata. Consequently, a genetic algorithm with the genepool of size n tests not only n different solutions but also a high number of different schemata at the same time. This is known as implicit parallelism in genetic algorithm and provides an explanation for their effectiveness and efficiency.

According to Holland, the number of instances m of a schema H that are contained in the population at generation $t+1$ can be determined as follows:

$$m(H, t+1) = m(H, t) \cdot \frac{\bar{f}(H)}{\bar{f}} \tag{2.5}$$

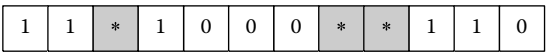


FIGURE 2.11: Example of a schema in GA.

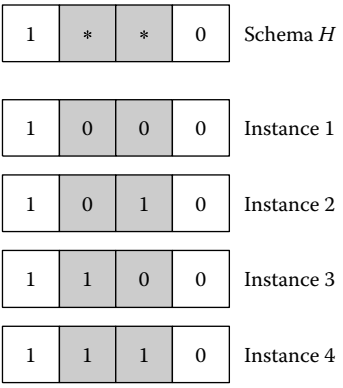


FIGURE 2.12: Example of a schema H and the instance it represents.

where:

H: schema or "building block" with at least one instance in the last generation
$m(H,t)$: number of instances of H at time t
$m(H,t+1)$: number of instances of H at time $t+1$
$\bar{f}(H)$: average fitness of the instances of schema H
$\bar{f}$: average fitness of the whole population

This is a simplified version of the schemata theorem because it does not take into account the effects of the crossover and the mutation operator. However, it is sufficient to demonstrate the basic idea. A more detailed description can be found, for example, in Goldberg (1989).

Suppose that a particular schemata H remains above average an amount $c \cdot \bar{f}$ with c being a constant factor. Equation 2.4 can be rewritten as follows:

$$m(H,t+1) = m(H,t) \cdot \frac{\bar{f} + c \cdot \bar{f}}{\bar{f}} = (1+c) \cdot m(H,t) \tag{2.6}$$

Assuming c is stationary and starts at $t = 0$, Equation 2.5 can be rewritten as follows:

$$m(H,t) = m(H,0) \cdot (1+c)^t \tag{2.7}$$

It can be seen that this equation is similar to the formula of interest: the number of instances of a schema H with a fitness above average grows exponentially to generation t. Hence, schemata with good fitness will survive, and those with a fitness below average will eventually die out. Therefore, the fitter building blocks (i.e., the better partial solution) will take over the genepool within finite time. However, the schemata theorem is controversial because it assumes that the factor c is constant over time.

2.2.6 Coding problem

Traditionally, genetic algorithms use binary strings. However, if an input variable is coded using standard binary coding, this can lead to the problem that a small change in the phenotype would require a large number of bits of the genotype to be inverted. An example of the coding problem is given in Figure 2.13.

As can be seen from the figure, a step from 3_{10} to 4_{10} requires flipping 3 bits in binary representation, whereas it changes the least significant digit in decimal representation by only one. One solution is to use gray code (see Figure 2.14), which has the advantage that only one bit changes between any two positions; it has a constant Hamming distance of one.

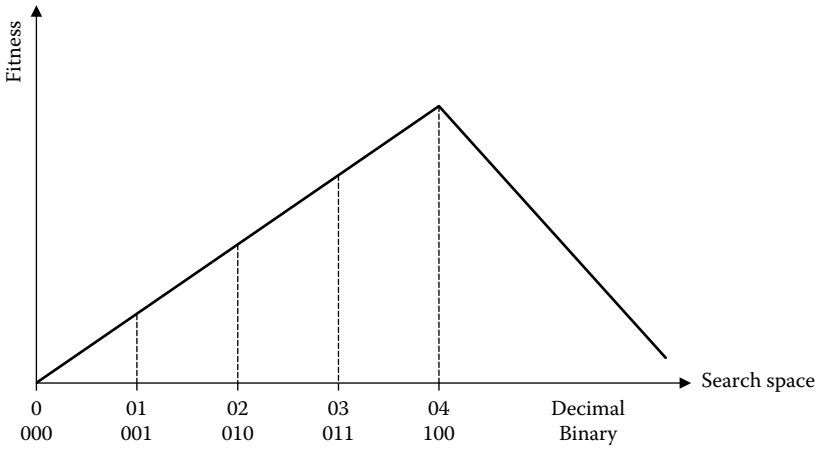


FIGURE 2.13: Differences between decimal and standard binary code.

Decimal	Binary	Gray code
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111

FIGURE 2.14: Gray code.

2.3 Genetic Programming

Genetic programming (GP) was introduced by Koza (1992) and is a machine learning technique that uses a genetic algorithm for the automated generation of computer programs. These programs model a system using sample data provided by that system. A typical application of genetic programming is, for example, symbolic regression.

In genetic programming, the programs are represented as tree structures in which a node represents an operator and a leaf represents an operand. Often, GP programs are coded in computer languages like List Processing (LISP) because they can be used straightforwardly to implement treelike structures.

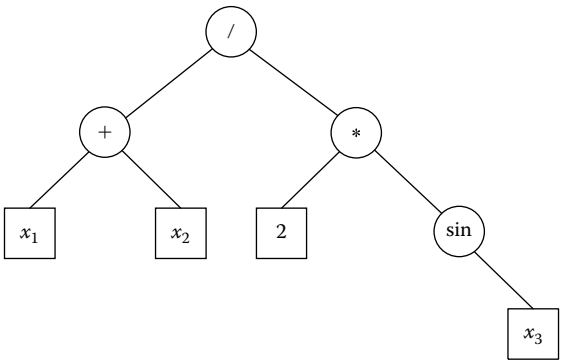


FIGURE 2.15: Example of a program represented by a tree.

The operators are chosen for a problem-specific function set, and the operands are chosen from a problem-specific terminal set. Figure 2.15 shows an example of a tree representing a program to calculate $y = f(x_1, x_2, x_3) = (x_1 + x_2)/2 * \sin(x_3)$.

The function set depends on the problem at hand. It may contain unary or binary operators and has to be chosen so that the function to be learned can be approximated as accurately as possible. The terminal set contains variables and constants and is again problem dependent. In the example above, the function set contains at least the operators $+$, $/$, $*$, and sin, and the terminal set contains at least the variables x_1, x_2, x_3 and the constant *2*. The quality of a program can be evaluated by applying the training data to the program and measuring either the average (often quadratic) error or by counting how many instances of the training set are reproduced correctly by the program.

Genetic programming works similarly to genetic algorithms. It works on a population of programs, the genepool. The individuals (i.e., the programs) are randomly created and then evaluated. The algorithm applies genetic operators to the population in each generation. These operators are selection, crossover, and mutation.

2.3.1 Selection

All the different selection methods from the GA domain (e.g., fitness proportional selection, tournament selection) can be employed in genetic programming.

2.3.2 Crossover

For each of the selected parents, a random node is chosen to be the crossover point. Figure 2.16 shows two individuals before crossover. The black-circled nodes are the randomly selected crossover points.

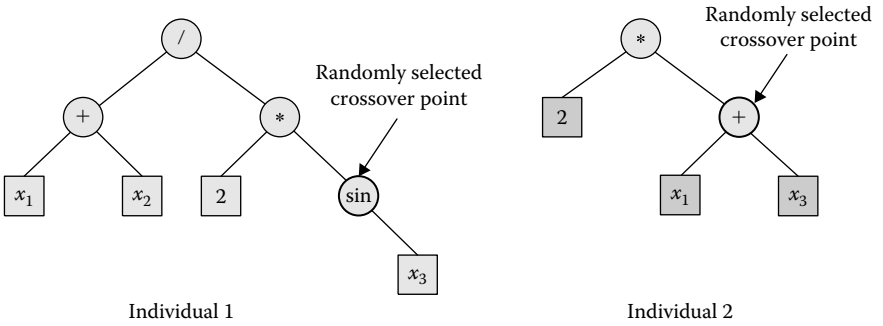


FIGURE 2.16: Individuals before crossover.

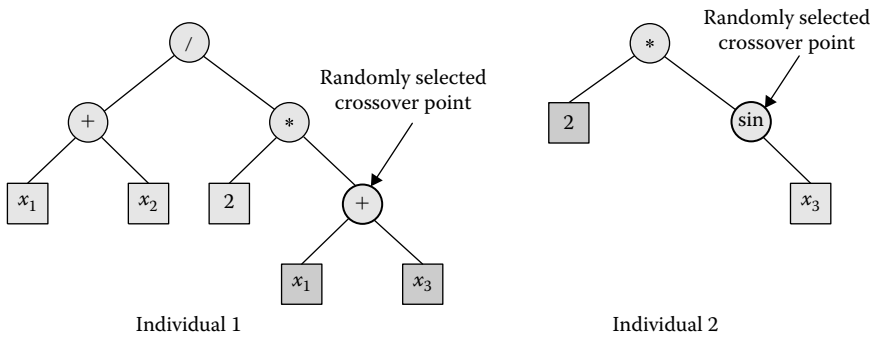


FIGURE 2.17: Individuals after crossover.

The nodes and their subtrees are subsequently swapped to create the offspring individuals (Figure 2.17).

As can be seen from Figures 2.16 and 2.17, the resulting individuals are quite different from their parent individuals, and hence the crossover operator is very effective. It can also be seen from the figures that the trees have all different sizes, which requires variable tree sizes. This is a major difference from genetic algorithms in which individuals usually have a fixed size.

2.3.3 Mutation

The other genetic operator commonly used in genetic programming is mutation. Here, either a random node is deleted from the tree or its content is replaced. In the latter case, it is important to maintain integrity. For example, if a binary operator would be replaced by a unary operator, one of the operands would become obsolete. Figure 2.18 shows an example of mutation.

It can be seen that the operator $-$ is replaced by the operator * in this example. Because both operators are binary, the integrity is maintained.

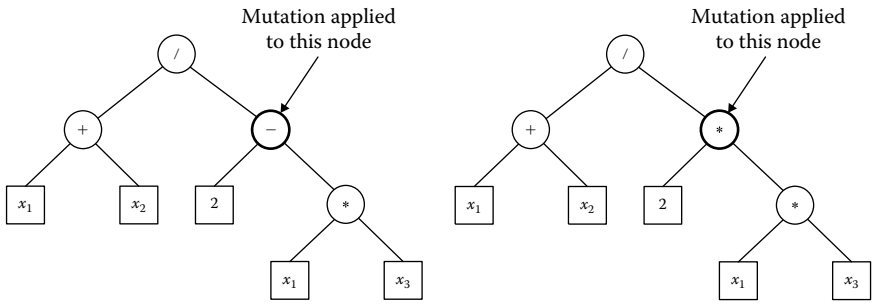


FIGURE 2.18: Example of mutation in GP. a) Individual before mutation. b) Individual after mutation.

2.4 Evolutionary Approaches for Medical Imaging

Evolutionary techniques can be applied to various medical imaging tasks. This section presents a few examples.

Once medical images have been captured and are available in digital form, they typically contain various types of noises and defects. A denoising step is hence usually necessary to remove these imperfections. A variational approach for image denoising based on genetic algorithms was introduced by Cannavo and colleagues (2006). Their approach can also run in parallel on multiple machines, which significantly reduces the computation time. An evolutionary approach to denoising color images was developed by Lukac, Plataniotis, and Venetsanopoulos (2006). They optimize the coefficients of weighted vector directional filters so as to match varying image and noise characteristics. A hybrid genetic algorithm–fuzzy logic–based noise removal technique was proposed by Lee, Guo, and Hsu (2005); the genetic algorithm is responsible for adjusting the parameters of an image knowledge base. Another hybrid technique for removing noise in medical images employs a genetic algorithm coupled with a cellular neural network (Su and Jhang, 2006).

Closely related to denoising is the concept of image enhancement, which aims to improve the perception or interpretability of all or part of an image. Genetic algorithms were used to select the optimal image operators for image enhancement by Pal, Bhandari, and Kundu (1994). A genetic algorithm for enhancing color images was developed by Shyu and Leou (1998). The authors form a set of generalized transforms for color image enhancement, while the task of the genetic algorithm is to identify the optimal set of generalized transforms from this set. In Poli and Cagnoni (1997), genetic programming was employed to adapt the color map in images, resulting in global contrast enhancement.

A crucial step of most medical image analysis approaches is image segmentation; its applications range from background subtraction to separation of anatomical structures to extraction of areas encompassing single

cells. Cagnoni, Dobrzeniecki, Poli, and Yanch (1999) developed a GA-based approach for segmenting medical 3-D images where, based on some initial segmentations performed by an expert, a genetic algorithm is used to employ a contour detector. Edge detection of MRI and CT data based on genetic algorithms was introduced by Gudmundsson, El-Kwae, and Kabuka (1998), who optimized edge maps by performing certain transformations on edge substructures. Genetic snakes, a GA-based approach of deformable model image segmentation, was proposed in Ballerini (1998) and applied to fundus images for detecting diabetic retinopathy. Hill and Taylor (1992) described the use of genetic algorithms for model-based image interpretation and applied them to delineate boundaries in echocardiograms.

Medical images often have to be compared to other medical images either of the same or a different patient, taken simultaneously or at different times, or using different modalities. For such a comparison to be effective, the two images must be geometrically aligned, which requires the application of an image registration algorithm. An early application of genetic algorithms for image registration was investigated in Fitzpatrick, Grefenstette, and Van Gucht (1984), who applied their technique to angiographic images. Garai and Chauduri (2002) presented a computationally more efficient genetic algorithm that starts with a chromosome of small length, representing possible solutions at low resolution, which are refined in a second stage. GA-based approaches to registering 3-D volume data were developed in, for example, Yamany, Ahmed, and Farag (1999) and He and Narayana (2002).

2.5 Conclusions

In this chapter, we provided a brief overview of evolutionary approaches to optimization. In particular, we discussed genetic algorithms and genetic programming and their subtasks of selection, crossover, mutation, and coding. Finally, we offered some pointers on the use of evolutionary techniques for common medical imaging tasks, such as image enhancement, image segmentation, and image registration.

References

Ballerini, L. (1998) Genetic snakes for medical image segmentation. In F. Prêteux, J. L. Davidson, and E. R. Dougherty (Eds.), *Mathematical Modeling and Estimation Techniques in Computer Vision*, pp. 284–95. Bellingham, Wash.: SPIE.

Cagnoni, A., Dobrzeniecki, A. B., Poli, R., and Yanch, Y. C. (1999) Genetic algorithm-based interactive segmentation of 3D medical images. *Image and Vision Computing*, 17: 881–95.

Cannavo, F., Nunnari, G., Giordano, D., and Spampinato, C. (2006) Variational method for image denoising by distributed genetic algorithms on GRID environment. In S. M. Reddy (Ed.), *15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 227–32. Los Alamitos, Calif.: IEEE Computer Society.

Edelbaum, T. N. (1962) Theory of maxima and minima. In G. Leitmann (Ed.), *Optimization Techniques with Applications to Aerospace Systems*. New York: Academic Press, pp. 1–32.

Fenton, N., and Hill, G. (1993) *Systems Construction and Analysis: A Mathematical and Logical Framework*. New York: McGraw-Hill.

Fitzpatrick, J. M., Grefenstette, J. J., and Van Gucht, D. (1984) Image registration by genetic search, *Proceedings of the IEEE Southeast Conference*, pp. 460–64.

Fonseca, C. M., and Fleming, P. J. (1995) An overview of evolutional algorithms in multiobjective optimization. *Evolutionary Computation*, 3: 1–16.

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*, Reading, Mass.: Addison-Wesley.

Gudmundsson, M., El-Kwae, E. A., and Kabuka, M. R. (1998) Edge detection in medical images using a genetic algorithm. *IEEE Transactions on Medical Imaging*, 17: 469–74.

He, R., and Narayana, P. A. (2002) Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images. *Computerized Medical Imaging and Graphics*, 26: 277–92.

Hill, A., and Taylor, C. J. (1992) Model-based image interpretation using genetic algorithms. *Image and Vision Computing*, 10: 295–300.

Holland, J. H. (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.

Koza, J. R. (1992) *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, Mass.: MIT Press.

Lee, C.-S., Guo, S.-M., and Hsu, C.-Y. (2005) Genetic-based fuzzy image filter and its application to image processing. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 35: 694–711.

Lukac, R., Plataniotis, K. N., and Venetsanopoulos, A. N. (2006) Color image denoising using evolutionary computation. *International Journal of Imaging Science and Technology*, 15: 236–51.

Pal, S. K., Bhandari, D., and Kundu, M. K. (1994) Genetic algorithms for optimal image enhancement. *Pattern Recognition Letters*, 15: 261–71.

Poli, R., and Cagnoni, S. (1997) Evolution of pseudocolouring algorithms for image enhancement, Technical Report: CSRP-97-5, University of Birmingham.

Schwefel, H.-P. (1995) *Evolution and Optimum Seeking*. New York: Wiley.

Shyu, M.-S., and Leou, J.-J. (1998) A genetic algorithm approach to color image enhancement. *Pattern Recognition*, 31: 871–80.

Su, T.-J., and Jhang, J.-W. (2006) Medical image noise reduction using cellular neural networks. In *Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 228–31.

Syswerda, G. (1989) Uniform crossover in genetic algorithms. In *Proceedings of the Third International Conference on Genetic Algorithms 1989* (ICGA89), pp. 2–9.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics*, 16: 97–159.

Yamany, S. M., Ahmed, M. N., and Farag, A. A. (1999) A new genetic-based technique for matching 3D curves and surfaces. *Pattern Recognition*, 32: 1817–20.

Chapter 3

Rough Sets in Medical Imaging: Foundations and Trends

Aboul Ella Hassanien, Ajith Abraham, James F. Peters, and Janusz Kacprzyk

Contents

This chapter presents a broad overview of rough sets theory and a review of the current literature on rough sets–based approaches to various problems in medical imaging, such as medical image segmentation and object extraction, feature extraction/reduction, and image classification. Rough

representation of the region of interest and rough image entropy are discussed. Hybridized rough set frameworks with other computational intelligence techniques, including neural networks, particle swarm optimization, genetic algorithms, support vector machines, and fuzzy sets, are presented. Challenges to be addressed and future directions of research are also presented.

3.1 Introduction

Computational intelligence techniques and approaches encompass various paradigms dedicated to approximately solving real-world problems in decision making, pattern classification, and learning. Prominent among these paradigms are fuzzy sets, neural networks, genetic algorithms (GAs), decision trees, rough sets, and a generalization of rough sets called near sets. Fuzzy sets [121] provide a natural framework for dealing with uncertainty. It offers a problem-solving tool between the precision of classical mathematics and the inherent imprecision of the real world. For example, imprecision in a segmented image can be represented and analyzed using fuzzy sets [55]. Neural networks [101] provide a robust approach to approximating real-valued, discrete-valued, and vector-valued functions. The well-known backpropagation algorithm that uses gradient descent to tune network parameters to best fit the training set with input–output pair has been applied as a learning technique for the neural networks. Genetic algorithms [43,47] are stochastic search techniques based on the principles of evolution. Extensive research has been performed exploiting the robust properties of genetic algorithms and demonstrating their capabilities across a broad range of problems. These evolutionary methods have gained recognition as general problem-solving techniques in many applications, including function optimization, image processing, classification and machine learning, training of neural networks, and system control. Other approaches, such as case-based reasoning and decision trees [22,95], are also widely used to solve data analysis problems. Each of these techniques has its own properties and features, including the ability to find important rules and information applicable to the medical field. Each technique contributes a distinct methodology for addressing problems in its domain in a cooperative rather than competitive manner. The result is a more intelligent and robust system providing a human-interpretable, low-cost, exact-enough solution, as compared to traditional techniques. Currently, there is great interest in the development of an efficient computer-aided diagnosis (CAD) system that assists the radiologist, the aim being not to replace the radiologist but to offer a second opinion [7,34]. Consequently, many efforts have been made to incorporate techniques for image processing. Recently, various published algorithms have been applied to build a computer-aided analysis system in

the medical field [20,21,34,35,94]. The most commonly used algorithms are neural networks, Bayesian classifier, genetic algorithms, decision trees, and fuzzy theory [93,98,106,107,110,116]. Unfortunately, the techniques developed have not been sufficient to introduce an efficient computer-aided analysis system for clinical use. A survey of the area can be found in Karssemeijer and others [34].

Rough sets theory, introduced by Zdzisaw Pawlak during the early 1980s [67–69], offers an approach to granular computing that is part of computational intelligence [80]. The basic idea underlying the rough sets approach to information granulation is to discover to what extent a given set of objects (e.g., pixel windows in an image) approximates another set of objects of interest. Objects are compared by considering their descriptions. An object description is modeled as a vector function, and values represent object features [70]. It is possible for an object feature to be represented by one or more functions, such as color represented by functions that measure luminance (intensity), type of color (hue), and purity of color (saturation). This is a fairly new intelligence technique for managing uncertainty that has been applied to the medical domain. It is used to discover data dependencies, evaluate the importance of features, discover the patterns of data, reduce all redundant objects and features, seek the minimum subset of features, and recognize and classify objects in medical imaging. Moreover, it is being used for the extraction of rules from databases. Rough sets have proven useful for representation of vague regions in spatial data. One advantage of the rough set is the creation of readable if-then rules. Such rules have the potential to reveal new patterns in the data material; furthermore, it also collectively functions as a classifier for unseen data sets. Unlike other computational intelligence techniques, rough sets analysis requires no external parameters and uses only the information presented in the given data. One of the nice features of rough sets theory is that it can tell, on the basis of the data itself, whether or not the data is complete. If the data are incomplete, it suggests more information about the objects needed to be collected in order to build a good classification model. On the other hand, if the data are complete, rough sets can determine whether there is more than enough or redundant information in the data and can find the minimum data needed for a classification model. This property of rough sets is very important for applications in which domain knowledge is very limited or data collection is very expensive or laborious because it makes sure the data collected are just good enough to build a good classification model without sacrificing the accuracy of the classification model or wasting time and effort to gather extra information about the objects [57,58,60,68].

The objective of this chapter is to present the state of the art in the rough sets applications to image processing and pattern recognition, particularly in medical imaging, and to motivate research in trendsetting directions. We review and discuss some representative methods to provide inspiring examples of how rough sets might be applied to resolve medical imaging problems

and how medical images could be analyzed, processed, and characterized by rough sets. These representative examples include (a) rough representation of a region of interest (ROI); (b) rough image entropy; (c) rough c-mean clustering; (d) and rough neural intelligent approach for image classification: a case of patients with suspected breast cancer.

Section 3.2 explains the basic framework of rough sets theory and defines some of the key concepts. Section 3.3 introduces rough image processing, including the rough image, rough representation of an ROI, rough image entropy, and rough-based medical image applications such as object extraction and medical image segmentation and clustering. Section 3.4 briefly reviews the use of rough sets in feature reduction and image classification. Section 3.5 reviews joint rough sets with other intelligent approaches, such as the rough neural network, rough fuzzy, and rough genetic algorithms. It also describes using rough neural networks in image mammogram breast cancer analysis. Section 3.6 discusses other applications of rough sets in the medical domain, such as for medical image retrieval and medical data mining and decision systems. Challenges and future trends are addressed in Section 3.7.

3.2 Rough Sets: Foundations

Rough sets theory is a new, intelligent mathematical tool proposed by Pawlak [57,58,67–69]. It is based on the concept of approximation spaces and models of the sets and concepts. In rough sets theory, the data are collected in a table, called a decision table. Rows of a decision table correspond to objects, and columns correspond to features. In the data set, we assume that a set of examples with a class label to indicate the class to which each example belongs are given. We call the class label a decision feature; the rest of the features are conditional. Let $\mathcal{O}, \mathcal{F}$ denote a set of sample objects and a set of functions representing object features, respectively. Assume that $B \subseteq \mathcal{F}, x \in \mathcal{O}$. Further, let $[x]_B$ denote

$$[x]_B = \{y : x \sim_B \ y)\}$$

Rough sets theory defines three regions based on the equivalent classes induced by the feature values: lower approximation $\underline{B}X$, upper approximation $\overline{B}X$, and boundary $BND_B(X)$. A lower approximation of a set X contains all equivalence classes $[x]_B$ that are subsets of X, and upper approximation $\overline{B}X$ contains all equivalence classes $[x]_B$ that have objects in common with X, while the boundary $BND_B(X)$ is the set $\overline{B}X \backslash \underline{B}X$ (i.e., the set of all objects in $\overline{B}X$ that are not contained in $\underline{B}X$). So, we can define a rough set as any set with a nonempty boundary.

The indiscernibility relation $\sim_B$ (or by Ind_B) is a mainstay of rough sets theory. Informally, $\sim_B$ is a set of all objects that have matching descriptions.

Based on the selection of B, $\sim_B$ is an equivalence relation that partitions a set of objects $\mathcal{O}$ into equivalence classes (also called elementary sets [67]). The set of all classes in a partition is denoted by $\mathcal{O}/\sim_B$ (also by $\mathcal{O}/Ind_B$). The set $\mathcal{O}/Ind_B$ is called the quotient set. Affinities between objects of interest in the set $X \subseteq \mathcal{O}$ and classes in a partition can be discovered by identifying those classes that have objects in common with X. Approximation of the set X begins by determining which elementary sets $[x]_B \in \mathcal{O}/\sim_B$ are subsets of X.

We provide a brief explanation of the basic framework of rough sets theory, along with some of the key definitions. A review of this basic material can be found in sources such as [56–58,60,68,69,122].

3.2.1 Information system and approximation

Definition 3.1 (Information system). *Information system is a tuple (U, A), where U consists of objects and A consists of features. Every $a \in A$ corresponds to the function $a : U \to V_a$, where V_a is a's value set. In applications, we often distinguish between conditional features C and decision features D, where $C \cap D = \emptyset$. In such cases, we define decision systems (U, C, D).*

Definition 3.2 (Indiscernibility relation). *Every subset of features $B \subseteq A$ induces indiscernibility relation:*

$$Ind_B = \{(x, y) \in U \times U : \forall_{a \in B}\, a(x) = a(y)\}$$

For every $x \in U$, there is an equivalence class $[x]_B$ in the partition of U defined by Ind_B.

Due to imprecision that exists in real-world data, there are sometimes conflicting classification of objects contained in a decision table. Conflicting classification occurs whenever two objects have matching descriptions but are deemed to belong to different decision classes. In that case, a decision table contains an inconsistency.

Definition 3.3 (Lower and upper approximations). *In rough sets theory, the approximations of sets are introduced to deal with inconsistency. A rough set approximates traditional sets using a pair of sets named the lower and upper approximations of the set. Given a set $B \subseteq A$, the lower and upper approximations of a set $Y \subseteq U$ are defined by Equations 3.1 and 3.2, respectively.*

$$\underline{B}Y = \bigcup_{x:[x]_B \subseteq X} [x]_B \tag{3.1}$$

$$\overline{B}Y = \bigcup_{x:[x]_B \cap X \neq \emptyset} [x]_B \tag{3.2}$$

Definition 3.4 (Lower approximation and positive region). *The positive region $POS_C(D)$ is defined by*

$$POS_C(D) = \bigcup_{X:X\in U/Ind_D} \underline{C}X$$

$POS_C(D)$ is called the positive region of the partition U/Ind_D with respect to $C \subseteq A$, that is, the set of all objects in U that can be uniquely classified by elementary sets in the partition U/Ind_D by means of C [69].

Definition 3.5 (Upper approximation and negative region). *The negative region $NEG_C(D)$ is defined by*

$$NEG_C(D) = U - \bigcup_{X:X\in U/Ind_D} \overline{C}X$$

the set of all objects that can be definitely ruled out as members of X.

Definition 3.6 (Boundary region). *The boundary region is the difference between upper and lower approximations of a set X consisting of equivalence classes with one or more elements in common with X. It given by the following formula:*

$$BND_B(X) = \underline{B}X - \overline{B}X \tag{3.3}$$

Figure 3.1 illustrates an example of approximation.

3.2.2 Reduct and core

An interesting question is whether there are features in the information system (feature–value table) that are more important than other features to the knowledge represented in the equivalence class structure. Often we wonder whether there is a subset of features that, by itself, can fully characterize the knowledge in the database. Such a feature set is called a reduct. Calculation of reducts of an information system is a key problem in rough sets theory [57, 58,60,68]. We must get reducts of an information system in order to extract rulelike knowledge from an information system.

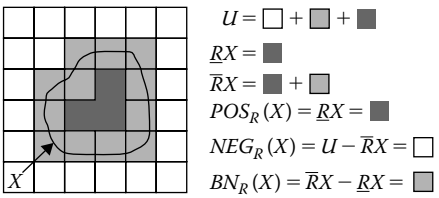


FIGURE 3.1: Illustrated example of approximation. (From S. Hirano and S. Tsumoto, 2005. *International Journal of Approximate Reasoning*, 40: 23–34. With permission.)

Definition 3.7 (Reduct). *Given a classification task related to the mapping $C \rightarrow D$, a reduct is a subset $R \subseteq C$ such that*

$$\gamma(C, D) = \gamma(R, D)$$

and no proper subset of R satisfies analogous equality.

Definition 3.8 (Reduct set). *Given a classification task mapping a set of variables C to a set of labeling D, a reduct set is defined with respect to the power set $P(C)$ as the set $R \subseteq P(C)$ such that $Red = \{A \in P(C) : \gamma(A, D) = \gamma(C, D)\}$. That is, the reduct set is the set of all possible reducts of the equivalence relation denoted by C and D.*

Definition 3.9 (Minimal reduct). *Given a classification task mapping a set of variables C to a set of labeling D, and R is the reduct set for this problem space, a minimal reduct R is the reduct such that $\|R\| \leq \|A\|, \forall A \in R$. That is, the minimal reduct is the reduct of least cardinality for the equivalence relation denoted by C and D.*

Definition 3.10 (Core). *Attribute $c \in C$ is a core feature with respect to D if and only if it belongs to all the reducts. We denote the set of all core features by Core(C). If we denote by $R(C)$ the set of all reducts, we can put*

$$Core(C) = \bigcap_{R \in R(C)} R \tag{3.4}$$

The computation of the reducts and the core of the condition features from a decision table is a way of selecting relevant features. It is a global method in the sense that the resultant reduct represents the minimal set of features necessary to maintain the same classification power given by the original and complete set of features. A more straightforward way to select relevant features is to assign a measure of relevance to each feature and choose the features with higher values. Based on the generated reduct system, we generate list of rules that will be used for building the classifier model of the new object with each object in the reduced decision table (i.e., reduct system) and classify the object to the corresponding decision class. The calculation of all the reducts is fairly complex (see [16,20,95,96]).

Nevertheless, in many practical applications, including medical imaging problems, it is unnecessary to calculate all the reducts. For example, the following heuristic procedure was used to obtain the most satisfactory reduct [103]. Starting from single features, the one with the greatest quality of classification is chosen; another feature is appended to the chosen feature, giving the greatest increase to the quality of classification for the pair of features; then, yet another feature is appended to the pair, giving the greatest increase to the quality of classification for the triple; and so on, until the maximal quality is reached by a subset of features. At the end of this procedure, it should be verified whether the obtained subset is minimal (i.e., whether elimination of any feature from this subset keeps the quality unchanged). Then, for further analysis, it is often efficient to consider a reduced data table in which the set Q of features is confined to the most satisfactory reduct.

3.2.3 Degree of dependency

Definition 3.11 (Degree of dependency). *The degree of dependency $\gamma(P,Q)$ of a set P of features with respect to a set Q of class labeling is defined as*

$$\gamma(P,Q) = \frac{|POS_P(Q)|}{|U|} \tag{3.5}$$

where $|S|$ denotes the cardinality of a set S.

The degree of dependency provides a measure of how important P is in mapping the data set examples into Q. If $\gamma(P,Q) = 0$, then classification Q is independent of the features in P, so the decision features are of no use to this classification. If $\gamma(P,Q) = 1$, then Q is completely dependent on P, so the features are indispensable. Values $0 < \gamma(P,Q) < 1$ denote partial dependency, which shows that only some of the features in P may be useful or that the data set was flawed to begin with. In addition, the complement of $\gamma(P,Q)$ gives a measure of the contradictions in the selected subset of the data set. It is now possible to define the significance of a feature by calculating the change of dependency when removing the feature from the set of considered conditional features.

3.2.4 Significance of the features

The significance of features enables us to evaluate features by assigning a real number from the closed interval [0, 1], expressing the importance of a feature in an information table. The significance of a feature a in a decision table DT can be evaluated by measuring the effect of removing feature a in C from feature set C on a positive region defined by the table DT. As shown in Definition 3.9, the number $\gamma(C,D)$ expresses the degree of dependency between features C and D or accuracy of approximation of U/D by C.

Definition 3.12 (Significance). *For any feature $a \in C$, we define its significance ζ with respect to D as follows:*

$$\zeta(a,C,D) = \frac{\left|POS_{C\setminus\{a\}}(D)\right|}{|POS_C(D)|} \tag{3.6}$$

Definitions 3.7 through 3.12 are used to express the importance of particular features in building the classification model. For a comprehensive study, refer to Swiniarski and Skowron [97]. As an importance measure, we can use frequency of occurrence of features in reducts. We can also consider various modifications of Definition 3.7—for example, approximate reducts, which preserve information about decisions only to some degree [95]. Further, positive region in Definition 3.4 can be modified by allowing for approximate satisfaction of inclusion $[x]_C \subseteq [x]_D$, as proposed, for example, in the variable precision rough sets (VPRS) model [122]. Finally, in Definition 3.2, the

meaning of $IND(B)$ and $[x]_B$ can be changed by replacing equivalence relation with similarity relation, especially useful in considering numeric features. For further reading, refer to Pawlak [57] and Polkowski [60].

3.2.5 Decision rules

In the context of supervised learning, an important task is the discovery of classification rules from the data provided in the decision tables. The decision rules not only capture patterns hidden in the data but can also be used to classify new, unseen objects. Rules represent dependencies in the data set and represent extracted knowledge that can be used when classifying new objects not in the original information system. When the reducts are found, the job of creating definite rules for the value of the decision feature of the information system is practically done. To transform a reduct into a rule, we simply bind the condition feature values of the object class from which the reduct originated to the corresponding features of the reduct. Then, to complete the rule, a decision part comprising the resulting part of the rule is added. This is done in the same way as for the condition features. To classify objects, which have never been seen before, rules generated from a training set are used. These rules represent the actual classifier. This classifier is used to predict to which classes new objects are attached. The nearest matching rule is determined as the one whose condition part differs from the feature vector of reimage by the minimum number of features. When there is more than one matching rule, we use a voting mechanism to choose the decision value. Every matched rule contributes votes to its decision value, which are equal to the t times the number of objects matched by the rule. The votes are added, and the decision with the largest number of votes is chosen as the correct class. Quality measures associated with decision rules can be used to eliminate some of the decision rules. Such quality measures include support, strength, accuracy, and coverage (see [59]).

3.3 Rough Image Processing

Rough image processing is the collection of all approaches and techniques that understand, represent, and process the images, their segments, and their features as rough sets [112]. In grayscale images, boundaries between object regions are often ill defined because of grayness and/or spatial ambiguities [98,99]. This uncertainty can be handled by describing the different objects as rough sets with upper (or outer) and lower (or inner) approximations. The concepts of upper and lower approximation can be viewed, respectively, as outer and inner approximations of an image region in terms of granules. Pal, Shankar, and Mitra [99] defined the rough image as follows:

Definition 3.13 (Rough image). *Let the universe U be an image consisting of a collection of pixels. Then, if we partition U into a collection of nonoverlapping windows of size mn, each window can be considered as a granule G. Given this granulation, object regions in the image can be approximated by rough sets.*

A *rough image* is a collection of pixels and the equivalence relation induced partition of an image into sets of pixels using within each non-overlapping window over the image. With this definition, the roughness of various transforms (or partitions) of the image can be computed using image granules for windows of different sizes.

3.3.1 Rough representation of a region of interest

An ROI is a selected subset of samples within an image identified for a particular purpose, such as the boundaries of an object in a 2-D image and the contours or surfaces outlining an object in a volume data set. The concept of an ROI is commonly used in medical imaging. For example, the boundaries of a tumor may be defined on an image or in a volume for the purpose of measuring its size. The endocardial border may be defined on an image, perhaps during different phases of the cardiac cycle—say, end systole and end diastole—for the purpose of assessing cardiac function.

Hirano and Tsumoto [26] introduced the rough direct representation of ROIs in medical images. The main advantage of this method is its ability to represent inconsistency between the knowledge-driven shape and image-driven shape of an ROI using rough approximations. The method consists of three steps including preprocessing: (1) derive discretized feature values that describe the characteristics of an ROI; (2) build up the basic regions in the image using all features so that each region includes voxels that are indiscernible on all features; and (3) construct an ideal shape of the ROI and approximate it by the basic categories according to the given knowledge about the ROI. Then the image is split into three regions, or sets of voxels, that are

1. Certainly included in the ROI (positive region)
2. Certainly excluded from the ROI (negative region)
3. Possibly included in the ROI (boundary region)

The ROI is consequently represented by the positive region associated with some boundary regions. In the experiments, we show the result of implementing a rough image segmentation system.

Hirano and Tsumoto [29,37] described the procedures for rough representation of *ROIs* under single and multiple types of classification knowledge. Usually, the constant variables defined in the prior knowledge, for example some threshold values, do not meet the exact boundary of images due to inter-image variances of the intensity. The approach tries to roughly represent the shape of the ROI by approximating the given shapes of the ROI by the primitive regions derived from features of the image itself. It is reported that the

simplest case is one in which we have information only about intensity range of the ROI. In this case, intensity thresholding is a conventional approach to obtain the voxels that fall into the given range. Let us denote the lower and upper thresholds by ThL and ThH, respectively. Then the ROI can be represented by

$$ROI = \{x(p) \mid Th_L \leq I(x)P \leq Th_P\} \tag{3.7}$$

where $x(p)$ denotes a voxel at location p, and $I(x(p))$ denotes intensity of voxel $x(p)$.

Figure 3.2 illustrates the concept of rough ROI representation. The left image is an original grayscale image. Suppose that the ROIs are three black circular regions: ROI_1, ROI_2, and ROI_3. Also suppose that we know the lower threshold value Th_L of the ROIs, derived from some knowledge base. With this knowledge, we can segment an ideal ROI $X_{R\hat{O}I}$ as follows:

$$X_{R\hat{O}I} = \{x(p) \mid Th_L \leq I(p)\} \tag{3.8}$$

However, $X_{R\hat{O}I}$ does not correctly match the expected ROIs because Th_L was too small to separate the ROIs. Th_L is a global threshold determined on the other sets; therefore, it should not be directly applied to this image. Then, represent the possible boundary of the ROIs according to the low-level feature of this image; for more details see Hirano and Tsumoto [26].

3.3.2 Rough image entropy

The entropy-based information theoretic approach has recently received considerable interest in image analysis, such as in image registration [30]. Previous work on entropic thresholding is based on Shannon entropy. The idea is to calculate Shannon entropy based on a co-occurrence matrix and use it as a criterion for selecting an appropriate threshold value. The approach using relative entropy for image thresholding has been shown very competitive compared to Pal and Pal's methods [124] (1991) in which the relative entropy is

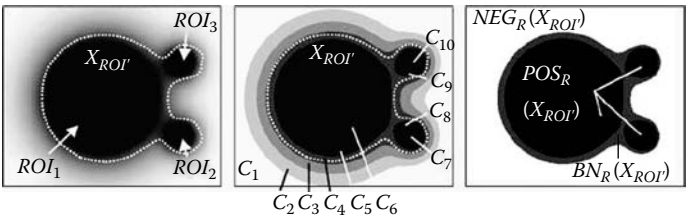


FIGURE 3.2: Rough ROI representation. Left: an original image. Middle: elementary categories C1C9. Right: roughly segmented ROI. (From S. Hirano and S. Tsumoto, 2005. *International Journal of Approximate Reasoning*, 40: 23–34. With permission.)

chosen to be a thresholding criterion of measuring mismatch between an image and a thresholded image. Currently, various published approaches use relative entropy and apply it to medical images, multispectral imagery, temporal image sequences, multistage thresholding, and segmentation.

Pal, Shankar, and Mitra [99] presented a new definition of image entropy in a rough sets theoretic framework and its application to the problem of object extraction from images by minimizing both object and background roughness. Granules carry local information and reflect the inherent spatial relation of the image by treating pixels of a window as indiscernible or homogeneous. Maximization of homogeneity in both object and background regions during their partitioning is achieved through maximization of rough entropy, thereby providing optimum results for object background classification.

Definition 3.14 (Rough image entropy). *Rough image entropy (R_IE) [99] is defined by*

$$R_IE = -\frac{e}{2}[R_{O_T} log_e(R_{O_T}) + R_{B_T} log_e(R_{B_T})] \tag{3.9}$$

Pal, Shankar, and Mitra [99] noted that the value of R_IE lies between 0 and 1 and has a maximum value of unity when $R_{O_T} = R_{B_T} = \frac{1}{e}$ and a minimum value of zero when $R_{O_T}, R_{B_T} \in \{0,\ 1\}$.

Figure 3.3 shows plot of rough entropy for various values of roughness of the object and background.

Pal, Shankar, and Mitra [99] reported that a maximization of homogeneity in both object and background regions during their partitioning is achieved through maximization of rough entropy, thereby providing optimum results for object-background classification. Also, maximization of the rough entropy measure minimizes the uncertainty arising from vagueness of the boundary region of the object. Therefore, for a given granule size, the threshold for object-background classification can be obtained through its maximization with respect to different image partitions. The rough entropy concepts may be suitable for many applications in image processing, particularly for feature extraction and image segmentation, such as feature extraction in mammogram images and identification of lymphomas by finding follicles in microscopy images.

3.3.3 Rough sets for object extraction

Identification of anatomical features is a necessary step for medical image analysis. Automatic methods for feature identification using conventional pattern recognition techniques typically classify an object as a member of a predefined class of objects but do not attempt to recover the exact or approximate shape of that object. Therefore, such techniques are usually not sufficient to identify the borders of organs when individual geometry varies in local detail, even though the general geometrical shape is similar.

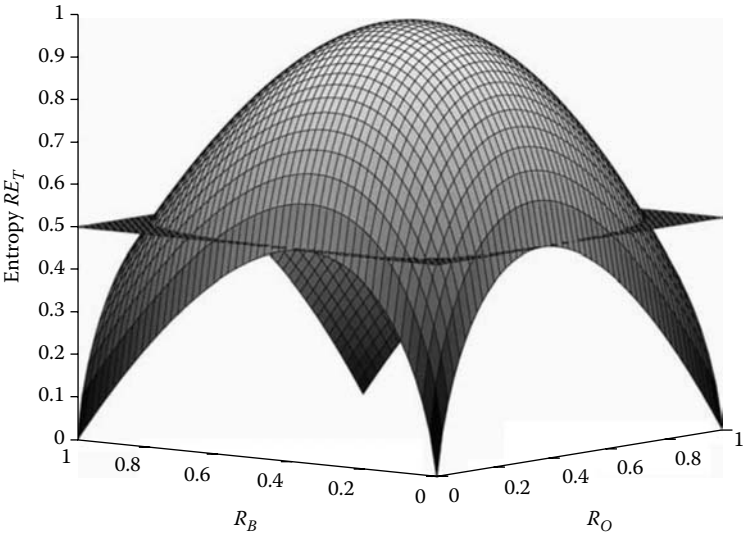


FIGURE 3.3: Rough entropy for various values. (From S. K. Pal, B. U. Shankar, and P. Mitra, 2005. *Pattern Recognition Letters*, 26(16): 2509–17. With permission.)

Pal, Shankar, and Mitra [99] demonstrated a new application of rough sets for object extraction from grayscale images. In grayscale images, boundaries between object regions are often ill defined. This uncertainty can be handled by describing the different objects as rough sets with upper (outer) and lower (inner) approximations. The set approximation capability of rough sets is exploited in the present investigation to formulate an entropy measure, called rough entropy, quantifying the uncertainty in an object-background image. They viewed the object and background as two sets with their rough representation by computing the inner approximation of the object ($\underline{Q}_T$), outer approximation of the object ($\overline{Q}_T$), inner approximation of the background ($\underline{B}_T$), and outer approximation of the background ($\overline{B}_T$) as follows:

$$\underline{Q}_T = \bigcup G_i | p_j > T, \forall j = 1, ..., mn \tag{3.10}$$

$$\overline{Q}_T = \bigcup G_i, \exists j, p_j > T, j = 1, ..., mn \tag{3.11}$$

$$\underline{B}_T = \bigcup G_i | P_j > T, \forall j = 1, ..., mn \tag{3.12}$$

$$\overline{B}_T = \bigcup G_i, \exists j, p_j \leq T, j = 1, ..., mn \tag{3.13}$$

where p_j is a pixel in G_i. The rough set representation of the image for a given I_{mn} depends on the value of T.

Pal, Shankar, and Mitra [99] define the roughness (R) of the object O_T and the background B_T as follows:

$$R_{O_T} = 1 - \frac{\left|\underline{Q}_T\right|}{\left|\overline{Q}_T\right|} \tag{3.14}$$

$$R_{B_T} = 1 - \frac{\left|\underline{B}_T\right|}{\left|\overline{B}_T\right|} \tag{3.15}$$

where $|\ |$ is the cardinality of the set.

The method may be suitable for many applications in image processing, such as for automatically identifying the myocardial contours of the heart, segmenting knee tissues in CT images, and segmenting brain tissues in MR images.

3.3.4 Rough sets in medical image segmentation and clustering

Image segmentation is one of the most critical steps in image analysis and understanding; therefore, it has been the subject of considerable research over the last four decades. During this time, we have witnessed a tremendous development of new, powerful instruments for detecting, storing, transmitting, and displaying images, but automatic segmentation has remained a challenging problem. This shortcoming is evident in medical applications, where image segmentation is particularly difficult due to restrictions imposed by image acquisition, pathology, and biological variation. Biomedical image segmentation is a sufficiently complex problem that no single strategy has proven to be completely effective. Due to the complex nature of biomedical images, it is practically impossible to select or develop automatic segmentation methods of a generic nature that can be applied to any type of images (e.g., to both micro- and macroscopic images, cytological and histological images, MRI and X-ray). Medical image segmentation is an indispensable process in the visualization of human tissues. However, medical images always contain a large amount of noise caused by operator performance, equipment, and environment, which leads to inaccuracy with segmentation. A robust segmentation technique is required.

The basic idea behind segmentation-based rough sets is that while some cases may be clearly labeled as being in a set X (called the positive region in rough sets theory), and some cases may be clearly labeled as not being in set X (the negative region), limited information prevents us from labeling all possible cases clearly. The remaining cases cannot be distinguished and lie in what is known as the boundary region. Some effort has been made to use rough sets in image segmentation, particularly in medical segmentation problems.

Peters and Borkowski [79] presented a new form of indiscernibility relation based on k-means clustering of pixel values. The end result is a partitioning of a set of pixel values into bins that represent equivalence classes. The proposed approach makes it possible to introduce a form of upper and lower approximation specialized relative to sets of pixel values. This approach is particularly relevant to a special class of digital images for power-line ceramic insulators. Until now, the problem of determining when a ceramic insulator needs to be replaced has relied on visual inspection. With the k-means indiscernibility relation, faulty ceramic insulators can be automatically detected. The contribution of this work is the introduction of an approach to classifying power-line insulators based on a rough sets method and k-means clustering in analyzing digital images. The indiscernibility relation based on k-means clustering of pixel values may prove useful in many medical imaging applications.

Among the many difficulties in segmenting MRI data, the partial volume effect (PVE) arises in volumetric images when more than one tissue type occurs in a voxel. In many cases, the voxel intensity depends not only on the imaging sequence and tissue properties but also on the proportions of each tissue type present in the voxel. Widz, Revett, and Ślęzak [114,115] discussed the PVE problem in the segmentation of magnetic resonance imaging data that entails assigning tissue class labels to voxels. They employ the rough sets to automatically identify the PVE, which occurs most often with low-resolution imaging with large voxels.

Kobashi, Kondo, and Hata [36] and Matsuura and others [46] introduced rough sets to treat nominal data based on concepts of categorization and approximation in medical image segmentation. The proposed clustering method extracts features of each pixel by using thresholding and labeling algorithms. Thus, the features are given by nominal data. The ability of the proposed method was evaluated by applying it to human brain MRI images.

An interesting strategy for color image segmentation using rough sets theory was presented by Mohabey and Ray [52]. A new concept of encrustation of the histogram, called histon, was proposed for the visualization of multidimensional color information in an integrated fashion, and its applicability in boundary region analysis was shown. The histon correlates with the upper approximation of a set such that all elements belonging to this set are clarified as possibly belonging to the same segment or segments showing similar color value. The proposed encrustation provides a direct means of segregating a pool of inhomogeneous regions into its components. Experimental results for various images were presented in their work. They extended their work to introduce a hybrid rough sets theoretic approximation and fuzzy c-means (FCM) algorithm for color image segmentation [53]. They segmented natural images with regions having gradual variations in color value. The technique extracts color information regarding the number of segments and the segments' center values from the image itself through rough sets theoretic approximations and presents it as input to an FCM block for the soft evaluation of the segments.

The performance of the algorithm has been evaluated on various natural and simulated images.

Although many clustering algorithms [40] have been developed and applied in medical imaging problems, most of them cannot process objects in hybrid numerical/nominal feature space or with missing values. In many of them, the number of clusters has to be manually specified while the clustering results are sensitive to the input order of the objects to be clustered. These limit applicability of the clustering and reduce the quality of clustering. To solve this problem, an improved clustering algorithm based on rough sets theory and entropy theory was presented by Chena and Wang [9]. It aims to avoid the need to pre-specify the number of clusters while also allowing clustering in both numerical and nominal feature space with the similarity introduced to replace the distance index. At the same time, the rough sets theory endows the algorithm with the function to deal with vagueness and uncertainty in data analysis. Shannon entropy was used to refine the clustering results by assigning relative weights to the set of features according to the mutual entropy values. A novel measure of clustering quality was also presented to evaluate the clusters. The experimental results confirm that performances of efficiency and clustering quality of this algorithm are improved.

Widz, Revett, and Ślęzak [114,115] introduced an automated multispectral MRI segmentation technique based on approximate reducts derived from the theory of rough sets. They utilized the $T1$, $T2$, and PD MRI images from a simulated brain database as a gold standard to train and test their segmentation algorithm. The results suggest that approximate reducts, used alone or in combination with other classification methods, may provide a novel and efficient approach to the segmentation of volumetric MRI data sets. Segmentation accuracy reaches 96% for the highest resolution images and 89% for the noisiest image volume. They tested the resultant classifier on real clinical data, which yielded an accuracy of approximately 84%.

3.3.5 Adaptation of C-means to rough sets theory

C-means clustering is an iterative technique used to partition an image into c clusters. FCM is one of the most commonly used fuzzy clustering techniques for degree-estimation problems, especially in medical image processing [1,3]. Lingras [43] described modifications of clustering based on genetic algorithms, the k-means algorithm, and Kohonen self-organizing maps (SOMs). These modifications make it possible to represent clusters as rough sets. In their work, Lingras and West established a rough k-means framework and extended the concept of c-means by viewing each cluster as an interval or rough set [41]. Here is a brief summary of their pioneer clustering work.

K-means clustering is one of the most popular statistical clustering techniques used in segmentation of medical images [8,18,54,79,113,118]. The name k-means originates from the means of the k clusters that are created from n objects. Let us assume that the objects are represented by m-dimensional

vectors. The objective is to assign these n objects to k clusters. Each of the clusters is also represented by an m-dimensional vector, which is the centroid or mean vector for that cluster. The process begins by randomly choosing k objects as the centroids of the k clusters. The objects are assigned to one of the k clusters based on the minimum value of the distance $d(v, x)$ between the object vector $v = (v_1, \ldots, v_j, \ldots, v_m)$ and the cluster vector $x = (x_1, \ldots, x_j, \ldots, x_m)$. After the assignment of all the objects to various clusters, the new centroid vectors of the clusters are calculated as

$$x_j = \frac{\sum_{v \in x} v_j}{SOC}, \quad \text{where } 1 \leq j \leq m \tag{3.16}$$

where SOC is the size of cluster x.

According to Lingras [43], incorporating rough sets into k-means clustering requires the addition of the concept of lower and upper bounds. Calculation of the centroids of clusters from conventional k-means must be modified to include the effects of lower as well as upper bounds. The modified centroid calculations for rough sets are then given by

$$cen_j = W_{low} \times \frac{\sum_{v \in \underline{R}(x)}}{|\underline{R}(x)|} + w_{up} \times \frac{\sum_{v \in (BN_R(x))}}{|\underline{BN_R}(x)|} \tag{3.17}$$

where $1 \leq j \leq m$. The parameters w_{lower} and $w_{(upper)}$ correspond to the relative importance of lower and upper bounds, and $w_{low} + w_{up} = 1$. If the upper bound of each cluster were equal to its lower bound, the clusters would be conventional clusters. Therefore, the boundary region $\underline{BN_R}(x)$ will be empty, and the second term in the equation will be ignored. Thus, the above equation will reduce to conventional centroid calculations. The next step in the modification of the k-means algorithms for rough sets is to design criteria to determine whether an object belongs to the upper or lower bound of a cluster.

Mitra [47] proposed an evolutionary rough c-means clustering algorithm. Genetic algorithms are employed to tune the threshold and relative importance of upper and lower approximations of the rough sets modeling the clusters. The Davies-Bouldin clustering validity index is used as the fitness function, which is minimized while arriving at an optimal partitioning. A comparative study of its performance is made with related partitive algorithms. The effectiveness of the algorithm is demonstrated on real and synthetic data sets, including microarray gene expression data from bioinformatics. In the same study, the author noted that the parameter threshold measures the relative distance of an object X_k from a pair of clusters having centroids $cen_i mi$ and cen_j. The smaller the value of threshold, the more likely is X_k to lie within the rough boundary (between upper and lower approximations) of a cluster. This implies that only those points that definitely belong to a cluster (lie close to the centroid) occur within the lower approximation. A large value of threshold implies a relaxation of this criterion, such that more patterns are allowed to belong to any of the lower approximations. The parameter w_{low}

controls the importance of the objects lying within the lower approximation of a cluster in determining its centroid. A lower w_{low} implies a higher w_{up} and hence an increased importance of patterns located in the rough boundary of a cluster toward the positioning of its centroid.

The rough c–k-mean is a new challenge for use in MR and mammogram image segmentation as an aid to small lesion diagnosis and can effectively remove the influence of tiny details and noise.

3.4 Rough Sets in Feature Reduction and Image Classification

In image processing, raw images represented by the gray levels of the pixels are usually transformed to features that can better capture the characteristics of the images in the preprocessing phase. Texture features are often used in image classification and segmentation. In particular, texture features proposed by Haralick [24] are typically computed from the gray-level co-occurrence matrices and then used to classify each pixel for its type. Feature selection is an important step in the preprocessing since there are numerous potential features, some of which might be irrelevant or unimportant. Not only can reducing features speed up the processing time and possibly improve the classification accuracy, it also allows us to use classification methods that are not good at processing high-dimensional data, such as neural networks and support vector machines (SVMs). Feature selection aims to determine a minimal feature subset from a problem domain while retaining a suitably high accuracy in representing the original features. Rough sets theory enables the discovery of data dependencies and the reduction of the number of features contained in a data set or extracted from images using the data alone, requiring no additional information (see [68]).

The computation of the core and reducts from a rough sets decision table is a way of selecting relevant features [16,20,95,96]. It is a global method in the sense that the resultant reducts represent the minimal sets of features necessary to maintain the same classification power given by the original and complete set of features. A more straightforward manner for selecting relevant features is to assign a measure of relevance to each feature and choose the features with higher values. Based on the reduct system, we can generate the list of rules that will be used for building the classifier model for the new objects. Reduct is an important concept in rough sets theory, and data reduction is a main application of rough sets theory in pattern recognition and data mining because finding the minimal reduct of an information system is often difficult.

Since rule induction methods generate rules whose lengths are the shortest for discrimination between given classes, they tend to generate rules too short

for medical experts. Thus, these rules are difficult for the experts to interpret from the viewpoint of domain knowledge. Tsumoto [106] introduced the characteristics of experts' rules and proposed a new approach to generate diagnostic rules using a rough sets and medical diagnostic model. The approach focuses on the hierarchical structure of differential diagnosis and consists of three procedures: (1) the characterization of decision features (given classes) is extracted from databases and the classes are classified into several generalized groups with respect to the characterization; (2) two kinds of subrules, classification rules for each generalized group and rules for each class within each group, are induced; and (3) those two parts are integrated into one rule for each decision feature. His proposed approach was evaluated on a medical database, the experimental results of which show that induced rules correctly represent experts' decision processes.

Many researchers have worked to develop efficient algorithms to compute useful feature extraction and reduction of information systems and to develop mutual information and discernibility matrix–based feature reduction methods. These techniques have been successfully applied to the medical domain [104,120].

Wojcika [123] approached the nature of a feature recognition process through the description of image features in terms of the rough sets. Since the basic condition for representing images must be satisfied by any recognition result, elementary features are defined as equivalence classes of possible occurrences of specific fragments existing in images. The names of the equivalence classes (defined through specific numbers of objects and numbers of background parts covered by a window) constitute the best lower approximation of window contents (i.e., names of recognized features). The best upper approximation is formed by the best lower approximation, its features, and its parameters, all referenced to the object fragments situated in the window. The rough approximation of shapes is resistant to accidental changes in the width of contours and lines and to small discontinuities and, in general, to possible positions or changes in shape of the same feature. The rough sets are utilized also on the level of image processing for noiseless image quantization. This initiative study may be useful in many areas of medical image processing, including filtering, segmentation, and classification.

Swiniarski and Skowron [109] presented applications of rough sets methods for feature selection in pattern recognition. They emphasize the role of the basic constructs of a rough sets approach in feature selection, namely, reducts and their approximations, including dynamic reducts. Their algorithm for feature selection is based on an application of a rough sets method to the result of principal components analysis (PCA) used for feature projection and reduction. Their work presents many experiments, including face and mammogram recognition experiments. In their study, 144 mammogram images were selected for recognition experiments using the Mammography Image Analysis Society (MIAS) Mini-Mammographic Database, with 1024×1024 pixel images [50]. The database contains three types of class-labeled images: normal,

benign (abnormal), and malignant (abnormal). For each abnormal image, the coordinates of center of abnormality and proximate radius (in pixels) of a circle enclosing the abnormality are given. For classifications, the center locations and radii apply to clusters rather than to the individual classifications. They have provided an experiment of recognition of normal and abnormal images (two category classification). This set was divided into 128 case training sets and 16 case test sets. From the original 1024-pixel grayscale mammographic image, they extracted a 64×64 pixel subimage around the center of abnormality (or at the average coordinate for normal cases). They concluded that the rough sets methods have the ability to significantly reduce the pattern dimensionality and to be viable image-mining techniques as a front end of neural network classifiers.

Wang [15,16] studied the relationship of the definitions of rough reduction in algebra view and information views. Some relationships, such as inclusion relationship under some conditions and equivalence relationship under other conditions, are presented. The inclusion relationship between the feature importance defined in algebra view and information view is presented also. For example, the reduction under algebra view is equivalent to the reduction under information view if the decision table is consistent. Otherwise, the reduction under information view will include the reduction under algebra view. These results will be useful for designing further heuristic reduction algorithms.

Hu, Yu, and Xie [17] proposed an information measure for computing discernibility power of a crisp equivalence relation or a fuzzy one, which is the key concept in classical rough sets and fuzzy-rough sets models. Based on the information measure, a general definition of significance of nominal, numeric, and fuzzy features is presented. They redefine the independence of a hybrid feature subset, reduct, and relative reduct. Then two greedy reduction algorithms for unsupervised and supervised data dimensionality reduction based on the proposed information measure are constructed. It is reported that the reducts founded by the proposed algorithms get a better performance than do classical rough sets approaches.

Lymphoma is a broad term encompassing a variety of cancers of the lymphatic system. Lymphoma is differentiated by the type of cell that multiplies and how the cancer presents itself. An exact diagnosis regarding lymphoma is important for determining the most effective treatments for the patient's condition. Milan and others [51] focused on the identification of lymphomas by finding follicles in microscopy images provided by the Laboratory of Pathology in the University Hospital of Tenerife, Spain. The study contains two stages: in the first stage, they did image preprocessing and feature extraction; and in the second stage, they used different rough sets approaches for pixel classification. These results were compared to decision tree results. The results they got are very promising and show that symbolic approaches can be successful in medical image analysis applications.

Pham [90] presented a rough sets–based medical decision support system for tumor, node, metastases (TNM, i.e., tumor characteristics, lymph node

involvement, and distant metastatic lesions) classification aiming to distinguish low- and high-risk patients. The system also explains the decision in the form of if-then rules and in this manner performs data mining and new knowledge discovery. The introduced system was reported to have the best classification performance, robust and not dependent on the database size and the noise. The accuracy was almost 80%, which is comparable to the accuracy of physicians and much better than obtained with more conventional discriminant analysis (62% and 67%).

Microcalcification on x-ray mammogram is a significant mark for early detection of breast cancer. Texture analysis methods can be applied to detect clustered microcalcification in digitized mammograms. In order to improve the predictive accuracy of the classifier, the original number of feature sets is reduced into a smaller set using feature reduction techniques. Thangavel, Karnan, and Pethalakshmi [111] introduced rough sets–based reduction algorithms such as decision relative discernibility–based reduction, a heuristic approach, the Hus algorithm, quick reduct (QR), and VPRS to reduce the extracted features. The performances of all the introduced algorithms are compared. The gray-level co-occurrence matrix (GLCM) is generated for each mammogram to extract the Haralick features as a feature set. The rough reduction algorithms are tested on 161 pairs of digitized mammograms from the MIAS database [50].

Cyran and Mrózek [33] showed how rough sets can be applied to improve the classification ability of a hybrid pattern recognition system. The system consists of a feature extractor based on a computer-generated hologram (CGH). Although extracted features are shift, rotation, and scale invariant, they can be optimized. Their work presented an original method of optimizing the feature extraction abilities of a CGH. The method uses rough sets theory to measure the amount of essential information contained in the feature vector. This measure is used to define an objective function in the optimization process. Since rough sets theory–based factors are not differentiable, they use a nongradient approach for a search in the space of possible solutions. Finally, rough sets theory is used to determine decision rules for the classification of feature vectors. The proposed method is illustrated by a system recognizing the class of speckle pattern images indicating the class of distortion of optical fibers.

Jiang and colleagues [31] developed a joining associative classifier (JAC) algorithm using the rough sets theory to mine digital mammography. The experimental results showed that the JAC performed at 77.48% of classifying accuracy, which is higher than the 69.11% obtained using associative classifier only. At the same time, the number of rules decreased significantly.

Slowinski, Stefanowski, and Siwinski [29] discussed a process of analyzing medical diagnostic data by means of the combined rule induction and rough sets approach. The first step of this analysis includes the use of various techniques for discretization of numerical features. Rough sets theory is applied to determine feature importance for the patients' classification. The novel

contribution concerns considering two different algorithms inducing either a minimum or satisfactory set of decision rules. Verification of classification abilities of these rule sets is extended by an examination of sensitivity and specificity measures, and a comparative study of these composed approaches against other learning systems is discussed. The approach is illustrated on a medical problem concerning anterior cruciate ligament (ACL) rupture in a knee. The patients are described by features coming from anamnesis, MR examinations, and verified by arthroscopy. Reportedly, the clinical impact of this research indicates two features, PCL index and age, and their specific values that could support a physician's decision not to perform arthroscopy for some patients.

3.5 Joint Rough Sets with Other Intelligent Approaches

Intelligent systems comprise various paradigms dedicated to approximately solving real-world problems (e.g., in decision making, classification, or learning); among these paradigms are fuzzy sets, neural networks, decision trees, and rough sets algorithms. Combination of computational intelligence techniques in pattern recognition, particularly in medical imaging problems, has become one of the most important areas of research [102].

3.5.1 Neural networks with rough sets

Neural networks demonstrates a strong ability to solve complex problems for medical image processing. But neural networks cannot tell the redundant information from huge amounts of data, which will easily lead to problems such as too complex a network structure, long training time, low converging speed, and excessive computation. Much successful research has focused on these problems. For example, Hassanien and Ślęzak [23] introduced a rough neural approach for rule generation and image classification. Algorithm 3.1 describes the steps of the rule generation. Hybridization of intelligent computing techniques has been applied to see their ability and accuracy to classify breast cancer images into malignant or benign cases. Algorithms based on fuzzy image processing are first applied to enhance the contrast of the whole original image, to extract the ROI, and to enhance the edges surrounding that region. Then, they extract features characterizing the underlying texture of the ROIs by using the GLCM. Then, the rough set approach to feature reduction and rule generation is presented. Finally, the rough neural network is designed for discrimination of different ROIs to test whether they represent malignant cancer or benign growth. The rough neural network is built from rough neurons, each of which can be viewed as a pair of subneurons corresponding to the lower and upper bounds. To evaluate performance of the

Algorithm 3.1 Rule generation.

Input: Decision system (U, C, D)
Decision reduct $R \subseteq C$; $R = \{a_1, ...a_m\}$; $m = |R|$
Output: The set of decision rules $RULES(R)$ generated for R
1: **for** $u \in U$ **do**
2: **for** $a_i \in R$ **do**
3: $v_i = a_i(u)$;
4: **end for**
5: $v_d = d(u)$;
6: $RULES(R) = RULES(R) \cup \{a_1 = v_1 \wedge \ldots \wedge \ a_m = v_m \rightarrow d = v_d\}$;
7: **end for**
8: Return $RULES(R)$

presented rough neural approach, they run tests over different mammogram images. In their experiments, results show that the overall classification accuracy offered by the rough neural approach is high compared with other intelligent techniques.

The rough neural networks [39,81,82] used in Hassanien and Ślęzak's study [23] consist of one input layer, one output layer, and one hidden layer. The input layer neurons accept input from the external environment. The outputs from the input layer neurons are fed to the hidden layer neurons. The hidden layer neurons feed their output to the output layer neurons, which send their output to the external environment.

The number of hidden neurons is determined by the following inequality [10,27]:

$$N_{hn} \leq \frac{N_{ts} * T_e * N_f}{N_f + N_o} \tag{3.18}$$

where N_{hn} is the number of hidden neurons, N_{ts} is the number of training samples, T_e is the tolerance error, N_f is the number of features, and N_o is the number of the output.

The output of a rough neuron is a pair of upper and lower bounds, while the output of a conventional neuron is a single value. Rough neuron was introduced in 1996 by Lingras [39]. It was defined relative to upper bound (U_n) and lower bound (L_n), and inputs were assessed relative to boundary values. Rough neuron has three types of connections: input–output connection to U_n, input–output connection to L_n, and connection between U_n and L_n.

Definition 3.15 (Rough neuron). *A rough neuron R_n is a pair of usual rough neurons $R_n = (U_n, L_n)$, where U_n and L_n are the upper rough neuron and the lower rough neuron, respectively.*

Let (Ir_{L_n}, Or_{L_n}) be the input/output of a lower rough neuron and (Ir_{U_n}, Or_{U_n}) be the input/output of an upper rough neuron. Calculation of

the input/output of the lower/upper rough neurons is given as follows:

$$Ir_{L_n} = \sum_{j=1}^{n} w_{L_{nj}} On_j \tag{3.19}$$

$$Ir_{U_n} = \sum_{j=1}^{n} w_{U_{nj}} On_j \tag{3.20}$$

$$Or_{L_n} = min(f(Ir_{L_n}), f(Ir_{U_n})) \tag{3.21}$$

$$Or_{U_n} = max(f(Ir_{L_n}), f(Ir_{U_n})) \tag{3.22}$$

The output of the rough neuron (O_{rn}) will be computed as follows:

$$O_{rn} = \frac{Or_{U_n} - Or_{L_n}}{avarge(Or_{U_n}, Or_{L_n})} \tag{3.23}$$

In Hassanien and Ślęzak's experiments [23], the segmentation performance is measured by the value of accuracy as defined below, and the average of segmentation accuracy achieved by the reported algorithm is 97%, which means that it is robust enough.

$$S_A = \frac{M_P}{T_{NP}} \tag{3.24}$$

where S_A, M_P, and T_{NP} are the segmentation accuracy, number of misclassified pixels, and total number of pixels, respectively.

The rule's importance measure R_I was used as an evaluation to study the quality of the generated rule. It is defined by

$$R_I = \frac{\tau_r}{\rho_r} \tag{3.25}$$

where τ_r is the number of times a rule appears in all reducts, and ρ_r is the number of reduct sets. The quality of rules is related to the corresponding reduct(s). Hassanien and Ślęzak [23] generate rules that cover the largest parts of the universe U. Covering U with more general rules implies a smaller sized rule set. They use the importance rule criteria introduced in Li and Cercone [38] to study the rules' importance (see also [2]).

Swiniarski and Hargis [108] described application of a rough sets method to feature selection and reduction as a front end of neural network–based texture images recognition. The methods applied include singular-value decomposition (SVD) for feature extraction, principal components analysis (PCA) for feature projection and reduction, and rough sets methods for feature selection and reduction. For texture classification, the feedforward backpropagation neural networks were applied. The numerical experiments showed the ability of rough sets to select a reduced set of a pattern's features, while providing better generalization of neural network texture classifiers (see also [104]).

Jiang and others [32] proposed rough neural network that integrates neural network with reduction of rough sets theory to classify digital mammography. It is reported that the combined method performed better than neural

networks alone in terms of complexity time, and it can achieve 92.37% classifying accuracy, which is higher than the 81.25% obtained using neural networks only.

3.5.2 Fuzzy with rough sets

Mao, Liu, and Lin [11] proposed a new fuzzy Hopfield-model network based on rough sets reasoning for the classification of multispectral images. The main purpose is to embed a rough sets learning scheme into the fuzzy Hopfield network to construct a classification system called a rough-fuzzy Hopfield net (RFHN). The classification system is a paradigm for the implementation of fuzzy logic and rough systems in neural network architecture. Instead of all the information in the image being fed into the neural network, the upper- and lower-bound gray levels, captured from a training vector in a multispectral image, are fed into a rough-fuzzy neuron in the RFHN. Therefore, only $2/N$ pixels are selected as the training samples if an N-dimensional multispectral image was used. In the simulation results, the proposed network not only reduces the consuming time but also reserves the classification performance.

Wang and colleagues [117] proposed a new fuzzy-rough nearest neighbor clustering (FRNNC) classification algorithm. First, they make every training sample fuzzy-roughness and edit the nearest neighbor algorithm to remove training sample points in class boundary or overlapping regions. Then they use the mountain clustering method to select representative cluster center points. Then the fuzzy-rough nearest neighbor (FRNN) algorithm is applied to classify the test data. The new algorithm is applied to hand gesture image recognition. The results show that it is more effective and performs better than other nearest neighbor methods. The algorithm is recommended for use in any type of medical image application, such as medical images sequences or finding similar tumor shapes.

Hassanien [19] introduced a hybrid scheme that combines the advantages of fuzzy sets and rough sets in conjunction with statistical feature extraction techniques. An application of breast cancer imaging was chosen, and a hybridization scheme was applied to examine its ability and accuracy in classifying breast cancer images into malignant and benign cases. The scheme starts with fuzzy image processing as preprocessing techniques to enhance the contrast of the whole image, to extract the ROI, and then to enhance the edges surrounding the region of interest. Subsequently extracted features from the segmented ROIs using the GLCM are presented. The rough sets approach for generation of all reducts that contain a minimal number of features and rules is introduced. The rules can then be passed to a classifier for discrimination for different ROIs to test whether they are cancerous or noncancerous. To measure the similarity, a new rough sets distance function is presented. The experimental results showed that the hybrid scheme applied in this study performs well, reaching over 98% in overall accuracy with a minimal number of generated rules.

Image clustering analysis is one of the core techniques for image indexing, classification, identification, and segmentation for medical image processing. Mitra, Banka, and Pedrycz [49] introduced a hybrid clustering architecture in which several subsets of patterns can be processed together to find a common structure. A detailed clustering algorithm is developed by integrating the advantages of both fuzzy sets and rough sets, and a measure of quantitative analysis of the experimental results is provided for synthetic and real-world data. The structure revealed at the global level is determined by exchanging prototypes of the subsets of data and by moving prototypes of the corresponding clusters toward each other. The required communication links are thereby established at the level of cluster prototypes and partition matrices without hampering the security concerns.

Petrosino and Salvi [89] presented a multi-scale method based on the hybrid-notion of rough fuzzy sets, coming from the combination of two models of uncertainty like vagueness by handling rough sets and coarseness by handling fuzzy sets. Marrying both notions leads us to consider, for instance, approximation of sets by means of similarity relations or fuzzy partitions. The most important features are extracted from the scale spaces by unsupervised cluster analysis to successfully tackle image processing tasks. Here, we report some results achieved by applying the method to multiclass image segmentation and edge detection, but it can be successfully applied to texture discrimination problems too. Mitra, Banka, and Pedrycz's [49] and Petrosino and Salvi's [89] approaches can be applied in many medical imaging clustering problems such as image segmentation in abdomen medical images that uses cluster filter bank response vectors to obtain a compact representation of the image structures found within an image quality verification of color retina images in diabetic retinopathy screening, for example.

3.5.3 Genetic algorithm with rough sets

Lingras [42] proposed an unsupervised rough sets classification algorithm using genetic algorithms. The proposed rough sets theoretic genetic encoding will be especially useful in unsupervised learning. A rough sets genome consists of upper and lower bounds for sets in a partition. The partition may be as simple as the conventional expert class and its complement or a more general classification scheme. Lingras provides a complete description of design and implementation of rough sets genomes. The algorithm was reported to have the best classification performance compared with other unsupervised techniques, which will be useful in classification and segmentation of medical images.

Mitra, Mitra, and Pal [48] described a way of designing a hybrid system for detecting the different stages of cervical cancer. Hybridization includes the evolution of knowledge-based subnetwork modules with genetic algorithms using rough sets theory and the iterative dichotomizer 3 (ID3) algorithm. Crude subnetworks for each module are initially obtained via rough sets theory and the ID3 algorithm. These subnetworks are then combined, and the final

network is evolved using genetic algorithms. The evolution uses a restricted mutation operator, which utilizes the knowledge of the modular structure, already generated, for faster convergence. The genetic algorithm tunes the network weights and structure simultaneously. The integration enhances the performance in terms of classification score, network size, and training time as compared to the conventional multilayer perceptron (MLP). This methodology also helps in imposing a structure on the weights, which results in a network more suitable for rule extraction.

3.5.4 Bayesian and particle swarm optimization with rough sets

In many applications in computer vision and signal processing, it is necessary to assimilate data from multiple sources. This is a particularly important issue in medical imaging, where information on a patient may be available from a number of different modalities. The original rough sets model is concerned primarily with algebraic properties of approximately defined sets. The VPRS model extends the basic rough sets theory to incorporate probabilistic information. As a result, there has been much recent research interest in this area. For example, Ślęzak and Ziarko [12] presented a nonparametric modification of the VPRS model called the Bayesian rough set (BRS) model, where the set approximations are defined by using the prior probability as a reference. Mathematical properties of BRS are investigated. It is shown that the quality of BRS models can be evaluated using probabilistic gain function, which is suitable for identification and elimination of redundant features. This is a promising algorithm for generating a segmented classification while improving reconstructions of a set of registered medical images or fusing CT and single-photon emission computed tomography (SPECT) brain scans.

Swiniarski [107] described an application of rough sets and Bayesian inference to a breast cancer detection using electropotentials. The statistical PCA and the rough sets methods were applied for feature extraction, reduction, and selection. The quadratic discriminant was applied as a classifier for breast cancer detection.

Das, Abraham, and Sarkar [93] presented a framework to hybridize the rough sets theory with a famous swarm intelligence algorithm known as particle swarm optimization (PSO). The hybrid rough-PSO technique has been used for grouping the pixels of an image in its intensity space. Medical images very often become corrupted with noise. Fast and efficient segmentation of such noisy images (which is essential for their further interpretation in many cases) has remained a challenging problem for years. In their work, Das, Abraham, and Sarkar treat image segmentation as a clustering problem. Each cluster is modeled with a rough set. PSO is employed to tune the threshold and relative importance of upper and lower approximations of the rough sets. The Davies-Bouldin clustering validity index is used as the fitness function, which is minimized while arriving at an optimal partitioning.

3.5.5 Support vector machines with rough sets

Support vector machines (SVMs) are a general algorithm based on guaranteed risk bounds of statistical learning theory. They have found numerous applications in image processing and pattern recognition, particularly in medical imaging problems such as classification of brain positron emission tomography (PET) images, detection of microcalcification (MC) clusters in digital mammograms, and lung cancer nodules extraction and classification. SVMs are now established as one of the standard computational intelligence tools. They achieve good classification performances and good fault-tolerance and generalization capabilities. To combine the merits of both rough sets theory and SVMs, a hybrid classifier called rough sets support vector machines (RS-SVMs) is proposed by Zhang, Cao, and Gu [13] to recognize radar emitter signals. Rough sets theory is used as a preprocessing step to improve the performances of SVMs. A large number of experimental results showed that RS-SVMs achieve lower recognition error rates than SVMs, and RS-SVMs have stronger classification and generalization capabilities than SVMs, especially when the number of training samples is small. RS-SVMs are greatly superior to SVMs.

SVMs are essentially binary classifiers. To improve their applicability, several methods have been suggested for extending SVMs for multiclassification, including one-versus-one (1-v-1), one-versus-rest (1-v-r), and decision-directed acyclic graph support vector machines (DAGSVM). Lingras and Butz [44] described how binary classification with SVMs can be interpreted using rough sets and how rough sets theory may help in reducing the storage requirements of the 1-v-1 approach in the operational phase. Their rough sets approach to SVM classification precludes exact classification and is especially useful when dealing with noisy data. By utilizing the boundary region in rough sets, they suggested two new approaches, extensions of 1-v-r and 1-v-1 to SVM multiclassifications that allow for an error rate. They explicitly demonstrate how their extended 1-v-r may shorten the training time of the conventional 1-v-r approach. In addition, they showed that their 1-v-1 approach may have reduced storage requirements compared to the conventional 1-v-1 and DAGSVM techniques. Their techniques provided better semantic interpretations of the classification process. The theoretical conclusions are supported by experimental findings involving a synthetic data set. The presented work is useful for soft margin classifiers in solving medical imaging problems, especially a multiclass classification system for medical images [5].

3.6 Challenges and Future Directions

Rough sets theory is a fairly new intelligent technique that has been applied to the medical domain and is used for discovering data dependencies, evaluating the importance of features, discovering the patterns of data, reducing

all redundant objects and features, and seeking the minimum subset of features. It is also being used for extracting rules from databases. Most of the current literature on rough sets–based methods for medical imaging addresses the classification and feature reduction issues. A few papers deal with medical imaging problems such as image segmentation, image filtering, and voxel representation. Medical imaging presents a challenging and fruitful area for rough sets to play crucial roles in resolving problems and providing solutions that understand, represent, and process the images, their segments, and their features as rough sets.

Near sets [25,61–66] represent a generalization of the rough set approach to the classification of objects introduced by Pawlak [67]. Near sets and rough sets are very much like two sides of the same coin. From a rough set point of view, the focus is on the approximation of sets with nonempty boundaries. By contrast, in a near sets approach to set approximation, the focus is on the discovery of near sets in cases where there is either a nonempty or an empty approximation boundary. There are a number of practical applications for the near sets approach, such as feature selection [65], object recognition in images [25,64], image processing [6], granular computing [74,78], and various forms of machine learning [71,73,75–78,86,87]. Near sets–based methods are another challenge that offer a generalization of traditional rough sets theory [67–69] and a new approach to classifying perceptual objects by means of features [70,71,76–79]. The near sets approach can be used to classify images that are qualitatively but not necessary quantitatively close to each other. This is essentially the idea expressed in classifying images in Henry and Peters [25] and Peters [64]. If one adopts the near sets approach in image processing, a byproduct of the approach is the separation of images into nonoverlapping sets of images that are similar to (*descriptively* near to) each other. This has recently led to an application of the near sets approach in 2-D and 3-D interactive gaming with a vision system that learns and serves as the backbone for an adaptive telerehabilitation system for patients with finger, hand, arm, and balance disabilities. Each remote node in the telerehabilitation system includes a vision system that learns to track the behavior of a patient [64].

A combination of computational intelligence techniques in application areas of pattern recognition, particularly in medical imaging problems, has become one of the most important focuses of research on intelligent information processing. Neural networks shows a strong ability to solve complex problems for medical image processing. From the perspective of the specific rough sets approaches that need to be applied, explorations into possible applications of hybrid rough sets with other intelligent systems (e.g., neural networks, genetic algorithms, fuzzy approaches) to image processing and pattern recognition could lead to new and interesting avenues of research. Finally, even though many rough-based approaches are being proposed for various applications in the medical field, their impact is mostly confined to academic circles. These methods are yet to find wide acceptance in industrial circles and to become incorporated in industrial products. This trend is also evident from

the very small number of industrial patents in this direction. Hence, the main challenge for rough sets researchers is to provide industry leaders a convincing demonstration of the superiority of their approaches over the traditional methods.

In conclusion, many successful algorithms applied in medical imaging have been reported in the literature, and the applications of rough sets in medical image processing have to be analyzed individually. Rough sets is a new way to deal with the issues that cannot be addressed by traditional image processing algorithms or by other classification techniques. By introducing rough sets, algorithms developed for medical imaging and pattern recognition often become more intelligent and robust, providing a human-interpretable, low-cost, exact-enough solution, as compared to traditional techniques. This chapter provided a focused survey on a range of rough sets and their applications to medical imaging. Our main purpose was to present to the rough sets and medical imaging research communities the state of the art in rough sets applications to image processing and pattern recognition, particularly in medical imaging, and to inspire further research and development on new applications and concepts in trend-setting directions and in exploiting rough sets.

References

[1] Ahmed, M. N., Yamany, S. M., Nevin, M., and Farag, A. A. (2003) A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 21(3): 193–99.

[2] Bazan, J., Nguyen, H. S., Nguyen, S. H., Synak, P., and Wrblewski, J. (2000) Rough set algorithms in classification problems. In L. Polkowski, S. Tsumoto, and T. Y. Lin (Eds.), *Rough Set Methods and Applications.* New York: Physica-Verlag, pp. 49–88.

[3] Bezdek, J. C., Ehrlich, R., and Full, W. (1984) FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, 10: 191–203.

[4] Bezdek, J. C., and Pal, S. K. (1992) *Fuzzy Models for Pattern Recognition.* New York: IEEE Press.

[5] Qiu, B., Xu, C. S., and Tian, Q. (2006) An automatic classification system applied in medical images. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1045–48.

[6] Borkowski, M., and Peters, J. F. (2006) Matching 2D image segments with genetic algorithms and approximation spaces. In J. F. Peters and A. Skonron (Eds.), *Transactions on Rough Sets, V*, Lecture Notes in Computer Science 4100, Berlin: Springer, pp. 63–101.

[7] Bovis, K. J., and Singh, S. (2002) Learning the optimal contrast enhancement of mammographic breast masses. In *Proceedings of the 6th International Workshop on Digital Mammography*, Bremen, Germany, 22–25 June.

[8] Chen, C. W., Luo, J. B., and Parker, K. J. (1998) Image segmentation via adaptive k-mean clustering and knowledge-based morphological operations with biomedical applications. *IEEE Transactions on Image Processing*, 7(12): 1673–83.

[9] Chena, C.-B., and Wang, L.-Y. (2006) Rough set-based clustering with refinement using Shannon's entropy theory. *Computers and Mathematics with Applications*, 52(10–11): 1563–76.

[10] Chiang, F., and Braun, R. (2004) Intelligent failure domain prediction in complex telecommunication networks with hybrid rough sets and adaptive neural nets. *3rd International Information and Telecommunication Technologies Symposium (I2TS '04)*, December 6–9, Sao Carlos, SP, Brazil, Sao Carlos Federal University.

[11] Mao, C.-W., Liu, S.-H., and Lin, J.-S. (2003) Classification of multispectral images through a rough-fuzzy neural network. *Optical Engineering*, 43(1): 103–12.

[12] Ślęzak, D., and Ziarko, W. (2005) The investigation of the Bayesian rough set model. *International Journal of Approximate Reasoning, Data Mining and Granular Computing*, 40(1–2): 81–91.

[13] Zhang, G., Cao, Z., and Gu, Y. (2005) A hybrid classifier based on rough set theory and support vector machines. In L. Wang, L. Jiao, G. Shi, X. Li, and J. Liu (Eds.), *Fuzzy Systems and Knowledge Discovery*, Lecture Notes in Computer Science 3613. Berlin: Springer, pp. 1287–96.

[14] Greco, S., Inuiguchi, M., and Slowinski, R. (2006) Fuzzy rough sets and multiple-premise gradual decision rules. *International Journal of Approximate Reasoning*, 41: 179–211.

[15] Wang, G. Y. (2001) Algebra view and information view of rough sets theory. In *Proceedings of SPIE: Data Mining and Knowledge Discovery: Theory, Tools, and Technology III*, Belur V, Vol. 4384, pp. 200–207.

[16] Wang, G. Y. (2003) Rough reduction in algebra view and information view. *International Journal of Intelligent Systems*, 18(6): 679–88.

[17] Hu, Q., Yu, D., and Xie, Z. (2006) Information-preserving hybrid data reduction based on fuzzy-rough techniques. *Pattern Recognition Letters*, 27: 414–23.

[18] Hartigan, J. A., and Wong, M. A. (1979) Algorithm AS136: A k-means clustering algorithm. *Applied Statistics*, 28: 100–108.

[19] Hassanien, A. E. (2006) Fuzzy-rough hybrid scheme for breast cancer detection. *Image and Computer Vision Journal*, 25(2): 172–83.

[20] Hassanien, A. E. (2004) Rough set approach for attribute reduction and rule generation: A case of patients with suspected breast cancer. *Journal of the American Society for Information Science and Technology*, 55(11): 954–62.

[21] Hassanien, A. E., Ali, J. M., and Hajime, N. (2004) Detection of spiculated masses in mammograms based on fuzzy image processing. L. Rutkowski, J. Siekmann, R. Tadeusiewicz, and L. A. Zadeh (Eds.) *Artificial Intelligence and Soft Computing—(ICAISC 2004)* LNAI 3070. Berlin: Springer, pp. 1002–1007.

[22] Hassanien, A. E. (2003) Classification and feature selection of breast cancer data based on decision tree algorithm. *International Journal of Studies in Informatics and Control*, 12(1): 33–39.

[23] Hassanien, A. E., and Ślęzak, D. (2006) Rough neural intelligent approach for image classification: A case of patients with suspected breast cancer, *International Journal of Hybrid Intelligent Systems*, 3(4): 205–18.

[24] Haralick, R. M. (1979) Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5): 786–804.

[25] Henry, C., and Peters, J. F. (2007) Image pattern recognition using approximation spaces and near sets. In *Proceedings of the 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2007), Joint Rough Set Symposium (JRS 2007),* LNAI 4482, pp. 475–82.

[26] Hirano, S., and Tsumoto, S. (2005) Rough representation of a region of interest in medical images. *International Journal of Approximate Reasoning*, 40: 23–34.

[27] Hu, X., Lin, T. Y., and Han J. (2004) A new rough sets model based on database systems. *Fundamenta Informaticae*, 59(2–3): 135–52.

[28] Karssemeijer, N. (1993) Adaptive noise equalization and image analysis in mammography. *Proceedings of the 13th International Conference on Information Processing in Medical Imaging.* Berlin: Springer, pp. 472–486.

[29] Slowinski, K., Stefanowski, J., and Siwinski, D. (2002) Application of rule induction and rough sets to verification of magnetic resonance diagnosis. *Fundamenta Informaticae*, 53(3–4): 345–63.

[30] Liang, J., and Shi, Z. (2004) The information entropy, rough entropy and knowledge granulation in rough set theory. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(1): 37–46.

[31] Yun, J., Zhanhuai, L., Yong, W., and Longbo, Z. (2005) Joining associative classifier for medical images. In *Proceedings of the Fifth International Conference on Hybrid Intelligent Systems HIS05*, 6–9 Nov. 2005, pp. 367–72.

[32] Yun, J., Zhanhuai, L., Yong, W., and Longbo, Z. (2006) A better classifier based on rough set and neural network for medical images. In *Proceedings of the Sixth IEEE International Conference on Data Mining*, pp. 853–57.

[33] Cyran, K. A., and Mrózek, A. (2001) Rough sets in hybrid methods for pattern recognition. *International Journal of Intelligent Systems*, 16(2): 149–68.

[34] Karssemeijer, N., Thijssen, M., Hendriks, J., and Van Erning, L. (Eds.). (1998) *Digital Mammography: Computational Imaging and Vision*. Dordrecht: Kluwer Academic Publishers.

[35] Kerre, E., and Nachtegael, M., eds. (2000) Fuzzy techniques in image processing: Techniques and applications. In *Studies in Fuzziness and Soft Computing*, Vol. 52. Heidelberg: Springer, pp. 137–71.

[36] Kobashi, S., Kondo, K., and Hata, Y. (2004) Rough sets based medical image segmentation with connectedness. In *Proceedings of the 5th International Forum on Multimedia and Image Processing (IFMIP)*, pp. 197–202.

[37] Kobashi, S., Kondo, K., and Hata, Y. (2003) A new approach of medical image segmentation based on rough sets, IEICE Technical Report. Institute of Electronics, Information and Communication Engineers, Japan, Vol. 103, No. 597, pp. 61–66.

[38] Li, J., and Cercone, N. (2005) A rough set based model to rank the importance of association rules. In *Proceedings of RSFDGrC '2005,* LNAI 2. Berlin: Springer, pp. 109–118.

[39] Lingras, P. (1996) Rough neural networks. In *Proceedings of the 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU96)*. London: Springer, pp. 1445–50.

[40] Lingras, P., Hogo, M., and Snorek, M. (2004) Interval set clustering of Web users using modified Kohonen self-organizing maps based on the properties of rough sets. *Web Intelligence and Agent Systems: An International Journal*, 2(3): 217–25.

[41] Lingras, P., and West, C. (2004) Interval set clustering of Web Users with rough k-means. *Journal of Intelligent Information Systems*, 23(1): 516.

[42] Lingras, P. (2001) Unsupervised rough set classification using GAs. *Journal of Intelligent Information Systems*, 16(3): 215–28.

[43] Lingras, P. (2007) Applications of rough set based k-means, Kohonen, GA clustering. *Transactions on Rough Sets*, VII: 120–39.

[44] Lingras, P., and Butz, C. (2007) Rough set based 1-v-1 and 1-v-r approaches to support vector machine multi-classification. *Information Sciences: An International Journal*, 177(18): 3782–98.

[45] MacKay, D. J. C. (2003) *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.

[46] Matsuura, T., Kobashi, S., Kondo, K., and Hata Y. (2002) An image classification based on new similarity measure in rough sets. In *Proceedings of the 1st International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Vol. 1, pp. 197–201, November.

[47] Mitra, S. (2004) An evolutionary rough partitive clustering. *Pattern Recognition Letters*, 25: 1439–49.

[48] Mitra, P., Mitra, S., and Pal, S. K. (2001) Evolutionary modular MLP with rough sets and ID3 algorithm for staging of cervical cancer. *Neural Computing and Applications*, 10(1): 67–76.

[49] Mitra, S., Banka, H., and Pedrycz, W. (2006) Rough-fuzzy collaborative clustering systems. *IEEE Transactions on Man and Cybernetics*, Part B, 36(4): 795–805.

[50] Mammographic Image Analysis Society: Mini-Mammography Database. (2003) http://www.wiau.man.ac.uk/services/MIAS/MIASmini.htm.

[51] Zorman, M., Kokol, P., Lenic, M., Sanchez de la Rosa, J. L., Sigut, J. F., and Alayon, S. (2007) Symbol-based machine learning approach for supervised segmentation of follicular lymphoma images. In *Proceedings of the 20th IEEE International Symposium on Computer-Based Medical Systems*, pp. 115–20.

[52] Mohabey, A., and Ray, A. K. (2000) Rough set theory based segmentation of color images. In *Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society, 2000 (NAFIPS)*, pp. 338–42.

[53] Mohabey, A., and Ray, A. K. (2000) Fusion of rough set theoretic approximations and FCM for color image segmentation. *IEEE International Conference on Systems, Man, and Cybernetics*, 2: 1529–34.

[54] Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., and Nowinski, W. L. (2006) Medical image segmentation using k-means clustering and improved watershed algorithm. In *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 61–65.

[55] Nachtegael, M., Van der Weken, M., Van DeVille, D., Kerre, D., Philips, W., and Lemahieu, I. (2001) An overview of classical and fuzzy-classical filters for noise reduction. In *Proceedings of the 10th International IEEE Conference on Fuzzy Systems (FUZZ-IEEE' 2001)*, Melbourne, Australia, pp. 3–6.

[56] Ning, S., Ziarko, W., Hamilton, J., and Cercone, N. (1995) Using rough sets as tools for knowledge discovery. In U. M. Fayyad and R. Uthurusamy (Eds.), *Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining (KDD'95)*, Montreal: Que, pp. 263–68.

[57] Pawlak, Z. (1991) Rough sets: *Theoretical Aspects of Reasoning about Data*. Dordrecht: Kluwer Academic Publishers.

[58] Pawlak, Z., Grzymala-Busse, J., Slowinski, R., and Ziarko, W. (1995) Rough sets. *Communications of the ACM*, 38(11): 88–95.

[59] Komorowski, J., Pawlak, Z., Polkowski, L., and Skowron, A. (1999) Rough sets: A tutorial. In S. K. Pal and A. Skowron (Eds.), *Rough Fuzzy Hybridization. A New Trend in Decision-Making*. Singapore: Springer, p. 398.

[60] Polkowski, L. (2003) *Rough Sets: Mathematical Foundations*. Heidelberg: Physica-Verlag.

[61] Peters, J. F. (2007) Near sets: Special theory about nearness of objects. *Fundamenta Informaticae*, 75(1–4): 407–33.

[62] Peters, J. F. (2007) Near sets: General theory about nearness of objects. *Applied Mathematical Sciences*, 1(53): 2609–29.

[63] Peters, J. F., Skowron, A., and Stepaniuk, J. (2007) Nearness of objects: Extension of approximation space model. *Fundamenta Informaticae*, 79(3): 497.

[64] Peters, J. F. (2007) Near sets. Toward approximation space-based object recognition. In Y. Yao, P. Lingras, W.-Z. Wu, et al. (Eds.), *Proceedings of the 2nd International Conference on Rough Sets and Knowledge Technology (RSKT07), Joint Rough Set Symposium (JRS07)*, LNAI Berlin: Springer, pp. 22–33.

[65] Peters, J. F., and Ramanna, S. (2008) Feature selection: Near set approach. In Z. W. Ras, S. Tsumoto, and D. A. Zighed (Eds.), *Proceedings of the 3rd International Workshop on Mining Complex Data (MCD '07)*, ECML/PKDD-2007, LNAI. Berlin: Springer.

[66] Peters, J. F., Skowron, A., and Stepaniuk, J. (2006) Nearness in approximation spaces. G. Lindemann, H. Schlilngloff, et al. (Eds.), *Proceedings Concurrency, Specification & Programming (CS&P 2006)*. Informatik-Berichte Nr. 206, Humboldt-Universität zu Berlin, pp. 434–45.

[67] Pawlak, Z. (1981) *Classification of Objects by Means of Attributes*. Institute for Computer Science, Polish Academy of Sciences, Report 429.

[68] Pawlak, Z. (1982) Rough sets. *International Journal of Computing and Information Sciences*, 11: 341–56.

[69] Pawlak, Z., and Skowron A. (2007) Rudiments of rough sets. *Information Sciences*, 177: 3–27.

[70] Peters, J. F. (2008) Classification of perceptual objects by means of features. *International Journal of Information Technology and Intelligent Computing*.

[71] Lockery, D., and Peters, J. F. (2007) Robotic target tracking with approximation space-based feedback during reinforcement learning. In *Proceedings of the 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2007), Joint Rough Set Symposium (JRS 2007),* LNAI 4482, Berlin: Springer, pp. 483–90.

[72] Peters, J. F. (2007) Classification of objects by means of features. In *Proceedings of the IEEE Symposium Series on Foundations of Computational Intelligence (IEEE SCCI 2007)*, Honolulu, Hawaii, 1–8.

[73] Peters, J. F. (2007) Granular computing in approximate adaptive learning. *International Journal of Information Technology and Intelligent Computing*, in press.

[74] Skowron, A., and Peters, J. F. (2008) Rough granular computing. In W. Pedrycz, A. Skowron, and V. Kreinovich (Eds.), *Handbook on Granular Computing*. New York: Wiley, pp. 285–329.

[75] Peters, J. F., and Henry, C. (2006) Reinforcement learning with approximation spaces. *Fundamenta Informaticae*, 71(2–3): 323–49.

[76] Peters, J. F., Borkowski, M., Henry, C., and Lockery, D. (2006) Monocular vision system that learns with approximation spaces. In A. Ella, P. Lingras, D. Sleßzak, and Z. Suraj (Eds.), *Rough Set Computing: Toward Perception Based Computing*. Hershey, Penn.: Idea Group Publishing, pp. 1–22.

[77] Peters, J. F., Borkowski, M., Henry, C., Lockery, D., Gunderson, D., and Ramanna, S. (2006) Line-crawling bots that inspect electric power transmission line equipment. In *Proceedings of the 3rd International Conference on Autonomous Robots and Agents 2006 (ICARA 2006)*, Palmerston North, New Zealand, pp. 39–44.

[78] Peters, J. F. (2008) Perceptual granulation in ethology-based reinforcement learning. In W. Pedrycz, A. Skowron, and V. Kreinovich (Eds.), *Handbook on Granular Computing*. New York: Wiley, pp. 671–88.

[79] Peters, J. F., and Borkowski, M. (2004) K-means indiscernibility relation over pixels. In *Proceedings of the 4th International Conference on Rough Sets and Current Trends in Computing (RSCTC 2004)*. Uppsala, Sweden, 1–5 June, pp. 580–85.

[80] Peters, J. F., and Pedrycz, W. (2008) Computational intelligence. In *EEE Encyclopedia*. New York: Wiley.

[81] Peters, J. F., Liting, H., and Ramanna, S. (2001) Rough neural computing in signal analysis. *Computational Intelligence*, 17(3): 493–513.

[82] Peters, J. F., Skowron, A., Liting, H., and Ramanna, S. (2000) Towards rough neural computing based on rough membership functions: Theory and application. In W. Ziarko and Y. Y. Yao (Eds.), *Second International Conference on Rough Sets and Current Trends in Computing, RSCTC'2000*, Lecture Notes in Computer Science 2005. Berlin: Springer, pp. 611–618.

[83] Peters, J. F., Skowron, A., Suraj, Z., Rzasa, W., and Borkowski, M. (2002) Clustering: A rough set approach to constructing information granules. In *Proceedings of 6th International Conference on Soft Computing and Distributed Processing (SCDP 2002)*, pp. 57–61.

[84] Peters, J. F. (2007) Granular computing in approximate adaptive learning. *International Journal of Information Technology and Intelligent Computing*, 2(4): 1–25.

[85] Peters, J. F., Henry, C., and Gunderson, D. S. (2007) Biologically-inspired approximate adaptive learning control strategies: A rough set approach. *International Journal of Hybrid Intelligent Systems*, 4(4): 203–16.

[86] Peters, J. F., Shahfar, S., Ramanna, S., and Szturm, T. (2007) Biologically-inspired adaptive learning: A near set approach. In *Proceedings of Frontiers in the Convergence of Bioscience and Information Technologies (FBIT07)*, IEEE, NJ, 11 pp. 403–408.

[87] Peters, J. F. (2007) Toward approximate adaptive learning. In M. Kkryszkiewicz, J. F. Peters, H. Rybinski, and A. Skowron (Eds.), *Rough Sets and Emerging Intelligent Systems Paradigms in Memoriam Zdzislaw Pawlak*, LNAI 4585, Berlin: Springer, pp. 57–68.

[88] Peters, J. F., and Henry, C. (2007) Approximation spaces in off-policy Monte Carlo learning. *Engineering Applications of Artificial Intelligence. The International Journal of Intelligent Real-Time Automation*, 20(5): 667–75.

[89] Petrosino, A., and Salvi, G. (2006) Rough fuzzy set based scale space transforms and their use in image analysis. *International Journal of Approximate Reasoning*, 41(2): 212–28.

[90] Pham, D. L. (2004) Sizing tumors with TNM classifications and rough sets method. In *Proceedings of the 17th IEEE Symposium on Computer-Based Medical Systems*, 24–25 June, pp. 221–23.

[91] Shankar, B. U. (2007) Novel classification and segmentation techniques with application to remotely sensed images. *Transactions on Rough Sets*, VII, LNCS 4400, pp. 295–80.

[92] Chandana, S. and Mayorga, R. V. (2006) RANFIS: Rough adaptive neuro-fuzzy inference system. *International Journal of Computational Intelligence*, 3(4): 289–95.

[93] Das, S., Abraham, A., and Sarkar, S. K. (2006) A hybrid rough set–particle swarm algorithm for image pixel classification. In *Proceedings of the 6th International Conference on Hybrid Intelligent Systems*, December 13–15, pp. 26–32.

[94] Setiono, R. (2000) Generating concise and accurate classification rules for breast cancer diagnosis. *Artificial Intelligence in Medicine*, 18(3): 205–19.

[95] Ślęzak, D. (2000) Various approaches to reasoning with frequency-based decision reducts: A survey. In L. Polkowski, S. Tsumoto, and T. Y. Lin (Eds.), *Rough Sets in Soft Computing and Knowledge Discovery: New Developments*. Heidelberg: Physica Verlag.

[96] Starzyk, J. A., Dale, N., and Sturtz, K. (2000) A mathematical foundation for improved reduct generation in information systems. *Knowledge and Information Systems Journal*, 2(2): 131–47.

[97] Swiniarski, R., and Skowron, A. (2003) Rough set methods in feature selection and recognition. *Pattern Recognition Letters*, 24(6): 833–49.

[98] Pal, S. K. (2001) Fuzzy image processing and recognition: Uncertainties handling and applications. *International Journal of Image Graphics*, 1(2): 169–95.

[99] Pal, S. K., Shankar, B. U., and Mitra, P. (2005) Granular computing, rough entropy and object extraction. *Pattern Recognition Letters*, 26(16): 2509–17.

[100] Pal, S. K., and Mitra, P. (2002) Multispectral image segmentation using rough set initialized EM algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 40: 2495–2501.

[101] Pal, S. K., Polkowski, L., and Skowron, A. (Eds.). (2004) *Rough-Neuro Computing: Techniques for Computing with Words*. London: Springer.

[102] Shen, Q., and Chouchoulas, A. (2002) A rough-fuzzy approach for generating classification rules. *Pattern Recognition*, 35(11): 2425–38.

[103] Slowinski, K., Slowinski, R., and Stefanowski, J. (1988) Rough set approach to analysis of data from peritoneal lavage in acute pancreatitis. *Medical Informatics*, 13: 143–59.

[104] Shang, C., and Shen, Q. (2002) Rough feature selection for neural network based image classification. *International Journal of Image and Graphics*, 2(4): 541–55.

[105] Hirano, S., and Tsumoto, S. (2002) Segmentation of medical images based on approximations in rough set theory, rough sets and current trends in computing. In *Proceedings of the 3rd International Conference on Rough Sets and Current Trends in Computing (RSCTC 2002)*, Malvern, PA, October 14–16, pp. 950–51.

[106] Tsumoto, S. (2004) Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model. *Information Sciences: An International Journal*, 162(2): 65–80.

[107] Swiniarski, R. W. (1008) Rough sets and Bayesian methods applied to cancer detection. In *Proceedings of the 1st International Conference on Rough Sets and Current Trends in Computing (RSCTC'98)*, Warsaw, Poland, pp. 609–16.

[108] Swiniarski, R. W., and Hargis, L. (2001) Rough sets as a front end of neural-networks texture classifiers. *Neurocomputing*, 36(1–4): 85–102.

[109] Swiniarski, R. W., and Skowron, A. (2003) Rough set methods in feature selection and recognition. *Pattern Recognition Letters*, 24: 833–49.

[110] Swiniarski, R. W., Lim, H. K., Shin, J. H., and Skowron, A. (2006) Independent component analysis, principal component analysis and rough sets in hybrid mammogram classification. In *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition, (IPCV '06)*, pp. 640–45.

[111] Thangavel, K., Karnan, M., and Pethalakshmi, A. (2005) Performance analysis of rough reduct algorithms in image mammogram. *ICGST International Journal on Graphics, Vision and Image Processing*, 8: 13–21.

[112] Wojcik, Z. (1987) Rough approximation of shapes in pattern recognition. *Computer Vision, Graphics, and Image Processing*, 40: 228–49.

[113] Voges, K. E., Pope, N. K. Ll., and Brown, M. R. (2002) Cluster analysis of marketing data: A comparison of k-means, rough set, and rough genetic approaches. In H. A. Abbas, R. A. Sarker, and C. S. Newton (Eds.), *Heuristics and Optimization for Knowledge Discovery*. Hershey, Penna.: Idea Group Publishing, pp. 208–16.

[114] Widz, S., Revett, K., and Ślęzak, D. (2004) Application of rough set based dynamic parameter optimization to MRI segmentation. In *Proceedings of the 23rd International Conference of the North American Fuzzy Information Processing Society (NAFIPS 2004)*. Banff, Canada.

[115] Widz, S., Ślęzak, D., and Revett, K. (2004) An automated multispectral MRI segmentation algorithm using approximate reducts. In *Proceedings of the 4th International Conference on Rough Sets and Current Trends in Computing (RSCTC 2004)*, LNAI 3066. Berlin: Springer, pp. 815–24.

[116] Widz, S., Ślęzak, D., and Revett, K. (2004) Application of rough set based dynamic parameter optimization to MRI segmentation. *IEEE Annual Meeting of the Information (Processing NAFIPS)*, 1: 440–45.

[117] Wang, X., Yang, J., Teng, X., and Peng, N. (2005) Fuzzy classifiers fuzzy rough set based nearest neighbor clustering classification algorithm. In *Fuzzy Systems and Knowledge Discovery*, Lecture Notes in Computer Science, Vol. 3613. Berlin: Springer, pp. 370–73.

[118] Yan, M. X. H., and Karp, J. S. (1004) Segmentation of 3D brain MR using an adaptive K-means clustering algorithm. *IEEE Conference on Nuclear Science Symposium and Medical Imaging*, 4(30): 1529–33.

[119] Yao, Y. Y. (1998) A comparative study of fuzzy sets and rough sets. *Journal of Information Sciences*, 109: 227–42.

[120] Zhang, Y., Tomuro, N., Raicu, D. S., and J. D. Furst, J. D. (2006) Investigation on feature selection to improve classification of abdominal organs in CT images. CTI Research Symposium, April 2006.

[121] Zadeh, L. A. (1965) Fuzzy sets. *Information and Control*, 8: 338–53.

[122] Ziarko, W. (1993) Variable precision rough sets model. *Journal of Computer and Systems Sciences*, 46(1): 39–59.

[123] Wojcik, Z. (1987) Rough approximation of shapes in pattern recognition. *Computer Vision, Graphics, and Image Processing*, 40(2): 228–49.

[124] Pal, N. R and, Pal, S. K. (1991) Entropy: A New Definition and its Applications. IEEE Trans. Systems man Cybernet. SMC-21, pp. 1260–1270.

Chapter 4

Early Detection of Wound Inflammation by Color Analysis

Peter Plassmann and Brahima Belem

Contents

Chronic wounds such as leg ulcers are a common problem, particularly for patients suffering from immunodeficiency, old age, side effects of medication, or circulatory problems. One of the main concerns regarding chronic wounds is their long healing time, during which they remain exposed to foreign bodies capable of triggering infection. In turn, infection delays healing even further, with several consequences, particularly the prolonged suffering for the patient and the financial resources required to provide treatment.

One of the complications that prevent a wound from healing is bacterial or viral infection. An episode of infection may last several weeks, is very painful, and is likely to significantly delay wound healing. There is an obvious relationship between wound color and wound inflammation that may be the result of an infection. This relationship is studied in this chapter.

In a first step, wound images and associated medical status information are collected. In a second step, the wound is identified manually in the images as the region of interest (ROI). Each ROI is then segmented automatically in chromaticity space into smaller ROIs, separating healthy and nonhealthy tissues. Statistical features are calculated for the different tissue types, and a "brute force" approach is used to select the most successful ones for subsequent classification.

Three classification engines are presented and contrasted in this study: logistic regression (LR), an artificial neural network (ANN), and a support vector machine (SVM). LRs and ANNs are widely recognized, while SVMs are an innovative approach providing a promising alternative. The classification engines are compared against each other and the clinicians' classification using the kappa agreement estimator. Results indicate that the best-performing classification engine, SVM, is more consistent in its judgment than are most clinical practitioners.

4.1 Introduction

Acute wounds of the human skin may be caused by trauma or surgical intervention. This type of wound usually heals quickly without complications. Chronic wounds, in contrast, may be caused by poor blood perfusion, a weakened immune system, old age, certain diseases of the vascular system, or as a side effect of medication such as steroids.

In an average hospital population, up to 8% of all bedridden or chair-bound patients suffer from pressure ulcers, causing pain for the patient [5]. Chronic wounds heal slowly: months as opposed to weeks [2] are considered "normal" for the healing process, resulting in enormous health care costs.

One way to improve wound healing is to constantly monitor wound size: if the physical dimensions of a wound are assessed at regular intervals, clinicians know if the patient is responding to a particular treatment regime and can change it if necessary. Early methods of wound measurement used alginate casts, transparent films, and rulers to measure volume and area. Modern methods are based on photography or stereophotogrammetry. Because wound healing, especially in chronic wounds, tends to be slow and measurements show only past changes in the healing rate, the measurement of wound size alone can lead to reactive but not to proactive treatment. Based on a previous pilot study [4], this chapter suggests that the measurement of wound color may enable a proactive approach.

4.1.1 Background

Repeated episodes of infection are a common complication in the treatment of chronic wounds; approximately half of all wounds undergo at least one period of infection. Each episode of infection can delay the progress of healing by several weeks.

One way to improve wound monitoring may be to measure its color properties: anecdotal evidence from clinicians suggests that an infected wound looks "angry," while a well-progressing lesion shows "nice" pink granulation tissue. If a wound is presenting angry colors, an infection is usually the underlying cause because the immune system's response to infection causes inflammation, which to the eye appears to have a red-purplish hue, distinctively different from the pink appearance of healthy granulation tissue. A 1996 pilot study [4] based on a case study in 1993 claimed that simple digital image processing techniques can correctly classify inflamed and not inflamed wounds in the majority of cases (80%). This early work was limited in both image capture and processing technology but provided the motivation for the work presented here.

Unfortunately, it is usually only at the late stage of an infection, when color changes are obvious and the patient is reporting pain, that treatment of the underlying infection commences. If the presence of infection could be detected at onset, wound deterioration could be potentially avoided, reducing treatment costs and patient discomfort.

In this context, it is important to note that although inflammation correlates with infection, it is not necessarily the sole underlying cause. In the first stages of a normal wound healing process, for example, inflammation is an expected reaction of the body to injury and is necessary to kick-start a chain of processes called the *wound healing cascade*. Other mechanisms, such as irritation due to chemical or physical stimulation, can also cause inflammation. Conversely, infection does not always result in inflammation if, for example, the body defense mechanisms are not functioning properly. Finally, infection should not be confused with bacterial or viral colonization of the wound. This is normal and it is only when the natural balance of microorganisms is disturbed or one particular agent multiplies aggressively that the term infection is appropriate and inflammation with its associated color changes takes place. Bearing this in mind, it is likely that an image processing technique capable of identifying and classifying subtle color changes should be able to detect the presence of inflammation in a chronic wound at an early stage and should therefore be a powerful tool for facilitating proactive wound treatment.

4.1.2 Contribution of this chapter

This chapter illustrates how data reduction techniques can be used to address the common problem of input data reduction for classification algorithms. Using the example of wound chromaticity, the usefulness of various color spaces and potential classification parameters is examined. We explain how useful parameters can be identified and how the number and type of

parameter combinations used as inputs to the classification algorithms can be reduced.

Three different classification methods (LR, ANN and SVM) are explored and applied to the wound classification problem used for demonstration. Then we addresses the "missing expert" problem common to many classification applications. In this instance, the problem is that at present there is no 100% correct classification system or infallible human expert for wound inflammation that can be used as a reliable benchmark. A methodology based on the kappa descriptor is therefore suggested that amalgamates a set of likelihood classification opinions into an artificial but reliable "expert."

4.2 Method Overview

One hundred fifty-one wound images were acquired under tightly controlled and repeatable lighting and image capture conditions and then subjected to a sequence of image preprocessing steps in order to reduce stochastic noise and undesired artifacts such as specular reflections to arrive at the raw data set shown at the top of Figure 4.1. In a second step, the wounds as ROIs are manually segmented from the surrounding healthy tissue. An ANN is trained on half of the data set in order to learn how to segment the wound ROI into two tissue types (i.e., sloughy and granulating), which are thought to respond differently to inflammation. In an iterative process using the other half of the data set, the performance of the ANN is analyzed, and the process is repeated until the segmentation error is below a predefined error value (DeltaE).

The wound segmentation process shown in Figure 4.1 uses a standard ANN and will not be expanded any further. Instead, this chapter focuses on two problem areas commonly found in computational decision making based on image content: information reduction and the subsequent selection of the most suitable decision-making engine.

For information reduction, the segmented wound images undergo a feature extraction process whereby the information contained in each segment is condensed into a set of 4 statistical moment classes (histogram mean, standard deviation, skewness, and kurtosis) using a range of color spaces. The result is a set of 37 features for each wound ROI.

The images thus preprocessed are then randomly split into two data sets, one for training or priming the three different classification algorithms used here and a second test set for evaluating the performance of the classification algorithms on new data. While some of the techniques applied have already been used in the context of wound healing research (e.g., Hoppe's segmentation routines [4]), others that have never been used on this type of problem (e.g., artificial neural network, logistic regression, and the support vector machine) are also implemented and tested.

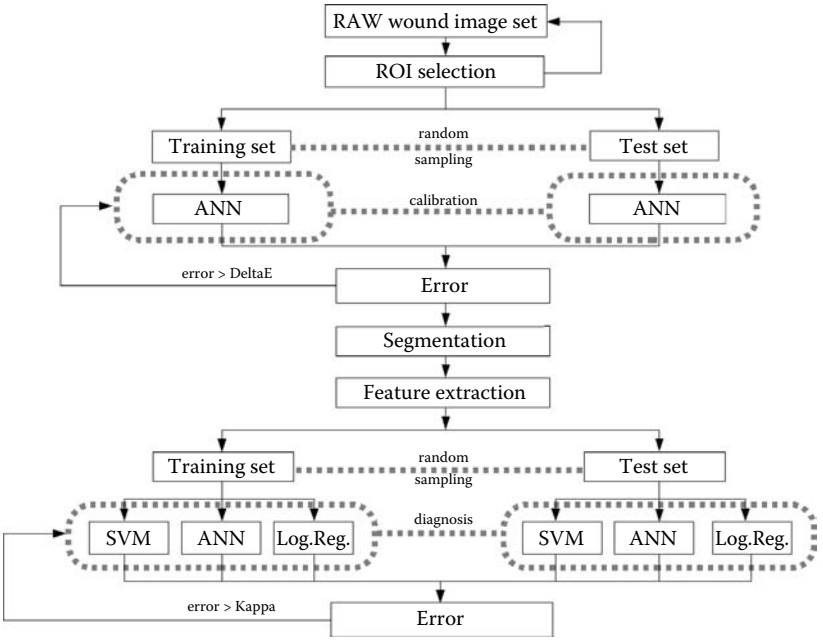


FIGURE 4.1: Flow of information from raw wound image data, via tissue type segmentation, to the extraction of features used as input to three different classification engines.

4.2.1 Information extraction and data reduction

This section introduces a variety of features that could subsequently be used by classification methods. In order to demonstrate the properties of the features and classification methods more clearly, a set of artificial wound images with precisely known characteristics is used first.

4.2.1.1 Definition of features

Unfortunately, the multidimensionality of the input data is close to the point of being impractical: features can be extracted from a variety of color spaces (the four used here are RGB, CIEL*u*v*, CIEL*C*h*, and CIEL*s*h*) with three components each (e.g., R, G, B), leading to a huge number of possible combinations that could be used as inputs to classification algorithms. This so-called curse of dimensionality means that the more features or feature combinations used, the larger the underlying learning database for classification must be. In a nontrivial case, the relationship between the number of features used and the number of images required in the learning database grows exponentially. Unless the learning database is large enough, the generalization capabilities of the classification algorithm may be compromised [5,6]. Due to the small size

sample available in this study, the number of images is consequently small. It is therefore necessary to reduce the number of input features to a smaller set. The underlying approach here is to use the chromaticity or color histogram to reduce the input data dimensionality.

The color histogram is an approximation of the color distribution and is considered as Gaussian. Therefore, the usual descriptors can be used to define the respective departure from the normal curve in characterizing the wound.

4.2.1.2 Mean and median

The mean of the histogram represents the center point of the distribution, separating the histogram into two equally probable subsets. The median is similar but not identical to the mean. It splits the range of the histogram into two equal sample sizes regardless of the probability. This can be an advantage over the mean in cases where the extreme probability would be artificially increased due to clipping. It is apparent from Figure 4.2 that neither the mean nor the median of the chromaticity distribution alone is adequate to discriminate between inflamed and noninflamed wounds. Therefore, further details of the histogram shape, such as the standard distribution, skewness, and/or kurtosis (respectively describing the width, the symmetry, and the flatness against the central point), are required.

4.2.1.3 Standard deviation

Standard deviation (std) represents the dynamics of the histogram, how widely around the mean the colors of the wound image are distributed. It is anticipated that noninfected wounds will have a higher standard deviation than infected wounds (see Figure 4.2). Standard deviation, or variance, can be generalized through the covariance matrix in the multidimensional case.

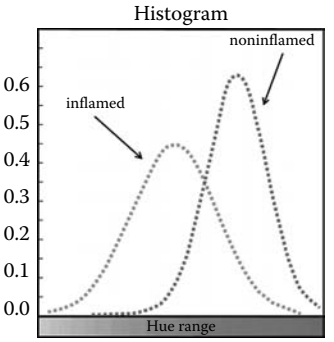


FIGURE 4.2: Change of hue distribution for inflamed and noninflamed wounds around the red part of the visible spectrum, as demonstrated by Hoppe [4]. The hue bar shown would range from a bright red, on the left, to deep red, in the middle, to purple, on the right.

4.2.1.4 Kurtosis

Kurtosis quantifies the flatness level of the distribution at the mean. Again, the expectation from Figure 4.2 is that the kurtosis for noninfected wounds should be smaller than that for infected wounds. Kurtosis is equal to 3 for a normal distribution. If kurtosis is lower than 3, then the distribution is said to be *platokurtic* (i.e., wide-peaked), and if kurtosis higher than 3, it is then said to be *leptokurtic* (i.e., narrow-peaked).

Figure 4.3 shows diagrams of 24 artificial wound images in a variety of color spaces. These wounds have been created deliberately in such a way that they fall into two clearly separable halves. Inflamed wounds are shown in dark gray circles and lines, and non-inflamed wounds are shown in lighter gray circles and dashed lines. For a further example, see Figure 4.4. Inflamed wounds are shown in red, and noninflamed are shown in green. These artificial images are subsequently used to test the classification methods. The clear separation between inflamed and noninflamed facilitates assessment of implementation methods and whether they are appropriate for application to real wound images.

The column on the right side of Figure 4.5 contains some features in textual form extracted from a single artificial wound. The top left graph shows the chromaticity diagram, the RGB primary of the device used, and the projection of the ROI pixel colors. The CIE v' graph to its right and the CIE u' graph in

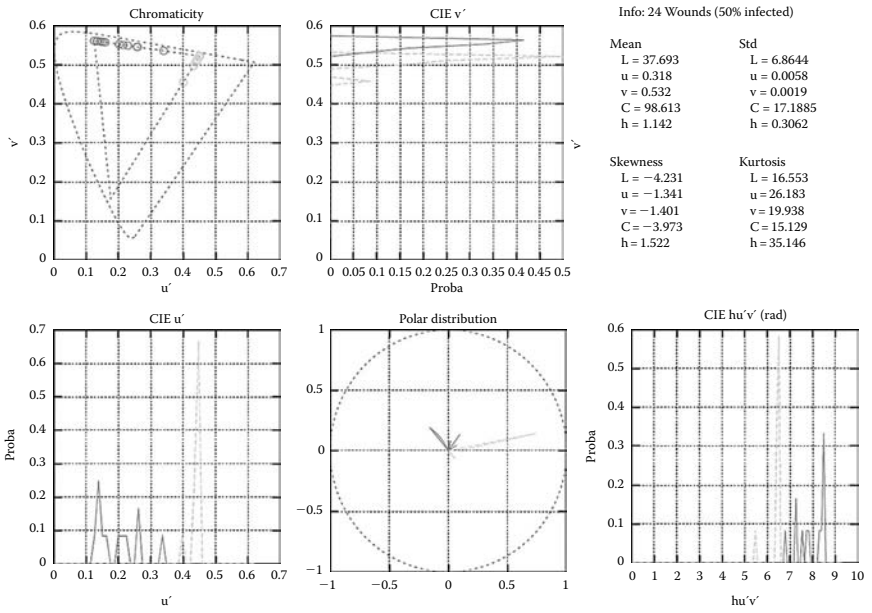


FIGURE 4.3: Features of an artificial wound set, with two chromatically distinct clusters (dark/light circles and continuous/dashed lines) in a variety of color spaces.

FIGURE 4.4: Three of the 24 artificial wound images used to create the diagrams in Figure 4.3.

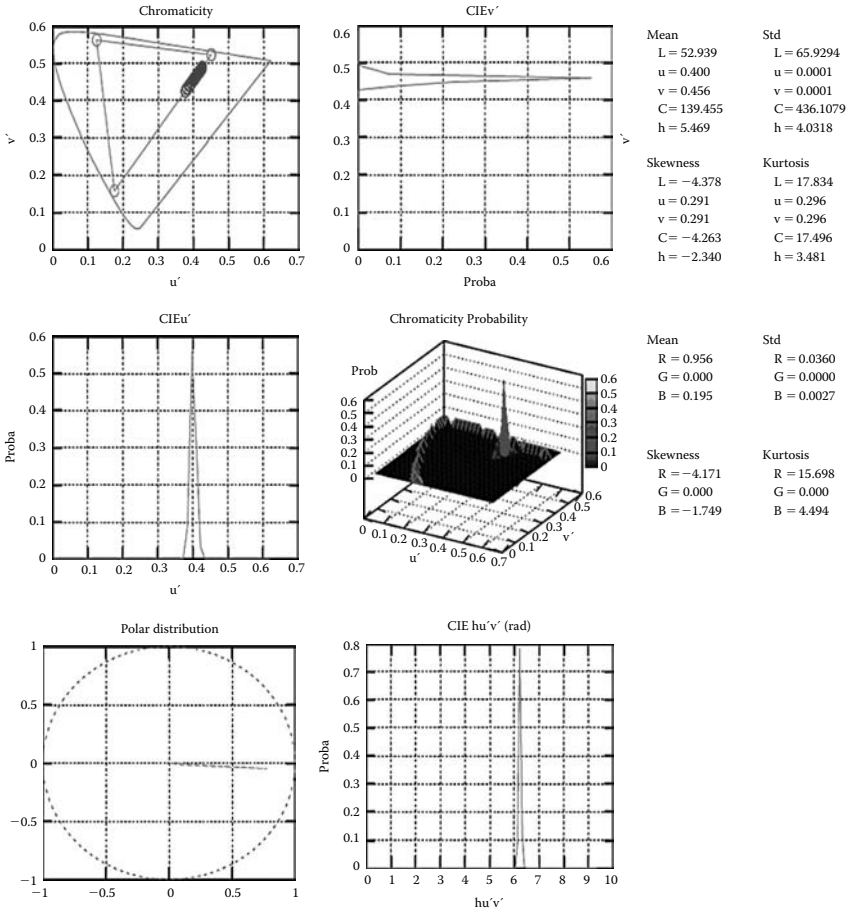


FIGURE 4.5: Feature details of one artificial wound image.

the next row on the left represent the marginal distribution. The 3-D plot in the center shows an alternative view of the chromaticity diagram with the ROI color distribution. The two graphs at the bottom depict the hue distribution: the left is the hue probability circle, that is, the probability for a specific hue angle (in radian).

4.2.2 Selection of features used by the classification algorithms

Having four statistical moments in four color spaces, each with three channels, results in a total of 37 possible inputs or features for the classification algorithms, as shown in Table 4.1. (There are not $48 = 4 \times 4 \times 3$ features because of the duplication of L* and h* channel in the L*u*v*, L*C*h*, and L*s*h* domains.) Obviously, each of these inputs could also be combined with each other to produce an additional input. If each input is combined with each other input, the result would be 666 possible combinations. The number of combinations is even higher if more than two inputs are combined. If 4 features of the possible 37 are combined, for example, the result would be 66,045 possible unique combinations. This is the curse of dimensionality, and clearly the number of inputs has to be reduced.

In this study, we examine two possible methods to reduce the number of features for classification. The first one employs a series of reduction steps using logical derivations based on assumptions; the second one uses a brute force approach.

4.2.3 Reduction of features using logical derivations

A first step in this reduction process is to limit the inputs under investigation to those containing color information. Inputs derived from luminance information are not used. The decision to ignore luminance information is based on the previously mentioned work by Hoppe [4], which suggests that there is no information about wound status in the luminance channel. Furthermore, under the experimental conditions employed in this study, luminance is correlated to environmental conditions during image capture. These are to a lesser degree the power of the light source, ambient lighting, and camera aperture settings because their variation is greatly reduced by the specially designed lighting system used for image acquisition. More influential is the automatic gain control circuitry of the camera, which cannot be switched off. In any case, we found that luminance information by itself contains little or no relevant information in discriminating between inflamed and noninflamed wounds.

TABLE 4.1: Feature index representation.

Features	**R**	**G**	**B**	**L***	**u***	**v***	**C***	**h***	**s***
Mean	2	3	4	14	15	16	26	27	34
Standard deviation	5	6	7	17	18	19	28	29	35
Skewness	8	9	10	20	21	22	30	31	36
Kurtosis	11	12	13	23	24	25	32	33	37

* The first index (not shown) is a constant required by some algorithms for computational reasons.

The two remaining channels, defining chromaticity, appear to be similar in response to wound inflammation; they can also be assembled under a single objective using the full chromaticity rather than just the marginals.

Through the exclusion of luminance information, the number of usable channels in the CIEL*u*v*, CIEL*C*h*, and CIEL*s*h* domains is reduced to two. The RGB color space is removed completely because the luminance information is present in all channels, adding noise. Additionally, all three channels are correlated to some degree, which further reduces the usefulness of this color space. The remaining three color spaces, L*u*v*, L*C*h*, and L*s*h*, describe positions in the chromaticity diagram using different geometrical representations. In other words, they do not provide extra information but show it from a different perspective, which may be more suitable for certain types of processing. In principle, only one of the color spaces needs to be investigated, especially considering that the classification algorithms can transform one color space into another anyway (using SVM, for instance). If we restrict ourselves now to the L*s*h* color space, there are only 8 features left to investigate. Figure 4.6 is produced using 80 images in which the granulation tissue (the one that medical evidence suggests is most likely to contain evidence of inflammation) was extracted. Circles represent the known inflamed wounds (probability of inflammation = 1), and crosses represent the known noninflamed (probability = 0). Ideally, we would like one of the classifiers to show all circles on one side of the graph and the crosses on the opposite side. Unfortunately, the graph shows that no feature, on its own, appears to be a good classifier at all.

If, however, we plot each feature against each other feature (without crossing channels), we obtain 12 graphs in which the plot of standard deviation against kurtosis of the hue channel (i.e., feature set 37) visually appears to perform best (see Figure 4.7), although far from perfectly. This would be the one to select for further evaluation.

At this point, it is worthwhile to consider the reduction steps performed so far: they are all based on probable assumptions, which, in the beginning, are most likely true and well understood in their consequences. As the reduction process continues, the underlying assumptions are more difficult to validate and become increasingly vague. As a result, the validity of this logical reduction process becomes increasingly subjective. For this reason, a second possibility of feature reduction is considered in the following section.

4.2.4 Reduction of features using a brute force approach

Modern computers allow the evaluation of combinations of each feature of each color space with one another in a reasonable amount of time. As already mentioned, we are investigating four color spaces with three channels each. Each channel contains four features: standard deviation, mean, kurtosis, and skewness.

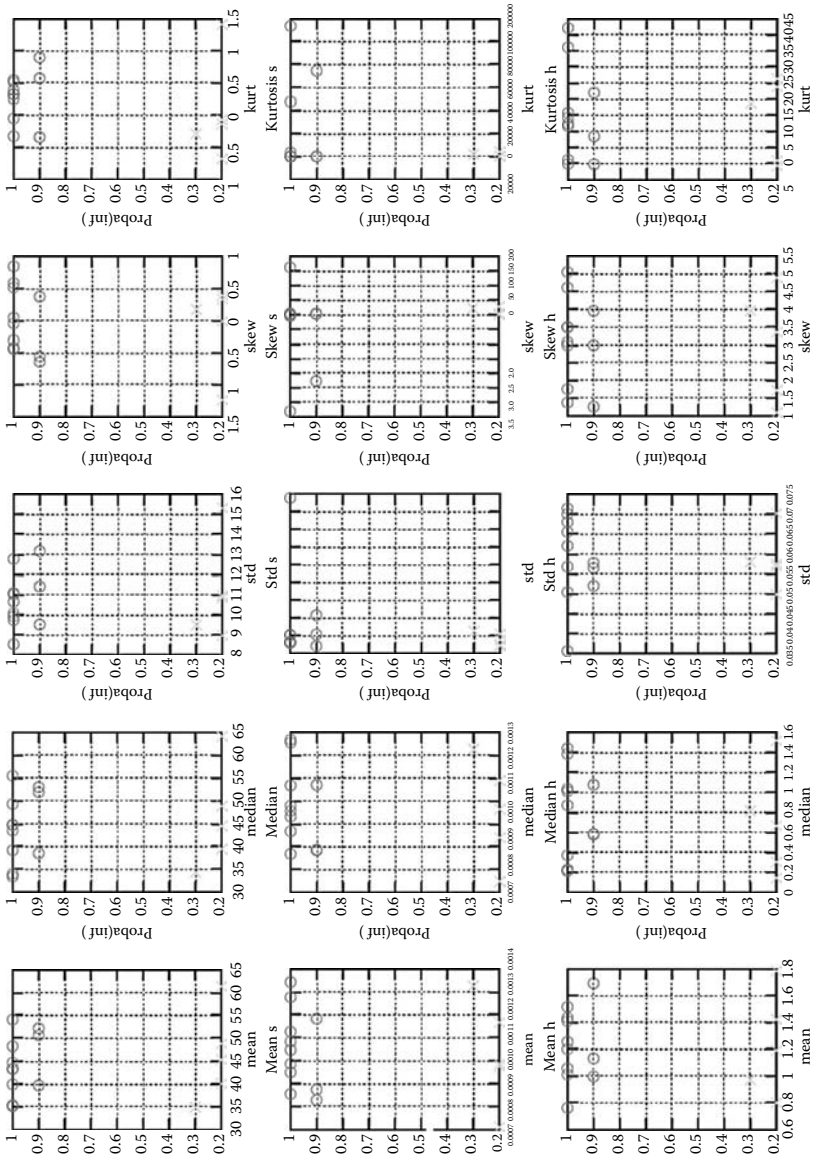


FIGURE 4.6: Probability of inflammation as a function of feature. Channels are in rows, statistical moments in columns.

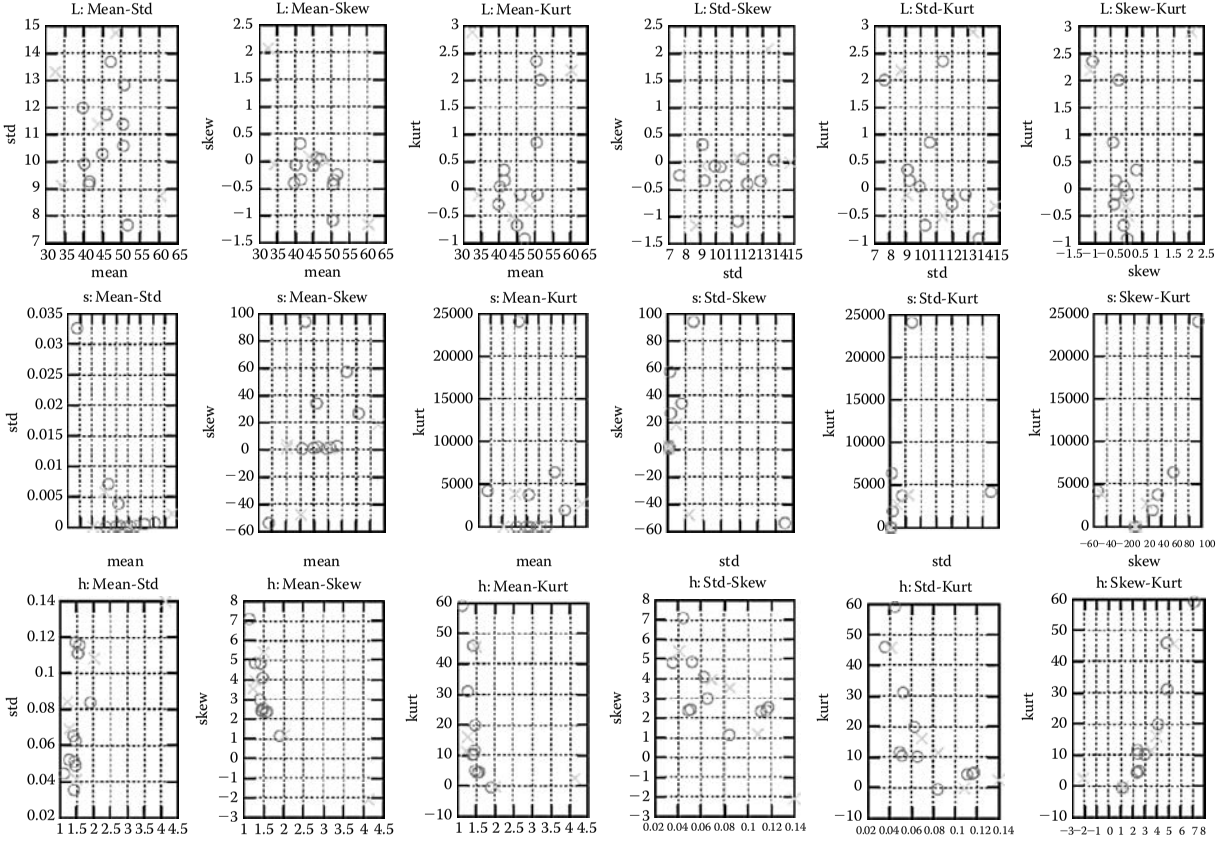


FIGURE 4.7: Characteristic map of one statistical moment against another for each color channel.

The 37 features can thus be combined in a variety of subsets of different sizes. Acha Pinero, Serrano, and Acha [7] used a neural network (ART-MAP). Starting with a single feature, new features were randomly added, one at a time, as inputs to the network and kept, provided they improved the classification result. The method is known as sequential forward selection. The researchers also tested the opposite approach, sequential backward selection, by starting with all features as inputs, then gradually and randomly removing those that improved the classification results by their removal. In both cases, the resultant sets of features were similar, adding credibility to this approach. Neural networks, however, are problematic in that it is difficult to determine whether the absolute optimum result in the decision plane has been reached or whether the method got stuck in a local minimum.

The approach chosen here selects the first 4 features out of the pool of 37. These features are now used to train the most rapid of the classification algorithms investigated (logistic regression) using the selected features on a training set of 76 images (half of the data set of 151 images). The trained algorithm is then used to classify the remaining 75 test images, and the results are compared with the classification produced by clinicians.

The performance of the algorithm with respect to the average judgment of all human classifiers is expressed in a kappa (κ) value. The κ value assesses the validity of individual judgments. It is obtained by using a contingency table in which two judges classify n observations into r categories. Kappa is then calculated using

$$\kappa = \frac{P_0 - P_e}{1 - P_e}$$

where P_0 is the agreement between judges, and P_e represents the expected random agreements under the null hypothesis that judges classify purely at random. For details, refer to Cohen [10]. This process is repeated with each possible combination of the 4 features, resulting in a total of 66,045 runs of this procedure. The total computing time on an Athlon™ 1800 processor was 15 minutes.

Figure 4.8 shows the result of this procedure using a single clinician as a reference. The two graphs on the left show results obtained on the full ROI (the entire wound area); the two graphs in the middle show the results on granulation tissue area only; and the two graphs on the right summarize the outcome from runs on slough tissues. In each of the three sets, the upper graphs shows the κ value for each of the 66,045 runs. The lower graphs show the κ distribution (i.e., the histogram) of the upper graphs.

A characteristic property shown in the three upper graphs is that there is a degree of periodicity in the κ values. Since the sets of 4 features are not randomly selected for each evaluation but are selected by a nested loop algorithm, specific combinations of features appear periodically. The reflection of this period in the graph is an initial indication that some combinations of features perform better than others. Otherwise, the distribution of peaks in the graphs would be completely random.

The higher the absolute κ value (plus or minus), the stronger the discriminatory power (the agreement between human and machine classification) of the underlying feature set. The bottom set of three graphs shows that by far the most κ values are very low and distributed closely around point zero. The feature sets represented by these values have no relevance. There is, however, a small number of feature combinations that produce relatively high κ values. Due to the normalization of the feature frequency graphs in Figure 4.8, however, they are not clearly apparent.

In order to visualize these important feature combinations, we extract the top performing 660 ones (i.e., 1% of the total) and sum the number of times a certain feature combination appears. The result is shown in Figure 4.9. It shows that for the above example, feature 37 appears in 14% of all the "successful"

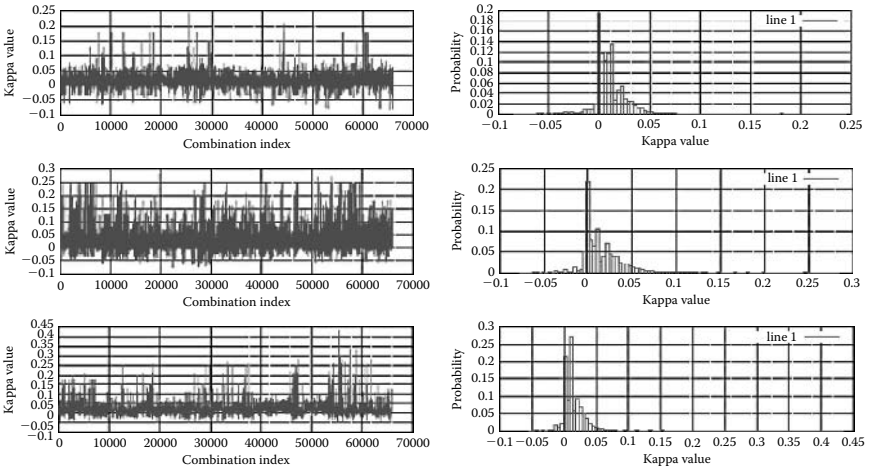


FIGURE 4.8: Kappa (κ) value list for all 66,045 combinations of 4 features out of the possible 37, using a logistic regression classifier.

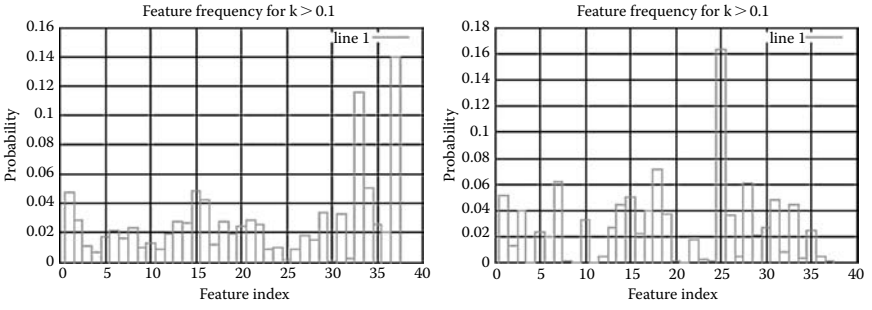


FIGURE 4.9: Probability of features appearing in successful feature combinations for granulation tissue (left) and sloughy tissue (right).

combinations (left side graph in Figure 4.9). Feature 33 appears with a probability of almost 12%, and the third runner-up is feature 15 with approximately 5%. Interestingly, feature 37 is also one of the two features produced as the result of the logical reduction process described in the previous section.

From these findings, it is clear that skewness and kurtosis are the most predictive features useful for classification. For granulation tissue–based classification, feature 32 (kurtosis of chroma C*) performs most effectively, followed by features 37, 30, and 36. For slough tissue–based classification, feature 21 (skewness of the u* channel) followed by 25, 12, and 11 perform best; and for their combination, feature 23 (kurtosis of L*) performs best. These four top features are used respectively to test their performance on three competing classification engines based on logistic regression, a neural network, and a support vector machine (SVM).

4.3 Classification Using the Selected Features

4.3.1 Classification by logistic regression

The difference between linear and logistic regression is the ability of the latter to regress ordered data such as binary inflammation/noninflammation information. Linear regression techniques cannot be used when there is a lack of continuity in the data. The logistic regression algorithm used here maps its prediction onto a probability value using the *logit* function, which computes the natural logarithm of the odds value and shows the probability separation achieved for standard deviation, skewness, mean, and kurtosis.

Classification is performed in two stages. In the first, the logistic regression learning machine is trained, and in the second, the classification is achieved. The machine is trained through randomly selecting a subset of 76 of the total 151 wound images. Each image is presented to the machine as a collection of the four features selected in the previous section. For each image, the four features are calculated for granulation tissue, slough, and a combination of both. Learning is completed in seconds. Following this, the trained machine is exposed to the remaining unknown 75 images from the database, resulting in three distinct verdicts of the wounds inflammation status. Table 4.2 compares the performance of the logistic regression machine against the clinicians participating in the experiment. Kappa values are calculated for the learning machine and the "mean clinician." The κ value of this mean clinician expresses the agreement between the mean of all clinician judgments with that of a "super-clinician." This super-clinician has been constructed using a weighted averaging process wherein the weights applied to individual clinicians' judgments are a function of their closeness to the median of all judgments. The super-clinician is defined to give the "true" judgment (i.e., the benchmark) against which both individual clinicians and classification engines are measured.

TABLE 4.2: Kappa agreement values quantifying the classification performance of the logistic regression machine against that of clinicians for granulating tissue, sloughy tissue, and the combination of both (unsegmented wound image).

Tissue	Logistic Regression	Mean Clinician
Granulation	0.509	0.683
Slough	0.531	
Both combined	0.127	

TABLE 4.3: Kappa values quantifying the classification performance of the ANN machine against that of clinicians for granulating tissue, sloughy tissue, and the combination of both (unsegmented wound image).

Tissue	ANN	Mean Clinician
Granulation	0.282	0.683
Slough	0.169	
Both combined	−0.018	

With the exception of the case in which both tissue types are combined, the machine produces a reasonable average classification that is slightly inferior to that of the mean of the nine participating clinicians but significantly better than that achieved by the "worst" clinicians. Note that the "best" clinicians are usually those with more experience and training, and those at the lower end tend to be younger and relatively inexperienced nurses currently undergoing training. The knowledge and understanding level of computers seems to have some degree of importance as well (e.g., spotting errors during data entry).

4.3.2 Classification by artificial neural network

The neural network used in this study implements a feedforward topology using a standard backpropagation learning algorithm $O(n^4)$, and was implemented using the SNNS software package interfaced with Octave2 (a free, open source MATLAB® clone) for testing.

As in the previous logistic regression section, Table 4.3 compares the ANN with the mean clinician. The neural network performs significantly worse than the logistic regression machine. While the logistic regression machine performance was at the center of two equally sized sets of clinicians, the ANN machine is located at the bottom.

4.3.3 Classification by support vector machine

As before, classification is performed in two stages using the same training and testing image sets. The SVM algorithm is slower than logistic regression, taking up to several tens of seconds. The algorithm speed is quite sensitive to the size of the data set, and computation time increases exponentially with the data set size due to the quadratic programming (minimization of quadratic problems), whereas the logistic regression algorithm performance is almost linear.

Table 4.4 shows the performance of the SVM to be better than that of logistic regression. There is, again, poor performance in the classification using the combination of both tissue types, whereas classification based on granulation tissue produces good results.

In the SVM, classification or separation between classes is made by defining a hyperplane between the classes. The main novelty in the SVM is the approach of finding the optimal hyperplane. It was originally developed by Vapnik [11] within his statistical learning theory. It has been successfully used in several real-world applications [12–14]. The theory behind SVM is not widely known. For a complete introduction refer to the original work by Vapnik [1] and its derivatives [12–16]. A brief introduction in three steps with increasing complexity of the problem solvable by the machine is provided below.

1. Linear hyperplane, linearly separable. This technique is useful only if we know the data are linearly separable, which is not the case in this work.

2. Linear hyperplane, nonlinearly separable. Although this method is already an improvement over the linear hyperplane used on a linearly separable problem, it is still not general enough for the classification task presented to this work. A further generalization step is therefore required.

3. Nonlinear hyperplane, nonlinearly separable. In this case, the classification is performed in higher feature space by replacing the inner product defining the hyperplane by a kernel function. The kernel maps the input

TABLE 4.4: Kappa values quantifying the classification performance of the SVM against that of clinicians for granulating tissue, sloughy tissue, and the combination of both (unsegmented wound image).

Tissue	SVM	Mean Clinician
Granulation	0.592	0.683
Slough	0.717	
Both combined	0.055	

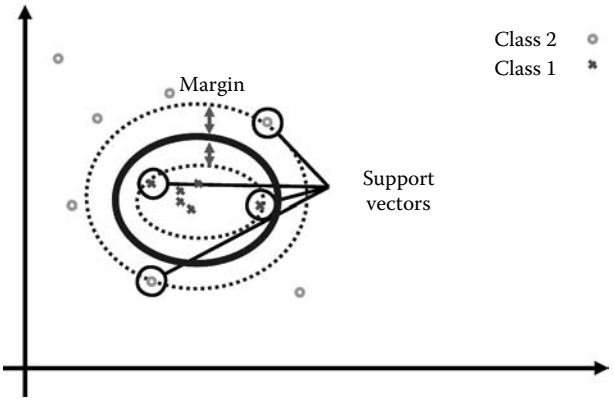


FIGURE 4.10: Quadratic hyperplane on nonlinearly separable data.

data from Rn into a higher feature space Rk (where $k > n$). In this new space, data are then more linearly separable. Figure 4.10 illustrates this principle. Interestingly, by being kernel based, the SVM can simulate other learning machines simply by changing the kernel function. Examples of kernel functions can be found in the literature [12–14].

4.4 Discussion

The classification performance of the three algorithms outlined in the previous section is of course not constant. It depends to some degree on the training and testing set used. The figures quoted in Tables 4.2, 4.3, and 4.4 represent typical results, and Table 4.5 summarizes these together with results for the nine clinicians participating in this study. Every time a different subset of 76 images is randomly selected as the learning set (the remaining 75 images are automatically assigned to the testing set), the results differ slightly. Importantly, however, the general ranking (i.e., the quality of classification) is consistent with the SVM and the logistic regression algorithms generally performing similarly well. The logistic regression machine is relatively easy to implement compared to the other algorithms. It is fast and the learning complexity is $O(nd^3)$, as a function of the number of variables d and linearity of the data size n. Testing/classification time increases linearly with variable dimension and data size. Computational performance, however, could be better. One of the problems results from the highly nonlinear nature of the feature set used. After their success with SVM, kernel functions have now become a popular approach to solve this problem. Using kernel functions with logistic regression could, in theory, greatly improve performance.

TABLE 4.5: Comparison of logistic regression, ANN, and SVM against clinician performance.

Clinician/System	**Individual/Tissue Type**	κ
Clinician	1	0.097
	2	0.178
	3	0.199
	4	0.331
	5	0.337
	6	0.532
	7	0.615
	8	0.677
	9	0.709
	Mean clinician	**0.683**
Logistic regression	Granulation	0.509
	Slough	0.531
	Both combined	0.127
ANN	Granulation	0.282
	Slough	0.169
	Both combined	−0.018
SVM	Granulation	0.592
	Slough	**0.717**
	Both combined	0.055
Super-clinician benchmark		1.000

Neural networks are slightly more complex to implement, but because they have been around for a while, there are some good off-the-shelf implementations available. SNNS, the software used here, implements a very large selection of networks, and it integrates very well with the remaining tools used in this project. Results are unfortunately not as expected.

The SVM general solution can be considered to be the most complex of the algorithms implemented. The main source of this complexity is the need for quadratic programming, with $O(n^3)$ on average, although some researchers have reported that the use of partial linear programming can improve this to be close to $O(n^2)$. When adding more features or data, speed decreases very rapidly; this is where a reduction of data, as demonstrated above, is appreciated. Fortunately, the SVM algorithm is asymmetric: training time and testing time are not directly linked. Testing (classification) time can be very short even when learning time is very long, because during the learning process, the SVM must look at all the input data in order to select the most relevant subset, the support vectors. The testing or classification phase uses only this subset and is thus much faster. Testing time depends mostly on the feature size and the kernel function. The implementation used here is adapted from the e1071 package, an interface of the libSvm library for the R language [8]. The implementation of SVMs suffers because it allows the use

of many types of mapping with their associated parameters ("fiddle factors"). The machine can also be very sensitive to these parameters although adequate sets of parameters can be found easily and empirically since they directly affect the mapping of the features where the classification takes place. There is, however, often an element of unease about the lack of rigor in the selection process. (This is mostly due to the lack of understanding of the problem itself rather than the kernel selection, which can be relatively well designed for specific needs. If those needs cannot be expressed clearly, the kernel function cannot be expressed clearly either.) Standard logistic regression (being a particular case of linear regression) does not require any external parameters but the knowledge of dependent and independent variables. The mathematical definition of the regression makes it easy to compare different models using standard statistical techniques [9], although in our case, kappa is more appropriate because it is independent of the classification algorithm.

In general, good ANN implementations rely on random numbers. In the case of the backpropagation algorithm, they are used in weight initialization, pattern permutation, and/or the sequence of neuron updating. This is not necessary, but it is recommended because it can help the system to escape from local optima and thus tends to increase the probability of reaching the optimal solution. Although neural networks generally converge toward an acceptable solution, it remains very difficult to compare two different networks. Even for the same network, with the same topology and the same learning rules, the neuron weights resulting from learning can theoretically be very different. In practice, they tend to converge toward the same optimum, although from a different path. For this reason, ANNs are often considered to be a black box where the main comparison and analysis is done using input/output pairs.

4.5 Future Research Directions

During this work, it became clear that there are a number of areas into which this research could be expanded in the future. In order to develop a simple, reliable, and easy-to-use wound status classification device, the following aspects will have to be considered.

- A robust, practical, and specular reflection-free illumination device using stable light sources such as light-emitting diodes (LEDs) will have to be developed.
- A camera featuring three charge-coupled devices (CCDs) (or a larger single CCD) and with lossless compression is definitely a requirement if the underlying image data are to be improved. Spatial resolution is less of a problem than dynamic resolution, stability, and the option to disable image compression.

- Some clinicians believe that the wound edge appears to hold interesting information about inflammation. Consequently, this could be an area of future interest. Comparing the skin at the edge of the wound with the skin further away (which appear to be more distinct in inflamed wounds) as well as the shape of the wound outline appears promising.
- The finding that both skewness and kurtosis are good features for classification points to the logical step of using the texture of wounds in the classification process. Wavelet or Gabor transforms may be a way forward in this direction.
- Currently, the amount of both image data and clinicians acting as experts is limited. More representative results should be obtainable by increasing the amount of input data. Wound types other than venous leg ulcers should be considered as well.
- Other classification engines, especially Bayes' probability approaches, should be considered.
- Calibration of equipment becomes an issue as soon as multiple acquisition systems are being used in a future trial. Much work still needs to be done to arrive at a reliable and simple calibration system that affects image data quality only minimally and is still easy to use.
- It is assumed that the success of classification on slough tissue may have been due to colonization by specific bacteria that tend to provoke an inflammatory response and also affect the color of the slough tissue itself. Further work, however, is required to confirm this assumption.
- The result of the SVM classification can possibly be improved by changing the kernel and its parameters. As the kernel's task is to map the features into a higher space, it is first necessary to understand the exact characteristics of this space.

4.6 Conclusions

This chapter investigates how to differentiate wound status by different classification algorithms. In a first step, the amount of possible features for use by the algorithms is reduced in order to minimize complexity and increase computational speed by using a logical approach that is confirmed by a brute force algorithm. The next step is to select the most efficient feature combination from the set by testing the performance of all possible combinations of 4 features out of a total of 37. As is to be expected, the best-performing features are different for each tissue type. Generally, however, the best features

employ skewness and kurtosis of channels rather than the mean and standard deviation. Due to the relatively small training database (76 images), classification performance of the learning machines depends on the respective subset of images used for training.

Increasing the database size should naturally reduce this sampling effect. In some cases, the SVM performs better than the most skilled clinician. This means that the classification of the SVM is in better agreement with the mean opinion of the most skilled individual clinicians. In most cases, the machine is at the top of the performance table. It is not considered a perfect decision, but it is consistently better than most clinicians and could therefore be part of a future expert training system.

The SVM is a powerful and flexible system. Changing the kernel function and/or even the type of machine could possibly improve the reliability of the diagnosis even further. The SVM has, for example, a "soft margin" to minimize the effect of overlapping instances in different classes. This overlapping would most likely increase with the error, which in turn would increase misclassification and subsequently affect the learning outcome.

In the case of ANNs, one possible improvement would be to try different topologies and learning algorithms. Splitting the network into more hidden layers would allow finer control over the black box. This would naturally result in an increased time consumption for the learning cycle, but the testing time would remain fairly small (in the order of seconds), virtually identical from the user point of view. ANNs, as mentioned earlier, do not allow the network designer direct control over the learning process, but the randomness inherent in the training process allows the network to extract itself from local optima.

Although the classification results achieved using SVM and logistic regression are already in some cases outperforming those of experienced clinicians, their performance could potentially be further improved by decreasing the number and degree of errors that invariably accumulate in the process that lead to the classification result. Even without such improvements, the system as presented here may already be useful as a training tool for medical students specializing in wound care because it provides consistent diagnostic output at the level of a reasonably well-trained wound care professional.

References

[1] J. C. Barbenel, M. M. Jordan, and S. M. Nicol. Incidence of pressure sores in the Greater Glasgow Health Board Area. *Lancet*, 2(8037): 548–550, 1977.

[2] Axel Perdrup. The healing rate of leg ulcers. *Acta Dermatovener (Stockholm)*, 52: 136–140, 1972.

[3] National Audit Office. Health service for physically disabled people aged 16 to 64. Technical Report, Comptroller and Auditor General (HMSO), 1992.

[4] Andreas Hoppe. Detection of possible infection in wounds by color image processing. Final year project, University of Glamorgan, Internal Report, 1996.

[5] Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1): 4–37, 2000.

[6] Vladimir Cherkassky and Filip Mulier. *Learning from Data: Concepts, Theory and Methods*. Wiley-Interscience, 1998.

[7] Begona Acha Pinero, Carmen Serrano, and Jose I. Acha. Segmentation of burn images using the L*u*v* space and classification of their depths by color and texture information. Vol. 4684, pp. 1508–1515. SPIE, 2002. Available at http://link.aip.org/link/?PSI/4684/1508/1.

[8] Ross Ihaka and Robert Gentleman. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3): 299–314, 1996.

[9] John O. Rawlings. *Applied Regression Analysis: A Research Tool*. Statistic/Probability series. Wadsworth & Brooks/Cole.

[10] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1): 37–46, 1960.

[11] Vladimir N. Vapnik. *The Nature of Statistical Learning Theory. Statistics for Engineering and Information Sciences*, 2nd ed. Springer-Verlag, 1999.

[12] Christopher J. C. Burges. A tutorial on support machine pattern recognition. In *Data Mining and Knowledge Discovery*, Vol. 2, pp. 121–167, 1998.

[13] Steve R. Gunn. Support vector machines for classification and regression. Technical Report, University of Southampton, May 1998.

[14] Edgar E. Osuna, Robert Freund, and Federico Girosi. Support vector machines: Training and applications. Technical Report AIM-1602, MIT, Cambridge, Mass., March 1997.

[15] Bernhard Schölkopf, Alexander Smola, and Klaus-Rober Müller. Advances in kernel methods: support vector learning (chapter on kernel principles). In *Component Analysis*, pp. 327–352. MIT Press, 1999.

[16] Alex J. Smola and Bernhard Schölkopf. A tutorial on support vector regression. Technical Report, ESPRIT Working Group in Neural and Computational Learning II, 1998.

Chapter 5

Analysis and Applications of Neural Networks for Skin Lesion Border Detection

Maher I. Rajab

Contents

Researchers in medical imaging and computational intelligence groups are keen to develop sophisticated and complex methods to meet the specialized needs of medical applications, including the associated medical data. One of the significant medical imaging applications is the accurate image segmentation of skin lesions, which is key to useful, early, and noninvasive diagnosis of coetaneous melanomas. Extracting the true border of a skin lesion is one of the most important features that reveals the global structure irregularity (indentations and protrusions), which may suggest excessive cell growth or

regression of a melanoma. In this chapter, we investigate the application of two developed approaches to the skin lesion segmentation problem: iterative segmentation (IS) and neural network edge detection (NNED). These approaches are compared for synthetic lesions at different image signal-to-noise ratios (SNRs) to test their ability to delineate the lesion border shape. The use of synthetic lesions is advantageous in initial analysis and verification because if we know the true position of the lesion border, the different methods can be quantitatively and more accurately compared.

5.1 Introduction

The principal aim in the evaluation of a skin lesion image segmentation technique is to study the effectiveness in distinguishing a lesion from the background. For example, Gao and others (1998), Hance and others (1996), and Xu and others (1999) used expert clinical delineations as the gold standard to evaluate the automated segmented images. The disadvantage of using expert delineations is that they are dependent on human subjectivity. For example, Xu and colleagues found that different clinical experts use different rules to segment an image. Some delineate the border where the lesion and nonaffected skin meet. Others locate the boundary where darker pigmentation and lighter pigmentation meet. An example of these variations made by four experts are contained in images provided by Xu and others (1999). The origin of these variations were analyzed (Xu et al., 1999) and referred to two main reasons. The first reason is the clarity of the lesion boundary—the boundary is not always perfect and visually clear; it may have some width, as shown in Figure 5.1. The second reason is the irregularity of the lesion boundary: experts' segmentations for benign lesions have less variation than for malignant lesions, as shown in Figure 5.2.

In this chapter, the automatic skin segmentation (ASS) method (Xu et al., 1998) is used as a comparison with the proposed iterative segmentation (IS) and neural network edge detection (NNED) methods. The variation in the location of the lesion border defined by different experts is therefore not ideal when making a quantitative comparison among different automatic segmentation methods. This study proposes that generating synthetic lesion images in the presence of different levels of noise would enable true quantitative comparisons of different segmentation techniques.

5.1.1 Segmentation of medical images

Current computing research in medical image analysis uses image segmentation as an intermediate step to isolate the object of interest from the background (Olabarriaga and Smeulders, 2001). In the area of biomedical image segmentation, most of the proposed techniques can be categorized

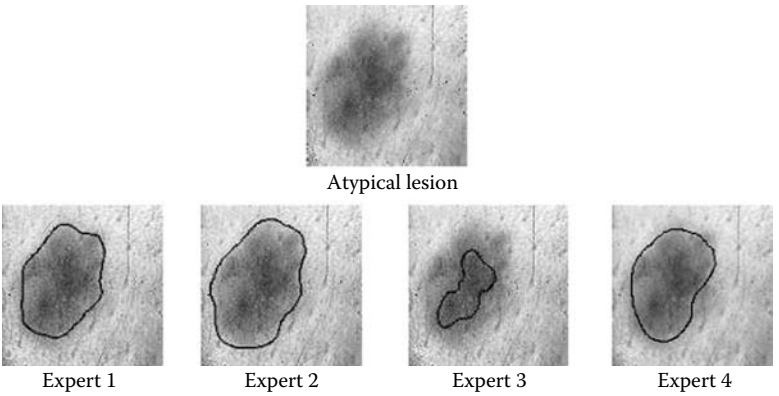


FIGURE 5.1: (See color insert following page 370.) Segmentation of atypical lesion by four clinical experts. (From Xu et al., 1999. *Image Visions Computing*, 17: 65–74. With permission.)

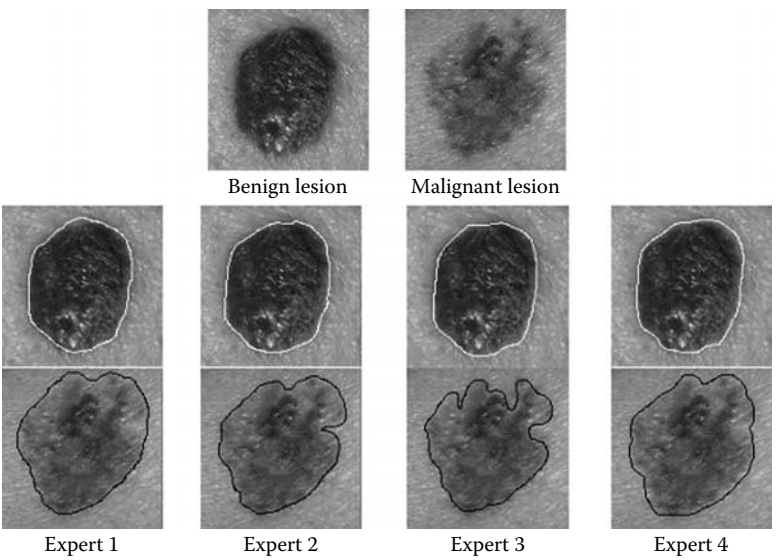


FIGURE 5.2: (See color insert following page 370.) Segmentation of a benign lesion (white border) and a malignant lesion (black border) by four clinical experts. (From Xu et al., 1999. *Image Visions Computing*, 17: 65–74. With permission.)

into three classes (Fu and Mui, 1981): (1) characteristic feature thresholding or clustering, (2) edge detection, and (3) region extraction tasks such as measurements (Giachetti et al., 2003), visualization (Campbell and Flynn, 2001), and registration (Maintz and Viergever, 1998). The need for an

efficient segmentation is of great importance because of its potential need in the identification of clinical morphological information about the patient's anatomy (e.g., Imeliska et al., 2000) and pathology (e.g., Brenner et al., 1981).

5.1.2 Segmentation of melanoma images

Malignant melanoma skin cancer is a potentially fatal disease, killing a disproportionate percentage of economically active younger people when compared with other human cancers. Early detection of melanoma is feasible because it is usually visible on the skin surface when in a curable stage (Weinstock, 2000). For instance, an epidemiological study for malignant melanoma in Scotland reported that survival fell in men from 93% for tumors thinner than 1.5 mm to 47% for tumors thicker than 3.5 mm, and in women from 97% to 55%, respectively (MacKie et al., 2002).

Sometimes physicians cannot visually distinguish between benign lesions and early malignant melanomas. Therefore, a biopsy is performed, and the pathologist diagnoses the lesion. However, several drawbacks of biopsies are encountered (Marquez, 2001): (a) return time for a biopsy may be significant; (b) it can be painful for a patient to have several questionable lesions excised, which in some cases can be disfiguring; and (c) incomplete biopsy might cause a spread or metastasis of melanoma cells to other parts of the body, and these may invade other organs. The well-recognized and established technique of dermatoscopy or epiluminescence microscopy (ELM) (Carli et al., 2000; Goldman, 1951; Pehamberger et al., 1993; Steiner et al., 1987) is a noninvasive, in vivo technique that would eliminate the need for unnecessary or incorrect biopsies because, with ELM, dermatologists can view and image the subsurface layers of the skin. This technique reveals morphological features, colors, and patterns not perceptible by the dermatologist during the clinical observation. Dermatologists trained in the use of ELM can improve their diagnostic accuracy of melanoma from about 65% using the unaided eye to approximately 80% with the benefit of ELM (Kanzler and Mraz-Gernhard, 2001). However, even with ELM, a trained dermatologist can be deceived at least 20% of the time by the appearance of a melanoma (Kanzler and Mraz-Gernhard, 2001).

Computer-aided diagnosis systems have been successfully applied for early detection of melanoma. For example, a recently developed PC-based pilot system (Binder et al., 2000) promises to automatically segment the digitized ELM images, measuring 107 morphological parameters. A neural network classifier trained with these features can differentiate between benign and malignant melanoma. Accurate delineation of the lesion border shape may improve the effectiveness of a computer-aided system, similar to the above, and make the lesion accessible for further analysis and classification (e.g., benign or malignant). Measurement of image features for diagnosis of malignant melanoma requires the first step to be the detection and localization of the lesion in an image. The accurate determination of lesion

boundaries is a primary aim in this work so that measurements of features (e.g., asymmetry, border irregularity, and diameters) can be accurately computed (Xu et al., 1999).

5.1.3 Objectives of the research

The development of the research is carried out in successive steps and is summarized as follows:

1. The application of two approaches to the skin lesion segmentation problem: IS (which is a nonneural network method) and NNED are investigated. The aim is to quantitatively analyze the error in locating the border due to the application of an automated segmentation method.

2. The IS algorithm is used in this work to aid the extraction of the true border that reveals the global structure irregularity (indentations and protrusions), which may suggest excessive cell growth or regression of a melanoma. The output of the algorithm is used to generate various synthetic data sets. Selected sets are used in the design of neural network training sets.

3. A limited number of noisy edge patterns is proposed to analyze the capability of NNED for both synthetic and real images. Some of the most successful training sets are chosen for experimenting on synthetic lesion images as an attempt to establish the basis for the NNED method (Section 5.3).

4. It may be hypothesized that the edges produced by different algorithms do not establish the correct maps. Thus, it may be possible to override some of the uncertainty in detecting the edges by comparison between these maps. The use of synthetic lesions is advantageous in initial analysis and verification because if the true position of the lesion border is known, the different methods can be quantitatively and more accurately compared (Section 5.4).

5. Variability among expert clinical delineations is dependent on human subjectivity and was not addressed in many evaluation studies of automated skin lesions segmentation methods. Moreover, in spite of these variations between clinical delineations, they were always considered the gold standard for evaluating skin segmentation methods. Therefore, our main objective of using real images is to aid the final stage of testing the derived methods.

Section 5.2 provides a background on topics relating to the research. Sections 5.3 and 5.4 present the practical implementation of the NNED, applying the analysis of this method to synthetic noisy lesions. In Section 5.5, IS and NNED are compared for synthetic lesions at different image SNRs. Overall conclusions drawn from the current work and suggestions for further work are presented in Sections 5.6 and 5.7.

5.2 Background

Skin lesions exhibit a variety of colors, the presence of which is not very useful in segmenting images (Xu et al., 1999). However, we can observe that there are changes in color when going from a lesion to its background (its surrounding healthy skin). When we notice color variations inside and outside the lesion, then a second observation can be carried out. Because the delineation of a skin lesion deals with the boundaries of that lesion, color changes belonging to the lesion boundaries are all necessarily important in this process. Therefore, we can ignore color changes both inside a lesion and in the background of the lesion. Xu and colleagues (1999) stated two main facts relevant to most skin lesion images:

- Skin lesions come in a variety of colors; however, changes in color from a lesion to its background (its surrounding skin) are similarly observed in all images.

- Skin lesions exhibit significant color variations, which may exist within a lesion or in the background.

Intensity variations in lesion boundaries are simulated by producing synthetic binary images of the lesions. Gaussian noise is chosen as a first approximation noise source with which to compare the robustness of border delineation between different techniques (Rajab, 2003; Rajab, Woolfson, and Morgan, 2004). Real lesion images may contain noise such as details from skin texture and hair that make it more difficult to localize the lesion border. In addition, the effect of color variations such as shadows and reflections (bright spots) tend to bias the color map when performing color segmentation. Further work needs to be done to address the influence of noise so as to increase the robustness of the investigated techniques.

5.2.1 Border irregularity of synthetic lesion

Irregularity of the lesion border is one of the important clinical features that aid in differentiating benign melanocytic nevi from malignant melanoma. There has been recent work in the development of more sophisticated techniques to measure border irregularity (e.g., Lee et al., 2003), which could offer extra features such as localization of the significant indentations and protrusions in a lesion. A measure like this is sensitive to structure irregularity and may be more useful for accurate classification of the lesion shape. Among numerous irregularity measures, the border irregularity used here (Gray, 1971; Rosenfeld, 1977) is given by

$$I = \frac{P^2}{4\pi A} \tag{5.1}$$

where P is the perimeter of the segmented lesion and A is its area measured in pixels ($\Delta A = 1$). The perimeter is approximated by the sums of all the infinitesimal arcs that define the curve enclosing a region S as (Nixon, 2002):

$$P(S) = \sum_i \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \tag{5.2}$$

where x_i and y_i represent the coordinate of the ith pixel forming the curve. Because pixels in the image are organized in a square grid, the summation can take only two values.

For example, and according to the irregularity measure in Equation 5.1, a circular border has a measure of unity irregularity: $I = P^2/4\pi A = (2\pi r)^2/4\pi(\pi r^2) = 1$, which reveals a perfectly regular border. However, due to small errors in the measured perimeter (Equation 5.2), it is not possible that an exact value of $2\pi r$ will be achieved for the perimeter of a circular region of radius r.

Golston and colleagues (1992) determined by the use of some preliminary test images that a threshold of $I = 1.8$ makes the best separation between regular and irregular borders; that is, a lesion with I greater than 1.8 was to be classified as an irregular lesion. According to Golston's classification threshold (1992), the three lesions shown in Figure 5.3 (Rajab, 2003) represent three types of lesion topology according to their border irregularity indices (Equation 5.1): (a) very regular, $I = 1.1$; (b) regular, $I = 1.8$; (c) irregular, $I = 2.8$. We can say that these three lesions are chosen as basic typical test lesions, which would represent three candidates of border irregularities at two extremes of the threshold of 1.8 and at the threshold itself. It should be noted that accurate segmentation methods such as the IS method, which is presented in this work, could provide accurate lesion boundaries. Therefore, we have adopted this method to support generating the synthetic lesions shown in Figure 5.3. Moreover, an accurate method like the IS could accurately delineate a boundary of a lesion that represents the structure irregularity of that lesion and may also suggest excessive cell growth or regression of a melanoma (Rajab, 2005).

To build an analysis environment for testing the capability of the segmentation methods to delineate the true border of a lesion, we extracted the borders of three real skin lesions to generate three corresponding typical synthetic lesions, shown in Figure 5.3. The three binary synthetic

FIGURE 5.3: Three cases of synthetic lesions: (a) very regular, $I = 1.1$, (b) regular, $I = 1.8$, and (c) irregular, $I = 2.8$.

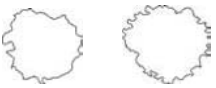


FIGURE 5.4: (a) IS segmentation result of a real lesion with irregularity index $I = 1.3$. (b) Border in (a) is manually modified such that $I = 1.8$.

lesions shown in Figure 5.3 were formed using the IS method (presented in Section 5.3), which is much more flexible in generating binary segments with a wide range of structure irregularities. To avoid the possibility that results using these synthetic lesions might be biased toward the IS method, we manually added an additional complexity to the border of one of the three lesions (Figure 5.3b) so that the final border structure differs from that determined by the IS method (see Figure 5.4) and can also be considered as if it were generated by another method.

5.2.2 Segmentation error

The error metric stated in Equation 5.3 has been used to quantitatively evaluate the result of segmentation (Gao et al., 1998; Hance et al., 1996; Xu et al., 1999). However, in those cases the manual segmentation produced by a clinical expert is used as the gold standard. The segmentation methods are tested on real clinical images and are quantitatively evaluated using the error metric developed by Hance and colleagues (1996):

$$e = \frac{A \otimes B}{A} \tag{5.3}$$

where the exclusive-or output image represents the total area that does not overlap between the segmented lesion B and the original reference synthetic lesion A. A zero normalized error means that the evaluated segmented lesion B completely overlaps with the reference A, while an error of 1.0 is obtained when there is no overlap. This error measure is always between 0.0 and 1.0 and is independent of the size of a lesion.

5.2.3 Conventional segmentation methods

The IS method (Rajab, 2003) and the ASS method (Xu et al., 1999) are two examples of conventional segmentation techniques. They are nonneural network algorithmic methods. Detailed theoretical background is presented in Rajab (2003). In lesion segmentation algorithms, the interest is mostly in changes that occur from the background to a lesion, or vice versa, then suppressing significant color variations inside a lesion or in the background are necessary. The requirements of boundary enhancement and suppression of color variations can be implemented by a mapping function as illustrated in Figure 5.5. The function tries to reduce image gradients corresponding to details belonging to the background and to a lesion; at the same time, the gradients of the intensities of pixels falling on lesion boundaries are increased

(the gradients of the boundary region are increased, as depicted in the shaded region of Figure 5.5).

Three steps are used in the IS method to segment the skin lesions: intensity mapping, application of an isodata algorithm, and object outlining. One of the advantages of the mapping function is that when mapping images containing low-intensity variations, small σ, in the background (e.g., ELM images; Rajab, 2005), then the function tends not to magnify those variations. Moreover, the selection of the standard deviation (σ) of this mapping function is automatically determined according to the estimated standard deviation of the background surrounding the lesion (Rajab, 2003). Figure 5.6 demonstrates various

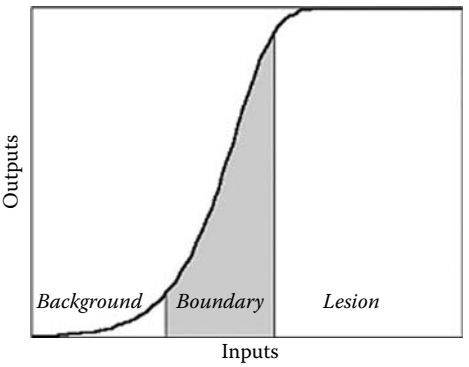


FIGURE 5.5: Property of a desirable mapping function.

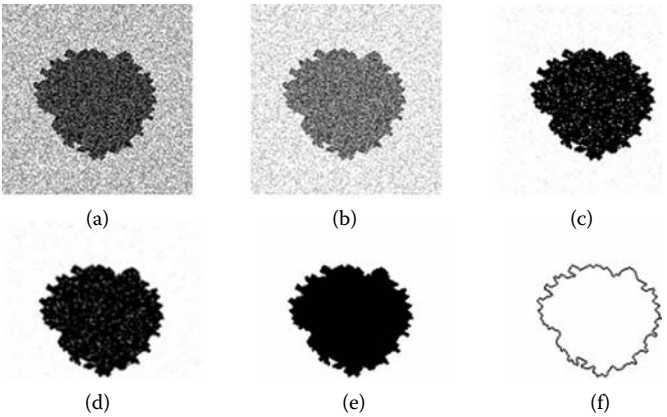


FIGURE 5.6: Iterative segmentation process: (a) noisy synthetic lesion (SNR = 3.0), (b) subtraction of median background, (c) intensity mapping, (d) smoothing by a 2-D Gaussian kernel of size 3×3 and standard deviation 0.6 pixel, (e) a binary segment at an optimal threshold produced by an isodata algorithm, and (f) a one-pixel-wide edge produced by object outlining.

outputs of imaging tasks of the iterative method to retrieve the true border of the noisy synthetic lesion.

5.3 Analysis of Neural Network Edge Detection (NNED)

Minimizing the complexity and ambiguity in the neural network training set rather than minimizing the convergence error of the trained neural network is a significant factor in the success of neural network recognition (Nakashima et al., 2001). The use of the SNR in characterizing the noisy patterns in the training set was one of the successful methods considered to analyze, study, and modify the NNED method. Further detailed background of the method used to analyze the behavior of various NNEDs is presented in Rajab (2005) and Rajab and Alhindi (2007).

In Section 5.5, we compare the robustness of segmentation errors to noise for the three segmentation methods (IS, ASS, and NNED) using the lesions shown in Figure 5.3 after they are corrupted by different realizations of Gaussian noise. Since we know that segmentation methods may yield slightly different errors when applied to different lesion structures, it is tedious and time consuming to generate numerous structures of synthetic binary lesions similar to those shown in Figure 5.3. We have instead taken three references synthetic images representing typical lesions with very different irregularity indices.

5.3.1 Structure and training of the NNED

This section presents the practical implementation of the NNED, utilizing the analysis of the NNED applied to real and synthetic noisy lesions. A novel neural network training set, consisting of a limited number of prototype edge patterns (see Figure 5.7), is proposed here to establish a base for a simple NNED scheme. Figure 5.7 illustrates four edge profiles in 3×3 pixel grids.

All the neural network training sets consist of a limited number of prototype edge patterns with 3×3 pixels similar to the patterns shown in Figure 5.7. We chose the smallest window size of 3×3 because it is sufficient to define all edge patterns encountered in this work. Various experiments have been performed using the neural network model with multilayer perceptron (MLP) architecture. The neural network is trained with the error backpropagation algorithm (Lippmann, 1987), using different sizes of training sets. When constructing and experimenting with various training sets, the total number of prototype edge patterns and their redundancy were considered. The number of prototype edge patterns used in this work is 42, incorporating five-edge

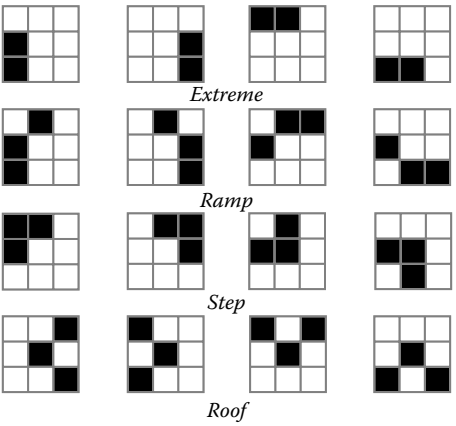


FIGURE 5.7: Four edge profiles in a 3×3 pixel grid at four orientations.

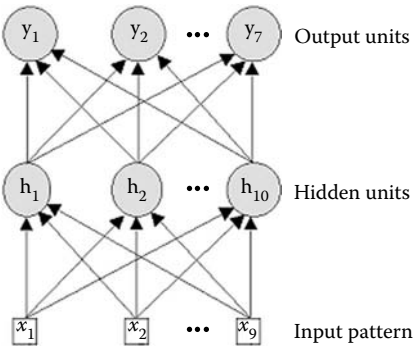


FIGURE 5.8: A 9-10-7 multilayer perceptron.

profiles with the possible orientations of each edge profile and also 100 redundant edge patterns (Section 5.4).

The training procedure is followed until a compromise solution is achieved and the neural network is capable of detecting object boundaries. The work presented here was based on a neural network model of the three layer MLP architecture. An example of 9-10-7 MLP is shown in Figure 5.8: there are 9 inputs in the input layer, 10 hidden nodes (units), and 7 nodes in its output layer.

Successful training was achieved when the neural network's mean square error converged to the value of 0.009 with a learning rate of 0.001 and momentum rate of 0.5 (Lippmann, 1987). The neural network was found to be highly capable of recognizing edge patterns in noisy synthetic lesions and less likely to recognize intensity variations surrounding the lesion. Figure 5.9a

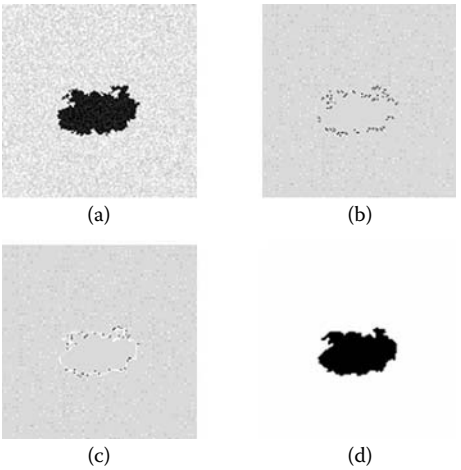


FIGURE 5.9: NNED method. (a) Noisy lesion at SNR = 12. (b) Highest intensities recognized edge patterns are the candidate points in the Gaussian curve fitting. (c) White intensity edge points representing the Gaussian curve. (d) Binary area under the edges in (c).

illustrates the NNED method when applied to a noisy synthetic lesion with SNR of 12. The edges found by the NNED are shown in Figure 5.9b. The discontinuities between edge patterns shown in Figure 5.9b can be overcome using the same Gaussian curve fitting method as used in the ASS method (Xu et al., 1999); see Figure 5.9c. Using this border outline, a binary image, shown in Figure 5.9d, can be easily formed.

It is important to mention that the NNED outputs only represent the decoded edge patterns that are expected to represent edge positions in the corresponding original input images. Synthetic images could easily provide a good test while experimenting with the NNED. Figure 5.10 compares the NNED outputs of real and synthetic lesions (the boundaries and the binary lesion are shown in Figure 5.9c–d) with the segmentation results of IS and ASS methods. Section 5.4 summarizes various experiments performed on the neural network to test this model's ability to locate edge maps.

5.4 NNED Test Experiments

The use of the SNR in characterizing noisy patterns in the training set was one of the successful methods considered to modify the technique of generating additive noise to the training set. In general, larger networks (with larger hidden nodes) show different behaviors (Rajab, 2005; Rajab and Alhindi, 2007). Some functioned as noise detectors only, and others detected

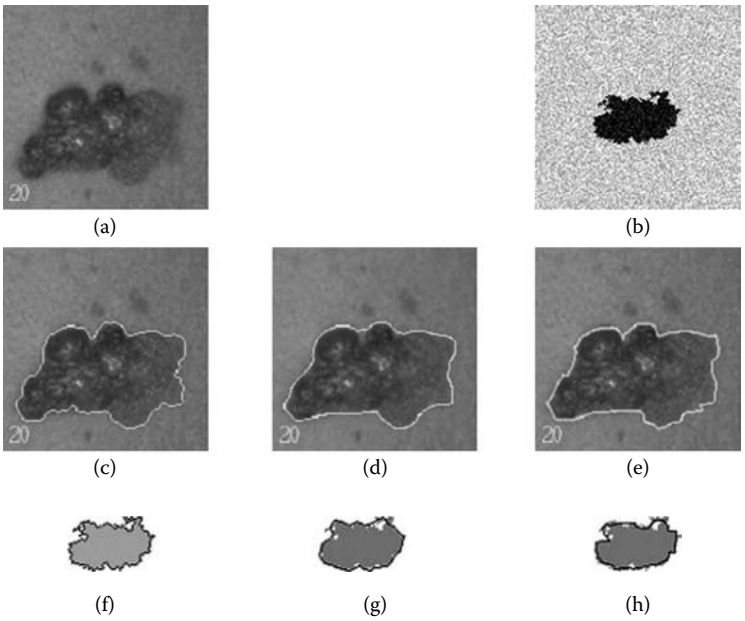


FIGURE 5.10: (See color insert following page 370.) (a) A real lesion. (b) Irregular noisy synthetic lesion (S/N = 3.0). (c–e) The lesion boundaries produced by IS and ASS, and NNED, respectively when applied to the lesion in (a). Similarly, (f–h) are the output results for IS, ASS, and NNED, respectively, when applied to a synthetic lesion in (b).

a few edge patterns in incorrect locations. Synthetic noisy images give an indication of their behavior in noisy environments. At lower SNRs, the NNED tends to detect fewer edge patterns. In this section, we present the experiments that are applied to the noise-free training sets. The successful training set that supports generalization for various synthetic images is a candidate in the study comparing segmentation errors obtained by the NNED method.

Here neural networks are trained with training sets containing sharp edges only (noise-free edge patterns). Table 5.1 compares the errors between four different sizes of training sets that are corrupted with noise and two successful noise-free training sets.

Neural network recognition improves remarkably when the number of noise-free prototype edge patterns is 42, incorporating five-edge profiles (Figure 5.7) with the possible orientations of each edge profile, and the repetition of patterns is equal to 150 times. The aim of increasing the repetition time in a noise-free training set is to provide sufficient patterns to the input layer of the neural network during the training phase. The criterion used here is that the candidate training set is the one that makes the NNED able to detect close edge patterns, which then support accurate curve fitting.

TABLE 5.1: Successful neural networks experimental results for six different sizes of training sets.

Training Set	Number of Edge Patterns	Hidden Nodes (H)	Learning Rate (η)	Momentum Rate (α)	Error (ε)	Iterations	NNED Recognize More	
							Edges	Noise
1	40	10	0.003	0.3	0.024	1700	✓	✓
2	36	10	0.003	0.3	0.020	1700	✓	✓
3	48	10	0.004	0.3	0.030	3000	✓	✓
4	60	9	0.003	0.2	0.027	1980	✓	✓
5	**20**	**9**	**0.001**	**0.5**	**0.001**	**373**	✓	
6	**42**	**10**	**0.001**	**0.5**	**0.001**	**1614**	✓	

5.5 Performance Evaluation of Segmentation Methods

In this section, we investigate the application of the IS and NNED approaches to the skin lesion segmentation problem. The aim is to quantitatively analyze the error in locating the border due to the application of an automated segmentation method. The ASS method (Xu et al., 1999) is also used here as a comparison with the two proposed methods. These approaches are compared for synthetic lesions at different image SNRs and are quantitatively evaluated using the error metric developed by Hance (Equation 2.3). Experiments are performed on digitized clinical photographs and also on pigmented networks captured with the ELM technique. A quantitative comparison is provided by comparison of the techniques on the three references of synthetic lesions shown in Figure 5.3. We also demonstrate the optimum technique on real lesions.

5.5.1 Test sets of synthetic skin lesions

Three sets of synthetic data are chosen to represent three different types of lesion irregularities. It is tedious and time consuming to generate numerous structures of synthetic binary lesions similar to those shown in Figure 5.3. Instead, we provide sufficient data sets by corrupting three candidate synthetic images (shown in Section 5.2, Figure 5.3) with different realizations of Gaussian noise (see Table 5.2). Figure 5.11 shows a sample of noisy synthetic lesions (for the case of very regular lesions), which are corrupted by 10 different realizations of Gaussian noise. These images represent only 1 out of 10 different realizations. Therefore, if we want to generate each lesion case at 10 different SNRs, for example, then in each case we will have a test set consisting of 100 noisy lesions. Table 5.2 demonstrates how we can use these three cases to generate a cluster of test images, which will finally produce a whole set of 300 test images. Such a simple and general model might provide a suitable preliminary test environment. Moreover, these synthetic images aid in the interpretation of different NNED outputs, especially at the preliminary stage of neural network edge detection experiments.

TABLE 5.2: Structure of a test set of synthetic lesions.

Contents of a Set	Number of Images
Noise levels	10
Realizations	10
Types of lesions	3
Whole set	300 images

Note: Whole set equals noise level times realizations times types of lesion.

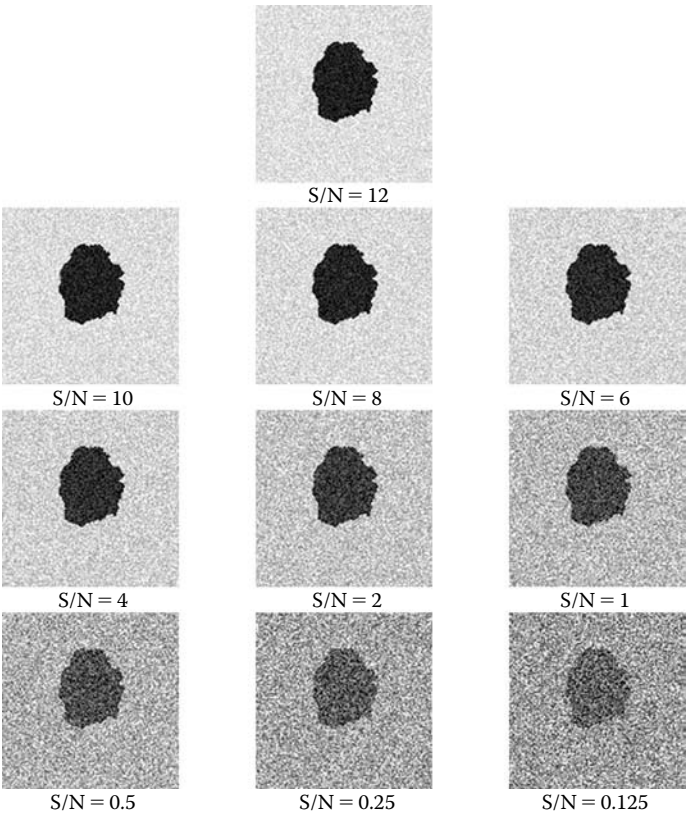


FIGURE 5.11: A very regular lesion corrupted with 10 different realizations of Gaussian noise.

The importance of good localization of lesion boundary by the result of a robust segmentation method is always considered key to the success of the evaluated method. Mean segmentation errors are computed and compared for the three synthetic lesions considered in this study (Rajab, 2003), using 12 discrete levels of Gaussian noise corresponding to SNRs between 12 and 0.125 (which correspond to 10.8 dB and −9.0 dB, respectively). Figure 5.12 shows the case of irregular lesion. It is interesting to note that when trying to estimate the SNRs for the set of 18 real high-quality SIAscope images (provided by Astron Clinica), it was found that the corresponding SNRs varied between 26.8 dB and 7.2 dB. However, lower SNRs are found in clinical photographs, for example, in a set of 20 images provided by Xu and colleagues' (1999) database of skin lesions; example images are shown in Figure 5.13. In Xu's database, the SNRs are found to vary between 13.1 dB and −6.7 dB.

Each individual point on the graph of Figure 5.12 uses 10 realizations of a random noisy lesion (at the same SNR) to compute the mean segmentation error and its standard deviation. In the case of the IS method, two calculations

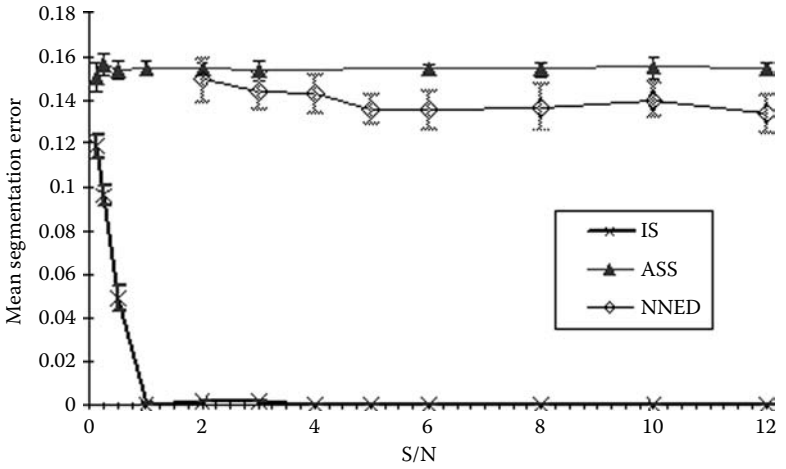


FIGURE 5.12: Comparison of mean error (%) between segmentation methods: case of irregular lesion.

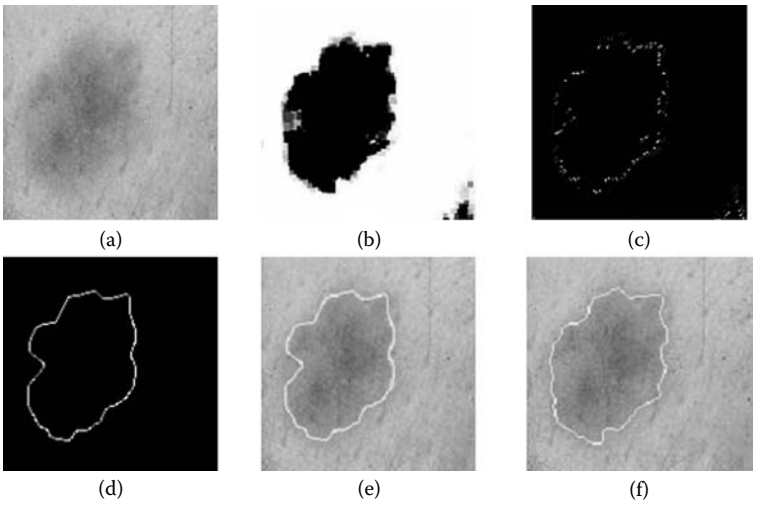


FIGURE 5.13: (a) A real lesion. (b) Intensity image produced by mapping function used in the IS algorithm. (c) Highest intensities recognized edge patterns by the neutral network. (d) The white intensity edge points by representing the Gaussian curve fitting of edge points in (c). (e) Final NNED result overlapped with the color lesion in (a). (f) Lesion boundaries produced by the IS for the color in lesion (a).

are carried out, with and without the mapping function (Section 5.2.3). In this way, we can assess the importance of this procedure when applying the IS method. At high SNR, the IS algorithm retrieves the complete lesion segment with highest precision. On the other hand, the ASS and the NNED methods have a significant level of error over the range of SNRs considered. At very low SNR, all techniques begin to break down, but the IS still achieves the lowest segmentation error. For synthetic lesions with different border irregularities, the IS method outperformed the others and should prove to be a valuable tool for skin cancer diagnosis. In the next section, the three methods are tested on real skin lesions.

The IS method has been demonstrated to be the most effective method of segmenting the three types of synthetic lesions considered (Rajab, 2003). One difference between the results from the ASS and IS methods is that the former approximates the boundary region by fitting an elastic curve between refined edge pixels, from approximated lesion boundaries, whereas the latter does not. In the IS method, instead of approximating the lesion boundary, a subsequent object-outlining operation is used to generate edge pixels from the resulting binary segment. The NNED result is also approximated by the same curve fitting as the ASS method. However, the NNED discrimination between edge and noise patterns is a critical issue, especially for the case of very noisy lesions, because the inputs to the NNED are the original noisy patterns, with no preprocessing like that used in the ASS and IS methods (i.e., intensity mapping and noise filtering). This reduction of image operations along with the simple NNED structure used has an advantage in reducing the execution time of the NNED segmentation method.

5.5.2 Real lesions

Real lesion images may contain noise such as details from skin texture and hair that make it more difficult to localize the lesion border. The first image preprocessing step in IS and NNED is to convert a real color image into its corresponding grayscale image. To ease image math calculations of the grayscale map, we adopted the technique used by Ganster and others (Ganster et al., 2001) to utilize one of the three grayscale channels in RGB color image. Because real skin images often contain narrow objects such as hair and other small objects, we added a grayscale morphological opening (Gray, 1971) operation as the first step of data reduction. Opening is a succession of two operations: erosion and dilation. The erosion of gray intensities means that each pixel is replaced with a minimum value in the structuring element, which may be the minimum value in the 3×3 neighborhood elements. Similarly, using the maximum value in a neighborhood achieves grayscale dilation. The erosion causes the thin and small objects to shrink or erode. Successively, dilation expands the eroded image by filling gaps in the lesion's contour and eliminating small holes. The remaining image operations are exactly the same as those explained in Section 5.2.3. Therefore, up to this stage, the imaging

TABLE 5.3: IS algorithm steps to delineate a color lesion.

Step 1:	{Source image} Source image = Blue channel of {R,G,B} color image
Step 2:	{Noise reduction} Gray morphology Subtract median background noise
Step 3:	{Lesion enhancement} Map intensities with appropriate function Smooth
Step 4:	{Optimal thresholding} Optimal thresholding
Step 5:	{delineate object} Outline binary object(s)
Step 6:	{Object analysis} Set minimum and maximum object size; MinSize and MaxSize Scan the binary image Until MinSize < Area < MaxSize

tasks that have been applied in the IS method to delineate a color lesion image could be summarized as shown in steps 1 through 5 in Table 5.3.

Experimental results reveal that when analyzing various real images and subtracting background noise, the median operation (Step 2) has a major drawback: it causes some holes to grow in an image. Therefore, to produce an image with a single lesion, we add another operation (Step 6). This process is useful when analyzing an image to correct errors caused in the delineation process, such as the delineation of images containing thick and dark hair. Scanning across the image is performed until a condition or a set of conditions is reached (Rajab, 2003).

When comparing the results of three segmentation methods for real lesions, consistency is found in the relative performances of these methods when applied to synthetic lesions. To illustrate this, Figure 5.10 shows how the IS, ASS, and NNED segmentation results are similar for both real and synthetic lesions. The similarity between the two categories of results indicates the success of the synthetic test model. For example, the IS method preserved the lesion's topology in both real (Figure 5.10c) and synthetic (Figure 5.10f) lesions better than the other two methods. Approximating the experts' delineations was one of the aims of the method of Xu and colleagues (1999), and accurate localization of a lesion boundary was of less importance in the Xu method. This would mostly yield regular lesion structures such as those shown in Figures 5.10d and 5.10g. However, the NNED has a different method to localize edge boundaries. The success of the trained NNED here is image dependent; that is, the NNED is trained to recognize only non-noisy sharp edges that also need an input image to contain at least similar patterns of

these sharp edges. For example, when a simple intensity-mapped image in Figure 5.13b of the real lesion shown in Figure 5.13a is applied to NNED, the output boundary produced in Figure 5.13e is similar to one obtained by the IS method in Figure 5.13f. This is mainly because the enhanced intensity-mapped image is easily learned by the NNED.

Digitized clinical photographs of skin lesions (from the database of Xu et al., 1999) were processed. These low-quality images were chosen to test the robustness of the algorithm to delineate images with clear skin texture. Figure 5.14 also compares more output results from the three segmentation methods when they are applied to the malignant lesions produced by Xu and colleague's skin lesion database. It is clearly shown that the IS method is better than the other two methods at preserving most of the lesion's topology. We also processed another set of 18 images, which were captured with the ELM technique. It was found that the ASS method produces relatively poor results when applied to these images, mainly because the class of ELM images contains very low variations between intensities, which will be incorrectly mapped

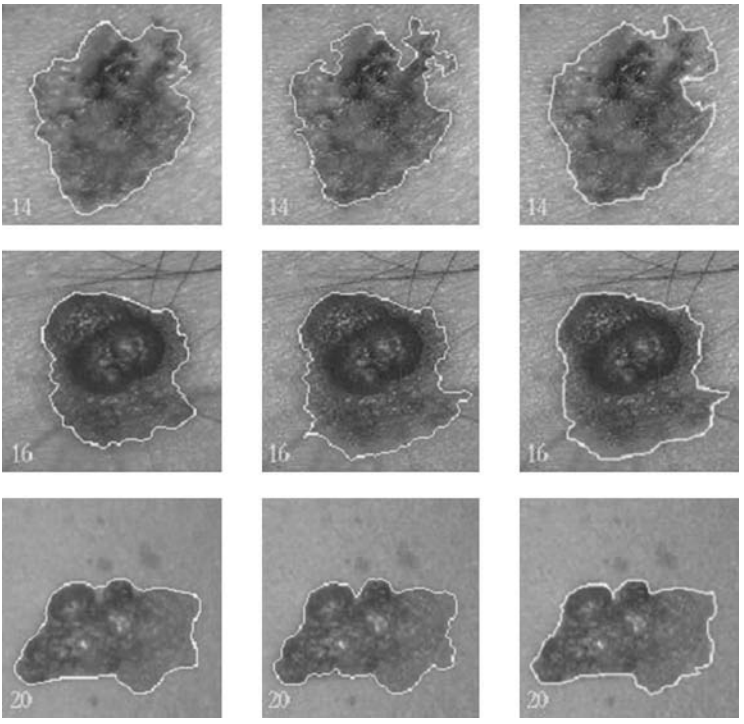


FIGURE 5.14: Final segmentation results (boundary delineations) for three malignant images by (a) ASS, (b) IS, and (c) NNED methods. Borders are super-imposed with the corresponding original malignant images.

by the ASS method. Its mapping function will include details of the background in the resulting mapped image. However, the high quality of this class of images makes the segmentation by IS and NNED methods an easy task.

5.6 Discussion

The generalization and flexibility to alter the training set structure (size, orientation, and also the amount of random noise added; see Figure 5.7) is one of the advantages to analyzing different structures of NNEDs. When constructing and experimenting on neural networks with various noise-free training sets, the total number of prototype edge patterns and their redundancy are considered. A remarkable improvement was shown in the neural network recognition when the total number of noise-free prototype edge patterns was 42, incorporating five-edge profiles with the possible orientations of each edge profile. NNED is evaluated by comparing its segmentation error with two nonneural network algorithms (IS and ASS). The synthetic lesions utilized in this work enabled a truly quantitative comparison of the three methods because the segmented images can be compared with the originals. Very low SNR was used in the simulation experiments to test the three segmentation methods in order to test them to their limits. Also, it may be possible to investigate the application of these methods to segment other types of images in which the SNR is lower than the case of real skin lesion images. The initial segmentation in the IS algorithm uses an accurate and comprehensive thresholding technique that iteratively analyzes the whole image, object and background, provided that both the lesion and its background have distinct, average gray levels. In contrast, the thresholding technique adopted in the ASS method is mostly suitable for the case of a lesion having gray levels distinct from those of the background. In practice, this case may not always exist, especially when considering a wide range of lesion scenes with widely different properties (e.g., effect of noise and image details such as skin texture and hair).

5.7 Future Research Directions

In future work, we intend to focus on the visual enhancement and delineation of skin lesions to make them accessible to further analysis and classification (e.g., using the IS method) and on the development of more sophisticated noise models to include image artifacts, which will also aid in the comparison of the various skin segmentation methods. Future NNED analysis study can include analyzing the behavior of different neural network

architectures. Using the SNR in characterizing the noisy patterns in the training set is a successful method of analyzing, studying, and modifying the NNED method (Rajab, 2005). The research of neural networks analysis implies to understand the internal behavior and functionality for various neural network architectures. The analysis study by Rajab and Alhindi (2007), for example, considered NNED's hidden units as templates, which were analyzed into three gradient components: low-pass or averaging, gradient, and second-order gradients.

5.8 Conclusions

In this chapter, a limited number of noisy edge patterns is proposed to analyze the capability of NNED for both synthetic images and real images. Various methods were applied to reduce the neural network noisy training set (e.g., variation of training set size). Using the SNR in characterizing the noisy patterns in the training set proved a successful way to modify the technique of generating additive noise to the training set.

In the next phase of the experiments, the successful generalization made by the noise-free training set is used to study the trend of the NNED method for the application to the evaluation of the segmentation errors when applied to similar noisy synthetic lesions. The principal aim in the evaluation of a skin lesion image segmentation technique is to study the effectiveness in distinguishing a lesion from the background. The NNED method is compared with two nonneural network methods: the IS and the ASS methods. The aim is to quantitatively analyze the error in locating the border due to the application of an automated segmentation method. The disadvantage of using expert delineations of actual lesions to evaluate the algorithms is that this procedure is dependent on human subjectivity. In particular, the variation in the location of the lesion border as defined by different experts is therefore not ideal when making a quantitative comparison among different automatic segmentation methods. This has motivated the generation of noise-free synthetic images with shape, irregularity, and size similar to real skin lesions. Gaussian noise is then added to the images, and the accuracy of the different techniques in segmenting the images in the presence of different levels of noise is evaluated. This enables a true quantitative comparison of different techniques. Mean segmentation errors are computed and compared for the three synthetic lesions considered. At SNR, the IS algorithm retrieves the complete lesion segment with highest precision. On the other hand, the ASS and the NNED methods have a significant level of error over the range of SNRs considered. At very low SNR, all techniques begin to break down, but the IS still achieves the lowest segmentation error. The IS technique is also tested on an image of a real skin lesion. In conclusion, for

synthetic lesions with different border irregularities, the IS method outperformed the others and should prove to be a valuable tool for skin cancer diagnosis.

References

Binder, M., Kittler, H., et al. 2000. Computer aided epiluminescence microscopy of pigmented skin lesions: the value of clinical data for the classification process. *Melanoma Research*, 10(6): 556–61.

Brenner, J. F., Lester, J. M. et al. 1981. Scene segmentation in automated histopathology: techniques evolved from cytology automation. *Pattern Recognition*, 13(1): 65–77.

Campbell, R. J., and Flynn, P. J. 2001. A survey of free-form object representation and recognition techniques. *Computer Vision and Image Understanding*, 81(2): 166–210.

Carli, P., De Giorgi, V., et al. 2000. Dermatoscopy in the diagnosis of pigmented skin lesions: a new semiology for the dermatologist. *Journal of the European Academy of Dermatology and Venereology*, 14: 353–69.

Cohn-Cedermark, G., Mansson-Brahme, E., et al. 2000. Trends in mortality from malignant melanoma in Sweden 1970–1996. *Cancer*, 89: 348–55.

Davis, L. 1975. Survey of edge detection techniques. *Computer Graphics and Image Processing*, 4: 248–70.

Fan, J., and Yau, D. K. Y. 2001. Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, 10(10): 1454–66.

Fu, K. S., and Mui, J. K. 1981. A survey on image segmentation. *Pattern Recognition*, 13(1): 3–16.

Ganster, H., Pinz, A., et al. 2001. Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, 20(3): 233–39.

Gao, J., Zhang, J., et al. 1998. Segmentation of dermatoscopic images by stabilized inverse diffusion equations. In *Proceedings of the International Conference on Image Processing '98*, Vol. 3. Los Alamitos, CA: IEEE Computer Society, pp. 823–27.

Giachetti, A., Frexia, F., et al. 2003. Distributed measurement and reporting system for surgical planning. *International Congress Series*, 1256: 828–33.

Goldman, L. 1951. Some investigative studies of pigmented nevi with cutaneous microscopy. *Journal of Investigative Dermatology*, 16: 407–26.

Golston, J. E., Stoecker, W. V., et al. 1992. Automatic detection of irregular borders in melanoma and other skin tumors. *Computerized Medical Imaging and Graphics*, 16(3): 199–203.

Gray, S. B. 1971. Local properties of binary images in two dimensions. *IEEE Transactions on Computers*, 20(5): 551–61.

Hance, G. A., Umbaugh, S. E., et al. 1996. Unsupervised color image segmentation with application to skin tumor borders. *IEEE Engineering in Medicine and Biology Magazine*, 15(1): 104–11.

Imeliska, C., Downes, M. S., et al. 2000. Semi-automated color segmentation of anatomical tissue. *Computerized Medical Imaging and Graphics*, 24(3): 173–80.

Jemal, A., Devesa, S. S., et al. 2001. Recent trends in melanoma incidence among whites in the United States. *Journal of the National Cancer Institute*, 93: 678–83.

Kanzler, M. H., and Mraz-Gernhard, S. 2001. Primary cutaneous malignant melanoma and its precursor lesions: diagnostic and therapeutic overview. *Journal of the American Academy of Dermatology*, 45(2): 260–76.

Lee, T. K., McLean, D. I., et al. 2003. Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. *Medical Image Analysis*, 7(1): 47–64.

Lippmann, R. P. 1987. 2002. An introduction to computing with neural nets. *IEEE Acoustic Speech and Signal Processing Magazine*, 4(2): 4–22.

MacKie, R. M., Bray C. A., et al. 2002. Incidence of and survival from malignant melanoma in Scotland: an epidemiological study. *Lancet*, 360(9333): 587–91.

Madisetti, V. K. and Williams D. B. *The digital signal processing handbook*, CRC Press, Boca Raton, FL, 1998.

Maintz, J. B. A., and Viergever, M. A. 1998. A survey of medical image registration. *Medical Image Analysis*, 2(1): 1–36.

Marquez G. Optical biopsy for the early detection of skin cancer. Doctoral dissertation, Texas A&M University, December 2001.

Marr, D., and Hildreth, E. 1980. Theory of edge detection. In *Proceedings of the Royal Society of London B*, vol. 207, pp. 187–217.

Marr, D., and Poggio, T. 1979. A computational theory of human stereo vision. In *Proceedings of the Royal Society of London B*, vol. 204, pp. 301–28.

Nakashima, A., Hirabayashi, A., et al. 2001. Error correcting memorization learning for noisy training examples. *Neural Networks*, 14: 79–92.

Nixon, M., and Aguado, A. 2002. *Feature Extraction and Image Processing*. Oxford: Newnes, pp. 278–79.

Olabarriaga, S. D., and Smeulders, A. W. M. 2001. Interaction in the segmentation of medical images: a survey. *Medical Image Analysis*, 5(2): 127–42.

Pehamberger, H., Binder, M., et al. 1993. In vivo epiluminescence microscopy: improvement of early diagnosis of melanoma. *Journal of Investigative Dermatology*, 100(suppl.): 356–62.

Prewitt, J. 1970. Object enhancement and extraction. In *Picture Processing and Psychopictures*. B. S. Lipkin and A. Rosenfeld, Eds. New York: Academic Press, pp. 75–149.

Rajab, M. I. 2003. Neural network edge detection and skin lesions image segmentation methods: analysis and evaluation. PhD thesis, University of Nottingham, Nottingham, U.K.

Rajab, M. I., Woolfson, M. S., and Morgan, S. P. 2004. Application of region based segmentation and neural network edge detection to skin lesions. *Computerized Medical Imaging and Graphics*, 28(1):61–8.

Rajab, M. I. 2005. Application of neural networks analysis in image feature extraction. *WSEAS Transactions on Information Science and Applications*, 2(7): 945–50.

Rajab, M. I. 2005. Feature extraction of epiluminescence microscopic images by iterative segmentation algorithm. *WSEAS Transactions on Information Science and Applications*, 2(8): 1127–30.

Rajab, M. I., and Alhindi, K. A. 2007. Analysis of neural network edge pattern detector in terms of domain functions. In *Proceedings of the International Conference on Modelling, Simulation and Applied Optimization (ICMSAO-07)*, Abu-Dhabi, March 24–27.

Rasband, W. 2000. Scion image for windows. e-mail: wayne@codon.nih.gov. http://www.scioncorp.com.

Roberts, L. G. 1965. Machine perception of three-dimensional solids. *Optical and Electro-Optical Information Processing*. J. T. Tippett et al., Eds. Cambridge, MA: MIT Press. pp. 159–97.

Rosenfeld, A. 1977. Extraction of topological information from digital images. University of Maryland, College Park, Md.: Computer Science Center, Tech. Rep. TR–547.

Steiner, A., Pehamberger, H., et al. 1987. In vivo epiluminescence of pigmented skin lesions. II. Diagnosis of small pigmented skin lesions and early detection of malignant melanoma. *Journal of the American Academy of Dermatology*, 17: 584–91.

Tarassenko, L. 1998. *A Guide to Neural Computing Applications*. London: NCAF.

Van der Hoek, D. P. 2000. Cancer in Australia 1997. [Online]. http://www.sunsmart.com.au/s/facts/skin_cancer_and_australians.htm.

van der Zwaag, B. J., Slump, C., et al. 2002. Process identification through modular neural networks and rule extraction. In *Computational Intelligent Systems for Applied Research: Proceedings of the 5th International FLINS Conference, Belgium 2002*. P. Da Ruan, et al, Eds. Singapore: World Scientific, pp. 268–77.

van der Zwang, B. J., and Slump, K. 2002. Analysis of neural networks for edge detection. In *Proceedings of the ProRISC Workshop on Circuits, Systems and Signal Processing, The Netherlands, 2002*. Utrecht: Dutch Technology Foundation STW, pp. 580–86.

van der Zwaag, B. J., Spaanenburg, L., Slump, C. H. 2002. Analysis of neural networks in terms of domain functions. In *Proceedings of the IEEE Benelux Signal Processing Symposium, Belgium, 2002*. Marc Moonen, Katholieke Universiteit Leuven, Dept. of Electrical Engineering, pp. 237–40.

Weinstock, M. A. 2000. Early detection of melanoma. *Journal of the American Medical Association*, 284(7): 886–89.

Xu, L., Jackowski, M., et al. 1999. Segmentation of skin cancer images. *Image Visions Computing*, 17: 65–74. http://www.cs.wright.edu/people/faculty/agoshtas/paper_fig.html.

Chapter 6

Prostate Cancer Classification Using Multispectral Imagery and Metaheuristics

Muhammad Atif Tahir, Ahmed Bouridane, and Muhammad Ali Roula

Contents

The introduction of multispectral imaging in pathology problems such as the identification of prostatic cancer is recent. Unlike conventional RGB color space, it allows the acquisition of a large number of spectral bands within the visible spectrum. The major problem arising in using multispectral data is high-dimensional feature vector size. The number of training samples used to design the classifier is small relative to the number of features. For such a high dimensionality problem, pattern recognition techniques suffer from the well-known curse of dimensionality. The aim of this chapter is to discuss and compare two tabu search–based computational intelligence algorithms proposed recently by authors for the detection and classification of prostatic tissues using multispectral imagery.

6.1 Introduction

Quantitative cell imagery in cancer pathology has progressed greatly in the last 25 years. The application areas are mainly those in which the diagnosis is still critically reliant on the analysis of biopsy samples. These include many common tumors such as prostate, colon, liver, breast, and cervix. These tumors account for 38% of the total incidence of all 200 types of cancers in the United Kingdom. In these types of tumors, biopsies remain the only conclusive method for making an accurate diagnosis of the disease. Biopsies are usually analyzed by a trained pathologist who, by looking at the biopsies under microscope, assesses the normality or malignancy of the samples submitted. The diagnosis is crucial for determining a course of treatment.

In December 1999, a study of more than 6,000 patients by Johns Hopkins researchers found that up to 2 out of every 100 people who come to large medical centers for treatment following a biopsy are given a diagnosis that is "totally wrong" [1]. The results suggested that second-opinion pathology examinations not only prevent errors but also save lives and money.

Over the last decade, prostate cancer has surpassed lung cancer as the most commonly diagnosed cancer in the male population with approximately 32,000 new cases diagnosed every year in the United Kingdom alone [2]. Prostate specific antigen (PSA) blood test and transrectal ultrasonography are commonly used methods of diagnosis. If a PSA-positive result is obtained, the urologist often advises a needle biopsy of the prostate in which a small sample of tissue is taken for analysis. A pathologist analyzes the textures and structures present in the samples to make a diagnosis. Different grades of malignancy correspond to different structural patterns as well as to apparent textures. This diagnosis applies whether the samples are analyzed at low-, medium-, or high-power magnification. In the case of the prostate gland, the following four major groups usually must be discriminated [3]. Figure 6.1 shows samples of the four classes.

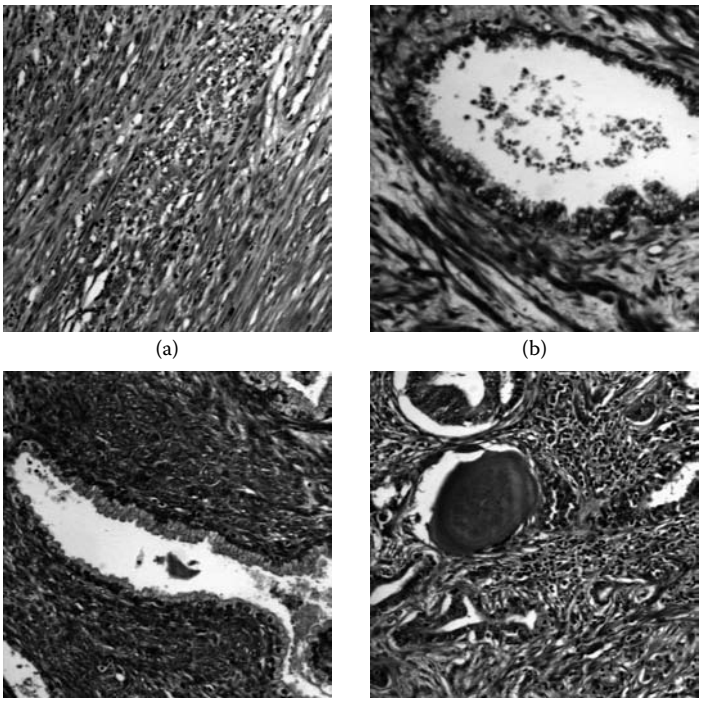


FIGURE 6.1: Images showing representative samples of the four classes: (a) stroma (STR), (b) benign prostatic hyperplasia (BPH), (c) prostatic intraepithelial neoplasia (PIN), and (d) prostatic carcinoma (PCa).

- Stroma: STR (normal muscular tissue).
- Benign prostatic hyperplasia: BPH (a benign condition).
- Prostatic intraepithelial neoplasia: PIN (a precursor state for cancer).
- Prostatic carcinoma: PCa (abnormal tissue development corresponding to cancer).

In standard clinical practice, the pathologist uses a microscope to visually navigate over the biopsy sample slides. The pathologist makes a decision about the grading and staging of the samples based on his or her expertise and experience. With the development in machine vision and intelligent image processing systems combined with advancement in computer hardware, the analysis of histopathological [1] images has become a reality. By providing quantitative measurements and objective-based decisions, machine vision can be a valuable help for pathologists and can contribute to reducing diagnosis

error cases. Such automated cancer diagnosis facilitates objective opinions for pathologists and provides a second opinion for patients [4].

Numerous investigations have been carried out using different approaches, such as morphology and texture analysis, for the classification of prostatic samples [5–10]. However, all these studies were performed using a color space that is limited either to gray-level images or to standard RGB channels. In both cases, the color sampling process causes the loss of a considerable amount of spectral information, which can be extremely valuable in the classification process.

The recent development of technologies such as high-throughput liquid crystal tunable filters (LCTF) has introduced multispectral imaging to pathology, enabling a complete, high-resolution optical spectrum to be generated for every pixel of a microscope image. Such an approach represents a completely novel way of analyzing pathological tissues. Nevertheless, a few pioneering investigations have been carried out; for example, Liu, Zahoa, and Zhang [11] used a large set of multispectral texture features for the detection of cervical cancer. Barshack and colleagues [12] used spectral morphometric characteristics on specimens of breast carcinoma cells stained with hematoxylin and eosin (H&E). Their analysis showed a correlation between specific patterns of spectra and different groups of breast carcinoma cells. Larsh and others [13] suggested that multispectral imaging can improve the analysis of pathological scenes by capturing patterns that are transparent to both human eye and standard RGB imaging. Zhang and others [14] presented a novel approach for automatic detection of white blood cells in bone marrow microscopic images. In their work, a multispectral imaging technique for image analysis is introduced.

Recently, researchers have described a novel approach in which additional spectral data are used for the classification of prostate needle biopsies [15–17]. The aim of their approach is to help pathologists reduce the diagnosis error rate. Instead of analyzing conventional grayscale or RGB color images, spectral bands are used in the analysis. Results show that the multispectral image classification using supervised linear discriminant analysis (LDA) [18] outperforms both RGB and gray-level-based classification.

Although an overall classification accuracy of 94% was achieved in their research, a principal component analysis (PCA) technique was used to reduce the high dimensionality of the feature vector. PCA has an obvious drawback: because each principal component is considered to be a linear combination of all other variables, the new variables may not have a clear physical meaning. A classification applied to PCA-reduced features may not be optimal because training may contain undesirable artifacts due to illumination, occlusion, or errors from the underlying data generation method. It is desirable not only to achieve dimensionality reduction but also to take into account the problems mentioned above in order to further improve the classification accuracy.

As discussed, the major problem arising in using multispectral data is high-dimensional feature vector size (>100). The number of training samples used

to design the classifier is small relative to the number of features. For such a high-dimensionality problem, pattern recognition techniques suffer from the well-known curse of dimensionality [19]. Recently, we proposed tabu search (TS)–based computational intelligence techniques. Section 6.2 discusses the background of the problem (i.e., curse of dimensionality, which arises from using multispectral data). Section 6.3 describes proposed TS-based computational intelligence methods for prostate cancer classification. Section 6.4 discusses sample preparation and image acquisition. In Section 6.5, experiments and results are described, followed by discussion in Section 6.6. Section 6.7 discusses future research directions, and Section 6.8 concludes the chapter.

6.2 Background

The major problem arising from multispectral data is related to the feature vector size. Typically, with 16 bands and 8 features in each band, the feature vector size is 128 [16]. For such a high-dimensionality problem, pattern recognition techniques suffer from the curse of dimensionality: keeping the number of training samples limited and increasing the number of features will eventually result in badly performing classifiers [19,20] for improving the classification accuracy of prostate cancer. The aim of this chapter is to discuss and compare two metaheuristics methods proposed recently for the detection and classification of prostatic tissues using multispectral imagery [48–50].

One way to overcome this problem is to reduce the dimensionality of the feature space. While a precise relationship between the number of training samples and the number of features is hard to establish, a combination of theoretical and empirical studies has suggested the following rule of thumb regarding the ratio of the sample size to dimensionality: the number of training samples per class should be greater than or equal to five times the features used [21]. For example, if we have a feature vector of dimension 20, then we need at least 100 training samples per class to design a satisfactory classifier.

PCA (a well-known unsupervised feature extraction method) has been used by Eble and Bostwick [3] on the large resulting feature vectors to reduce the dimensionality to a manageable size. The classification tests were carried out using the supervised LDA [18]. A classification accuracy of 94% was achieved in their experiments.

Another way to reduce the dimensionality of the feature space is by using feature selection methods. The term *feature selection* refers to the selection of the best subset of the input feature set. The methods used in the design of pattern classifiers have three goals: (1) to reduce the cost of extracting the features, (2) to improve the classification accuracy, and (3) to improve the reliability of the estimation of the performance, since a reduced feature set requires fewer training samples in the training procedure of a pattern

classifier [19,22]. Feature selection produces savings in the measuring features (since some of the features are discarded), and the selected features retain their original physical interpretation [19]. This feature selection problem can be viewed as a multiobjective optimization problem because it involves minimizing the feature subset and maximizing classification accuracy J.

Mathematically, the feature selection problem can be formulated as follows. Suppose X is an original feature vector with cardinality n, and $\bar{X}$ is the new feature vector with cardinality $\bar{n}$, $\bar{X} \subseteq X$, and $J(\bar{X})$ is the selection criterion function for the new feature vector $\bar{X}$. The goal is to optimize $J()$.

This feature selection problem is NP-hard [23,24]. Therefore, the optimal solution can only be achieved by performing an exhaustive search in the solution space [32]. However, exhaustive search is feasible only for small n. A number of algorithms have been proposed for feature selection to obtain near-optimal solutions [19,22,25–29]. The choice of an algorithm for selecting the features from an initial set depends on n. The feature selection problem is said to be of small scale, medium scale, or large scale according to whether n belongs to the intervals [1, 19], [20, 49] or [50, ∞], respectively [22, 27]. Sequential forward selection (SFS) [30] is the simplest greedy sequential search algorithm and has been used for landmine detection using multispectral images [31]. Other sequential algorithms, such as sequential forward floating search (SFFS) and sequential backward floating search (SBFS), are more efficient than SFS and usually find fairly good solutions for small- and medium-scale problems [26]. However, these algorithms suffer from the deficiency of converging to local optimal solutions for large-scale problems when n is greater than 100 [22,27]. Recent iterative heuristics such as TS and genetic algorithms have proven effective in tackling this category of problems, which are characterized by having an exponential and noisy search space with numerous local optima [27,28,33,34].

In previous work [48–50], we proposed TS-based computational intelligence techniques to solve this curse of dimensionality. In the following sections, a number of computational intelligence techniques recently proposed for the detection and classification of prostatic tissues are described with emphasis on their impact of classification accuracy and simplicity of implementation.

6.3 Computational Intelligence Techniques for Prostate Cancer Classification

Two different methods have been proposed to improve the classification accuracy using tabu search [48–50]. The first method uses a simple multiclass approach; the other method uses a round-robin approach. In this section, we discuss both methods. We first discuss feature selection using tabu search since it has been used in both methods.

6.3.1 Feature selection using tabu search

6.3.1.1 Overview of tabu search

Tabu search was introduced by Glover [35,36] as a general iterative metaheuristic for solving combinatorial optimization problems. Tabu search is conceptually simple and elegant. It is a form of local neighborhood search that starts from an initial solution and then examines feasible neighboring solutions. It moves from a solution to its best admissible neighbor, even if this causes the objective function to deteriorate. To avoid cycling, solutions that were recently explored are declared forbidden, or tabu, for a number of iterations. The tabu status of a solution is overridden when certain criteria (aspiration criteria) are satisfied. Sometimes intensification and diversification strategies are used to improve the search. In the first case, the search is accentuated in promising regions of the feasible domain. In the second case, an attempt is made to consider solutions in a broad area of the search space. The TS algorithm is given in Figure 6.2.

6.3.1.2 Fuzzy objective function

In this section, we present a TS algorithm in which the quality of a solution is characterized by a fuzzy logic rule expressed in linguistic variables of the problem domain. Fuzzy set theory has recently been applied in many areas of science and engineering. In the most practical situations, one is faced with several concurrent objectives. Classic approaches usually deal with such difficulty by computing a single utility function as a weighted sum of the individual objectives, where more important objectives are assigned higher weights [50]. Balancing different objectives by weight functions is at best controversial. Fuzzy logic is a convenient vehicle for trading of different objectives. It allows the mapping of values of different criteria into linguistic values that characterize the level of satisfaction of the system designer with the numerical value of objectives and operation over the interval [0, 1] defined by the membership functions for each objective.

Three linguistic variables are defined to correspond to the three component objective functions: number of features f_1, number of incorrect predictions f_2, and average classification error rate f_3. One linguistic value is defined for each component of the objective function. These linguistic values characterize the degree of satisfaction of the designer with the values of objectives $f_i(x)$, $i = \{1, 2, 3\}$. These degrees of satisfaction are described by the membership functions $\mu_i(x)$ on fuzzy sets of the linguistic values, where $\mu(x)$ is the membership value for solution x in the fuzzy set. The membership functions for the minimum number of features, the minimum number of incorrect predictions, and the low classification error rate are easy to build. They are assumed to be nonincreasing functions because the smaller the number of features $f_1(x)$, the number of incorrect predictions $f_2(x)$, and the classification error rate $f_3(x)$, the higher is the degree of satisfaction $\mu_1(x)$, $\mu_2(x)$, and $\mu_3(x)$ of the expert

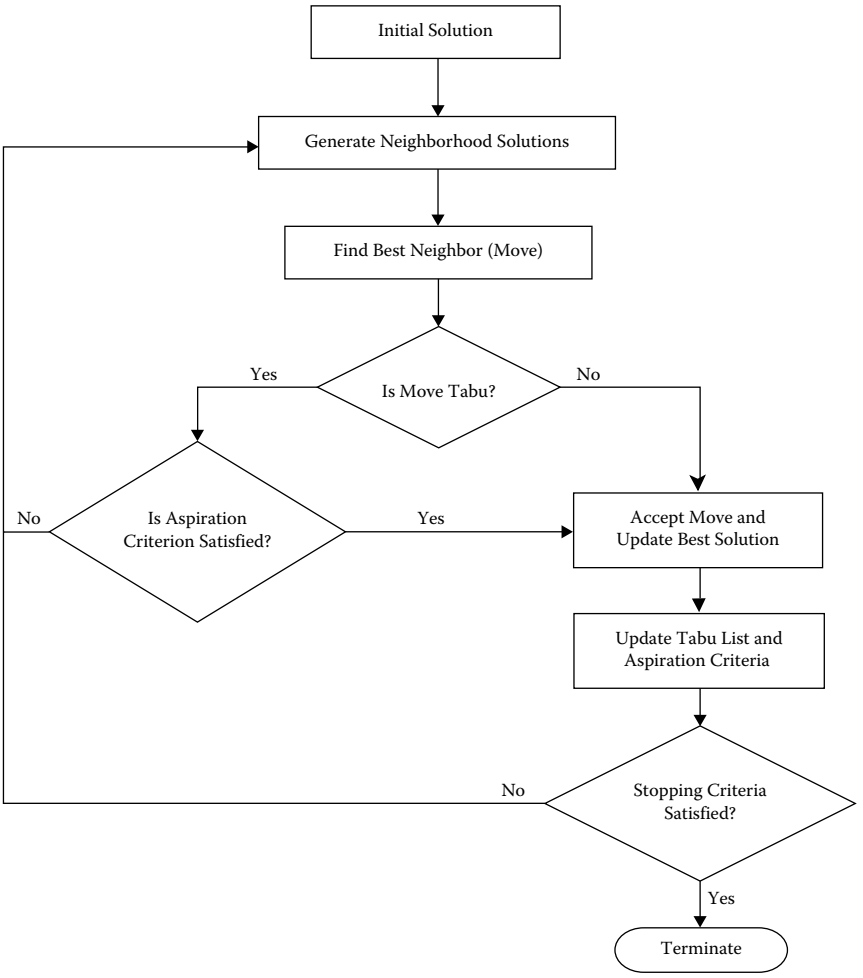


FIGURE 6.2: Flow chart of a short-term tabu search.

system (see Figure 6.3). The fuzzy subset of a good solution is defined by the following fuzzy logic rule:

> IF a solution has a small number of features AND a small number of incorrect predictions AND low classification error rate, THEN it is a good solution.

According to the AND/OR like ordered-weighted-averaging logic [40,41], the above rule corresponds to the following:

$$\mu(x) = \gamma \times min(\mu_i(x)) + (1 - \gamma) \times \frac{1}{3} \sum_{i=1}^{3} \mu_i(x) \tag{6.1}$$

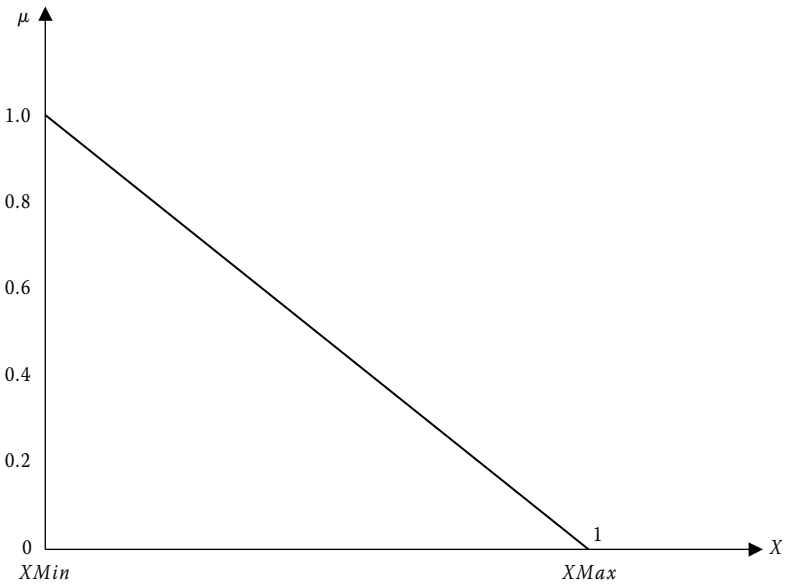


FIGURE 6.3: Membership function for fuzzy subset X, where, in this application, X is the number of features F, the number of incorrect predictions P, or the classification error rate E.

where γ is a constant in the range [0, 1]. The shape of the membership function $\mu(x)$ is shown in Figure 6.3. Membership of data in a fuzzy set is defined using values in the range [0, 1]. The membership values for the number of features F, the number of incorrect predictions P, and the classification error rate E are computed using Equations 6.2, 6.3, and 6.4 respectively.

$$\mu_1(x) = \begin{cases} 1 & \text{if} \quad F \leq FMin \\ \dfrac{FMax - F}{FMax - FMin} & \text{if} \quad FMin \leq F \leq FMax \\ 0 & \text{if} \quad FMax \leq F \end{cases} \tag{6.2}$$

$$\mu_2(x) = \begin{cases} 1 & \text{if} \quad P \leq PMin \\ \dfrac{PMax - P}{PMax - PMin} & \text{if} \quad PMin \leq P \leq PMax \\ 0 & \text{if} \quad PMax \leq P \end{cases} \tag{6.3}$$

$$\mu_3(x) = \begin{cases} 1 & \text{if} \quad E \leq EMin \\ \dfrac{EMax - E}{EMax - EMin} & \text{if} \quad EMin \leq E \leq EMax \\ 0 & \text{if} \quad EMax \leq E \end{cases} \tag{6.4}$$

The maximum number of features ($FMax$) is the size of the feature vector and the minimum number of features ($FMin$) is 1. The maximum number of incorrect predictions ($PMax$) and the maximum classification error rate ($EMax$) are determined by applying a 1NN classifier [37] for the initial solution. The minimum number of incorrect predictions ($PMin$) is 0, while the minimum classification error rate ($EMin$) is 0%. Neighbors are calculated using a squared Euclidean distance defined as:

$$D(x, y) = \sum_{i=1}^{m} (x_i - y_i)^2 \tag{6.5}$$

where x and y are two input vectors and m is the number of features.

6.3.1.3 Initial solution

The feature selection vector is represented by a 0/1 bit string, where 0 indicates that the feature is not included in the solution, and 1 indicates that it is. All features are included in the initial solution.

6.3.1.4 Neighborhood solutions

Neighbors are generated by randomly adding or deleting a feature from the feature vector of size n. For example, if 11001 is the current feature vector, then the possible neighbors with a candidate list size of 3 might be 10001, 11101, 01001. Among the neighbors, the one with the best cost (i.e., the solution that results in the minimum value of Equation 1.1) is selected and considered as a new current solution for the next iteration.

6.3.1.5 Tabu moves

A tabu list is maintained to avoid returning to previously visited solutions. Using this approach, if a feature (move) is added or deleted at iteration i, then adding or deleting the same feature (move) for T subsequent iterations (tabu list size) is tabu.

6.3.1.6 Aspiration criterion

Aspiration criterion is a mechanism used to override the tabu status of moves. It temporarily overrides the tabu status if the move is sufficiently good. In our approach, if a feature is added or deleted at iteration i, and this move results in a best cost for all previous iterations, then this feature is allowed to add or delete even if that feature is in the tabu list.

6.3.1.7 Termination rule

The most commonly used stopping criteria in tabu search are

- After a fixed number of iterations.

- After some number of iterations when there has been no increase in the objective function value.
- When the objective function reaches a prespecified value.

In this algorithm, the termination condition is implemented using the fixed number of iterations criterion.

6.3.1.8 Intensification

For intensification, the search is concentrated in the promising regions of the feasible domain. Intensification is based on some intermediate-term memory. Since the solution space is extremely large (with initial feature vector $n > 100$), the search in the promising regions is intensified by removing poor features from the search space. The following steps, where the values of M and N are determined empirically, are proposed for this purpose:

STEP 1: Store the M best solutions in intermediate memory for T_1 iterations.

STEP 2: Remove features that are not included in the best M solutions for N times.

STEP 3: Rerun the tabu search with the reduced set of features for another T_2 iterations.

STEP 4: Repeat steps 1–3 until there is improvement in the objective function.

As an example, assume that the following $M = 5$ best solutions, as shown in Figure 6.4, are found by tabu search during T_1 iterations. Feature f_1 is always used, while feature f_5 is never used in good solutions. For $N = 2$, the reduced feature set comprises only f_1, f_2, f_3, f_6, and f_8. The search space is now reduced from 2^8 to 2^5. Thus, tabu search will search for the near-optimal solutions in a reduced search space, avoiding visits to nonpromising regions.

6.3.2 Approach 1: Multiclass classification using tabu search

In this approach, the classifier treats all classes as one multiclass problem. Figure 6.5 illustrates a multiclass (four-class) learning problem where one classifier (TS/1NN classifier in this study) separates all classes.

6.3.3 Approach 2: Round-robin classification using tabu search

In Approach 1, we addressed the high-input dimensionality problem by selecting the best subset of features using an intermediate-memory tabu search

$f1$	$f2$	$f3$	$f4$	$f5$	$f6$	$f7$	$f8$
1	0	1	0	0	1	0	1
1	0	1	1	0	1	1	1
1	1	0	0	0	0	1	1
1	1	1	0	0	1	0	0
1	1	1	0	0	1	0	0
Σ	**5**	**3**	**4**	**1**	**0**	**4**	**2**	**3**

FIGURE 6.4: An example showing intensification steps for tabu search. $\sum$ is the number of occurrences of each feature in the best solutions.

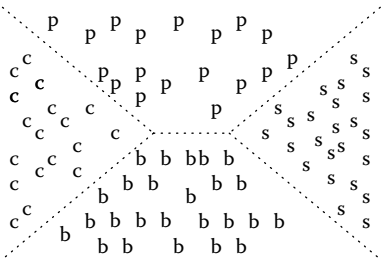


FIGURE 6.5: Multiclass learning. p, PIN; c, PCa; b, BPH; s, STR.

with a classification process using a 1NN classifier. In this approach, the classifier treats all classes as one multiclass process. We have proposed another scheme [48] in which the multiclass problem is solved using round-robin classification in which the classification problem is decomposed into a number of binary classes. The key point is that it is then possible to design simpler and more efficient binary classifiers, as described below.

6.3.3.1 Round-robin classification

Johannes [38] defines round-robin classification as follows:

> The round robin or pairwise class binarization transforms a c-class problem into $c(c-1)/2$ two-class problems $<i,j>$, one for each set of classes $\{i,j\}$, $i = \{1.....c-1\}$, $j = \{i+1.....c\}$. The binary classifier for problem $<i,j>$ is trained with examples of classes

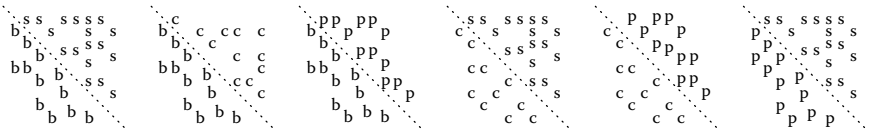


FIGURE 6.6: Round-robin learning. p, PIN; c, PCa; b, BPH; s, STR.

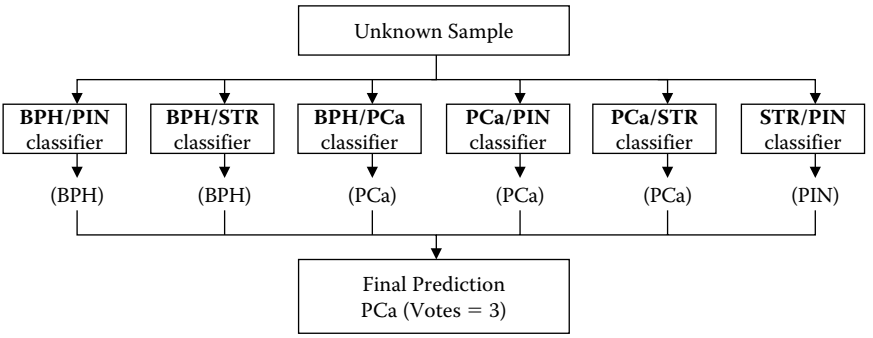


FIGURE 6.7: Simple voting scheme.

i and j, whereas examples of classes $k \neq i, j$ are ignored for this problem.

Figure 6.6 shows round-robin learning with $c(c-1)/2$ classifiers. For a four-class problem, the round robin trains six classifiers, one for each pair of classes. Each class is trained using a feature selection algorithm based on the intermediate-memory TS/1NN classifier, as discussed in Approach 1. A simple voting technique [38] is then used to combine the predictions of the pairwise classifiers, thereby computing the final result. In the case of tie, a distance metric (squared Euclidean distance) is used for the final prediction. Figure 6.7 illustrates the simple voting scheme. When classifying an unknown new sample, each classifier (1NN in this case) determines to which of its two classes the sample is more likely to belong.

6.4 Sample Preparation, Image Acquisition, and Data Sets Description

Methods for data collection [16,17] are reviewed briefly here. Entire tissue samples were taken from prostate glands. Sections $5\,\mu$m thick were extracted and stained using H&E. These samples were routinely assessed by

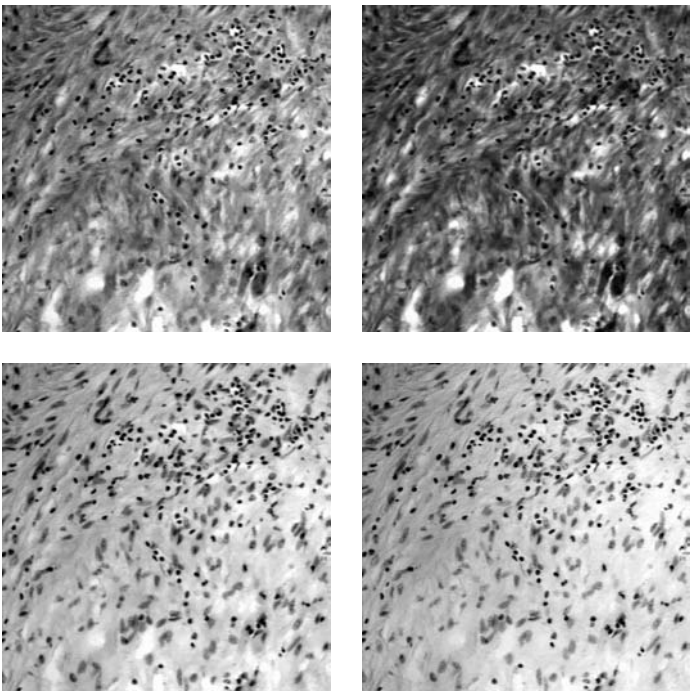


FIGURE 6.8: Images showing different subbands of multispectral image of type Stroma.

two experienced pathologists and graded histologically as showing STR, BPH, PIN, and PCa.

From these samples, whole section subimages were captured using a classical microscope and CCD camera. An LCTF (VariSpec^{TM}) was inserted in the optical path between the light source and the chilled charge-coupled device (CCD) camera. The LCTF has a bandwidth accuracy of 5 nm. The wavelength is controllable through the visible spectrum (from 400 nm to 720 nm). This allowed the capture of different multispectral images of the tissue samples at different spectral frequencies. In order to show the impact of multispectral imaging, experiments were carried out for varying numbers of bands [16,17]. It was shown that the classification accuracy increases with the number of spectral bands. Figures 6.8 through 6.11 show the thumbnails of eight bands of multispectral images of type STR, PCa, BPH, and PIN, respectively.

6.4.1 Data sets description

The round-robin TS/1NN classifier was tested on two data sets [16,17]. In order to offset any bias due to the different range of values for the original features, the input feature values are normalized over the range [1, 11]

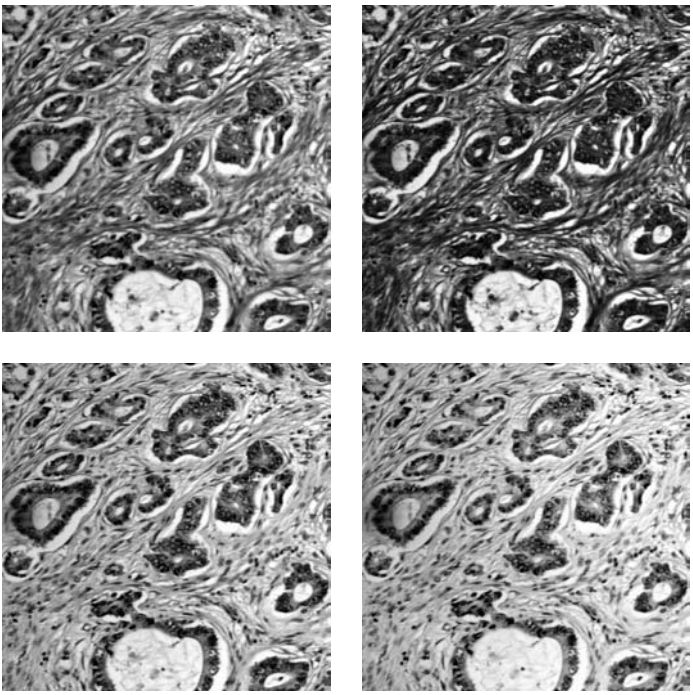


FIGURE 6.9: Images showing different subbands of multispectral image of type PCa.

using Equation 6.6 [44]. Normalizing the data is important to ensure that the distance measure allocates equal weight to each variable.

Without normalization, the variable with the largest scale will dominate the measure.

$$x'_{i,j} = \left(\frac{x_{i,j} - min_{k=1\ldots n} x_{(k,j)}}{max_{k=1\ldots n} x_{(k,j)} - min_{k=1\ldots n} x_{(k,j)}} * 10 \right) + 1 \tag{6.6}$$

where $x_{i,j}$ is the j^{th} feature of the i^{th} pattern, $x'_{i,j}$ is the corresponding normalized feature, and n is the total number of patterns.

The first data set consists of textured multispectral images taken at 16 spectral channels (from 500 nm to 650 nm) [16]. Five hundred ninety-two different samples (multispectral images) of size 128×128 were used to carry out the analysis. The samples are seen at low power ($\times$40 objective magnification) by two highly experienced, independent pathologists and labeled into four classes: 165 cases of STR, 106 cases of BPH, 144 cases of PIN, and 177 cases of PCa. The size of the feature vector is 128: 16 Bands $\times$ 8 Features (1 Statistical + 2 Structural + 5 Haralick).

The second data set is derived from prostatic nuclei extracted from prostate tissue [17]. Nuclei are imaged under high power ($\times$100 objective

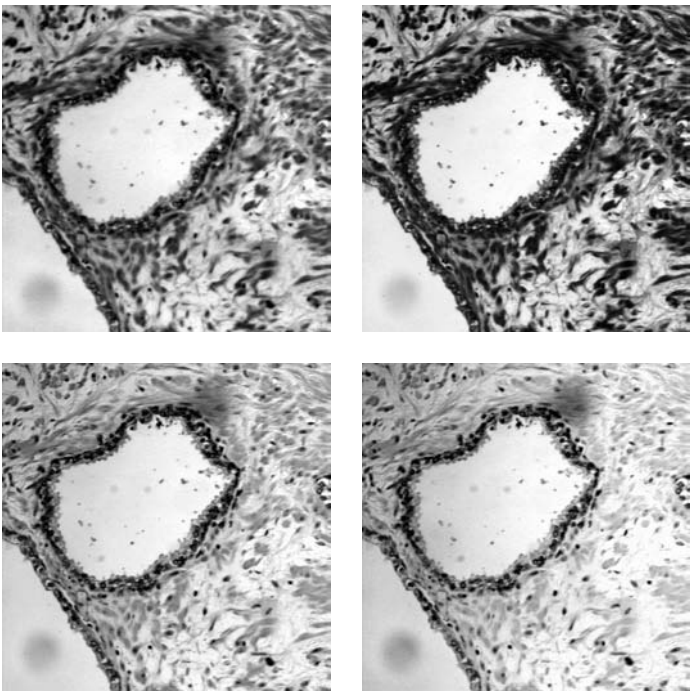


FIGURE 6.10: Images showing different subbands of multispectral image of type BPH.

magnification). These nuclei are taken at 33 spectral channels (from 400 nm to 720 nm). Two hundred thirty different images of size 256×256 are used to carry out the analysis. The samples are labeled into three classes: 63 cases of BPH, 79 cases of PIN, and 88 cases of PCa. The size of feature vector is 266: (33 Bands $\times$ 8 Features (3 Statistical + 5 Haralick) + 2 Morphology Features).

The explanation of the choice of features is discussed in the works of Roula and others [16,17]. The following subsections briefly review the features.

6.4.1.1 Data set 1

The following features are used in data set 1 [16].

Statistical feature: Statistical feature (variance) is calculated for each band in a multispectral image and added to the feature vector.

Haralick features: The five Haralick features [39] (dissimilarity, contrast, angular second moment, entropy, and correlation) are calculated for each band in a multispectral image and added to the feature vector.

Structural features: Roula and others [16] showed that statistical and Haralick features are not enough to capture the complexity of the patterns in prostatic neoplasia. BPH and PCa present more complex structures, as both

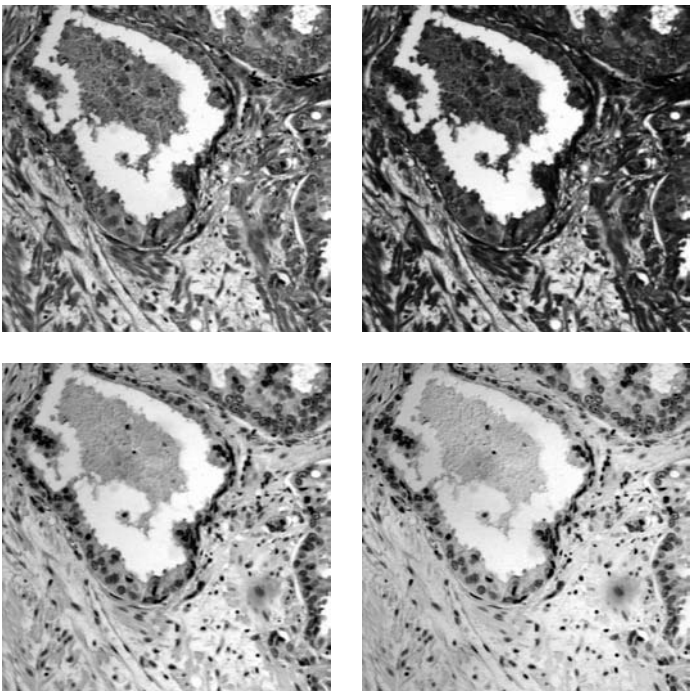


FIGURE 6.11: Images showing different subbands of multispectral image of type PIN.

contain glandular areas in addition to nuclei clusters. Accurate classification requires the quantification of these differences. Quantification first requires segmenting the glandular and the nuclear areas using their color differences: glandular areas are lighter than the surrounding tissue, while nuclear clusters are darker. Two structural features (one with the number of pixels classified as glandular area and the other with number of pixels classified as nuclear area) are computed for each band in a multispectral image and added to the feature vector.

6.4.1.2 Data set 2

The following features are used in data set 2 [17].

Statistical features: Three statistical features (mean, standard deviation, and geometric moment) are calculated for each band in a multispectral image and added to the feature vector.

Haralick features: Five Haralick features [39] (dissimilarity, contrast, angular second moment, entropy, and correlation) are calculated for each band in a multispectral image and added to the feature vector.

Morphology features: Two morphology features (nuclei area and nuclei round factor) [7] are calculated and added to the feature vector.

6.5 Experiments and Results

The leave-one-out method is used for cross-validation [19]. In this method, a classifier is designed using $(s-1)$ samples and is evaluated on the one remaining sample; the process is repeated s times, with different training sets of size $(s-1)$.

Table 6.1 shows the confusion matrix for data set 1. From these results, it can be observed that a round-robin-based classification yields better results than the multiclass approach. The overall classification error has been reduced to 1.23% from 5.71% and 2.90% using PCA/LDA and 1NN/TS respectively. Table 6.2 shows that the overall classification error has been reduced to 0% from 5.1% and 0.91% using PCA/LDA and 1NN/TS, respectively, for data set 2.

Table 6.3 shows the comparison of RR-TS/1NN with bagging and boosting. Bagging [42] and boosting [43] are also well-known ensemble design techniques to improve the prediction of classifier systems. It is clear from the table that

TABLE 6.1: Confusion matrix for data set 1.

PCA/LDA					
C	**BPH**	**PCa**	**PIN**	**STR**	**E(%)**
BPH	100	3	2	1	5.7
PCA	1	174	2	0	1.7
PIN	1	4	133	6	7.6
STR	11	1	1	152	7.9
O					5.7
Multiclass TS/1NN					
C	**BPH**	**PCa**	**PIN**	**STR**	**E(%)**
BPH	101	1	0	4	4.7
PCA	1	174	2	0	1.7
PIN	0	2	141	2	2.8
STR	2	2	0	161	2.4
O					2.9
Round-Robin TS/1NN					
C	**BPH**	**PCa**	**PIN**	**STR**	**TM**	**E(%)**
BPH	104	0	0	2	2	1.9
PCA	0	175	2	1	2*	1.1
PIN	0	1	143	0	1	0.7
STR	1	1	0	163	2	1.2
O						1.2

Note: One misclassified sample is common in binary classifiers (PCa vs. PIN and PCa vs. Stroma).
C, classified as; STR, stroma; TM, total misclassified; O, overall.

TABLE 6.2: Confusion matrix for data set 2.

	PCA/LDA				Multiclass TS/1NN				Round-Robin TS/1NN			
C	**BPH**	**PCa**	**PIN**	**E(%)**	**BPH**	**PCa**	**PIN**	**E(%)**	**BPH**	**PCa**	**PIN**	**E(%)**
BPH	60	2	1	4.8	62	1	0	1.6	63	0	0	0
PCa	3	82	3	6.8	1	87	0	1.1	0	88	0	0
PIN	1	2	76	3.8	0	0	79	0	0	0	79	0
O				5.1				0.9				0

C, classified as; O, overall.

TABLE 6.3: Classification accuracy (%) using various ensemble techniques.

	C4.5				Nearest Neighbor		
	C4.5	Bagging	Boosting	NN	Bagging	Boosting	Round-Robin TS
Data set 1	91.6	93.2	95.4	88.0	89.2	88.1	98.9
Data set 2	80.9	83.9	89.1	90.9	91.7	90.4	100

the proposed round-robin ensemble technique using TS/1NN outperformed both bagging and boosting ensemble design techniques.

Tables 6.4 and 6.5 show the number of features used by various data reduction techniques for data set 1 and data set 2, respectively. Different numbers of features have been used by the various binary classifiers, producing an overall increase in the classification accuracy. F_c represents those features that are common in two or more different binary classifiers. Although the total number of features has increased in our proposed round-robin technique, the number of features used by each binary classifier is never greater than that used in any multiclass method with either PDA/LDA or TS/1NN. Consequently, multispectral data are better utilized by using a round-robin technique since the use of more features means more information is captured and used in the classification process. Furthermore, simple binary classes are also useful for analyzing features and are extremely helpful for pathologists in distinguishing various patterns such as BPH, PIN, STR, and PCa.

TABLE 6.4: Number of features used by different classifiers.

i	DR Technique		Features Used (F)
	PDA/LDA	Multiclass	20
	TS/1NN	Multiclass	16
1	TS/1NN	Binary class (BPH, STR)	13
2		Binary class (PCa, PIN)	15
3		Binary class (PIN, BPH)	10
4		Binary class (PIN, STR)	13
5		Binary class (PCa, BPH)	16
6		Binary class (PCa, STR)	11
	TS/1NN	Round robin	$\sum_{i=1}^{6} F - F_c$ $= (78 - 13) = 65$

Total number of features: 128. F_c, common features; DR, data reduction.

TABLE 6.5: Number of features used by different classifiers.

i	DR Technique		Features Used (F)
	PDA/LDA	Multiclass	25
	TS/1NN	Multiclass	13
1	TS/1NN	Binary class (PCa, PIN)	7
2		Binary class (BPH, PCa)	10
3		Binary class (PIN, BPH)	6
	TS/1NN	Round robin	$\sum_{i=1}^{3} F - F_c$ $= (23 - 1) = 22$

Total number of features: 266. F_c, common features; DR, data reduction.

6.6 Discussion

The algorithms discussed in this chapter are generic and can be used for the diagnosis of other diseases, such as lung and breast cancer. A key characteristic of the round-robin approach is that different features are captured and used for each binary classifier in the four-class problem, thus producing an overall increase in classification accuracy. In contrast, in a multiclass problem, the classifier tries to find those features that distinguish all four classes at once. Furthermore, the inherent curse-of-dimensionality problem, which arises in multispectral data, is also resolved by the round-robin TS/1NN classifiers, since each classifier is trained to compute and use only those features that distinguish its own binary class.

6.6.1 Quality of solutions produced by tabu search

Figure 6.12 shows the value of an objective function versus the number of iterations when searching the solution search space using tabu search for binary classifier BPH versus STR. The objective functions are fuzzy membership functions, the number of incorrect predictions, the classification error rate, and the number of features. All figures show how well focused the tabu search is on the good solution space. The graphs also show that the tabu search rapidly converges to the feasible/infeasible region border for all of these

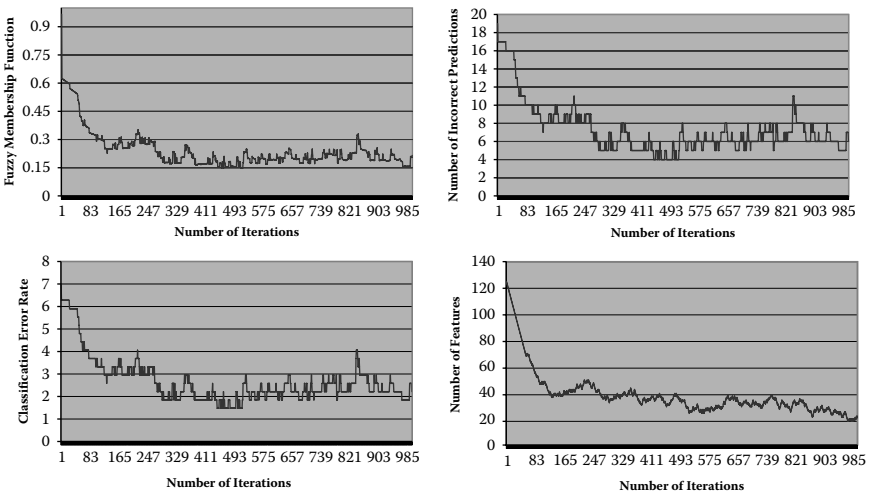


FIGURE 6.12: Objective function versus number of iterations for BPH versus STR before intensification.

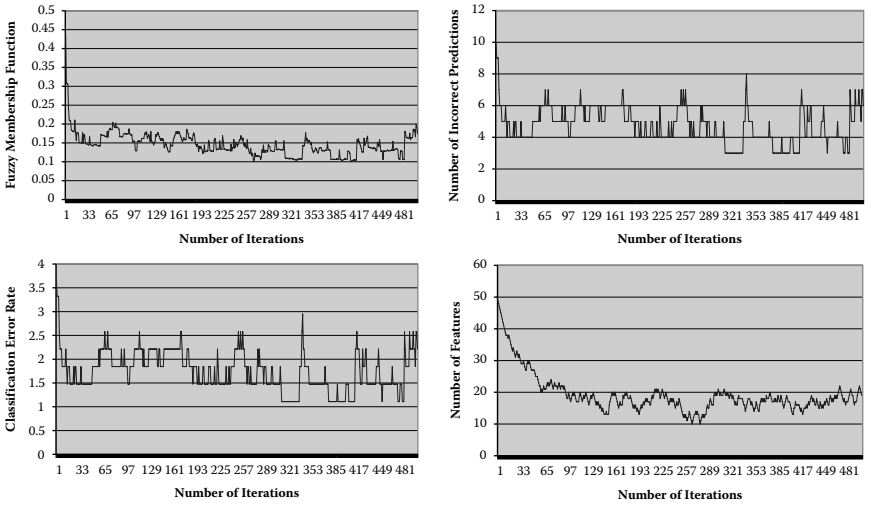


FIGURE 6.13: Objective function versus number of iterations for BPH versus STR after intensification.

TABLE 6.6: Tabu runtime parameters.

	Before Intensification	After Intensification
V^*	10	8
T	10	8
I	1000	500

V^*, number of neighborhood solutions; T, tabu list size; I, number of iterations.

objectives. Figure 6.13 depicts the value of objective functions versus the number of iterations after reducing the size of the feature set by using an intensification technique discussed in Section 6.3.1. The graphs show that the search for the best solutions is now limited only to the good solution space (i.e., the membership function is in the range 0.1 to 0.15 for most of the iterations, while the same membership function was in the range 0.15 to 0.3 without the inclusion of an intensification process). Similarly, the number of features is in the range 10 to 20 for most of the iterations, while the number of features was in the range 20 to 40 without intensification.

6.6.2 Runtime parameters for tabu search

Table 6.6 shows the tabu runtime parameters chosen after experimentation with different values. The values of M and N, mentioned in Section 6.3.1.8, are 100 and 10, respectively.

TABLE 6.7: Computation time comparison.

		M (msec)	**C (msec)**	**T (msec)**
Multiclass	Data set 1	48.2	0.11	48.31
	Data set 2	65.3	0.061	65.361
Round robin	Data set 1	74.0	0.49	74.49
	Data set 2	109.34	0.14	109.48

M, measuring features cost; C, classification; T, total time.

6.6.3 Computation time

The training of round-robin classifiers using TS/1NN is an offline procedure used to find the best subset of features while keeping the classification error rate low for each binary classifier. Once the tabu search finds the best subset of features, a 1NN classifier is used to determine the class of the new sample (a multispectral image), which provides an online diagnosis decision for the pathologist.

Table 6.7 shows the computation times for multiclass and binary class classifiers when determining the class of the new sample. The execution times using round-robin classifiers are higher than those for the multiclass classifier. The cost of measuring features is increased because more features are required in the round-robin approach. Thus, the classification accuracy using the round-robin technique has been improved for multispectral imagery at a cost of 1.54 and 1.68 times the execution time for data set 1 and data set 2, respectively.

6.7 Future Research Directions

Due to the pioneering nature of this area of research, there are many directions for future investigations. We believe that the novel use of multispectral imagery in pathology must be explored further, especially for other cancer data types. Issues that can be investigated in the future include the following:

- *Improvement of the classification rates:* The tabu search algorithm with only short and intermediate (intensified) memory was used in this chapter. Diversification techniques (use of long-term memory in tabu search) have not been explored. Future research should explore diversification techniques to enhance the classification accuracy using tabu search for prostate cancer classification.

- *Computational challenge:* With the emerging new technology of scanscope [45], pathologists will no longer be restricted to viewing a very small portion of the tissue samples. Entire slides can be viewed at a very high resolution, and larger data sets will become available to pathologists. Consequently, improving the computation will become critical. The identification of entire samples at high resolution (very large

images) requires using medium- and high-power identification techniques along with low-power detection of regions of interest. The analyses at low and high power are performed separately. Scanscope images allow accessing the same areas viewed at low power but with a high-power resolution. It would be interesting to optimally combine the results obtained using these two approaches similarly to what is already applied in remote sensing imagery [46]. The low power could be used to draw a probability map of areas of interest from which the high-power analysis could be directed to specific regions, and the probabilities obtained from both magnifications could then be combined to give a final decision.

- *Visible spectrum and beyond:* The same company that developed the visible light LCTF, VariSpec, is now commercializing a new model for mid-infrared (MIR) imaging: VariSpec MIR. This new model allows the capture of spectral images with wavelengths up to 2,000 nm. The visible spectrum is a relatively narrow band within the spectrum of frequencies that can practically be used in microscopy. There is increasing evidence that the area beyond the visible spectrum can be interesting to study from a pathological point of view. For example, recent research [47] suggests that the MIR bands might even contain richer information content than the visible bands. The authors applied the extracted one-dimensional spectrum of adenocarcinoma and crypts in colorectal tissue and showed a correlation between spectral peak patterns and the presence of biochemicals such as carbonyl, Amid II, and lipid/protein ratio. Applying the methods used in this chapter to MIR multispectral data could be very useful for understanding the biochemistry of prostate cancer and its diagnosis.

6.8 Conclusions

The introduction of multispectral imaging in pathology problems such as the identification of prostatic cancer is recent. In this chapter, a potential curse of dimensionality related to using multispectral data for prostate cancer diagnosis was addressed. Four major groups of prostate cancer were successfully discriminated: stroma, benign prostatic hyperplasia, prostatic intraepithelial neoplasia, and prostatic carcinoma. Two metaheuristics methods recently proposed for the detection and classification of prostatic tissues using multispectral images were described with emphasis on their impact of classification accuracy and simplicity of implementation. Results indicated a significant increase in the classification accuracy. The algorithms discussed in this chapter are generic and can be used for the diagnosis of other diseases, such as lung and breast cancer. Furthermore, the proposed tabu search–based metaheuristic progressively zoomed toward a better solution subspace as time elapsed, a desirable characteristic of iterative heuristics.

References

[1] J. D. Kronz, W. H. Westra, and J. I. Epstein. Mandatory second opinion: Surgical pathology at a large referral hospital. *Cancer*, 86(11): 2426–2435, 1999.

[2] Cancer Research UK. http://www.prostate-cancer.org.uk/news/news.asp. September 25, 2006.

[3] J. N. Eble and D. G. Bostwick. *Urologic Surgical Pathology*. St. Louis: Mosby-Year Book, 1996.

[4] C. Demir and B. Yener. Automated cancer diagnosis based on histopathological images: A systematic survey. Rensselaer Polytechnic Institute, TR-05–09, Department of Computer Science, March, 2005.

[5] P. H. Bartels et al. Nuclear chromatin texture in prostatic lesions: I PIN and adenocarcinoma. *Analytical and Quantitative Cytology and Histology*, 20(15): 389–396, 1998.

[6] T. D. Clark, F. B. Askin, and C. R. Bagnell. Nuclear roundness factor: A quantitative approach to grading in prostate carcinoma, reliability of needle biopsy tissue, and the effect of tumour stage on usefulness. *Prostate*, 10(3): 199–206, 1987.

[7] R. Christen et al. Chromatin texture features in hematoxylin and eosin-stained prostate tissue. *Analytical and Quantitative Cytology and Histology*, 15(6): 383–388, 1993.

[8] J. L. Mohler et al. Nuclear morphometry in automatic biopsy and radical prostatectomy specimens of prostatic carcinoma. A comparison. *Analytical and Quantitative Cytology and Histology*, 16(6): 415–420, 1994.

[9] C. Minimo et al. Importance of different nuclear morphologic patterns in grading prostatic adenocarcinoma. An expanded model for computer graphic filters. *Analytical and Quantitative Cytology and Histology*, 16(5): 307–314, 1994.

[10] D. E. Pitts, S. B. Premkumar, A. G. Houston, R. J. Badaian, and P. Troncosa. Texture analysis of digitized prostate pathologic cross-sections *Medical Imaging: Image Processing SPIE*, pp 456–470, 1993.

[11] Y. Liu, T. Zahoa, and J. Zhang. Learning multispectral texture features for cervical cancer detection. *IEEE International Symposium on Biomedical Imaging*, Washington, D.C., pp. 169–172, 2002.

[12] I. Barshack, J. Kopolovic, Z. Malik, and C. Rothmann. Spectral morphometric characterization of breast carcinoma cells. *British Journal of Cancer*, 79(9–10): 1613–1619, 1999.

[13] P. Larsh, L. Cheriboga, H. Yee, and M. Diem. Infrared spectroscopy of humans cells and tissue: Detection of disease. *Technology in Cancer Research and Treatment*, 1(1): 1–7, 2002.

[14] H. Zhang et al. A novel multispectral imaging analysis method for white blood cell detection advances in natural computation. *Lecture Notes in Computer Science*. Berlin: Springer, 3611: 210–213, 2005.

[15] M. A. Roula. Machine vision and texture analysis for the automated identification of tissue patterns in prostatic tumours. PhD Thesis, Queens University of Belfast, 2004.

[16] M. A. Roula, J. Diamond, A. Bouridane, Paul Miller, and A. Amira. A multispectral computer vision system for automatic grading of prostatic neoplasia. *IEEE International Symposium on Biomedical Imaging*, 2002.

[17] M. A. Roula, A. Bouridane, and P. Miller. A quadratic classifier based on multispectral texture features for prostate cancer diagnosis. *Proceedings of the Seventh International Symposium on Signal Processing and Its Applications*, Paris, France, 2003.

[18] R. O. Duda, P. E. Hart, D. G. Stork. *Pattern Classification*. New York: Wiley, 2001.

[19] A. K. Jain, R. P. W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1): 4–37, 2000.

[20] L. O. Jimenez and D. A. Landgrebe. Supervised classification in high dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions on System, Man, and Cybernetics*, Part C, 28(1): 39–54, 1998.

[21] K. Fukunaga. *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.

[22] M. Kudo and J. Sklansky. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33:, 25–41, 2000.

[23] E. Amaldi and V. Kann. On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. *Theoretical Computer Science*, 209: 237–60, 1998.

[24] S. Davies and S. Russell. NP-completeness of searches for smallest possible feature sets. In *Proceedings of the AAAI Fall Symposium on Relevance.* Menlo Park, CA: AAAI Press, pp. 37–39, 1994.

[25] A. K. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2): pp. 153–158, 1997.

[26] P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15: 1119–1125, 1994.

[27] H. Zhang and G. Sun. Feature selection using tabu search method. *Pattern Recognition*, 35: 701–711, 2002.

[28] W. Siedlecki and J. Sklansy. A note on genetic algorithms for large-scale feature selection. *Pattern Recognition Letters*, 10(11): 335–347, 1989.

[29] S. B. Serpico and L. Bruzzone. A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7): 1360–1367, 2001.

[30] A. W. Whitney. A direct method of nonparametric measurement selection. *IEEE Transactions on Computers*, 20(9): 1100–1103, 1971.

[31] G. A. Clark et al. Multispectral image feature selection for land mine detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(1): 304–311, 2000.

[32] T. M. Cover and J. M. Van Campenhout. On the possible orderings in the measurement selection problem. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(9): 657–661, 1977.

[33] S. M. Sait and H. Youssef. *General Iterative Algorithms for Combinatorial Optimization*. Los Alamitos, CA: IEEE Computer Society, 1999.

[34] S. Yu, S. D. Backer, and P. Scheunders. Genetic feature selection combined with composite fuzzy nearest neighbor classifiers for hyperspectral satellite imagery. *Pattern Recognition Letters*, 23: 183–190, 2002.

[35] F. Glover. Tabu search I. *ORSA Journal on Computing*, 1(3): 190–206, 1989.

[36] F. Glover (1990). Tabu search II. *ORSA Journal on Computing*, 2(1): 4–32, 1990.

[37] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21–27, 1967.

[38] J. Furnkranz. Round robin classification. *Journal of Machine Learning Research*, 2: 721–747, 2002.

[39] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6): 610–621, 1973.

[40] H. J. Zimmerman. *Fuzzy Set Theory and Its Application* 3rd ed. Boston: Kluwer, 1996.

[41] S. A. Khan, S. M. Sait, and H. Youssef. Topology design of switched enterprise networks using fuzzy simulated evolution algorithm. *Engineering Applications of Artificial Intelligence*, 327–340, 2002.

[42] L. Breiman. Bagging predictors. *Machine Learning*, 24: 123–140, 1996.

[43] Y. Freund and R. Schapire. Experiments with a new boosting algorithm. *Proceedings of the Thirteenth International Conference on Machine Learning*. San Francisco: Morgan Kaufmann, pp. 148–156, 1996.

[44] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Kuhn, A. K. Jain. Dimensionality reduction using genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 4(2): 164–172, 2000.

[45] Aperio Technology. URL: http://www.aperio.com.

[46] D. Cohran, D. Sinno, and A. Clausen. Source detection and localization using a multi-mode detector: A Bayesian approach. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing ICASSP*, 1999.

[47] P. Larsh, L. Cheriboga, H. Yee, and M. Diem. Infrared spectroscopy of humans cells and tissue: Detection of disease. *Technology in Cancer Research and Treatment*, 1(1): 1–8, 2002.

[48] M. A. Tahir et al. Novel round-robin tabu search algorithm for prostate cancer classification and diagnosis using multispectral imagery. *IEEE Transactions on Information Technology in Biomedicine*, 10(4): 782–793, 2006.

[49] M. A. Tahir, A. Bouridane, and F. Kurugollu. Simultaneous feature selection and feature weighting using hybrid tabu search/k-nearest neighbor classifier. *Pattern Recognition Letters*, 28(4): 438–446, 2007.

[50] M. A. Tahir et al. A novel prostate cancer classification technique using intermediate memory tabu search. *Eurasip Journal on Applied Signal Processing, Special Issue: Advances in Intelligent Vision Systems: Methods and Applications*, 14: 2241–2249, 2005.

Chapter 7

Intuitionistic Fuzzy Processing of Mammographic Images

Ioannis K. Vlachos and George D. Sergiadis

Contents

Symbols used in chapter

$\tilde{A}$	Denotes a fuzzy set
A	Denotes an Atanassov's intuitionistic fuzzy set
FS	Abbreviation for fuzzy set
A–IFS	Abbreviation for Atanassov's intuitionistic fuzzy set
μ_A	Membership function of an Atanassov's intuitionistic fuzzy set
ν_A	Nonmembership function of an Atanassov's intuitionistic fuzzy set
π_A	Intuitionistic fuzzy index or hesitancy margin
DSM	Distribution separation measure
TBC_s	Target-to-background contrast ratio using variance
TBC_e	Target-to-background contrast ratio using entropy
CII	Contrast improvement index
D	Combined enhancement measure
g	Gray level of a digital image
L	Maximum number of quantized intensity levels in an image

The main purpose of this chapter is to extend and adapt the intuitionistic fuzzy image processing (IFIP) framework in the context of mammographic imaging. Intuitionistic fuzzy sets, favored for their ability to cope with the imprecise and imperfect nature of information, successfully handle the intrinsic uncertainty carried by mammographic images. Analytical models are presented for analyzing/synthesizing the images to and from their intuitionistic fuzzy components, based on different notions of intuitionistic fuzzy entropy and intuitionistic defuzzification schemes. Moreover, extensions of fuzzy approaches for contrast enhancement in the intuitionistic fuzzy setting demonstrate the efficiency of the IFIP framework. Finally, experimental results are assessed by objective enhancement evaluation measures that further verify the potential of using intuitionistic fuzzy sets for mammographic image processing.

7.1 Introduction

Mammography is a specific type of medical imaging for breast cancer diagnosis that uses low-dose X-ray sources and high-resolution films. Mammograms often suffer from low contrast due to the small difference of the X-ray attenuation between normal glandular tissues and malignant masses [24],

which leads in displaying only about 3% of the information to be detected [6]. Additionally, the complicated structured background, as well as the presence of artifacts originating from identification labels, radiopaque markers, and wedges, may drastically influence the contrast enhancement process. It is therefore important for the diagnostic procedure to enhance the contrast of mammograms in order for the human expert, or the computer-aided diagnostic (CAD) system, to discriminate between normal and suspected malignant tissue. Moreover, due to all the aforementioned factors, the uncertainty associated with mammographic images is increased.

The framework of intuitionistic fuzzy image processing (IFIP) was introduced by Vlachos and Sergiadis [69] and further developed in other works [59, 60,62–68,70]. This work exploits and adjusts the IFIP framework in the context of mammography by considering its particular characteristics. Mammographic images are susceptible to bearing uncertainties associated with the intensity levels of pixels. The origin of this type of uncertainty can be traced back to the acquisition chain or to the noisy environment and the imaging mechanisms. An overview of the problem of analyzing/synthesizing the mammographic image to and from its corresponding intuitionistic fuzzy components is presented, and analytical ways for solving it are briefly described. Intuitionistic fuzzy approaches to contrast enhancement are also presented, and their performance is evaluated both subjectively and objectively.

This chapter is organized as follows. Section 7.2 provides an overview of fuzzy and nonfuzzy approaches to mammographic image processing and outlines their relative drawbacks. In Section 7.3, a brief description of the concepts and definitions of Atanassov's notion of intuitionistic fuzzy sets theory is given, while Section 7.4 outlines the different notions of intuitionistic fuzzy entropy and their properties. The IFIP framework is then presented in full detail in Section 7.5. In order to provide a quantitative way of assessing the performance of the different contrast enhancement approaches, objective enhancement evaluation measures are discussed in Section 7.6, and the MiniMIAS data set [48] used for experimental assessment is described in Section 7.7. The proposed algorithms of intuitionistic fuzzy histogram hyperbolization and intuitionistic fuzzy contrast intensification for mammographic image processing are given in Sections 7.8 and 7.9, respectively. Finally, discussion and future research directions are provided in Sections 7.10 and 7.11, and conclusions are drawn in Section 7.12.

7.2 Related Work

Contrast enhancement techniques constitute an inextricable part of mammographic imaging systems. Their objective is to enhance the mammogram under processing in order to improve the image quality, either for manual

reading or to enhance the efficiency of CAD systems, by improving the subsequent image segmentation and feature extraction [46]. In this context, a variety of methods have been applied, ranging from simple enhancement approaches, like the histogram equalization (HEQ) technique [18,47,73], to methods employing Gaussian filters [28], adaptive background correction [11,71], tree-structured filtering [41], adaptive neighborhood contrast enhancement [42], and density weighted contrast enhancement [37–39].

Fuzzy sets (FSs) theory [74] has been successfully applied to many medical imaging problems. The application of fuzzy Hough transform for feature extraction was presented by Philip and others [40], and fusion of FSs for processing of multiplanar tomographic images was demonstrated by Vial and others [58]. Especially for mammographic image processing, FSs have been used in many CAD systems. Automated detection of tumors in mammograms was performed using fuzziness-driven approaches, such as fuzzy pyramid linking [7] and fuzzy binary decision trees [27]. Furthermore, Saha and colleagues [45] carried out breast tissue density quantification based on the notion of fuzzy connectedness; Hassanien, Ali, and Nobuhara [21] demonstrated detection of spiculated masses in mammograms using FSs; and Hassanien and Ali [20] demonstrated mammogram segmentation using pulse coupled neural networks. Cheng, Lui, and Freimanis [13] used membership functions to denote, under the scope of FSs theory, the degree of being a microcalcification in terms of intensity values, in view of performing microcalcification detection. Finally, Hassanien and Badr [22] carried out a comparative study on digital mammography enhancement algorithms using FSs, and Rick and colleagues [44] presented fuzzy techniques in mammographic imaging. Due to their flexible framework, fuzzy-based contrast enhancement approaches [5,36] often perform better than nonfuzzy ones. Recently, Singh and Bovis [46] studied different fuzzy and nonfuzzy techniques on synthetic and real mammographic images, using objective enhancement evaluation measures and demonstrated the efficiency of the fuzzy approach.

However, and despite their vast impact on the design of algorithms for mammographic image processing, FSs are often characterized by a number of limitations mostly associated with the requirement of precisely defining the membership function of the FS itself. As a result, FSs are not always able to efficiently model the uncertainties arising out of the imprecise and/or imperfect nature of information. These limitations and drawbacks were identified and described by Mendel and Bob John [31], who traced their sources back to the uncertainties that are present in most ordinary fuzzy logic systems (FLSs). The very meaning of words used in the antecedents and consequents of FLSs can be uncertain, since words often mean different things to different people. Additionally, extracting the knowledge to construct the rule-base from a group of experts who do not all agree leads to consequents having a histogram of values associated with them. Moreover, data presented as input to an FLS, as well as data used for its tuning, are often noisy, thus bearing a degree of uncertainty. Consequently, these uncertainties translate into additional uncertainties about FS membership functions.

The aforementioned observations led researchers to seek more efficient ways to express and model imprecision and uncertainty. Among the various notions of higher-order FSs, Atanassov's intuitionistic fuzzy sets (A–IFSs) [1,3] are found to provide a flexible mathematical and intuitive framework to cope with the imperfect and imprecise nature of digital images. Finally, it should be mentioned that Atanassov and colleagues [4] stated and proved that there exists a fundamental justification for applying methods based on higher-order FSs to deal with real-world situations.

7.3 Elements of Intuitionistic Fuzzy Sets Theory

A–IFSs are described using two characteristic functions expressing corresponding degrees of membership (belongingness) and nonmembership (nonbelongingness) of elements of the universe to the A–IFS. It is this additional degree of freedom that allows for the efficient modeling of imprecise or/and precise information. In this section, we briefly present fundamental concepts of A–IFSs theory and connections with Zadeh's FSs.

Definition 7.1 *An FS $\tilde{A}$ defined on a universe X may be given as [74]*

$$\tilde{A} = \{\langle x, \mu_{\tilde{A}}(x)\rangle | x \in X\} \tag{7.1}$$

where $\mu_{\tilde{A}}(x) : X \to [0,1]$ is the membership function of $\tilde{A}$.

The membership function of $\tilde{A}$ describes the *degree of belongingness* of $x \in X$ in $\tilde{A}$.

Definition 7.2 *An A–IFS A defined on a universe X is given by [1–3]*

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in X\} \tag{7.2}$$

where

$$\mu_A(x) : X \to [0,1] \quad \textit{and} \quad \nu_A(x) : X \to [0,1]$$

with the condition

$$0 \leqslant \mu_A(x) + \nu_A(x) \leqslant 1 \tag{7.3}$$

for all $x \in X$.

The values of the functions $\mu_A(x)$ and $\nu_A(x)$ denote the *degree of belongingness* and the *degree of nonbelongingness* of x to A, respectively. For an A–IFS A in X, we call the *intuitionistic fuzzy index* of an element $x \in X$ in A the following expression:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \tag{7.4}$$

FSs can also be represented using the notation of A–IFSs. An FS $\tilde{A}$ defined on X can be represented as the following A–IFS:

$$A = \{\langle x, \mu_A(x), 1 - \mu_A(x)\rangle | x \in X\} \tag{7.5}$$

with $\pi_A(x) = 0$ for $x \in X$.

Definition 7.3 *The complementary set A^c of A is defined as*

$$A^c = \{\langle x, \nu_A(x), \mu_A(x)\rangle | x \in X\} \tag{7.6}$$

Finally, throughout this paper, by $\mathscr{IFS}(X)$ we denote the set of all A–IFSs defined on X. Correspondingly, $\mathscr{FS}(X)$ is the set of all FSs on X, while 2^X denotes the set of all crisp sets on X.

7.4 Notions of Intuitionistic Fuzzy Entropy

Entropy plays an important role in digital image processing. Therefore, it comes as no surprise that the notion of entropy constituted a fundamental aspect from the beginning of the development of FSs theory. De Luca and Termini [30] were the first to introduce an axiomatic skeleton of a nonprobabilistic entropy in the setting of FSs that captured our intuition regarding the very essence of fuzzy entropy. As a natural consequence, the quest for entropy measures in the context of A–IFSs was a very interesting topic that intrigued many researchers working in this field.

Burillo and Bustince [9] were the first to state and propose an axiomatic skeleton of entropy for A–IFSs and interval-valued fuzzy sets.

Definition 7.4 (Burillo and Bustince [9]). *A real function $E_{BB}: \mathscr{IFS}(X) \to \mathbb{R}^+$ is called an entropy on $\mathscr{IFS}(X)$, if E_{BB} has the following properties:*

- *$E_{BB}(A) = 0$ if and only if $A \in \mathscr{FS}(X)$.*
- *$E_{BB}(A) = Cardinal(X)$ if and only if $\mu_A(x) = \nu_A(x) = 0$ for all $x \in X$.*
- *$E_{BB}(A) = E_{BB}(A^c)$ for all $A \in \mathscr{IFS}(X)$.*
- *$E_{BB}(A) \geqslant E_{BB}(B)$ if $\mu_A(x) \leqslant \mu_B(x)$ and $\nu_A(x) \leqslant \nu_B(x)$ for all $x \in X$.*

Moreover, they proposed the following entropy measure for A–IFSs, satisfying the aforementioned system of axioms:

$$E_{BB}(A) = \sum_{i=1}^{|X|} \pi_A(x_i) \tag{7.7}$$

which can be consider as the degree of intuitionism of an A–IFS. In the above formula, $|X|$ denotes the cardinality of the finite universe X.

Motivated by De Luca and Termini's set of axiomatic requirements, Szmidt and Kacprzyk [51] proposed an alternative interpretation of entropy, accompanied by a different set of axioms.

Definition 7.5 (Szmidt and Kacprzyk [51]). *A real function* $E_{SK} : \mathscr{IFS}(X) \to \mathbb{R}^+$ *is called an entropy on* $\mathscr{IFS}(X)$*, if* E_{SK} *has the following properties:*

- $E_{SK}(A) = 0$ *if and only if* $A \in 2^X$.
- $E_{SK}(A) = 1$ *if and only if* $\mu_A(x) = \nu_A(x)$ *for all* $x \in X$.
- $E_{SK}(A) = E_{SK}(A^c)$ *for all* $A \in \mathscr{IFS}(X)$.
- $E_{SK}(A) \leqslant E_{SK}(B)$ *if*
  $\mu_A(x) \leqslant \mu_B(x)$ and $\nu_A(x) \geqslant \nu_B(x)$ for $\mu_B(x) \leqslant \nu_B(x)$
  or
  $\mu_A(x) \geqslant \mu_B(x)$ and $\nu_A(x) \leqslant \nu_B(x)$ for $\mu_B(x) \geqslant \nu_B(x)$
  for all $x \in X$.

The aforementioned definition degenerates to De Luca and Termini's definition when FSs are considered.

Along with their axiomatic skeleton, Szmidt and Kacprzyk proposed a new measure of entropy by considering distances between A–IFSs using a geometrical representation [50], defined as

$$E_{SK}(A) = \frac{1}{|X|} \sum_{i=1}^{|X|} \frac{\min\{\mu_A(x_i), \nu_A(x_i)\} + \pi_A(x_i)}{\max\{\mu_A(x_i), \nu_A(x_i)\} + \pi_A(x_i)} \tag{7.8}$$

Szmidt and Kacprzyk introduced a generalized framework of Definition 7.5 [55]. Finally, it should be mentioned that a connection between the different concepts of entropy for A–IFSs was explored and proved by Vlachos and Sergiadis [61].

Recently, Hung and Yang [23] proposed a different axiomatic framework for defining the fuzzy entropy of A–IFSs by exploiting the concept of probability.

Definition 7.6 (Hung and Yang [23]). *A real function* $E_{HY} : \mathscr{IFS}(X) \to \mathbb{R}^+$ *is called an entropy on* $\mathscr{IFS}(X)$ *if* E_{HY} *has the following properties:*

- $E_{HY}(A) = 0$ *if and only if* $A \in 2^X$.
- $E_{HY}(A)$ *assumes a unique maximum if* $\mu_A(x) = \nu_A(x) = \pi_A(x) = \frac{1}{3}$ *for all* $x \in X$.

- $E_{HY}(A) = E_{HY}(A^c)$ *for all* $A \in \mathscr{IFS}(X)$.
- $E_{HY}(A) \leqslant E_{HY}(B)$ *if*
  $\mu_A(x) \leqslant \mu_B(x)$ *and* $\nu_A(x) \leqslant \nu_B(x)$ *for* $\max\{\mu_B(x), \nu_B(x)\} \leqslant \frac{1}{3}$
  or
  $\mu_A(x) \geqslant \mu_B(x)$ *and* $\nu_A(x) \geqslant \nu_B(x)$ *for* $\min\{\mu_B(x), \nu_B(x)\} \geqslant \frac{1}{3}$
  for all $x \in X$.

Finally, motivated by Renyi's entropy [43], they introduced the following parametric entropy measure

$$E_{HY}(A) = \sum_{i=1}^{|X|} \frac{1}{1-\beta} \log\left(\mu_A^k(x_i) + \nu_A^k(x_i) + \pi_A^k(x_i)\right) \tag{7.9}$$

with $0 < k < 1$.

7.5 Intuitionistic Fuzzy Image Processing

FSs have found a fertile ground for application in the context of digital image processing, mainly due to their ability to model notions and definitions in a qualitative rather than quantitative manner. Consequently, properties of an image such as the "brightness" of a pixel, the "edginess" of a boundary, or the "homogeneity" of a region, characterized by an intrinsic vagueness, can be successfully modeled by suitable membership functions corresponding to FSs.

However, attributing a single membership degree to an element of the universe to describe its belongingness to a particular FS is by itself constraining. Therefore, FSs are characterized by a number of limitations mostly associated with the requirement of precisely defining the membership function describing the FS itself. This drawback introduces additional uncertainty to any FLS [31]. This observation led Vlachos and Sergiadis [69] to explore the potential of an IFIP framework.

7.5.1 Intuitionistic fuzzy model of an image

Let us consider an image A of size $M \times N$ pixels, having L gray levels g ranging between 0 and $L-1$. When applying FSs for image processing, an image can be considered as an array of fuzzy singletons. Each element of the array denotes the membership value $\mu_{\tilde{A}}(g_{ij})$ of the gray level g_{ij} corresponding to the (i,j)-th pixel with respect to a predefined image property such as brightness, edginess, or homogeneity [33–35].

As a generalization of this approach, we introduce the following representation of an image A in the intuitionistic fuzzy environment.

Definition 7.7 *An image A is described by the A–IFS*

$$A = \{\langle g_{ij}, \mu_A(g_{ij}), \nu_A(g_{ij})\rangle | g_{ij} \in \{0, \ldots, L-1\}\} \tag{7.10}$$

with $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N\}$, where $\mu_A(g_{ij})$ and $\nu_A(g_{ij})$ denote respectively the degrees of membership and nonmembership of the (i, j)-th pixel to the set A associated with an image property.

The functions μ_A and ν_A describe the *membership* and *nonmembership* components of the image. If instead of A–IFSs we consider FSs, then Definition 7.7 reduces to the one given by Pal and King [33–35].

7.5.2 Fuzzification: From gray levels to membership degrees

The first stage of IFIP, as illustrated in Figure 7.1, involves the assignment of membership degrees to image pixels by defining a suitable membership function $\mu_{\tilde{A}}$. To carry out this task, different fuzzification schemes can be applied, from simple gray-level normalization to more complex techniques [12].

7.5.3 Intuitionistic fuzzification: From fuzzy to intuitionistic fuzzy sets

Intuitionistic fuzzification is probably the most important stage in the chain of operations involved in the IFIP framework. In this stage, the gray levels of the image are rendered as an A–IFS by attributing to image pixels corresponding degrees of membership and nonmembership.

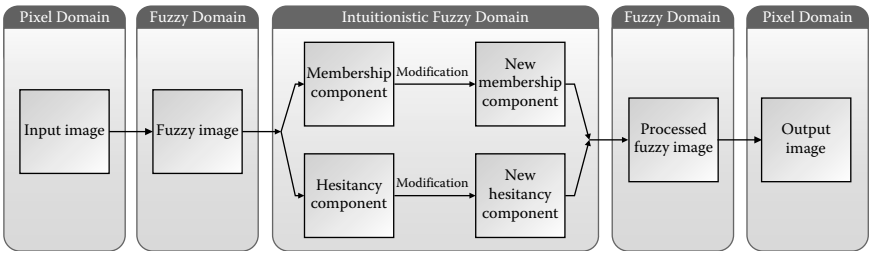


FIGURE 7.1: Overview of the IFIP framework.

Vlachos and Sergiadis [69] introduced a family of monoparametric membership and nonmembership functions based on the concepts of intuitionistic fuzzy generators [8,10], described by

$$\mu_A(g;\lambda) = 1 - (1 - \tilde{\mu}_A(g))^{\lambda} \tag{7.11}$$

and

$$\nu_A(g;\lambda) = (1 - \tilde{\mu}_A(g))^{\lambda(\lambda+1)} \tag{7.12}$$

with $\lambda \geqslant 0$.

By varying the parameter λ, different A–IFSs can be generated, yielding different representations of gray levels in the intuitionistic fuzzy domain (IFD). In order to select the optimal parameter λ_{opt} that efficiently models the gray levels, the intuitionistic fuzzy entropy of the image is employed. The aforementioned optimization criterion can be formulated as follows:

$$\lambda_{opt} = \arg\max_{\lambda \geqslant 0} \{E(A;\lambda)\} \tag{7.13}$$

where E is an intuitionistic fuzzy entropy measure. This optimization procedure is called the *maximization of intuitionistic fuzzy entropy principle* [69]. The optimal representation of the image in the IFD is given by the A–IFS

$$A_{opt} = \{\langle g, \mu_A(g;\lambda_{opt}), \nu_A(g;\lambda_{opt})\rangle \mid g \in \{0,\ldots,L-1\}\} \tag{7.14}$$

Figure 7.2 illustrates a mammographic image from the MiniMIAS database [48], along with its corresponding optimal membership, nonmembership, and hesitancy components. The optimization criterion is shown in Figure 7.3b, where the entropy curve is illustrated as a function of parameter λ, while Figures 7.3c and 7.3d depict the optimal membership, nonmembership, and hesitancy functions of the A–IFS A_{opt} describing the brightness of pixels in the image under processing. For more details on the derivation of the aforementioned functions, refer to Vlachos and Sergiadis [69]. It should be stressed that for mammographic image processing, we restrict λ to the [0,3] interval.

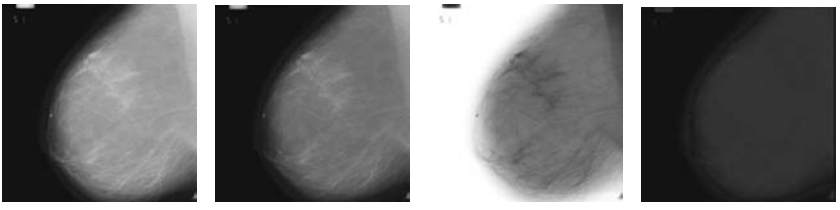


FIGURE 7.2: (a) Mammographic image (MiniMIAS database name: `mdb141`) with its corresponding optimal (b) membership, (c) nonmembership, and (d) hesitancy components.

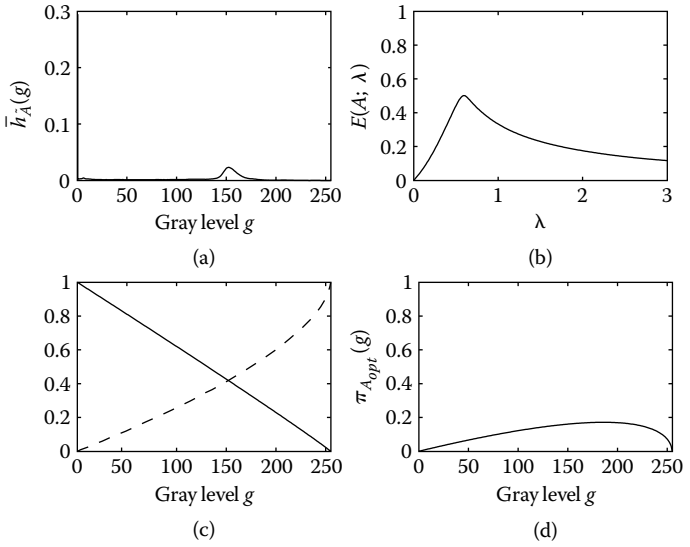


FIGURE 7.3: (a) Normalized histogram $\bar{h}_{\tilde{A}}$ for the mammogram of Figure 7.2a after fuzzification. (b) Intuitionistic fuzzy entropy curve, used for obtaining the optimal parameter λ_{opt} for transferring the image into the IFD. (c) Membership (*dashed line*), nonmembership (*solid line*), and (d) hesitancy optimal functions describing the brightness of gray levels in the image.

7.5.4 Modification in the intuitionistic fuzzy domain

In general, this stage involves the modification of the membership and nonmembership components of the image, obtained using the maximum intuitionistic fuzzy entropy principle, according to the desired image processing task one wants to carry out. In this work, we focus on the contrast enhancement of mammographic images by applying the concepts of histogram hyperbolization and contrast intensification, discussed in full detail in Sections 7.8 and 7.9, respectively.

7.5.5 Intuitionistic defuzzification: Embedding hesitancy

From the representation of the processed image in the IFD, we must optimally combine the information carried by both the membership and nonmembership components in order to obtain its representation in the fuzzy domain (FD). In the context of this work, we assess the performance of the different intuitionistic defuzzification schemes proposed by Vlachos and Sergiadis [63,69] in the particular context of mammographic image processing.

7.5.5.1 Maximum index of fuzziness intuitionistic defuzzification

The representation of images is based on the intensity levels of the corresponding pixels. From the overview of the IFIP framework illustrated in Figure 7.1, we may observe that in order to obtain the processed image in the pixel domain (PD), both the membership and nonmembership components of the image in the IFD must be efficiently rendered into a single membership component describing the degree of pixel brightness in terms of FSs. Atanassov [3] proposed an operator, *Atanassov's operator*, for the deconstruction of an A–IFS into an FS by flexibly handling its hesitancy index; it is defined as follows.

Definition 7.8 *If* $A \in \mathscr{IFS}(X)$, *then* $D_\alpha : \mathscr{IFS}(X) \to \mathscr{FS}(X)$, *where*

$$D_\alpha(A) = \{\langle x, \mu_A(x) + \alpha\pi_A(x), \nu_A(x) + (1-\alpha)\,\pi_A(x)\rangle \,|\, x \in X\} \tag{7.15}$$

with $\alpha \in [0,1]$.

Szmidt and Kacprzyk [49] investigated the properties of the aforementioned operator for constructing the FS corresponding to an A–IFS.

By varying the parameter α of Atanassov's operator, the hesitancy can be embedded in the membership component in different portions. In the context of IFIP, this fact is translated as the generation of a number of different possible representations of the image in the FD. Therefore, an optimization scheme must be employed in order to select that value of the parameter α that corresponds to an optimal degeneration of the image with respect to some meaningful criterion.

For contrast enhancement, it is sometimes desirable to increase the grayness ambiguity by increasing the number of gray levels in an image [57] since images with higher fuzziness values are considered to be more suitable in terms of human brightness perception. The most common measure of the intrinsic fuzziness of a set is the *linear index of fuzziness* [25]. In the case of an FS $\tilde{A}$, the linear index of fuzziness is given by

$$\gamma_l(\tilde{A}) = \frac{1}{2|X|}\sum_{i=1}^{|X|} \min\{\mu_{\tilde{A}}(x_i), 1-\mu_{\tilde{A}}(x_i)\} \tag{7.16}$$

where $|X|$ is the cardinality of the finite universe X. A modified index of fuzziness is obtained by substituting the min t-norm with the product operator:

$$\gamma(\tilde{A}) = \frac{1}{4|X|}\sum_{i=1}^{|X|} \mu_{\tilde{A}}(x_i)\,(1-\mu_{\tilde{A}}(x_i)) \tag{7.17}$$

Therefore, in our quest for the optimal parameter α_{opt} of Atanassov's operator, the maximization of the index of fuzziness will serve as the optimization

criterion for deconstructing the modified A–IFS A'_{opt} into its optimal FS. This optimization criterion can be formulated as

$$\alpha_{opt} = \max_{\alpha\in[0,1]} \left\{\gamma\left(D_\alpha\left(A'_{opt}\right)\right)\right\} \tag{7.18}$$

where

$$\gamma\left(D_\alpha(A'_{opt})\right) = \frac{1}{4MN}\sum_{g=0}^{L-1} h_{\tilde{A}}(g)\mu_{D_\alpha(A'_{opt})}(g)\left(1-\mu_{D_\alpha(A'_{opt})}(g)\right) \tag{7.19}$$

It should be stressed that Equation 7.18 can be solved analytically, yielding

$$\alpha'_{opt} = \frac{\sum_{g=0}^{L-1} h_{\tilde{A}}(g)\pi_A(g;\lambda_{opt})\left(1-2\mu_A(g;\lambda_{opt})\right)}{2\sum_{g=0}^{L-1} h_{\tilde{A}}(g)\pi_A^2(g;\lambda_{opt})} \tag{7.20}$$

with $h_{\tilde{A}}$ being the histogram of the fuzzified image $\tilde{A}$. By considering the second derivative of Equation 7.19, it easy to verify that its extremum is a global maximum.

However, Equation 7.20 does not guarantee that α'_{opt} will lie in the [0,1] interval since Equation 7.19 may attain a maximum outside of [0,1]. Therefore, the optimal parameter α_{opt} used to transfer the image from the IFD to the FD is obtained as

$$\alpha_{opt} = \begin{cases} 0, & \text{if } \alpha'_{opt} < 0 \\ \alpha'_{opt}, & \text{if } 0 \leqslant \alpha'_{opt} \leqslant 1 \\ 1, & \text{if } \alpha'_{opt} > 1 \end{cases} \tag{7.21}$$

Gray-level transformations for contrast enhancement must in most cases satisfy an additional critical constraint rendered as the monotonicity of the transformation employed, which is required to be increasing. Such transformations preserve the ordering of the intensity levels, insuring that no additional artifacts will be introduced in the enhanced image. Considering Atanassov's operator of Equation 7.15, we obtain

$$\frac{d\mu_{D_\alpha(A)}(g)}{dg} = (1-\alpha)\frac{d\mu_A(g)}{dg} - \alpha\frac{d\nu_A(g)}{dg} \tag{7.22}$$

Since $\alpha \in [0,1]$ and $\mu_A(g)$ and $\nu_A(g)$ are increasing and decreasing functions of g, respectively, it follows immediately that $\mu_{D_\alpha(A)}(g)$ is also an increasing function of the intensity levels.

Vlachos and Sergiadis [69] proved the following property, which guarantees that the constraint of monotonicity is satisfied for Atanassov's operator.

Proposition 7.1 *Let $A = \{x, \mu_A(x), \nu_A(x)\,|x \in X\}$ be an A–IFS, with $\mu_A(x)$ and $\nu_A(x)$ being increasing/decreasing and decreasing/increasing functions of x, respectively. Then, the membership function of the FS $D_\alpha(A)$ retains the monotonicity of $\mu_A(x)$.*

7.5.5.2 Generalized intuitionistic defuzzification

Seeking more flexible alternatives for rendering the membership and non-membership components of an image into membership values in the FD leads us toward the exploitation of the aforementioned validity condition.

A more generic form of Atanassov's operator, the *Atanassov's point operator*, was proposed by Burillo and Bustince [8] and is defined as follows.

Definition 7.9 *If* $A \in \mathscr{IFS}(X)$*, then* $D_{\alpha_x}\colon \mathscr{IFS}(X) \to \mathscr{FS}(X)$ *where*

$$D_{\alpha_x} = \{\langle x, \mu_A(x) + \alpha(x)\pi_A(x), \nu_A(x) + (1-\alpha(x))\pi_A(x)\rangle | x \in X\} \quad (7.23)$$

with $\alpha(x) \in [0, 1]$.

Considering Atanassov's point operator, Equation 7.22 is generalized as

$$\frac{d\mu_{D_{\alpha_g}(A)}(g)}{dg} = (1-\alpha(g))\frac{d\mu_A(g)}{dg} - \alpha(g)\frac{d\nu_A(g)}{dg} + (1-\mu_A(g)-\nu_A(g))\frac{d\alpha(g)}{dg} \quad (7.24)$$

One may observe that if $\alpha = const$ for all $g \in \{0, \ldots, L-1\}$, then Equation 7.24 degenerates to 7.22. From Equation 7.24, it follows immediately that if $\alpha(g)$ is an increasing function of the intensity levels and additionally $\alpha(g) \in [0, 1]$ for all $g \in \{0, \ldots, L-1\}$, then $\mu_{D_{\alpha_g}(A)}(g)$ is also a proper membership modification function and thus a valid gray-level transformation function since it holds that $\frac{d\mu_{D_{\alpha_g}(A)}(g)}{dg} \geqslant 0$.

In this work we are going to utilize the cumulative histogram of the image for calculating the parameter of Atanassov's point operator, using the following scheme

$$\alpha_{h_{\tilde{A}}}(g) = \frac{1}{MN}\sum_{i=0}^{g} h_{\tilde{A}}(i) \quad (7.25)$$

for all $g \in \{0, \ldots, L-1\}$. The cumulative nature of the function ensures that the parameter will be an increasing function of the gray levels. Moreover, such a function has the effect of keeping darker levels dark while simultaneously brightening higher intensity levels, resulting in an enhancement of the contrast of the image under processing.

7.5.6 Defuzzification: From fuzzy sets to grayscale images

The last stage of the IFIP framework involves the transition of the image from the FD to the PD. Different defuzzification schemes can be applied, described in general as

$$g' = t\left[\mu_{\tilde{A}'}(g)\right] \quad (7.26)$$

where g' are the modified gray levels of the image, $\mu_{\tilde{A}'}$ the modified membership function, and $t[\cdot]$ a transformation function $t : [0, 1] \to \{0, \ldots, L-1\}$.

7.6 Enhancement Evaluation Measures

In order to assess the performance of various contrast enhancement methods, apart from the subjective visualization of the results, we applied quantitative objective measures for evaluating the efficiency of different contrast enhancement algorithms. The underlying idea of the described measures is based on the fact that contrast enhancement methods for digital mammography should aim to improve the contrast between regions of interest (ROIs) and their surrounding background, thus leading to a more accurate estimation of the target boundaries and location during the segmentation process. Bovis and Singh [5] presented enhancement evaluation measures for digitized mammograms. These measures were successfully applied for assessing the performance of a fuzzy logic–based scheme for detection of spiculated masses in mammograms [20].

It should be mentioned that the described measures are based on the assumption that the target and background areas are accurately specified by the human expert. In the following description, O and E denote the original and enhanced images, while T and B refer to the target and background regions of the images, respectively.

7.6.1 Distribution separation measure

The key objective of enhancement techniques, especially when applied to digitized mammograms, is to maximize the distance between two normal probability density functions (PDFs), used to describe the target and the background, in order to make the target more visible against its background [5]. The best decision boundary between the target and the background in the original image is given by

$$D'_O = \frac{m_B^O \sigma_T^O + m_T^O \sigma_B^O}{\sigma_T^O + \sigma_B^O} \tag{7.27}$$

while for the enhanced image,

$$D'_E = \frac{m_B^E \sigma_T^E + m_T^E \sigma_B^E}{\sigma_T^E + \sigma_B^E} \tag{7.28}$$

where m_T^O, m_B^O, m_T^E, and m_B^E are the mean values of the normal distributions for the target and the background in the original and the enhanced image, and $\sigma_T^O, \sigma_B^O, \sigma_T^E$, and σ_B^E are the corresponding standard deviations.

Singh and Bovis [46] proposed an alternative approximation of D'_O and D'_E based on the cutting score [19]. Assuming that the groups are representative

of the population, a weighted average of the group centroids will provide an optimal cutting score. Consequently,

$$D_O = \frac{m_B^O N_T^O + m_T^O N_B^O}{N_B^O + N_T^O} \tag{7.29}$$

and

$$D_E = \frac{m_B^E N_T^E + m_T^E N_B^E}{N_B^E + N_T^E} \tag{7.30}$$

where N_B^O and N_T^O are the number of samples in the background and target prior to enhancement, and N_B^E and N_T^E are the corresponding sample numbers in the enhanced image. Thus, based on Equations 7.29 and 7.30, the following *distribution separation measure* (DSM) is defined as [46]

$$DSM = \left(|m_T^E - m_B^E|\right) - \left(|m_T^O - m_B^O|\right) \tag{7.31}$$

which is a distance measure between the decision boundaries and the means of the target and background before and after enhancement and serves as a measure of the quality of the enhancement procedure.

7.6.2 Target-to-background contrast ratio using variance

In mammographic images, the target is characterized by a higher mean intensity level compared to the one corresponding to its surrounding background. The following measure assesses the performance of a contrast enhancement algorithm by measuring the increase/decrease of mean intensity level of the target/background in the original and enhanced images, as well as the change of target gray-level spread, and is given by [5]

$$TBC_s = \frac{\frac{m_T^E}{m_B^E} - \frac{m_T^O}{m_B^O}}{\frac{\sigma_T^E}{\sigma_T^O}} \tag{7.32}$$

7.6.3 Target-to-background contrast ratio using entropy

In a similar manner as the TBC_s, this measure evaluates the capability of an enhancement algorithm by considering, instead of standard deviation, the change of the entropies ε_T^E and ε_T^O associated with the target area in the original and the enhanced images, respectively, and is defined as [5]

$$TBC_e = \frac{\frac{m_T^E}{m_B^E} - \frac{m_T^O}{m_B^O}}{\frac{\varepsilon_T^E}{\varepsilon_T^O}} \tag{7.33}$$

7.6.4 Contrast improvement index

The *contrast improvement index* (CII) was introduced by Laine and colleagues [26] as a quantitative measure of contrast improvement and is given by

$$CII = \frac{C^E}{C^O} \tag{7.34}$$

where

$$C^E = \frac{m_T^E - m_B^E}{m_T^E + m_B^E} \tag{7.35}$$

is the optical definition of contrast of the enhanced image [32]. Similarly, the optical definition of contrast C^O for the original image is computed.

Larger values of the aforementioned measures correspond to better separation between the distributions of the target and the background, indicating a better performance of the contrast enhancement algorithm employed.

7.6.5 Combined enhancement measure

Seeking a way to quantitatively rank enhancement techniques for a particular mammographic image, Singh and Bovis [46] combined the DSM, TBC_s, and TBC_e measures into a single quantitative value by defining the *combined enhancement measure* D_{SB}. To carry out this task, Singh and Bovis represented for an image each of the aforementioned enhancement measures with a point in a three-dimensional Euclidean space by scaling each measure within the range [0,1]. Using this representation, a high-performance contrast enhancement technique will have a concentration of points close to the (1, 1, 1) vertex of the unit cube. Consequently, the combined measure D_{SB} is calculated based on the Euclidean distance between the corresponding point and the vertex (1, 1, 1) representing the best possible enhancement and is defined as

$$D_{SB} = \sqrt{(1 - \overline{DSM})^2 + (1 - \overline{TBC_s})^2 + (1 - \overline{TBC_e})^2} \tag{7.36}$$

where $\overline{DSM}$, $\overline{TBC_s}$, and $\overline{TBC_e}$ denote the normalized versions of the corresponding measures. To calculate the combined measure, each of the enhancement measures is given the same weight, and a simple min–max normalization is applied. In the proposed approach, in order to also take into account

the CII index, the following modified combined enhancement measure was utilized:

$$D = \sqrt{(1 - \overline{DSM})^2 + (1 - \overline{TBC_s})^2 + (1 - \overline{TBC_e})^2 + (1 - \overline{CII})^2} \quad (7.37)$$

Finally, it should be mentioned that the smaller the value of D, the better the performance of the enhancement technique.

7.7 The MiniMIAS Database

In order to evaluate the efficiency of the proposed intuitionistic fuzzy framework for contrast enhancement of mammographic images, we considered the MiniMIAS database [48], which consists of left and right breast images of size 1024×1024 pixels with 8 bits-per-pixel gray-tone resolution obtained from 161 patients. The set of images contains digitized mammograms depicting different types of abnormalities, such as architectural distortions, calcifications, circumscribed masses, stellate lesions, examples of benign and malignant conditions, as well as a comprehensive set of normals. Moreover, mammograms are also classified according to the breast density as *fatty*, *glandular*, and *dense*. For the purpose of evaluating the performance of the proposed IFIP framework, 28 images containing spiculated/circumscribed masses were selected. For each image in the MiniMIAS database containing an abnormality, the spatial coordinates of the center of the abnormality along with the radius of the circle enclosing it were also provided.

The enhancement evaluation measures require the accurate definition of the ROI bounding the target, as well as specifying its corresponding background [5, 46]. In the MiniMIAS database, the ROI supplied by the radiologist serves as a *rough* indication in order for the user to be directed to the area of the image where the abnormality exists [5]. To calculate the measures, we selected as the ROI the escribed square corresponding to the circle having half the radius of the one specified by the MiniMIAS database. This was chosen in order to ensure that the ROI is accurately located within the target. The surrounding background was defined as the area between the escribed square corresponding to the circle with radius of 130% of the one proposed by the radiologist and the square corresponding to the ROI. The selection of the set of 28 images from the MiniMIAS database was based on the constraint that the ROIs and background regions, defined as described above, be defined as accurately as possible in order to guarantee the trusted performance of the enhancement evaluation measures. This fact can be verified by observing the considered images shown in Figure 7.4.

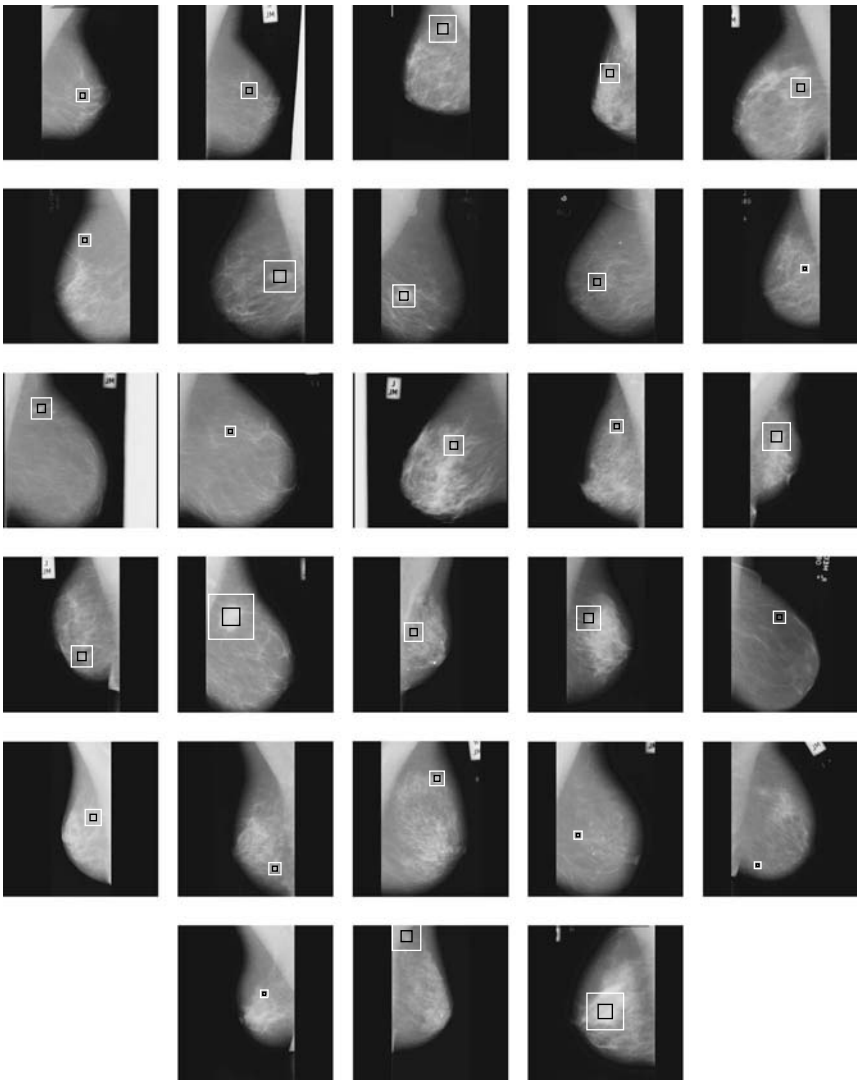


FIGURE 7.4: Images from the MiniMIAS database containing well-defined spiculated/circumscribed masses used for performance evaluation of contrast enhancement algorithms. Target and background regions are annotated as black and white squares, respectively (MiniMIAS database names: `mdb10`, `mdb12`, `mdb15`, `mdb17`, `mdb19`, `mdb23`, `mdb25`, `mdb28`, `mdb69`, `mdb91`, `mdb132`, `mdb142`, `mdb145`, `mdb175`, `mdb178`, `mdb181`, `mdb184`, `mdb186`, `mdb188`, `mdb190`, `mdb191`, `mdb199`, `mdb202`, `mdb204`, `mdb206`, `mdb207`, `mdb270`, `mdb315`).

7.8 Intuitionistic Fuzzy Histogram Hyperbolization

Contrast enhancement techniques, in general, aim at modifying the intensity levels of image pixels, with respect to a gray-level transformation function, in order for the image to be more suitable in terms of human brightness perception [17]. Among the various approaches to contrast enhancement, HEQ is widely applied, mainly due to its efficient performance and simple implementation. In HEQ, we seek an intensity transformation in order for the histogram of the resulting image to approximate to a uniform distribution of the gray levels. However, due to the nonlinear nature of human brightness perception, the requirement of a uniform histogram is not always adequate. Exploiting this particular characteristic of brightness perception mechanisms, Frei [16] proposed the idea of *histogram hyperbolization* in which the underlying model for equalization is logarithmic.

A straightforward extension of histogram hyperbolization method in the fuzzy setting was proposed by Tizhoosh and Fochem [56]. In the fuzzy histogram hyperbolization (FHH) approach, the membership values $\mu_{\tilde{A}}(g)$ of gray levels of an image A are modified using a logarithmic function to produce the new intensity levels according to

$$g' = \frac{L-1}{e^{-1}-1}\left[e^{-\left(\mu_{\tilde{A}}(g)\right)^{\beta}} - 1\right] \tag{7.38}$$

where g' is the new gray level, and β is the fuzzifier controlling the enhancement procedure. It should be stressed that the selection of β is image-dependent and greatly affects the enhanced image. Typical values of β are within the range of [0.5, 2]. Small values of β produce brighter images, while as β increases the resulting images become darker.

Vlachos and Sergiadis [69] proposed a generalization of the FHH approach, using the elements of A–IFSs theory. The intuitionistic fuzzy histogram hyperbolization (IFHH) method exploits the IFIP framework presented in Section 7.5. In the first stage, different fuzzification schemes may be applied to assign membership values to the image pixels. In the proposed approach for mammographic image processing, we consider an S-function to generate the membership values given by

$$\mu_{\tilde{A}}(g) = \begin{cases} 0, & \text{if } g \leqslant a \\ 2\left(\dfrac{g-a}{c-a}\right)^2, & \text{if } a < g \leqslant b \\ 1 - 2\left(\dfrac{g-c}{c-a}\right)^2, & \text{if } b < g \leqslant c \\ 1, & \text{if } g > c \end{cases} \tag{7.39}$$

The advantage of using the S-function as a fuzzifier is mainly due to its ability to pre-enhance certain intensity ranges, depending on the selection of parameters a, b, and c. Wirth and Nikitenko [72] evaluated the performance of various fuzzy contrast enhancement algorithms for mammographic imaging and proposed the following selection of parameters for the S-function:

$$a = \min_{\substack{i\in\{0,\ldots,M\}\\ j\in\{0,\ldots,N\}}}\{g_{ij}\}, \qquad c = \max_{\substack{i\in\{0,\ldots,M\}\\ j\in\{0,\ldots,N\}}}\{g_{ij}\}, \qquad \text{and} \qquad b = \frac{a+c}{2}$$

By examining the fuzzy approach to histogram hyperbolization, one may observe that the membership values are modified according to

$$\mu'_{\tilde{A}}(g) = (\mu_{\tilde{A}}(g))^{\beta} \tag{7.40}$$

for all $g \in \{0, \ldots, L-1\}$. De, Biswas, and Roy [14] proposed an operator $A^n : \mathscr{IFS}(X) \to \mathscr{IFS}(X)$ given by

$$A^n = \{\langle x, (\mu_A(x))^n, 1-(1-\nu_A(x))^n\rangle | x \in X\} \tag{7.41}$$

where n is any positive real number. It is evident that in the case of A being an FS, the operator of Equation 7.41 coincides with that of Equation 7.40 used for modifying the membership values in the FHH approach. As a straightforward extension for the IFHH method, the modification of the A–IFSs describing the image in the IFD is performed using the operator of Equation 7.41 according to

$$A'_{opt} = \left\{\langle g, \left(\mu_{A_{opt}}(g)\right)^{\beta}, 1-\left(1-\nu_{A_{opt}}(g)\right)^{\beta}\rangle | g \in \{0,\ldots,L-1\}\right\} \tag{7.42}$$

To transfer the image back to the FD, one of the intuitionistic defuzzification schemes described in Section 7.5.5 is employed to produce the modified membership values $\mu_{\tilde{A}'_{opt}}(g)$. Finally, the image in the PD is obtained as

$$g' = \frac{L-1}{e^{-1}-1}\left[e^{-\mu_{\tilde{A}'_{opt}}(g)} - 1\right] \tag{7.43}$$

As already stated, in contrast enhancement, the employed transformation of the gray levels should be an increasing function in order to preserve their ordering. Since the operator of Equation 7.41 preserves monotonicity of its arguments and the functions 7.11 and 7.12, used to obtain the image in the IFD, are increasing and decreasing functions of g, respectively, then from Proposition 7.1, it is ensured that the produced FS, from which the gray-level transformation will result, preserves the ordering of intensity levels. An overview of the IFHH approach is given in Algorithm 7.1.

Algorithm 7.1 Intuitionistic fuzzy histogram hyperbolization (IFHH).

Input: Grayscale mammographic image A of size $M \times N$ pixels with L gray levels g

begin

Calculate the parameters a, b, and c of the S-function

Fuzzify image according to $\mu_{\tilde{A}}(g) = S(g; a, b, c)$, where $S(\cdot; a, b, c)$, denotes the S-function

Select intuitionistic fuzzy entropy measure E and set the corresponding parameters

Set parameter n of the A^n operator

for $\lambda \in [0, 3]$ **do**

$\mu_A(g; \lambda) \leftarrow 1 - (1 - \mu_{\tilde{A}}(g))^{\lambda}$

$\nu_A(g; \lambda) \leftarrow (1 - \mu_{\tilde{A}}(g))^{\lambda(\lambda+1)}$

Calculate $E(A; \lambda)$

Select $\lambda_{opt} \leftarrow \arg\max_{\lambda} \{E(A; \lambda)\}$

$\mu_{A_{opt}}(g) \leftarrow 1 - (1 - \mu_{\tilde{A}}(g))^{\lambda_{opt}}$

$\nu_{A_{opt}}(g) \leftarrow (1 - \mu_{\tilde{A}}(g))^{\lambda_{opt}(\lambda_{opt}+1)}$

Obtain modified membership and nonmembership values $\left(\mu_{A_{opt}}(g)\right)^{\beta}, 1 - \left(1 - \nu_{A_{opt}}(g)\right)^{\beta}$

Obtain modified membership function $\mu_{\tilde{A}'_{opt}}(g)$ using intuitionistic defuzzification

Obtain modified gray levels according to $g' = \frac{L-1}{e^{-1}-1}\left[e^{\mu_{\tilde{A}'_{opt}}(g)} - 1\right]$

end

Output: Contrast-enhanced mammogram

7.8.1 Algorithm evaluation

The maximization of intuitionistic fuzzy entropy principle of Equation 7.13 allows for the selection of different entropy measures, with difference properties, for obtaining the optimal parameter λ_{opt} in order to represent the image in the IFD. To fully exploit the potential of the proposed IFHH technique, we evaluated its performance by considering different entropy measures in the intuitionistic fuzzification stage and different intuitionistic defuzzification approaches. Due to the particular underlying properties of the various entropy measures, an analysis is carried out for selecting the optimal one for mammographic contrast enhancement on the basis of specific characteristics of the mammographic image under processing and the breast density characteristics. A description of the alternate variations of the IFHH, and the IFIP in general that we considered, are given in Table 7.1. For the intuitionistic fuzzy entropy E_{HY}, we set $k = 0.5$.

In order to provide a quantitative assessment of the considered alternatives, we employed the enhancement evaluation measures of Section 7.6. Within

TABLE 7.1: Variations of the IFIP framework with respect to the intuitionistic fuzzy entropy measures and intuitionistic defuzzification schemes employed (MIFID, maximum index of fuzziness intuitionistic defuzzification; GID, generalized intuitionistic defuzzification).

Algorithm	**Alg1**	**Alg2**	**Alg3**	**Alg4**	**Alg5**	**Alg6**
Entropy Measure	E_{SK}	E_{SK}	E_{BB}	E_{BB}	E_{HY}	E_{HY}
Intuit. Defuzzification	MIFID	GID	MIFID	GID	MIFID	GID

TABLE 7.2: Enhancement evaluation measures for the variations of the IFHH approach for different breast types.

Breast Type	***Measures***	***Algorithm***					
		Alg1	**Alg2**	**Alg3**	**Alg4**	**Alg5**	**Alg6**
Fatty	*DSM*	18.8950	20.9283	17.0816	19.4430	11.7785	0.9122
	TBC_s	0.0279	0.0702	−0.0573	−0.1212	−5.3760	$-\infty$
	TBC_e	0.1953	0.2569	0.0577	0.0156	−25.0694	$-\infty$
	CII	1.4570	1.5448	1.2345	1.2786	1.0063	0.7477
	D	0.5736	**0.5097**	0.7144	0.6726	0.8852	1.1176
Glandular	*DSM*	12.6989	13.5402	18.7608	22.0836	5.2836	4.6038
	TBC_s	0.1609	0.1735	0.1148	0.1201	0.0225	−0.0006
	TBC_e	0.3390	0.3839	0.2192	0.2410	0.0369	0.0010
	CII	1.6393	1.7050	1.4317	1.4653	0.4232	0.3453
	D	0.5393	**0.5038**	0.5887	0.5451	1.2839	1.3369
Dense	*DSM*	15.9365	17.1358	17.1629	19.2083	2.1915	−0.5119
	TBC_s	0.1292	0.1370	0.0984	0.1010	−0.0573	$-\infty$
	TBC_e	0.2572	0.2860	0.1564	0.1565	−0.0150	$-\infty$
	CII	1.4448	1.4791	1.2878	1.2832	0.2391	0.1853
	D	0.6083	**0.5757**	0.6825	0.6714	1.4167	1.4659
All images	*D*	0.5600	**0.5139**	0.6543	0.6175	1.2029	1.3086

this study, the different variations of IFHH were also tested against three breast types: fatty, glandular, and dense. For the assessment, we considered the set of 28 mammograms from the MiniMIAS database containing spiculated/circumscribed masses with well-defined ROIs and background regions, as shown in Figure 7.4. The results are given in Table 7.2.

Table 7.2 shows that the IFHH approach with the entropy E_{SK} of Equation 7.8 employed in the intuitionistic fuzzy stage and the generalized intuitionistic defuzzification scheme based on the cumulative histogram outperforms the other variations of IFHH. For the fatty breast type, all measures indicate that Alg2 performs better, a fact also verified by the combined enhancement measure D. For the glandular and dense breast types, Alg4 performs slightly

better with respect to the DSM criterion, while for the other three criteria, as well as the combined enhancement measure D, Alg2 exhibits the best performance. Finally, the selection of the best algorithm was based on the value of the combined enhancement measure over the entire dataset of the 28 selected mammograms of the MiniMIAS database, shown as the last row of Table 7.2.

7.8.2 FHH versus IFHH: Experimental evaluation

The main purpose of this work is to extend and adapt the IFIP framework in the context of mammographic imaging. In order to evaluate the efficiency of the framework, we compared the proposed IFHH method with its fuzzy counterpart. Figure 7.5 illustrates a sample set of images from the MiniMIAS database, along with their processed versions. The first column contains the initial mammograms, as provided by the database, while subsequent columns correspond to images obtained using the HEQ, the FHH, and the proposed IFHH method, respectively. The target and background regions were also annotated in the initial image as black and white squares. From the visual assessment of the images, we can see that the proposed intuitionistic fuzzy extension of FHH successfully enhances their contrast.

The efficiency of the proposed intuitionistic fuzzy framework was further examined in a more objective way using the measures described in Section 7.6. The measures were computed for all of the 28 selected images and algorithms. The comparison graphs are depicted in Figure 7.6. It is evident that the proposed method delivers better results than the HEQ and the conventional fuzzy approaches to histogram hyperbolization in terms of producing larger values for the DSM, TBC_s, TBC_e, and CII criteria. This denotes the accentuation of the target against its surrounding background. Additionally, Figure 7.7 illustrates the mean values of the aforementioned evaluation measures for all the considered images of the database. A significant improvement of the contrast offered by the proposed intuitionistic fuzzy approach is demonstrated. Moreover, the combined enhancement measure D is also plotted for all images in Figure 7.8, which additionally supports the superiority of the IFHH approach.

To further demonstrate the efficiency of the proposed IFHH technique, we constructed the line profiles corresponding to the images of Figure 7.5, illustrated in Figure 7.9. The line profiles were taken across the middle of the square denoting the background surrounding the target for constant y coordinate. From Figure 7.9 and for all images considered, the proposed method (thick black line) successfully improves the contrast inside the ROI, as can be seen from the increased dynamic range, while the particular morphological characteristics of the target and its surrounding area are preserved. Finally,

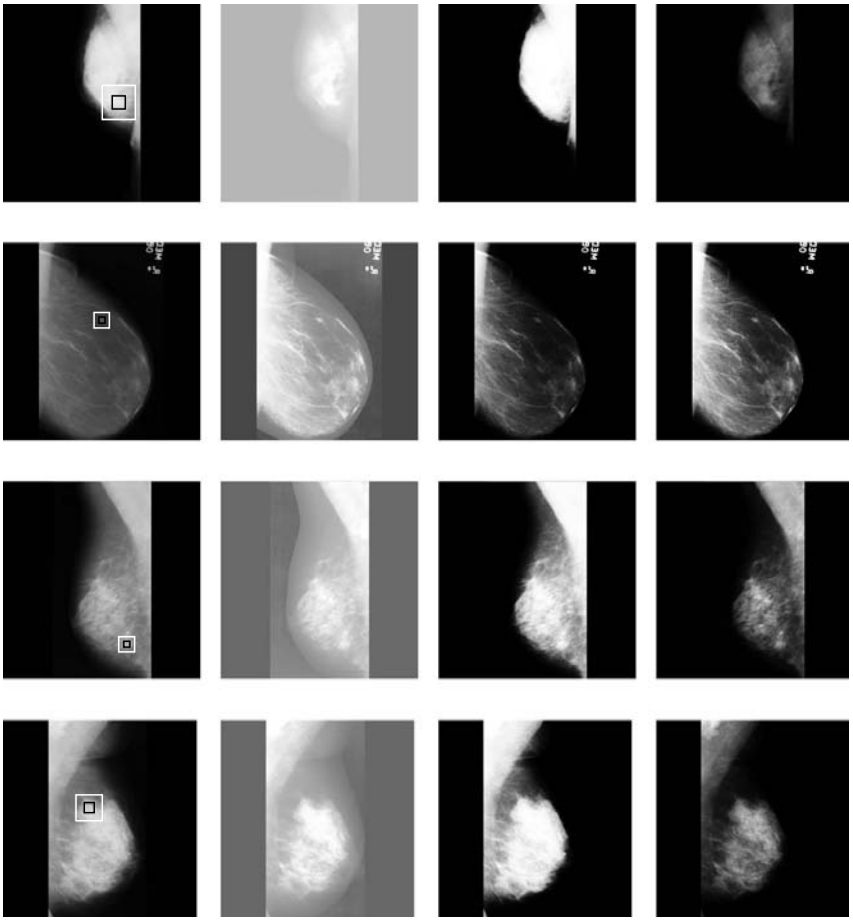


FIGURE 7.5: Sample images from the MiniMIAS database used for experimental evaluation. First column contains initial mammogram (from top to bottom, `mdb179, mdb190, mdb199, mdb244`), while subsequent columns correspond to images processed by the HEQ, FHH, and proposed IFHH approaches. The target and background regions are annotated as black and white squares, respectively.

the same conclusions can be drawn from Figure 7.10, which depicts a three-dimensional representation of the intensity of the target and its corresponding background for the second image of Figure 7.5 (MiniMIAS database name: `mdb190`) and its contrast-enhanced versions delivered by HEQ, FHH method, and its proposed intuitionistic fuzzy extension.

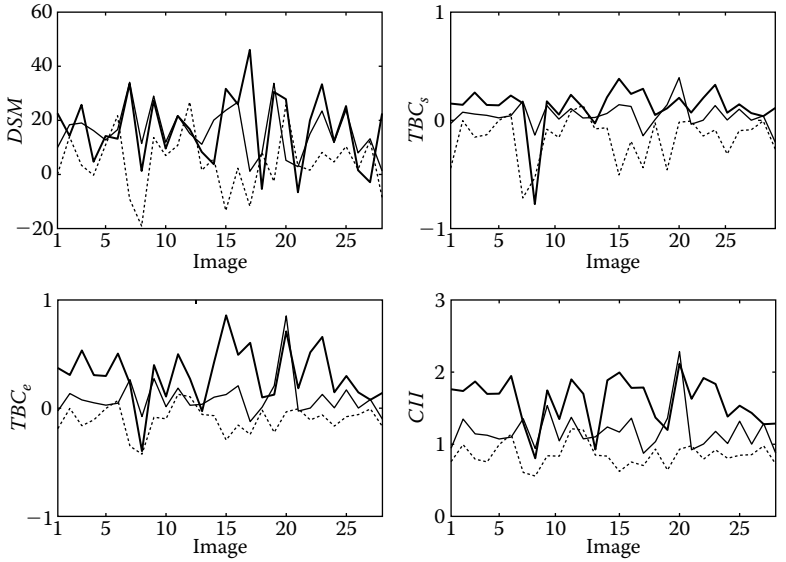


FIGURE 7.6: Enhancement evaluation measures (a) DSM, (b) TBC_s, (c) TBC_e, and (d) CII for the 28 images selected from the MiniMIAS database containing spiculated/circumscribed masses for the HEQ (*gray line*), FHH (*thin black line*), and IFHH method (*thick black line*). The set of test images is shown in Figure 7.4. Larger values of the measures correspond to better separation between the distributions of the target and the background, indicating a better performance of the contrast enhancement algorithm employed.

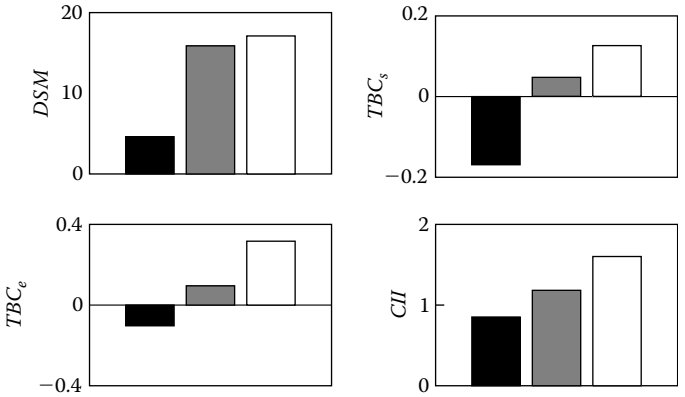


FIGURE 7.7: Mean values for the enhancement measures (a) DSM, (b) TBC_s, (c) TBC_e, and (d) CII of Figure 7.6 for all considered images and algorithms. Evaluated algorithms include HEQ (*black bar*), FHH (*gray bar*), and IFHH (*white bar*).

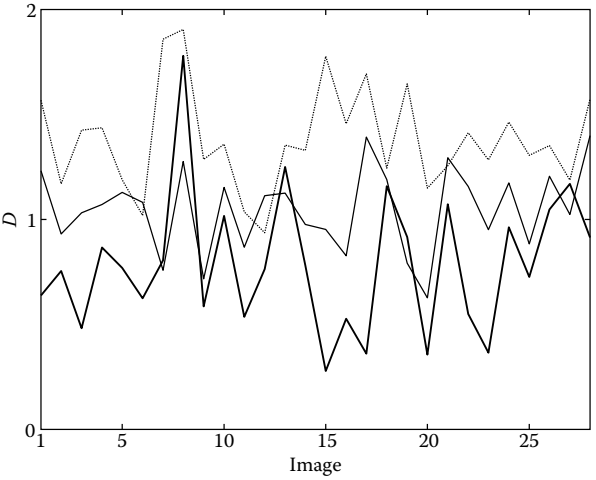


FIGURE 7.8: Combined enhancement measures D for the evaluation of HEQ (*gray line*), FHH (*thin black line*), and IFHH (*thick black line*).

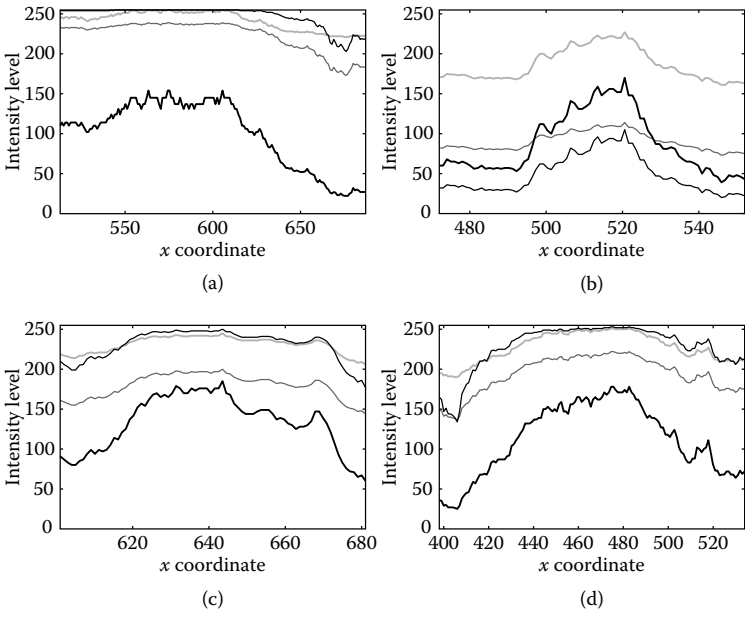


FIGURE 7.9: Line profiles for the images of Figure 7.5 taken across the middle of the square denoting the surrounding background of the target for constant y coordinate. (a) Image `mdb179`, (b) Image `mdb190`, (c) Image `mdb199`, and (d) Image `mdb244`. For each line profile, curves corresponding to the initial image (*thin gray line*) and images processed by HEQ (*thick gray line*), FHH (*thin black line*), and IFHH (*thick black line*) algorithms are plotted against the x coordinate.

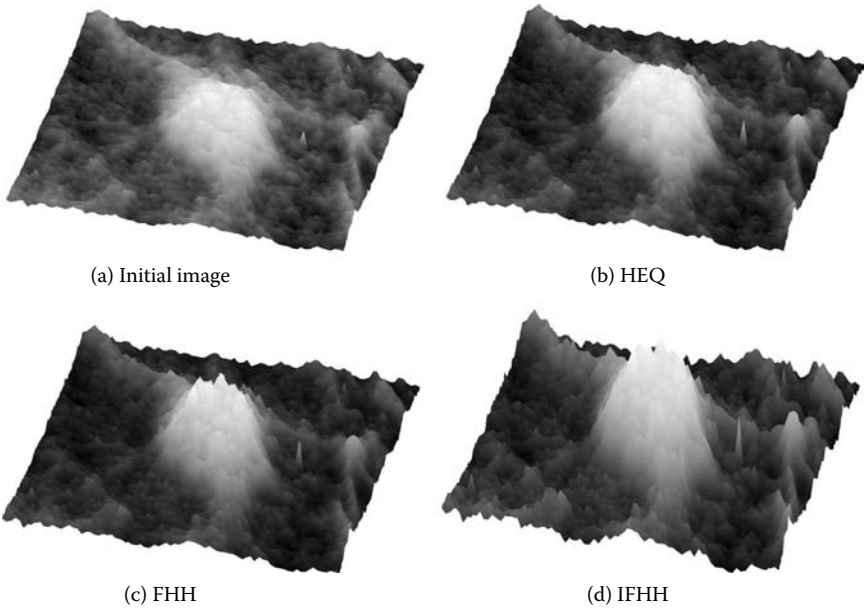


FIGURE 7.10: Three-dimensional representation of the intensity of the target and its background for (a) the second image of Figure 7.5 (Mini-MIAS database name: `mdb190`) and its contrast-enhanced versions delivered by (b) the HEQ technique, (c) the FHH method, and (d) the proposed IFHH approach.

7.9 Intuitionistic Fuzzy Contrast Intensification

7.9.1 Minimization of fuzziness contrast enhancement

Image fuzziness is a measure of the grayness ambiguity associated with image pixels. Reducing the amount of fuzziness carried by a digital image, it is sometimes required to enhance the contrast between bright and dark regions. A contrast enhancement algorithm based on the minimization of fuzziness was proposed by Pal and King [34]. In this approach, the image is fuzzified according to

$$\mu_{\tilde{A}}(g) = \left(1 + \frac{g_{max} - g}{F_d}\right)^{-F_e} \tag{7.44}$$

where g_{max} denotes the maximum gray level desired, and F_e, F_d are the exponential and denominational fuzzifiers, respectively, which control the

ambiguity in the fuzzy plane. Fuzzifier F_d can be determined by the crossover point as

$$F_d = \frac{g_{max} - g_\circ}{\left(\frac{1}{2}\right)^{\frac{-1}{F_e}} - 1} \tag{7.45}$$

with $g_\circ$ being the gray level associated with the crossover point. The modification of the membership values is carried out using the intensification operator [75] given by

$$T_1(\mu_{\tilde{A}}(g)) = \begin{cases} 2\left(\mu_{\tilde{A}}(g)\right)^2, & \text{if } 0 \leqslant \mu_{\tilde{A}}(g) \leqslant \frac{1}{2} \\ 1 - 2\left(1 - \mu_{\tilde{A}}(g)\right)^2, & \text{if } \frac{1}{2} \leqslant \mu_{\tilde{A}}(g) \leqslant 1 \end{cases} \tag{7.46}$$

Successive applications of the intensification operator T are possible, according to the following scheme:

$$T_r(\mu_{\tilde{A}}(g)) = T_1\left\{T_{r-1}(\mu_{\tilde{A}}(g))\right\} \tag{7.47}$$

where $r = 1, 2, \ldots$, which results in further decreasing the fuzziness of the image. In the limiting case of $r \to \infty, T_r$ produces a two-level (binary) image.

After the modification of the membership values has taken place, the defuzzification is performed according to

$$g' = \begin{cases} 0, & \text{if } \bar{g}' < 0 \\ \bar{g}', & \text{if } 0 \leqslant \bar{g}' \leqslant 255 \\ 255, & \text{if } \bar{g}' > 255 \end{cases} \tag{7.48}$$

where g' is the new gray level and $\bar{g}'$ is obtained from the inverse of the membership function as

$$\bar{g}' = g_{max} - F_d\left(\left(\mu'_A(g)\right)^{\frac{-1}{F_e}} - 1\right) \tag{7.49}$$

In order to exploit the ambiguity present in digital images and the difficulties in precisely defining the membership function, an intensification-like operator for A–IFSs was introduced by Vlachos and Sergiadis [69], and its properties were stated and proved.

7.9.2 Intuitionistic fuzzy intensification operator

In order to extend the minimization of fuzziness contrast enhancement, or fuzzy contrast intensification (FCI), algorithm into the intuitionistic fuzzy setting, Vlachos and Sergiadis [69] introduced an intensification-like operator for A–IFSs and investigated its properties.

Definition 7.10 *Let a set $A \in \mathscr{IFS}(X)$. Then,*

$$\boxdot A^m = \{\langle x, \mu_{\boxdot A^m}(x), \nu_{\boxdot A^m}(x)\rangle | x \in X\} \tag{7.50}$$

where

$$\mu_{\boxdot A^m}(x) = \begin{cases} \dfrac{\mu_A^m(x)}{\nu_A^{m-1}(x)}, & \text{if } \mu_A(x) \leqslant \nu_A(x) \\ 1 - \dfrac{(1-\mu_A(x))^m}{(1-\nu_A(x))^{m-1}}, & \text{if } \mu_A(x) \geqslant \nu_A(x) \end{cases} \tag{7.51}$$

and

$$\nu_{\boxdot A^m}(x) = \begin{cases} 1 - \dfrac{(1-\nu_A(x))^m}{(1-\mu_A(x))^{m-1}}, & \text{if } \mu_A(x) \leqslant \nu_A(x) \\ \dfrac{\nu_A^m(x)}{\mu_A^{m-1}(x)}, & \text{if } \mu_A(x) \geqslant \nu_A(x) \end{cases} \tag{7.52}$$

with $m \geqslant 0$ being a real parameter, is an A–IFS.

An interesting property of the $\boxdot$ operator was also stated and proved [69].

Proposition 7.2 *For any set $A \in \mathscr{IFS}(X)$, it holds that*

$$E_{SK}(\boxdot A^m) \leqslant E_{SK}(A) \tag{7.53}$$

with $m \geqslant 1$.

7.9.3 An intuitionistic extension of fuzzy contrast intensification

Let us consider an image A of size $M \times N$ pixels having L gray levels g ranging between 0 and $L-1$. The image is transferred to the FD using the following membership function:

$$\mu_{\tilde{A}}(g) = \frac{g - g_{min}}{g_{max} - g_{min}} \tag{7.54}$$

where g_{min} and g_{max} are the minimum and maximum gray levels of the image, respectively. To construct the optimal A–IFS A_{opt} associated with the gray levels of the image, we employ the optimization of intuitionistic fuzzy entropy method. Similarly to the FCI approach, the modification of the image in the IFD is carried out by the following recursive scheme:

$$\boxdot_r A_{opt}^m = \boxdot_1 \{\boxdot_{r-1} A_{opt}^m\} \tag{7.55}$$

where $r = 1, 2, \ldots$, denotes the number of successive applications of the $\boxdot$ operator. To transfer the image back to the FD, a suitable defuzzification

Algorithm 7.2 Intuitionistic fuzzy contrast intensification (IFCI).

Input: Grayscale mammographic image A of size $M \times N$ pixels with L gray levels g

begin

Calculate minimum and maximum gray levels g_{min} and g_{max} of the image

Fuzzify image according to $\mu_{\tilde{A}}(g) = (g - g_{min})/(g_{max} - g_{min})$

Select intuitionistic fuzzy entropy measure E and set the corresponding parameters

Set number of iterations *iter*

Set parameter m of $\boxdot$ operator

for $\lambda \in [0, 3]$ **do**

$\mu_A(g; \lambda) \leftarrow 1 - (1 - \mu_{\tilde{A}}(g))^{\lambda}$

$\nu_A(g; \lambda) \leftarrow (1 - \mu_{\tilde{A}}(g))^{\lambda(\lambda+1)}$

Calculate $E(A; \lambda)$

Select $\lambda_{opt} \leftarrow \arg\max_{\lambda} \{E(A; \lambda)\}$

$\mu_{A_{opt}}(g) \leftarrow 1 - (1 - \mu_{\tilde{A}}(g))^{\lambda_{opt}}$

$\nu_{A_{opt}}(g) \leftarrow (1 - \mu_{\tilde{A}}(g))^{\lambda_{opt}(\lambda_{opt}+1)}$

for $r = 1$ **to** *iter* **do**

$\boxdot_r A^m_{opt} \leftarrow \boxdot_1 \left\{\boxdot_{r-1} A^m_{opt}\right\}$

Obtain modified membership and nonmembership functions $\mu_{\boxdot iter} A^m_{opt}$ and $\nu_{\boxdot iter} A^m_{opt}$

Obtain modified membership function $\mu'_{\tilde{A}}(g)$ using intuitionistic defuzzification

Obtain modified gray levels according to $g' = (L - 1)\,\mu'_{\tilde{A}}(g)$

end

Output: Contrast-enhanced mammogram

method is applied to obtain the FS with membership function $\mu'_{\tilde{A}}(g)$. Finally, the contrast-enhanced image in the PD is obtained as

$$g' = (L - 1)\,\mu'_{\tilde{A}}(g) \tag{7.56}$$

where g' are the transformed gray levels of the image. An overview of the IFCI approach is provided in Algorithm 7.2.

7.9.4 Algorithm evaluation

Similarly to the IFHH approach, we evaluated the proposed IFCI method by considering the different intuitionistic fuzzy entropy measures and intuitionistic defuzzification schemes, summarized in Table 7.1. The enhancement evaluation criteria for the variations of the IFCI approach are listed in Table 7.3 for different breast types. For the intuitionistic fuzzy entropy E_{HY}, we set $k = 0.5$, while we have considered only one iteration of the algorithm: $r = 1$.

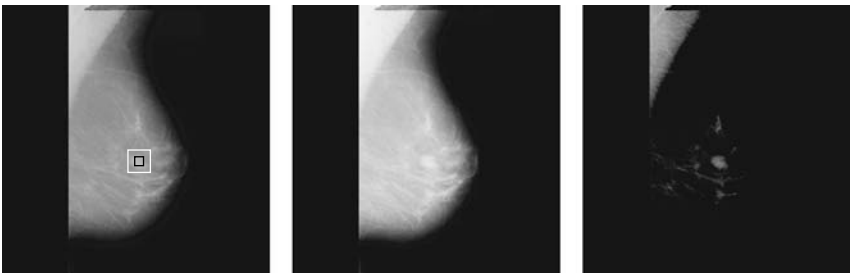
TABLE 7.3: Enhancement evaluation measures for the variations of the IFCI approach for different breast types.

Breast Type	*Measure*	*Algorithm*					
		Alg1	**Alg2**	**Alg3**	**Alg4**	**Alg5**	**Alg6**
Fatty	*DSM*	20.9716	21.9597	10.8661	5.1848	−4.3828	−8.6248
	TBC_s	0.0332	−0.2816	−0.1046	$-\infty$	$-\infty$	$-\infty$
	TBC_e	0.1311	−0.1636	−0.0287	$-\infty$	$-\infty$	$-\infty$
	CII	1.3190	1.2754	1.0204	0.8847	0.6584	0.5857
	D	**0.5176**	0.5347	0.7583	0.8833	1.0830	1.1707
Glandular	*DSM*	22.4714	26.2006	15.3028	13.1693	1.0626	−2.8799
	TBC_s	0.1451	0.1255	0.0370	−0.1135	$-\infty$	$-\infty$
	TBC_e	0.3211	0.2809	0.0611	−0.0400	$-\infty$	$-\infty$
	CII	1.6302	1.5674	1.1196	0.9593	0.7289	0.6502
	D	0.3557	**0.3461**	0.6704	0.7772	0.9971	1.0758
Dense	*DSM*	19.9833	19.9401	9.2446	4.0150	−5.2913	−9.5020
	TBC_s	0.1139	0.0530	−0.0229	−1.2101	−17.0839	$-\infty$
	TBC_e	0.1980	0.1283	−0.0021	−1.0273	−87.7796	$-\infty$
	CII	1.3436	1.2305	0.9976	0.8527	0.6926	0.6232
	D	**0.5121**	0.5745	0.7830	0.9022	1.0603	1.1401
All images	*D*	**0.4421**	0.4589	0.7250	0.8391	1.0407	1.1235

Table 7.3 shows that the IFCI approach driven by the E_{SK} entropy and the maximum index of fuzziness intuitionistic defuzzification schemes delivers the best results with respect to the combined enhancement measure D in case of fatty and dense breast tissues, while using the same entropy and the generalized intuitionistic defuzzification approach is the optimal combination for the glandular breast type.

7.9.5 FCI versus IFCI: Experimental evaluation

For assessing the performance of the IFIP framework, we compared the proposed IFCI approach with its fuzzy counterpart. A critical point in the fuzzy approach is the selection of the gray level g_o associated with the crossover point, which soundly affects the enhancement procedure. In order to demonstrate the importance of the g_o point, we processed a mammographic image with various values of g_o. Figures 7.11b and 7.11c illustrate the images obtained using the FCI approach for two extreme values of $g_o = 10$ and $g_o = 240$. It is easy to verify the impact of the parameter g_o in the final result. Considering that mammographic images are characterized by the variability in tissue density and its inhomogeneous nature suggests that a safe initial selection of g_o would be the midpoint $g_o = 128$ of the gray-level range. Moreover, for experimental evaluation, we considered that $F_e = 2$ and one iteration. It should be stressed that in contrast with the FCI, the proposed intuitionistic fuzzy extension is totally automated, and no parameter must be set manually.



(a) (b) (c)

FIGURE 7.11: (a) Initial mammographic image (MiniMIAS database name: `mdb010`) with fatty background tissue and a well-defined circumscribed benign mass. Images obtained using the minimization of fuzziness algorithm with $F_e = 2, g_{max} = 255$, and one iteration for (b) $g_{\text{o}} = 10$ and (c) $g_{\text{o}} = 240$.

Figure 7.12 depicts images from the sample set of the MiniMIAS database and their enhanced versions derived by the HEQ, the FCI, and the proposed IFCI approach, where different intuitionistic fuzzification–defuzzification configurations for glandular and fatty/dense breast types were used. From visual comparison of the images, we see that the IFCI method produces images where different regions and high-frequency edges are drastically enhanced and are more suitable for examination by the human expert. Moreover, in the corresponding ROIs, the target is radically exposed against its surrounding background. Additionally, the enhancement evaluation measures of Section 7.6 were employed to provide, in Figure 7.13, an objective criterion for assessing the performance of the different approaches. Figure 7.14 illustrates the mean values of the aforementioned evaluation measures for all the considered images of the database. The combined enhancement measure is also plotted for all images in Figure 7.15. Finally, line profiles and three-dimensional representations of the ROIs were also used, and are shown in Figures 7.16 and 7.17, to further demonstrate the efficiency of the proposed IFCI method.

7.10 Discussion

The purpose of this work was to extend existing fuzzy algorithms for contrast enhancement into the intuitionistic fuzzy setting and evaluate the relative merits of this extension. The main objective of contrast enhancement techniques for mammographic imaging is often to enhance the contrast in the ROI that will lead to an increased separation between the distributions of the mass and its surrounding background. This is expected to enhance the performance

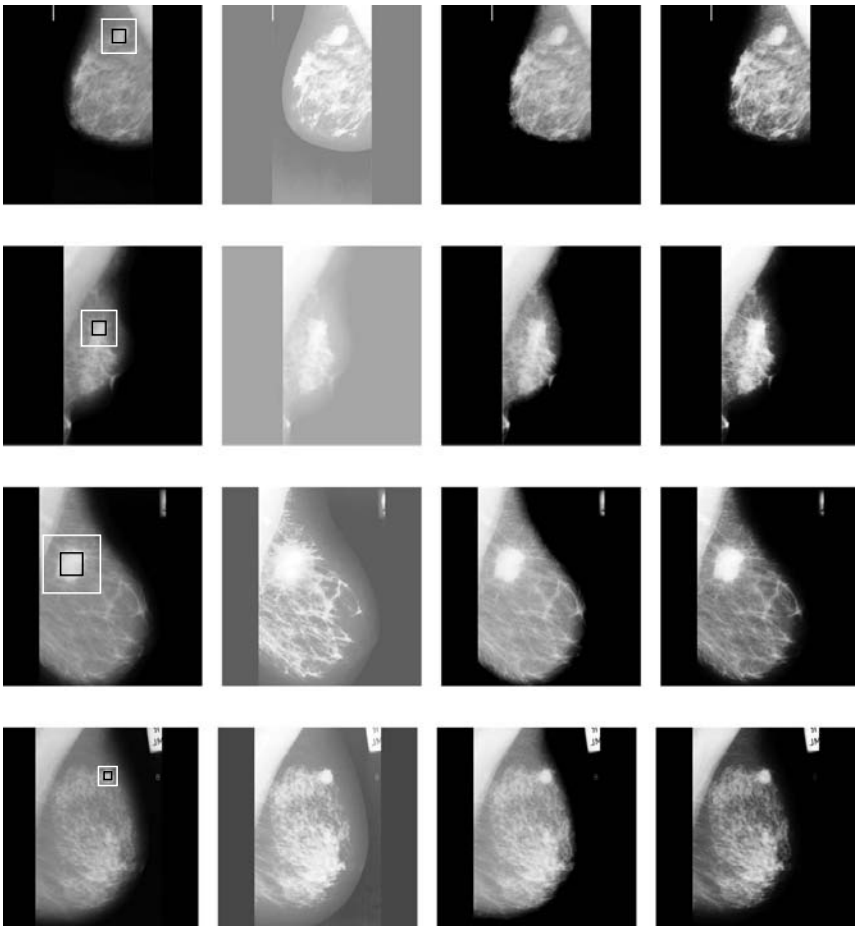


FIGURE 7.12: Sample images from the MiniMIAS database used for experimental evaluation. First column contains initial mammogram (from top to bottom: `mdb015, mdb178, mdb184, mdb202`), while subsequent columns correspond to images processed by the HEQ, FCI, and proposed IFCI approaches. The target and background regions are annotated as black and white squares, respectively.

of the mass detection, and extraction schemes that are to be applied. Driven by this objective, 28 images were selected from the MiniMIAS database that contained well-defined speculated/circumscribed masses with known spatial coordinates. From the location and the size of the mass, the corresponding ROIs were calculated along with their respective surrounding backgrounds. For the MiniMIAS database, data were collected by mammograms digitized at 200 micron pixel edge. To quantitatively rank the enhancement methods,

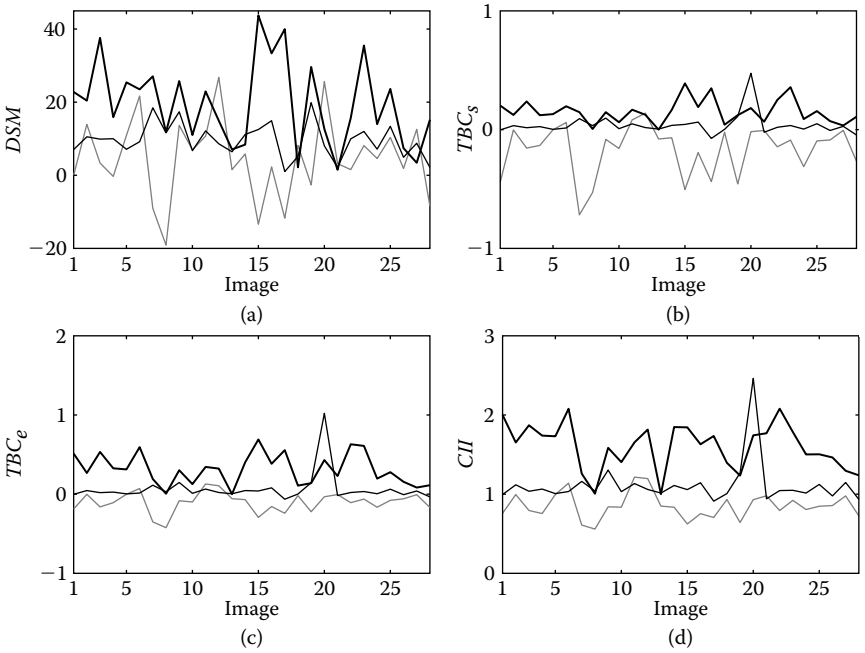


FIGURE 7.13: Enhancement evaluation measures for the 28 images selected from the MiniMIAS database containing spiculated/circumscribed masses for the HEQ (*gray line*), FCI (*thin black line*), and IFCI method (*thick black line*). The set of test images is shown in Figure 7.4. Larger values of the measures correspond to better separation between the distributions of the target and the background, indicating a better performance of the contrast enhancement algorithm employed.

the experimental setup was equipped with objective enhancement evaluation measures assessing the separation between the target and background classes.

The experimental evaluation of the algorithms was carried out using a two-step procedure. In the first step, both the IFHH and IFCI approaches were assessed, considering all possible combinations of intuitionistic fuzzification and defuzzification schemes. Moreover, to provide a more detailed analysis, the different variations of each approach were also tested against different breast types, categorized as fatty, glandular, and dense. The overall optimal combination of intuitionistic fuzzification and defuzzification schemes was derived according to the performance of the combined enhancement measure in the specified ROIs. Finally, the second step of the evaluation procedure involved the comparison of the proposed intuitionistic fuzzy approaches with their fuzzy counterparts, as well as with the well-known HEQ technique.

Experimental results demonstrated both objectively and subjectively the efficiency of the proposed intuitionistic fuzzy approaches to successfully

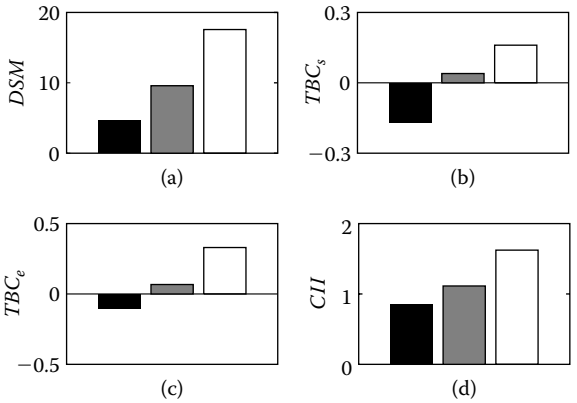


FIGURE 7.14: Mean values for the enhancement measures of Figure 7.13 for all considered images and algorithms. Evaluated algorithms include HEQ *(black bar)*, FCI *(gray bar)*, and IFCI *(white bar)*.

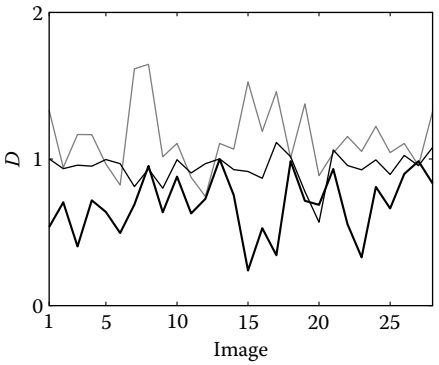


FIGURE 7.15: Combined enhancement measures D for the evaluation of HEQ *(gray line)*, FCI *(thin black line)*, and IFCI *(thick black line)*.

enhance the target (mass) against its surrounding background. Objective assessment involved the calculation of the enhancement evaluation and combined enhancement measures, as well as the comparison of their respective mean values. To further support the potential of the IFIP framework, the line profiles across the ROI were taken into account to demonstrate the increase in the dynamic range between the target and its background. In the same spirit, three-dimensional representations of the intensity of the mass were also considered.

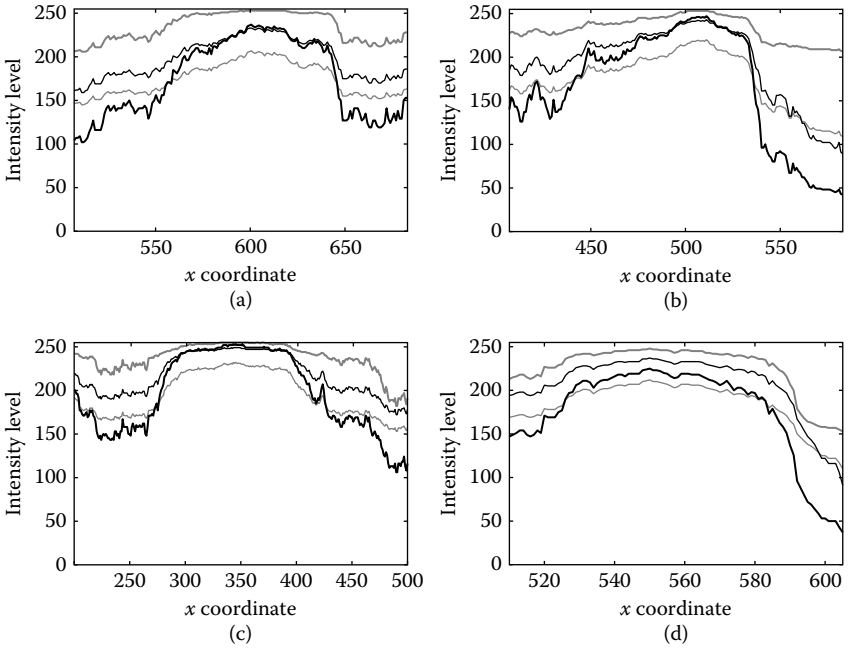


FIGURE 7.16: Line profiles for the images of Figure 7.11 taken across the middle of the square denoting the surrounding background of the target for constant y coordinate. (a) Image `mdb015`, (b) Image `mdb178`, (c) Image `mdb184`, and (d) Image `mdb202`. For each line profile, curves corresponding to the initial image *(thin gray line)* and images processed by HEQ *(thick gray line)*, FCI *(thin black line)*, and IFCI *(thick black line)* algorithms are plotted against the x coordinate.

7.11 Future Research Directions

The IFIP architecture and the theory of A–IFSs in general provide a flexible framework for dealing with the uncertainty characterizing medical images. Moreover, recent publications show that the field of A–IFSs is beginning to shift from theory to application. Therefore, as a newly introduced approach, IFIP constitutes a prominent open field for future research.

Two of the most important stages of IFIP are the processes of analyzing and synthesizing images to and from their intuitionistic fuzzy components. An open issue that deserves to be further investigated is the construction of alternative adaptive intuitionistic fuzzification and defuzzification schemes that take into account the particular characteristics of medical images as well as their generation mechanisms.

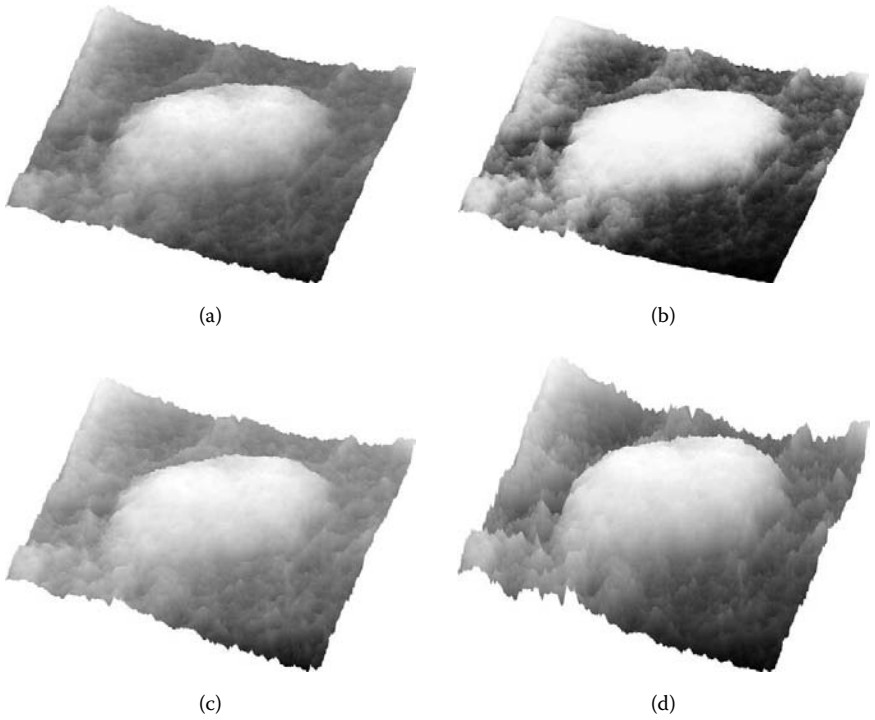


FIGURE 7.17: Three-dimensional representation of the intensity of the target and its background for (a) the first image of Figure 7.11 (MiniMIAS database name: `mdb015`) and its contrast-enhanced versions delivered by (b) the HEQ technique, (c) the FCI method, and (d) the proposed IFCI approach.

This work showed that the application of A–IFSs theory to existing fuzzy approaches to contrast enhancement of mammographic images significantly enhances their performance. This observation reveals promising ways in which the IFIP framework can be applied. Experimental results suggest that the presented architecture can successfully replace fuzzy, or even non-fuzzy, approaches in the context of digital mammography or medical imaging in general. The fact that many image-related definitions, such as the "boundary" of a region, the "shape" of a mass, or the "size" of a microcalcification, are themselves fuzzy in nature, thus bearing an appreciable amount of uncertainty and hesitancy, allows us to assert that the IFIP architecture needs to be further explored.

Finally, an emerging trend in the intuitionistic fuzzy setting is the application of A–IFSs in supporting medical diagnostic reasoning using distance and similarity measures [29, 52–54] or the concept of intuitionistic fuzzy relations [15]. This is justified by the fact that A–IFSs are a flexible tool for reasoning under imperfectly defined facts and imprecise knowledge in a more human-consistent manner than FSs. In the context of medical imaging and

diagnosis, imperfect information springs out of the lack of knowledge or the discrimination limitations of the measuring tools. Consequently, CAD systems can be favored by the implementation of an intuitionistic fuzzy reasoning module, able to express and model aspects of imperfect information.

7.12 Conclusions

In this chapter, we presented a thorough analysis of the IFIP framework in the context of mammographic image processing. The properties of different intuitionistic fuzzy entropy measures and intuitionistic defuzzification schemes were evaluated, considering the particular characteristics of digitized mammograms. Under this scope, intuitionistic fuzzy algorithms, namely the IFHH and IFCI, were presented. To provide an objective and quantitative assessment of the proposed algorithms, enhancement evaluation measures were applied that demonstrated the efficiency of the IFIP framework. Finally, we showed that using the concepts of A–IFSs, one can enhance the performance of existing fuzzy approaches, yielding satisfactory and promising results.

References

[1] Atanassov K.T. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 20: 87–96, 1986.

[2] Atanassov K.T. More on intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 33: 37–45, 1989.

[3] Atanassov K.T. *Intuitionistic Fuzzy Sets: Theory and Applications*. Studies in Fuzziness and Soft Computing. Berlin: Springer-Verlag, 1999.

[4] Atanassov K.T., Koshelev M., Kreinovich V., Rachamreddy B., and Yasemis H. Fundamental justification of intuitionistic fuzzy logic and of interval-valued fuzzy methods. *Notes on Intuitionistic Fuzzy Sets*, 4: 42–46, 1995.

[5] Bovis K. and Singh S. Enhancement technique evaluation using quantitative measures on digital mammograms. In M. Yaffe (Ed.), *Proc. 5th International Workshop on Digital Mammography*, pp. 547–553. Toronto: Medical Physics Publishing, 2000.

[6] Brodie I. and Gutcheck R.A. Radiographic information theory and application to mammography. *Med. Phys.*, 9: 79–95, 1982.

[7] Brzakovic D., Luo X.M., and Brzakovic P. An approach to automated detection of tumors in mammograms. *IEEE Trans. Med. Imag.*, 9: 233–241, 1990.

[8] Burillo P. and Bustince H. Construction theorems for intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 84: 271–281, 1996.

[9] Burillo P. and Bustince H. Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst.*, 78: 305–316, 1996.

[10] Bustince H., Kacprzyk J., and Mohedano V. Intuitionistic fuzzy generators—Application to intuitionistic fuzzy complementation. *Fuzzy Sets Syst.*, 114: 485–504, 2000.

[11] Chang H.P., Wei D., Helvie M.A., Sahiner B., Adler D.D., Goodsitt M., and Petrick N. Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space. *Phys. Med. Biol.*, 40: 857–876, 1995.

[12] Cheng H.D. and Chen J.R. Automatically determine the membership function based on the maximum entropy principle. *Inform. Sci.*, 96: 163–182, 1997.

[13] Cheng H.D., Lui Y.M., and Freimanis R.I. A novel approach to microcalcification detection using fuzzy logic technique. *IEEE Trans. Med. Imag.*, 17: 442–450, 1998.

[14] De S.K., Biswas R., and Roy A.R. Some operations on intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 114: 477–484, 2000.

[15] De S.K., Biswas R., and Roy A.R. An application of intuitionistic fuzzy sets in medical diagnosis. *Fuzzy Sets Syst.*, 117: 209–213, 2001.

[16] Frei W. Image enhancement by histogram hyperbolization. *Comput. Graphics Image Process.*, 6: 286–294, 1977.

[17] Gonzalez R.C. and Woods R.E. *Digital Image Processing.* Upper Saddle River, NJ: Prentice Hall, 3rd ed., 2006.

[18] Gupta R. and Undrill P.E. The use of texture analysis to delineate masses in mammography. *Phys. Med. Biol.*, 40: 835–855, 1995.

[19] Hair J., Anderson R., and Tatham R. *Multivariate Data Analysis*, Upper Saddle River, NJ: Prentice Hall, 1998.

[20] Hassanien A.E. and Ali J.M. Digital mammogram segmentation algorithm using pulse coupled neural networks. In *Proc. Third International Conference on Image and Graphics (ICIG '04)*, 2004.

[21] Hassanien A.E., Ali J.M., and Nobuhara H. Detection of spiculated masses in mammograms based on fuzzy image processing. In L. Rutkowsk, J. Siekmann, R. Tadeusiewicz, and L.A. Zadeh (Eds.), *Artificial Intelligence and Soft Computing—ICAISC 2004*, Lecture Notes in Computer Science, vol. 3070, pp. 1002–1007. Berlin: Springer-Verlag, 2004.

[22] Hassanien A.E. and Badr A. A comparative study on digital mammography enhancement algorithms based on fuzzy theory. *Studies in Informatics and Control*, 12: 21–31, 2003.

[23] Hung W.L. and Yang M.S. Fuzzy entropy on intuitionistic fuzzy sets. *Int. J. Intell. Sys.*, 21: 443–451, 2006.

[24] Johns P.C. and Yaffe M.J. X-ray characterization of normal and neoplastic breast tissues. *Phys. Med. Biol.*, 32: 675–695, 1987.

[25] Kaufmann A. *Introduction to the Theory of Fuzzy Subsets—Fundamental Theoretical Elements*, vol. 1. New York: Academic Press, 1975.

[26] Laine A.F., Schuler S., Fan J., and Huda W. Mammographic feature enhancement by multiscale analysis. *IEEE Trans. Med. Imag.*, 13: 725–740, 1994.

[27] Li H.D., Kallergi M., Clarke L.P., Jain V.K., and Clark R.A. Markov random field for tumor detection in digital mammography. *IEEE Trans. Med. Imag.*, 14: 565–576, 1995.

[28] Li L. and Clarke L.P. Digital mammography: Computer assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms. *Acad. Radiol.*, 4: 725–731, 1997.

[29] Liang Z. and Shi P. Similarity measures on intuitionistic fuzzy sets. *Pattern Recognit. Lett.*, 24: 2687–2693, 2003.

[30] Luca A.D. and Termini S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf. Control*, 20: 301–312, 1972.

[31] Mendel J.M. and John R.I.B. Type-2 fuzzy sets made simple. *IEEE Trans. Fuzzy Syst.*, 10: 117–127, 2002.

[32] Morrow W.M. and Paranjape R.B. Region-based contrast enhancement of mammograms. *IEEE Trans. Med. Imag.*, 11: 392–406, 1992.

[33] Pal S.K. and King R.A. Image enhancement using fuzzy set. *Electron. Lett.*, 16: 376–378, 1980.

[34] Pal S.K. and King R.A. Image enhancement using smoothing with fuzzy sets. *IEEE Trans. Syst., Man, Cybern.*, 11: 495–501, 1981.

[35] Pal S.K. and King R.A. A note on the quantitative measure of image enhancement through fuzziness. *IEEE Trans. Pattern Anal. Machine Intell.*, 4: 204–208, 1982.

[36] Pal S.K. and Majumder D.K.D. *Fuzzy Mathematical Approach to Pattern Recognition*. New York: Wiley, 1986.

[37] Petrick N., Chan H.P., Sahiner B., and Helvie M.A. Combined adaptive enhancement and region growing segmentation of breast masses on digitized mammograms. *Med. Phys.*, 26: 1642–1654, 1999.

[38] Petrick N., Chan H.P., Sahiner B., and Wei D. An adaptive density weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Trans. Med. Imag.*, 15: 59–67, 1996.

[39] Petrick N., Chan H.P., Wei D., Sahiner B., Helvie M.A., and Adler D.D. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. *Med. Phys.*, 23: 1685–1696, 1996.

[40] Philip K.P., Dove E.L., McPherson D.D., Gotteiner N.L., Stanford W., and Chandran K.B. The fuzzy Hough transform—Feature extraction in medical images. *IEEE Trans. Med. Imag.*, 13: 235–240, 1994.

[41] Qian W., Clarke L.P., Song D., and Clark R.A. Digital mammography: Hybrid four channel wavelet transform for microcalcification segmentation. *Acad. Radiol.*, 5: 355–364, 1998.

[42] Rangayyan R.M., Shen L., Shen Y., Desautels J.E.L., Byrant H., Terry T.J., Horeczko N., and Rose M.S. Improvement of sensitivity of breast cancer diagnosis with adaptive neighborhood contrast enhancement of mammograms. *IEEE Trans. Inf. Technol. Biomed.*, 1: 161–169, 1997.

[43] Renyi A. On measures of entropy and information. In *Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547–561. Berkeley: University of California Press, 1960.

[44] Rick A., Bouchon-Meunier B., Muller S., and Rifqi M. *Fuzzy Techniques in Mammographic Image Processing.* Heidelberg: Physica-Verlag, 2000.

[45] Saha P.K., Udupa J.K., Conant E.F., Chakraborty D.P., and Sullivan D. Breast tissue density quantification via digitized mammograms. *IEEE Trans. Med. Imag.*, 20: 792–803, 2001.

[46] Singh S. and Bovis K. An evaluation of contrast enhancement techniques for mammographic breast masses. *IEEE Trans. on Information Technology in Biomedicine*, 9: 109–119, 2005.

[47] Singh S. and Bovis K.J. *Advanced Algorithmic Approaches in Medical Image Segmentation: State-of-the-Art Application in Cardiology, Mammography and Pathology*, chap. Digital Mammography Segmentation, pp. 440–540. London: Springer-Verlag, 2001.

[48] Suckling J., Parker J., Dance D., Astley S., Hutt I., Boggis C., Ricketts I., Stamatakis E., Cerneaz N., Kok S., Taylor P., Betal D., and Savage J. The mammographic image analysis society digital mammogram database. *International Congress Series*, vol. 1069, pp. 375–378, 1994.

[49] Szmidt E. and Kacprzyk J. A fuzzy set corresponding to an intuitionistic fuzzy set. Int. *J. Uncertainty Fuzziness Knowl.-Based Syst.*, 6: 427–435, 1998.

[50] Szmidt E. and Kacprzyk J. Distances between intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 114: 505–518, 2000.

[51] Szmidt E. and Kacprzyk J. Entropy for intuitionistic fuzzy sets. *Fuzzy Sets Syst.*, 118: 467–477, 2001.

[52] Szmidt E. and Kacprzyk J. Intuitionistic fuzzy sets in intelligent data analysis for medical diagnosis. In *Proc. International Conference on Computational Science—Part II*, pp. 263–271. San Francisco, CA, 2001.

[53] Szmidt E. and Kacprzyk J. Intuitionistic fuzzy sets in some medical applications. In *Proc. 7th Fuzzy Days on Computational Intelligence: Theory and Applications*, pp. 148–151. Dortmund, Germany, 2001.

[54] Szmidt E. and Kacprzyk J. A similarity measure for intuitionistic fuzzy sets and its application in supporting medical diagnostic reasoning. In *Proc. 7th International Conference Artificial Intelligence and Soft Computing*, pp. 388–393. Zakopane, Poland, 2004.

[55] Szmidt E. and Kacprzyk J. Entropy and similarity of intuitionistic fuzzy sets. In *Proc. Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 2375–2382. Paris, France, 2006.

[56] Tizhoosh H.R. and Fochem M. Image enhancement with fuzzy histogram hyperbolization. In *Proc. 3rd European Congress on Intelligent Techniques and Soft Computing, EUFIT 1995*, vol. 3, pp. 1695–1698, 1995.

[57] Tizhoosh H.R., Krell G., and Michaelis B. λ-enhancement: Contrast adaptation based on optimization of image fuzziness. In *Proc. IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 1998*, vol. 2, pp. 1548–1553. Alaska, 1998.

[58] Vial S., Gibon D., Vasseur C., and Rousseau J. Volume delineation by fusion of fuzzy sets obtained from multiplanar tomographic images. *IEEE Trans. Med. Imag.*, 20: 1362–1372, 2001.

[59] Vlachos I.K. and Sergiadis G.D. Towards intuitionistic fuzzy image processing. In *Proc. International Conference on Computational Intelligence for Modelling, Control and Automation, 2005*. Vienna, Austria, 2005.

[60] Vlachos I.K. and Sergiadis G.D. A heuristic approach to intuitionistic fuzzification of color images. In *Proc. 7th International FLINS Conference on Applied Artificial Intelligence*, pp. 767–774. Genova, Italy, 2006.

[61] Vlachos I.K. and Sergiadis G.D. Inner product based entropy in the intuitionistic fuzzy setting. *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 14: 351–366, 2006.

[62] Vlachos I.K. and Sergiadis G.D. Intuitionistic fuzzy information—Applications to pattern recognition. *Pattern Recognit. Lett.*, 28: 197–206, 2006.

[63] Vlachos I.K. and Sergiadis G.D. On the intuitionistic defuzzification of digital images for contrast enhancement. In *Proc. 7th International FLINS Conference on Applied Artificial Intelligence*, pp. 759–766. Genova, Italy, 2006.

[64] Vlachos I.K. and Sergiadis G.D. *Applications of Fuzzy Sets Theory*, Lecture Notes in Computer Science, vol. 4578, chap. Intuitionistic Fuzzy Histogram Hyperbolization for Color Images, pp. 328–334. Berlin: Springer, 2007.

[65] Vlachos I.K. and Sergiadis G.D. *Applications of Fuzzy Sets Theory*, Lecture Notes in Computer Science, vol. 4578, chap. A Two-Dimensional Entropic Approach to Intuitionistic Fuzzy Contrast Enhancement, pp. 321–327. Berlin: Springer, 2007.

[66] Vlachos I.K. and Sergiadis G.D. *Foundations of Fuzzy Logic and Soft Computing*, Lecture Notes in Computer Science, vol. 4529, chap. The Role of Entropy in Intuitionistic Fuzzy Contrast Enhancement, pp. 104–113. Berlin: Springer, 2007.

[67] Vlachos I.K. and Sergiadis G.D. *Foundations of Fuzzy Logic and Soft Computing*, Lecture Notes in Computer Science, vol. 4529, chap. Intuitionistic Fuzzy Histograms of an Image, pp. 86–95. Berlin: Springer, 2007.

[68] Vlachos I.K. and Sergiadis G.D. Hesitancy histogram equalization. In *Proc. IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2007*, pp. 1685–1690. London, 2007.

[69] Vlachos I.K. and Sergiadis G.D. *Soft Computing in Image Processing: Recent Advances*, Studies in Fuzziness and Soft Computing, vol. 210, chap. Intuitionistic Fuzzy Image Processing, pp. 385–416. Berlin: Springer, 2007.

[70] Vlachos I.K. and Sergiadis G.D. Intuitionistic Fuzzy Image Processing, *Encyclopedia of Artificial Intelligence*, Information Science Reference, pp. 967–974, 2008.

[71] Wei D., Chan H.P., Petrick N., Sahiner B., Helvie M.A., Adler D.D., and Goodsitt M.M. False positive reduction technique for detection of masses on digital mammograms: Global and local multiresolution texture analysis. *Med. Phys.*, 24: 903–914, 1997.

[72] Wirth M.A. and Nikitenko D. Quality evaluation of fuzzy contrast enhancement algorithms. In *Proc. 2005 Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 436–441, 2005.

[73] Yin F.F., Giger M.L., Vyborny C.J., Doi K., and Schmidt R.A. Comparison of bilateral subtraction techniques in the computerized detection of mammographic masses. *Investigative Radiol.*, 28: 473–481, 1993.

[74] Zadeh L.A. Fuzzy sets. *Inf. Control*, 8: 338–353, 1965.

[75] Zadeh L.A. A fuzzy-set-theoretic interpretation of linguistic hedges. *J. Cybern.*, 2: 4–34, 1972.

Chapter 8

Fuzzy C-Means and Its Applications in Medical Imaging

Huiyu Zhou

Contents

8.1 Introduction

Image segmentation can be defined as the grouping of similar pixels in a parametric space, where they are associated with each other in the same or different images. Classical image segmentation methodologies include thresholding, edge detection, and region detection. However, many of these systems have demonstrated less effectively in the presence of neighboring structures or noise.

In recent years, numerous efforts of c-means/k-means segmentation methods with faster computation and more flexible capabilities than the classical techniques have emerged. The conventional hard c-means (HCM) and fuzzy c-means (FCM) algorithms are two clustering based segmentation techniques. In contrast to HCM, FCM allows us to reduce the uncertainty of pixels belonging to one class and therefore provides improved segmentation. In addition, multiple classes with varying degrees of membership can be continuously updated.

In this chapter, we summarize several fizzy c-means based segmentation techniques, namely conventional FCM, fast FCM with random sampling, and fast generalized FCM, and also present a new anisotropic mean shift based fizzy c-means algorithm. The proposed clustering method incorporates a mean field term within the standard fuzzy c-means objective function. Since mean shift can quickly and reliably find cluster centers, the entire strategy is capable of optimally segmenting clusters within the image.

8.2 Related Work

The image segmentation can be defined as the grouping of similar pixels in a parametric space, where they are associated with each other in the same or different images. Classical image segmentation methodologies include thresholding (e.g., Sahoo et al. [26]), edge detection (e.g., Nevatia [22]), region detection (e.g., Ito, Kamekura, and Shimazu [16]), and connectivity-preserving relaxation.

Thresholding methods are relatively straightforward but lack sensitivity and specificity for accurate segmentation in the presence of different objects with similar intensities or colors [26]. Edge-based methods, quite similar to the contour detection, are fast but sensitive to noise in the background and fail to link together broken contours. Region detection is superior to the thresholding and edge-based methods in terms of stability and consistency. Nevertheless, these approaches need further modifications in order to effectively handle clutters, which commonly exist in real scenarios. Connectivity-preserving relaxation methods, also called active contour models (e.g., Kass, Witkin, and Terzopoulos [17]), start from an initial contour shape to which shrink or expansion operations are applied according to a defined object function. Unfortunately, the convergence of the computation is readily affected by local minima of the function.

Recently, computational intelligence schemes, such as neural networks and genetic algorithms, have been substantially studied. The motivation of exploring these new schemes is that the classical strategies demand predefined similarities of image attributes for segmentation and hence lack generality, robustness, and flexibility. Using the computational intelligence schemes, we expect to hold improvements of groupings of similar characteristics, depending

on the correspondence between the learned geometric topology and the real structures. In this chapter, we highlight only some of the established computational intelligence approaches.

Neural networks do not need any well-described properties for an image to be segmented. They train a preliminary model and then refine the required parameters progressively. Ahmed and Farag [1] presented a volume segmentation based on an unsupervised clustering capability of neural networks. They assigned each voxel a feature pattern with a scaled family of differential geometrical invariant features. These features were then linked to a region using a two-stage neural network scheme. The first stage was a self-organizing principal component analysis, which projected the feature vector onto its leading principal axes. The second stage utilized a self-organizing feature map so as to cluster the inputs into different regions. Iscan and others [15] reported an improved incremental self-organizing map network using automatic thresholding to segment ultrasound images. This method eliminated the thresholding problem in the classical methods. A voxel-based multispectral image segmentation was introduced by Wismüller and colleagues [33], which reduced a class of similar function approximation problems to the explicit supervised one-shot training of a single data set. Similar work using neural networks can be found in Wang and others [31] and Wang and Xu [32].

Genetic algorithm is a highly self-adapted algorithm of best search. It is especially suited to handling traditional algorithms that cannot effectively search for an appropriate solution to nonlinear problems. For example, Chun and Yang [8] discovered fuzzy genetic algorithm fitness functions for mapping a region-based segmentation onto the binary string that represents an individual while demonstrating the evolution of potential segmentation. In addition, Maulic and Bandyopadhyay [21] reported a genetic algorithm-based clustering that searched for appropriate cluster centers in the feature space such that a similarity metric of the resulting cluster is optimized. The centers of the clusters were encoded by the chromosomes by strings of real numbers.

FCM has gained wide attention in recent years, and its applications are found throughout the areas of image segmentation and clustering. FCM is important enough to be explored because of its multiclass division and automatic evolution of class states capabilities. Pham [25] claimed an improved FCM objective function integrating a penalty function in the membership functions. To reduce the computational time, Chen and Zhang [5] proposed two invariants so that the neighborhood term of the objective function discovered by Ahmed and others [2] could be simplified. These two variants enhanced the efficiency by considering mean-filtered and median-filtered images. More recently, an enhanced fuzzy c-means (EnFCM) algorithm [28] was proposed to accelerate the segmentation. The motivation of EnFCM was driven by the fact that the number of gray level is usually much smaller than the image size. This size constraint helped formulate a linearly weighted sum image, and hence we could achieve clustering in the intensity histogram instead of every single pixel.

In comparison to the established neural networks and genetic algorithms, which require intensive pretraining for tuning parameters, FCM does not demand such a training stage and hence shows improvements of efficiency. We do not attempt to prove that the proposed FCM algorithm outperforms those neural networks and genetic algorithms in computational efficiency and accuracy. Instead, our goal is to demonstrate the computational efficiency of the proposed technique against the established FCM approaches.

8.3 Classical FCM Strategies

In this section, those established frameworks with respect to FCM are summarized to identify the advantages and disadvantages of different approaches. Possible improvements can be launched on the basis of the introduction of these state-of-the-art schemes.

8.3.1 Original FCM clustering

Let $X = (x_1, x_2, \ldots, x_N)$ be an image with N pixels to be assigned to C clusters and x_i denote multispectral data. The clustering algorithm is an iteration process that minimizes the object function as follows:

$$f = \sum_{j=1}^{N} \sum_{i=1}^{C} u_{ij}^{k} ||x_j - s_i||^2 \tag{8.1}$$

where u_{ij} is the membership of pixel x_j in the ith cluster, s_i is the ith cluster center, $||\cdot||$ is a norm metric, and k is a constant, which dominates the fuzziness of the resulting partition.

This object function can be of a global minimum when pixels near the centroid of corresponding clusters are assigned higher membership values, while lower membership values are assigned to pixels with data far from the centroid [7]. In this text, the membership function is proportional to the probability that a pixel belongs to a specific cluster. Here, the probability is only dependent on the distance between the image pixel and each independent cluster center. The membership function and the cluster center are updated by

$$u_{ij} = \frac{1}{\sum_{m=1}^{C} (||x_j - s_i|| / ||x_j - s_m||)^{2/(k-1)}} \tag{8.2}$$

with

$$s_i = \frac{\sum_{j=1}^{N} u_{ij}^{k} x_j}{\sum_{j=1}^{N} u_{ij}^{k}} \tag{8.3}$$

An initial guess for each cluster center is required. The FCM can converge to a local minimization solution. Unless the difference between two neighboring iterations is less than a fixed threshold (e.g., one pixel distance), the iteration of the FCM will not be terminated. Algorithm 8.1 tabulates the classical FCM algorithm [3]. Figure 8.1 presents a cryosectioned brain image with its segmentation result using this classical FCM scheme. Figure 8.1b shows more than three significant intensity peaks, indicating that a proper clustering number is 4 or higher (4 is used in this example). Figure 8.1c demonstrates that the subregions can be well partitioned using the classical FCM method. The entire computation for the segmentation takes 99.4 seconds on a PC with a 1.5 GHz Intel® Pentium® CPU and MATLAB® implementation (this setup is also used for the evaluation Section 8.5). The efficiency of the classical FCM is investigated in Hu and Hathaway [14]. To effectively handle this inefficiency problem, we introduce a fast FCM algorithm with random sampling.

Algorithm 8.1 The classical FCM image segmentation algorithm.

1: Initialize s and record the cluster centers. Set $j = 0$.
2: Initialize the fuzzy partition matrix $u(j = 0)$ with Equation 8.2.
3: Increment $j = j + 1$. Compute s by using Equation 8.3.
4: Compute $u(j)$ by using Equation 8.2.
5: Iterate steps 3−4 until the stopping criterion is met.

8.3.2 Fast FCM algorithm with random sampling

To combat a computationally intensive process, Cheng, Goldgof, and Hall [6] proposed a multistage random-sampling FCM strategy. This method has fewer feature vectors and requires fewer iterations to converge. The concept is to randomly sample and obtain a small subset of the full data set for approximating cluster centers of the full data set. This approximation is then used to reduce the number of converging iterations. The entire multistage random-sampling FCM algorithm consists of two phases, whereas a multistage iterative process of a modified FCM is performed in phase 1. Phase 2 is a standard FCM with the cluster centers approximated by the final cluster centers from phase 1.

Before introducing the details of this method, we must determine the following parameters. The first one is the size of the subsamples, which is $\Delta\%$ of the n_s samples in X, denoted by $X_{(\Delta\%)}$. The second one is the number of the stages, n. After phase 1, the data set includes $({n_s}^*\Delta\%)$ samples, which is denoted as $X_{(n_s^*\Delta\%)}$. The third one is the stopping criterion for phase $1, \epsilon_1$, and the last one is the stopping condition for phase $2, \epsilon_2$. The entire algorithm is summarized in Algorithm 8.2.

Evidence has shown that this improved FCM with random sampling can reduce the computation required in the classical FCM method. To

Algorithm 8.2 The multistage random-sampling FCM algorithm.

Phase 1: Randomly initialize the cluster matrix s.

Step 1: Select $X_{(\Delta\%)}$ from the set of the original feature vectors matrix $(z = 1)$.

Step 2: Initialize the fuzzy partition matrix u by using Equation 8.2 with $X_{(z^*\Delta\%)}$.

Step 3: Compute the stopping condition $\epsilon = \epsilon_1 - z^*((\epsilon_1 - \epsilon_2)/n_s)$.

Step 4: Set $j = 0$.

Step 5: Set $j = j + 1$.

Step 6: Compute the cluster center matrix $s_{(z^*\Delta\%)}$ by Equation 8.3.

Step 7: Compute $u_{(z^*\Delta\%)}$ by Equation 8.2.

Step 8: If $\left\| u^j_{(z^*\Delta\%)} - u^{j-1}_{(z^*\Delta\%)} \right\| \geq \epsilon$, then go to Step 5.(2) Else if $z \leq n_s$, then select another $X_{(\Delta\%)}$ and merge it with the current $X_{(z^*\Delta\%)}$. Set $z = z + 1$. Go to Step 2. (3) Else move to phase 2.

Phase 2: This is the main structure of the whole algorithm due to the computation issue.

Step 9: Initialize u using the results from phase 1, $s_{(ns^*\Delta\%)}$, with Equation 8.3 for the full data set.

Step 10: Follow steps 3–4 of the classical FCM algorithm and satisfy the stopping criterion ϵ_2.

demonstrate the performance of this approach, we use the same image as illustrated in Figure 8.1. This new proposal requires less computational time; see Figure 8.2. Similar platforms, as variants of this multistage random-sampling FCM framework, can be found in Eschrich and others [12] and Kolen and Hutcheson [19].

The success of this development verifies that the original FCM can still be improved in terms of efficiency of convergence. However, is this the only means to deal with the inefficiency problem? In the next subsection, some different strategies are attempted.

8.3.3 Fast generalized FCM scheme

Ahmed and colleagues [2] introduced an alternative to the classical FCM by adding a term that enables the labeling of a pixel to be associated with its neighborhood. As a regulator, the neighborhood term can change the solution to piecewise, homogeneous labeling.

As a further extension of Ahmed's work, Szilágyi and colleagues [28] proposed the well-known EnFCM algorithm to speed up the segmentation process for black-and-white images. In order to reduce the computational complexity, a linearity-weighted sum image g is formed from the original image, and the

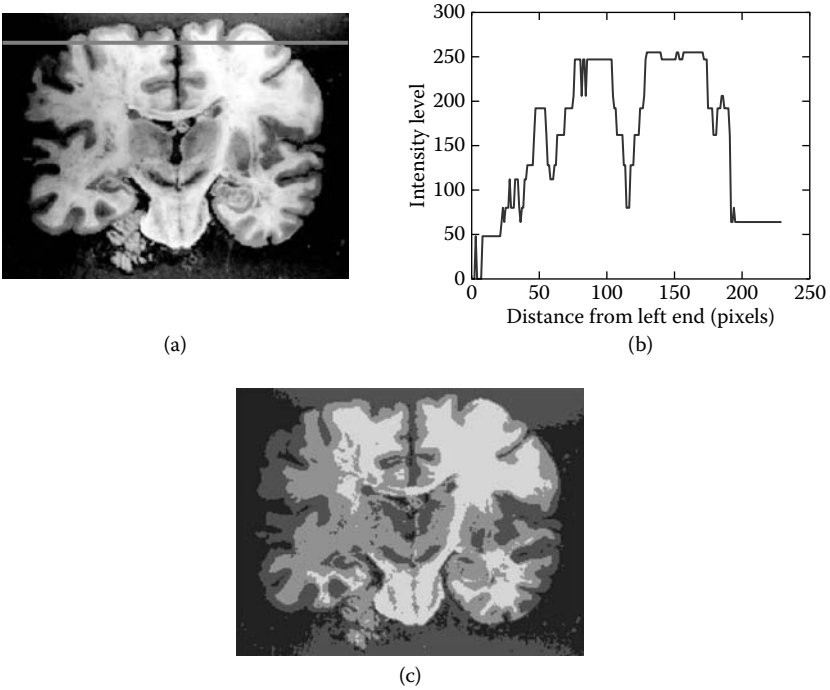


FIGURE 8.1: Classical FCM segmentation: (a) original image superimposed by a sample line, (b) intensity profile of the sample line, and (c) FCM segmentation result, where the cluster number is 4. Note that image size is 230×175. (From A. Toga, K. Ambach, B. Quinn, M. Hutchin, J. Burton, 1994. *J. Neurosci. Meth.* 54: 239–252. With permission.)

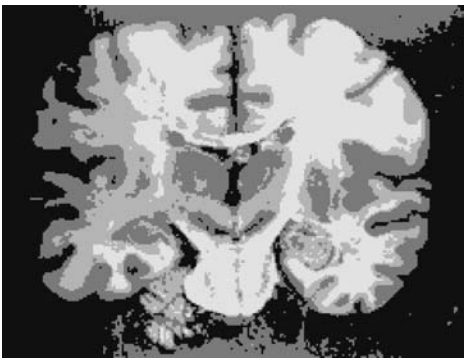


FIGURE 8.2: Multistage random-sampling FCM segmentation. The computational time was reduced to 93.7 seconds.

local neighbor average image regarding

$$g_m = \frac{1}{1+\alpha}\left(x_m + \frac{\alpha}{N_R}\sum_{j\in N_r} x_j\right) \tag{8.4}$$

where g_m denotes the gray value of the mth pixel of the image g, x_j represents the neighbors of x_m, N_R is the cardinality of a cluster, and N_r stands for the set of neighbors falling into a window around x_m and $\sum_{j\in N_r}$. The object function used for fast segmenting image g is defined as

$$J = \sum_{i=1}^{C}\sum_{i=1}^{q_c}\gamma_l u_{ij}^m (g_l - s_i)^2 \tag{8.5}$$

where q_c denotes the number of the gray image, and γ_l is the number of the pixels having the intensity equal to l, where $l = 1, 2, \ldots, q_c$. Then, we should hold

$$\sum_{l=1}^{q_c}\gamma_l = N \tag{8.6}$$

This is under the constraint that $\sum_{i=1}^{C} u_{ij} = 1$ for any l. Finally, we can obtain the following expression [4]:

$$u_{il} = \frac{(g_l - s_i)^{-2/m-1}}{\sum_{j=1}^{C}(g_l - s_j)^{-2/m-1}} \tag{8.7}$$

and

$$s_i = \frac{\sum_{l=1}^{q_c}\gamma_l u_{il}^m g_l}{\sum_{l=1}^{q_c}\gamma_l u_{il}^m} \tag{8.8}$$

In fact, EnFCM is different from other FCM schemes in the areas where the former considers a number of pixels with similar intensity as a weight. Thus, this process may accelerate the convergence of searching for global similarity. On the other hand, to avoid image blur during the segmentation, which may lead to inaccurate clustering/grouping, Cai, Chen, and Zhang [4] introduced a novel factor S_{ij}, which incorporates the local spatial relationship S_{ij}^s and the local gray-level relationship S_{ij}^g. This factor replaces the parameter α in Equation 8.4:

$$S_{ij} = \begin{cases} S_{ij}^s \times S_{ij}^g, & j \neq i \\ 0, & j = i \end{cases} \tag{8.9}$$

Here, the following equation is held:

$$S_{ij}^{s} = \exp\left(\frac{-\max(|p_{cj} - p_{ci}|, |q_{cj} - q_{ci}|)}{\lambda_s}\right) \tag{8.10}$$

where (p_{ci}, q_{ci}) is a spatial coordinate of the ith pixel, and λ_s represents the scale factor of the spread of S_{ij}^{s}.

$$S_{ij}^{g} = \exp\left(\frac{-||x_i - x_j||^2}{\lambda_g \times \sigma_{gi}^2}\right) \tag{8.11}$$

where σ_g is a global scale factor of the spread of S_{ij}^{s}, and λ_g is similar to λ_s.

Hence, the new generated image g is updated as

$$g_i = \frac{\sum_{j \in N_i} S_{ij} x_j}{S_{ij}} \tag{8.12}$$

where g_i is the gray value of the ith pixel of the image g, x_j is the gray value of the neighbors of x_i (window center), N_i is the set of neighbors falling in the local window, and S_{ij} is the local similarity measure between the ith pixel and the jth pixel. The value of g_i is restricted to [0, 255] due to the denominator.

In summary, the fast generalized FCM algorithm [4] is illustrated in Algorithm 8.3, where steps 4 and 5 are not terminated until the following criterion is met: $|s_{new} - s_{old}| < \epsilon$ and $s = [s_1, \ldots, s_C]$ are the vectors of cluster prototypes.

The performance of this fast generalized FCM approach is demonstrated in Figure 8.3, which shows more computational time consumption than that of the multistage random-sampling FCM in this example. This time consumption most likely is attributable to the similarity measures undertaken during the iterations. On the other hand, the image segmentation outcomes by the fast generalized FCM are close to the results by the classical and the multistage random-sampling FCMs (in Section 8.5, we provide a metric to measure the similarity of the segmentation). Similar efforts to improve the computational efficiency and robustness have also been reported by Chen and Zhang [5] and Leski [20].

Algorithm 8.3 The fast generalized FCM algorithm.

Step 1: (a) Set the cluster number C to be $C_{\max}$ (by a priori knowledge); (b) initialize randomly the parameters and set $\epsilon > 0$ to be a small value.

Step 2: Compute the local similarity measures S_{ij} using Equation 8.9 for all neighbor windows over the image.

Step 3: Compute linearly weighted summed image g via Equation 8.12.

Step 4: Update the partition matrix using Equation 8.7.

Step 5: Update the prototypes using Equation 8.8.

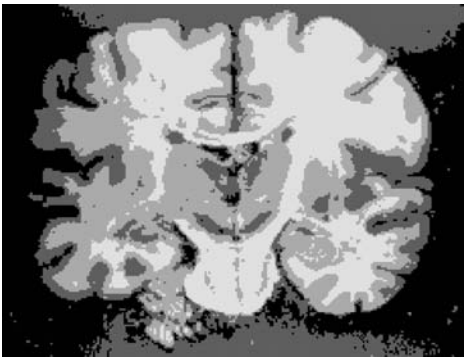


FIGURE 8.3: Fast generalized FCM segmentation. The computational time was 95.5 seconds.

8.4 Proposed FCM-Based Scheme

The state-of-the-art techniques have exhibited effectiveness in dealing with the compromise between efficiency and robustness. Is there any potential to further enhance the performance of these existing algorithms? In this section, we look at the possibility of applying a mean shift strategy to further improve the computational efficiency while keeping segmentation accuracy.

8.4.1 Anisotropic mean shift with FCM

8.4.1.1 Preliminaries

Mean shift-based techniques can estimate the local density gradients of similar pixels. These gradient estimates are iteratively performed so that all pixels can find similar pixels in the same image [9,10]. A standard mean shift method uses radially symmetric kernels. Unfortunately, the temporal coherence is reduced in the presence of irregular structures and noise buried in the image [30]. This reduced coherence may not be properly detected by the radially symmetric kernels. Thus, an improved mean shift platform, the anisotropic kernel mean shift, is under investigation.

We start with a brief description of the classical mean shift methodology: the image points are iteratively moved forward along the gradient of the density function before they become stationary. Those points gathering in an outlined area are treated as the members of the same segment. A kernel density estimate is hence defined by

$$\tilde{f}(x) = \frac{1}{n} \sum_{i=1}^{n} K(x - x_i) \tag{8.13}$$

where n is the point number, x_i stands for a sample from unknown density f, and

$$K(x) = |H|^{-0.5} K(H^{-0.5} x) \tag{8.14}$$

where $K(\cdot)$ is the d-variate kernel function with compact support satisfying the regularity constraints, and H is a symmetric positive definite $d \times d$ bandwidth matrix. Usually, we have $K(x) = k_e(\phi)$, where $k_e(\phi)$ is a convex decreasing function (e.g., for a Gaussian kernel)

$$k_e(\phi) = c_t e^{-\phi/2} \tag{8.15}$$

and for an Epanechnikov kernel,

$$k_e(\phi) = c_t \max(1 - \phi, 0) \tag{8.16}$$

where c_t is a normalizing constant. If a single global spherical bandwidth is applied, $H = h^2\mathbf{I}$ ($\mathbf{I}$ is identity matrix), then we have

$$\tilde{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{8.17}$$

Since the kernel can be divided into two different radially symmetric kernels, we have the kernel density estimate as follows:

$$\tilde{f}(x) = \sum_{i=1}^{n} \frac{1}{n(h^\alpha)^p (h^\beta)^q} k^\alpha\left(||(x^\alpha - x_i^\alpha)/h^\alpha||^2\right) k^\beta\left(||(x^\beta - x_i^\beta)/h^\beta||^2\right) \tag{8.18}$$

where p and q are two ratios, and α and β denote the spatial and temporal components, respectively [30].

8.4.1.2 Anisotropic mean shift segmentation

The mean shift utilizes symmetric kernels that may experience a lack of temporal coherence in the regions where the intensity gradients exist with a slope relative to the evolving segment. The anisotropic kernel mean shift links with every data point by an anisotropic kernel. This kernel, associated with a pixel, can update its shape, scale, and orientation. The density estimator is represented by

$$\tilde{f}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h^\beta (H_i^\alpha)^q} k^\alpha(d(x^\alpha, x_i^\alpha, H_i^\alpha)) k^\beta\left(||(x^\beta - x_i^\beta)/(h^\beta(H_i^\alpha))||^2\right) \tag{8.19}$$

where $d(x^\alpha, x_i^\alpha, H_i^\alpha)$ is the Mahalanobis metric in space:

$$d(x^\alpha, x_i^\alpha, H_i^\alpha) = (x_i^\alpha - x^\alpha)^T H_i^{\alpha-1} (x_i^\alpha - x^\alpha) \tag{8.20}$$

Anisotropic mean shift is intended to modulate the kernels during the mean shift procedure. The target is to keep reducing the Mahalanobis distance so as to group similar points as much as possible. First, the anisotropic bandwidth matrix H_i^{α} is estimated with the following constraints:

$$\begin{cases} k_e^{\alpha}\left(d\left(x,\, x_i,\, H_i^{\alpha}\right)\right) < 1 \\ k_e^{\beta}\left(\left|\left|\left(x - x_i\right)/h^{\beta}(H_i^{\alpha})\right|\right|^2\right) < 1 \end{cases} \tag{8.21}$$

The bandwidth matrix is decomposed to be

$$H_i^{\alpha} = \lambda V A V^T \tag{8.22}$$

where λ is a scalar, V is a matrix of normalized eigenvectors, and A is a diagonal matrix of eigenvalues whose diagonal elements a_i satisfy

$$\prod_{i=1}^{p} a_i = 1 \tag{8.23}$$

The bandwidth matrix is updated by adding more and more points to the computational list: if these points are similar in intensity or color, then the Mahalanobis distance by Equation 8.20 will be consistently reduced. Otherwise, the Mahalanobis distance is increased, and these points will not be considered in the computation.

8.4.1.3 Proposed strategy

The combination of FCM and anisotropic mean shift segmentation starts from the former stage. This process allows the latter stage to have prior knowledge (point position, neighborhood, etc.) before the latter is activated. The significant difference between this proposal and other similar methods is that our approach continuously inherits and updates the states, based on the mutual correction of FCM and mean shift. The entire proposed algorithm is depicted in Algorithm 8.4. One example of the proposed mean shift–based FCM algorithm is shown in Figure 8.4.

8.5 Experimental Results

To evaluate the effectiveness of the proposed algorithm in comparison with the state-of-the-art, we conducted several experiments using publicly accessed images. The entire evaluation is divided into three steps: (a) subjective comparison of image segmentation; (b) numerical statistics of computational time consumption; and (c) numerical statistics of similarity between the segmented images. The whole implementation (MATLAB coding) was run on a PC with a 1.5 GHz Intel Pentium CPU.

Algorithm 8.4 The fusion of FCM and anisotropic mean shift algorithms.

Step 1: Given the cluster number by observations, randomly initialize FCM parameters (e.g., point position and weights).
Step 2: Initialize s and record the cluster centers; set $j = 0$.
Step 3: Initialize the fuzzy partition matrix $u(j = 0)$ with Equation 8.2.
Step 4: Increment $j = j + 1$; compute s by using Equation 8.3.
Step 5: Compute $u(j)$ by using Equation 8.2.
Step 6: For each point x_i, determine anisotropic kernel and related color radius by Equations 8.21 and 8.22; note that mean shift works on the outcome image of FCM.
Step 7: Calculate the mean shift vector and then update the mean shift until the mean shift, $M^+(x_i) - M^-(x_i)$, is less than a pixel considering the previous position and a normalized position change:

$$M^+(x_i) = \nu M^-(x_i) + (1-\nu) \times \frac{\sum_{j=1}^{n}(x_j - M^-(x_i))||(M^-(x_i^\beta) - x_j^\beta)/(h^\beta H_j^\alpha)||^2}{\sum_{j=1}^{n}||(M^-(x_i^\beta) - x_j^\beta)/(h^\beta H_j^\alpha)||^2}$$

where $\nu = 0.5$.
Step 8: Merge pixels with similar colors.
Step 9: Repeat steps 4–8.

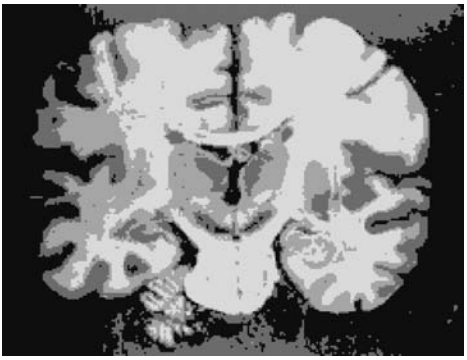


FIGURE 8.4: Mean shift FCM segmentation. The computational time was 92.1 seconds.

The experimental group consists of four color images (see Figure 8.5): heart, head 1, head 2, and lesion. The resolution of each image is tabulated in Table 8.1. The image shown in Figure 8.1 is here named as "image 5." Four segmentation algorithms are applied to the tests: (a) original FCM ("FCM"); (b) multistage random FCM ("MRFCM"); (c) fast generalized FCM ("FGFCM"); and (d) our new approach ("MSFCM"). Before running any

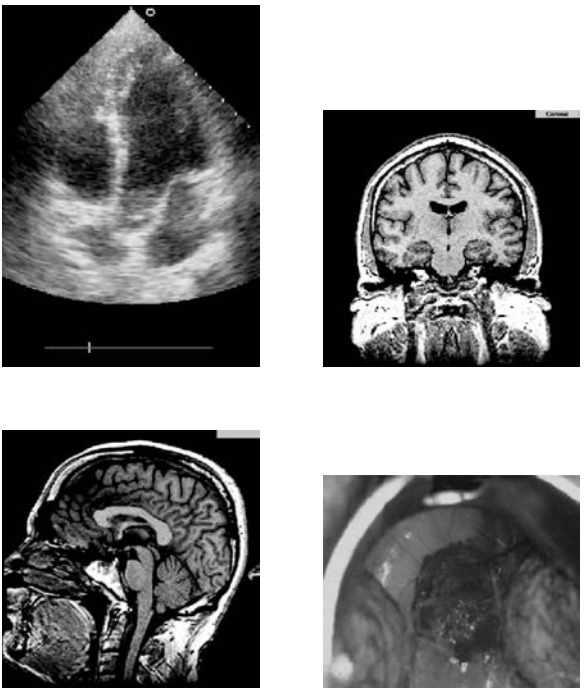


FIGURE 8.5: Original images used in this entire evaluation.

TABLE 8.1: Computational time of different approaches for image segmentation (unit: seconds).

Types	Image Resolution (Pixels)	FCM	MRFCM	FGFCM	MSFCM	Clustering Number
Image 1	240×339	416.8	389.9	402.4	364.4	4
Image 2	256×256	266.5	241.9	247.7	214.4	4
Image 3	300×300	511.8	492.4	500.6	465.8	4
Image 4	175×212	69.8	66.0	68.6	64.9	3
Image 5	230×175	99.4	93.7	95.5	92.1	4

FCM algorithm onto an image, we need to define a clustering number for this specific image. Normally, the criterion to determine this clustering number is based on possible grouping of similar intensity or color, which can be directly observed from the image.

To evaluate the engaged computational time of each algorithm, we calculate the time period between the start and end of an entire segmentation process. In order to obtain a valid comparison of the used computational time, we expect that different algorithms should end up with the same or approximate segmentation outcomes. If this condition cannot be satisfied, the

analysis of the computational time becomes meaningless. For this reason, we utilize Shannon entropy, or information entropy, which as an index reflects image contents. The Shannon entropy of a discrete random variable Z that has values $z_1, \ldots, z_2$ is

$$H(Z) = E(I(Z)) = -\sum_{i=1}^{n} p(z_i) \log_2 p(z_i) \tag{8.24}$$

where $I(Z)$ is the information content of Z, and $p(z_i)$ is the probability mass function of Z. If the entropy values of two images are close enough, these images must contain approximate contents.

8.5.1 Segmentation comparison

In this subsection, we mainly focus on evaluating the outcomes of segmentation by the four algorithms. By directly observing the outcomes of image segmentation, we can report whether or not the investigated images have been properly segmented. This test is carried out through the subjective judgment of a professional expert.

In Figure 8.6, we attempt to separate four heart chambers from the background in an ultrasonic image. The intensity of the heart wall, to some extent, is similar to that of the background. This may result in difficult discrimination in these regions. Figure 8.6a–d show that the overall algorithms are able to correctly separate the two upper chambers, and MSFCM clearly outlines the two lower chambers with homogeneous surroundings due to the gradient computation in the anisotropic mean shift. However, the other three algorithms end up with transitive regions between the chambers and individual heart walls. This uncertainty is unexpected and needs to be reduced.

The original image of Figure 8.7 exhibits a magnetic resonance head image: due to the complexity of the head structure, it is a very difficult task to discriminate among gray/white matter, tissue layers, and surroundings. In Figure 8.7, only MSFCM enables the separation of white matter from gray matter. The other algorithms cannot do this, indicating MSFCM shows the best clustering performance in this example. The failure of the three algorithms results from the developing segmentation with no appropriate spatial restriction.

The challenge in Figure 8.8 is that there are so many structures that may strongly affect the segmenting performance. These structures consist of white/gray matter, tissues, bones, and muscles. Similar to Figure 8.7, MSFCM permits the gray/white matter to be separated. However, the other three algorithms fail to do so. Interestingly, individual structures in the cemporal of the head image have been clearly segmented by MSFCM, while the other approaches fail to achieve this result. The possible cause is the same as that of Figure 8.7. Figure 8.9 presents the segmentation results of a lesion image. Individual subfigures show that the results of different algorithms are quite similar.

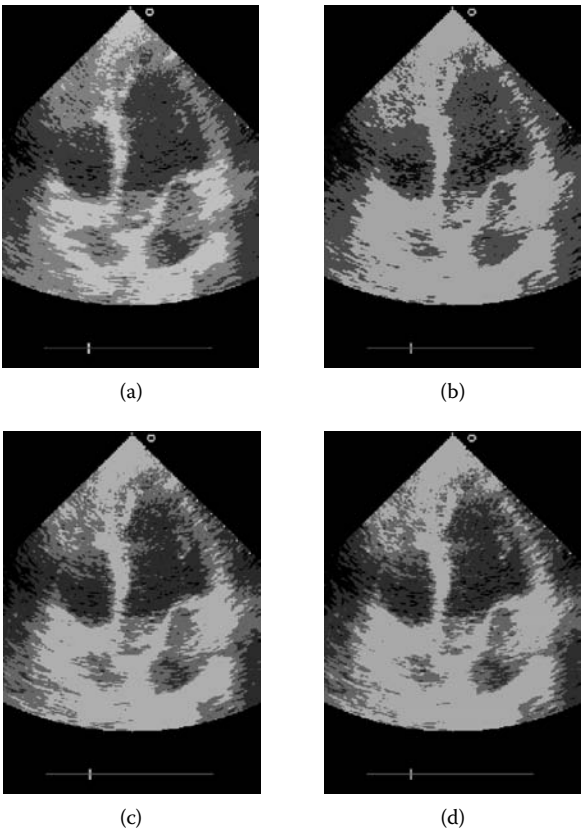


FIGURE 8.6: Image 1 (heart): performance comparison of different segmentation methods. (a) FCM, (b) MRFCM, (c) FGFCM, and (d) MSFCM. (Image courtesy of Huiyu Zhou [34].)

8.5.2 Computational time issues

Computational time is the major concern of this performance evaluation. We expect that our new segmentation algorithm not only results in a reasonably good clustering but also shows better performance efficiency. In this subsection, when interpreting the experimental results, we also refer to statistics of nonnormalized Shannon entropy for individual images. Statistics of the time consumption are obtained by calculating the mean value of eight estimates for each image.

In Table 8.1, we observe that the proposed algorithm (MSFCM) always has the least computational time. The success of our algorithm in efficiency is due to the interaction between FCM and mean shift, leading to fast convergence in local energy search: the output of FCM brings a better prediction for mean shift, and vice versa. We also discover that the three improved FCM

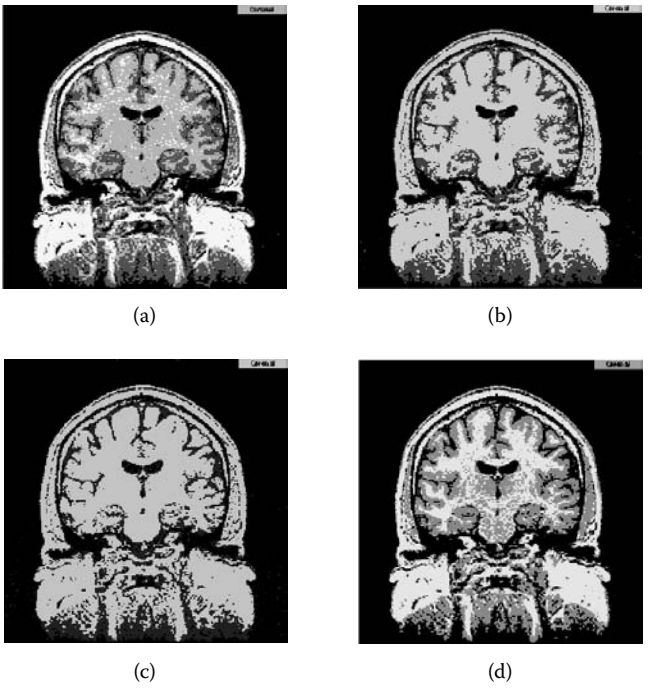


FIGURE 8.7: Image 2 (head 1): performance comparison of different segmentation methods. (a) FCM, (b) MRFCM, (c) FGFCM, and (d) MSFCM. (Image courtesy of Lauren O'Donnell at MIT, USA.)

schemes normally have faster convergence than the original FCM. Furthermore, Table 8.1 shows that the larger the image size, the more computational time is spent.

Now, we take a look at the Shannon entropy. This parameter can be used to evaluate the similarity of two images. The purpose of using this parameter here is to provide a criterion to justify the validity of the above analysis. Figure 8.10 clearly shows that the averaged relative errors between the segmented images by any two algorithms are less than 10%. This indicates that for each testing group (e.g., "image 1" group), the necessary condition (mentioned in the beginning of this experimental section) is fully satisfied. That is, although different algorithms take different computational time, they actually result in similar segmentation consequences.

Finally, we conduct efficiency tests at different clustering numbers. These tests target finding the relationship between the number of clustering and the computational time. For demonstration purpose, we study the engaged time consumption of images 2 and 3. The other images hold similar results. The used computational time in segmenting image 2 is denoted in Figure 8.11, where we observe that (a) the more clusters placed, the more time

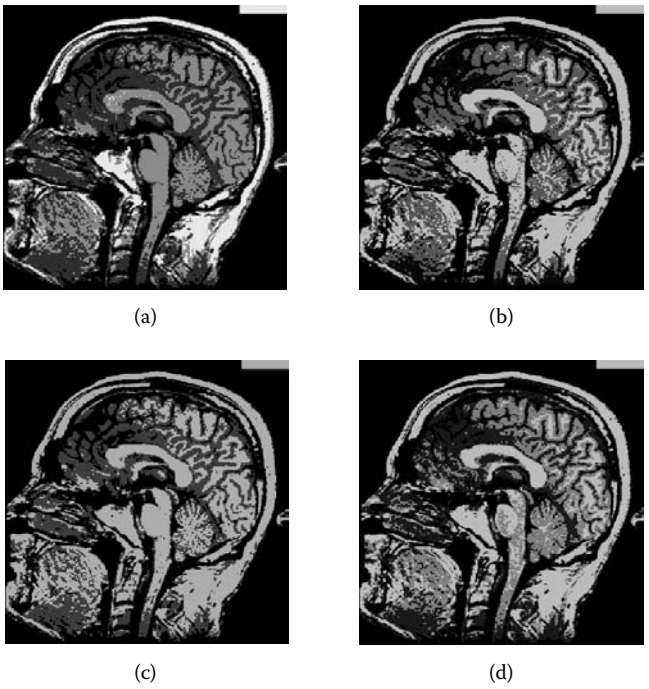


FIGURE 8.8: Image 3 (head 2): performance comparison of different segmentation methods. (a) FCM, (b) MRFCM, (c) FGFCM, and (d) MSFCM. (Image courtesy of Lauren O'Donnell at MIT, USA.)

consumption for the overall algorithms is required (not significantly though), and (b) the proposed MSFCM method has the shortest processing time among these four algorithms. Similar observations can be found in Figure 8.12.

8.6 Discussion

Efficiency issues in image segmentation by FCM strategies are discussed in this chapter. The literature has shown that classical FCM algorithms suffered from the inefficiency problem [5,25]. To enhance the efficiency of the original FCM, approaches have been established, which to some degree demonstrate promising progress. These methods include EnFCM [28], fast generalized FCM [4], and multistage random-sampling FCM [6]. Despite some success, these established approaches still demonstrate unstable efficiency performance (e.g., Figure 8.11).

The multistage random-sampling FCM starts with random sampling of a small subset of the full image. This step is intended to explore the cluster

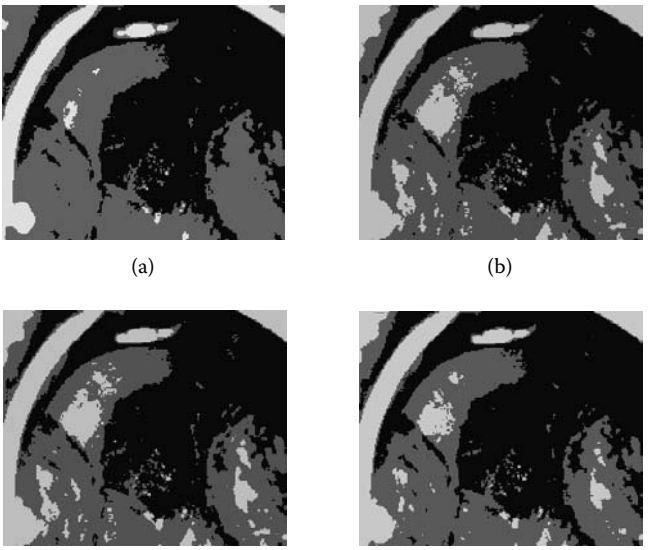


FIGURE 8.9: Image 4 (lesion): performance comparison of different segmentation methods. (a) FCM, (b) MRFCM, (c) FGFCM, and (d) MSFCM. (Original image courtesy of STI-Medical Systems, USA.)

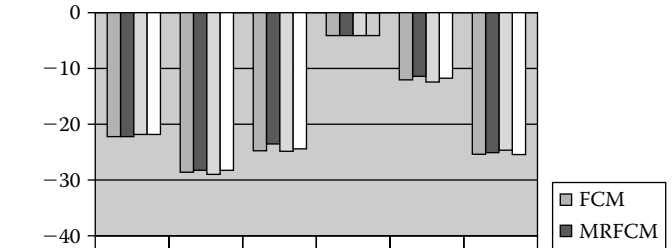


	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6
FCM	−22.35	−28.65	−24.73	−4.11	−11.95	−25.35
MRFCM	−22.29	−28.23	−23.57	−4.06	−11.31	−25.25
FGFCM	−21.96	−28.86	−24.55	−4.11	−12.16	−24.89
MSFCM	−21.8	−28.15	−24.16	−4.09	−11.64	−25.3

FIGURE 8.10: Nonnormalized Shannon entropy of different segmented images. Horizontal: images; vertical: Shannon entropy ($\times 10^9$).

centers of the entire image. Ideally, this subsampling scheme enables us to maximally reduce the converging iterations if the sampled subset contains characteristics similar to those of the entire image. Nevertheless, this

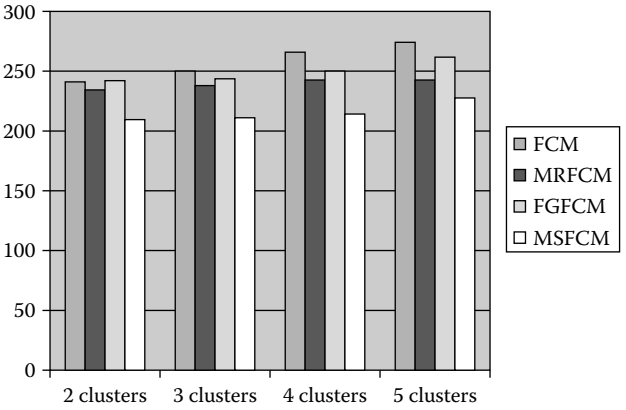


FIGURE 8.11: Computational time of image 2 against different clustering numbers. Horizontal: cluster numbers; vertical: time (seconds).

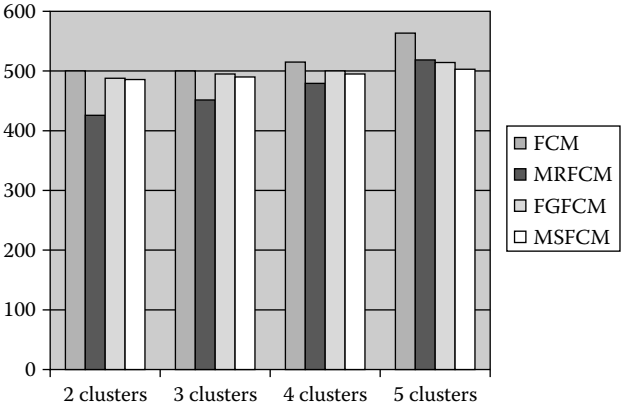


FIGURE 8.12: Computational time of image 3 against different clustering numbers. Horizontal: cluster numbers; vertical: time (seconds).

requirement cannot be fully guaranteed in a complex image with different structures or clusters. Consequently, the prior subsampling may involve an incomplete learning process, resulting in a lengthy segmentation time for the entire image. For example, the result of four or five clusters in Figure 8.11 shows that FCM has the largest time consumption compared to the other methodologies. This is partly due to the complicated structures (i.e., clutters) in image 2, and hence the algorithm makes much more effort to discriminate between the neighboring intensity and color regions of more details.

On the other hand, EnFCM or fast generalized FCM techniques add a term to the original clustering function so as to associate a pixel with its

neighborhood. To consider the effect of the neighbors, a window must be defined beforehand. Determination of the window size is image dependent. This parameter may affect the final outcomes of segmentation efficiency and accuracy. For example, Table 8.1 reveals that FGFCM is usually faster than the classical FCM but slower than MRFCM. Possible reasons for this consequence are that (a) the small sampling strategy in MRFCM helps faster convergence than FCM and FGFCM, and (b) the window size in FGFCM may not be well set up, so this algorithm causes lengthy computation time.

Comparatively, the proposed mean shift FCM has held the least computation needs in segmentation (e.g., Table 8.1). This method leads to similar segmentation outcomes as others (e.g., Figures 8.1 through 8.4). Similarity tests shown in Figure 8.10 validate the performance of the proposed algorithm in different testing images. The new scheme is successful because the anisotropic kernel used allows us to dynamically update the state parameters and achieve fast convolution by the anisotropic kernel function. In spite of its superiority, the proposed segmentation algorithm only exhibited small enhancements in some circumstances (e.g., the results by MRFCM and MSFCM shown in Table 8.1). Possibly, the proposed MSFCM algorithm experienced "oscillation" in the areas where rich similarity of intensity or color confused the evolution of the kernel. As a result, the FCM failed to efficiently converge. Unfortunately, such questions have not yet been appropriately answered. One of our interests is to study in the future how this latency is developed and how we can best handle this situation.

8.7 Future Research Directions

Computational intelligence has demonstrated its successful achievements in solving complex problems. As the applications of computational intelligence methods become more widespread, researchers and practitioners pay more attention to the efficiency and automation of the engaged implementation. As a particular example, this chapter presented several FCM methodologies for segmentation of medical images. This study is motivated by the fact that the iteration process underlying a similarity search may be reduced so that the computational efficiency of the entire implementation will be enhanced. This technical success may lead to a number of opportunities for improvement in medical diagnosis.

First of all, as the computation becomes more efficient, clinical diagnosis can be achieved faster than with the established process. As image segmentation proceeds faster, we can integrate more complicated algorithms, targeting performance improvement of segmentation in accuracy and consistency. Once a system becomes stable with a high accuracy, we can start to explore an automatic segmentation scheme based on currently available techniques.

For example, research can be directed toward automatic selection of stopping criteria.

Second, the success of the integration of FCM and an isotropic mean shift scheme suggests that a further investigation on the characteristics of the isotropic mean shift is necessary. This chapter demonstrated that using this mean shift scheme, we can improve the convergence speed. Further studies may consists of (a) whether or not this scheme can be used with other computation intelligence methods (e.g., neural networks, genetic algorithms); (b) further testing of the proposed approaches; (c) analysis of the computational complexity. These research topics will lead to fundamental discoveries in the fast image segmentation field.

In spite of its success, the FCM algorithm introduced in this chapter still faces relative challenges in computational efficiency. The experimental results show that there is still room for the computation to be accelerated. For example, Figures 8.11 and 8.12 illustrate that as the cluster numbers increase, the computation converges more slowly. A future study can be launched on how to keep the computation approximate no matter how many cluster numbers are applied. This is an important issue to be explored because a complex image may contain more than six clusters that may result in a huge amount of computation for segmentation.

8.8 Conclusions

In this chapter, we proposed a new approach to tackle the efficiency problem of the FCM algorithm while maintaining optimal performance in accuracy. The issue of improving efficiency is important because of the requirement of real-time computation in real applications. The proposed method is based on the idea that the FCM method enables faster evolution of mean shift points, and vice versa. In the proposed scheme, we adopt an improved FCM algorithm, incorporating a mean field term within the standard FCM objective function. This mean field is expected to minimize the effects of multiple structures or objects and hence accelerate the clustering. To derive this mean field, we utilize an established anisotropic mean shift method that uses a kernel function to convolve with the investigated image. Comparisons of the proposed scheme with three well-established segmentation methods showed that the proposed method is consistently faster than the other schemes.

Even though the proposed approach shows good reliability, it is limited by the fact that the current algorithm has not been able to determine the cluster number before any FCM practice is allowed. This might be resolved by extracting the most vivid and distinguishable colors [18]. These colors and the corresponding centers represent potential clusters to be formulated. In addition, we plan to extend the proposed method to three-dimensional image segmentation, which is of more value to the community.

Acknowledgment

The author is grateful to the anonymous reviewers for their constructive comments, which helped improve the quality of this manuscript.

References

[1] M. Ahmed, A. Farag. Two-stage neural network for volume segmentation of medical images. *Pattern Recogn. Lett.* 18 (11–13) (1997) 1143–1151.

[2] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, T. Moriaty. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Trans. Med. Imaging* 21 (2002) 193–199.

[3] J. Bexdek. A convergence theorem for the fuzzy isodata clustering algorithms. *IEEE Trans. Pattern Anal. Machine Intell.* 2 (1980) 1–8.

[4] W. Cai, S. Chen, D. Zhang. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recogn.* 40 (3) (2007) 825–838.

[5] S. Chen, D. Zhang. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Trans. on Systems, Man and Cybernetics*—Part B 34 (2004) 1907–1916.

[6] T. Cheng, D. Goldgof, L. Hall. Fast fuzzy clustering. *Fuzzy Sets and Systems* 93 (1998) 49–56.

[7] K. Chuang, S. Tzeng, H. Chen, J. Wu, T. Chen. Fuzzy c-means clustering with spatial information for image segmentation. *Comput. Med. Imag. and Graph.* 30 (2006) 9–15.

[8] D. Chun, H. Yang. Robust image segmentation using genetic algorithm with a fuzzy measure. *Pattern Recogn.* 29 (7) (1996) 1195–1211.

[9] D. Comaniciu, P. Meer. Mean shift analysis and applications. *IEEE Int. Conf. Computer Vision (ICCV'99)*, Kerkyra, Greece (1999) 1197–1203.

[10] D. Comaniciu, P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on PAMI* 24 (2002) 603–619.

[11] J. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *J. Cybern.* 3 (1973) 32–57.

[12] S. Eschrich, J. Ke, L. Hall, D. Goldgof. Fast accurate fuzzy clustering through data reduction. *IEEE Trans. Fuzzy Sys.* 11 (2003) 262–270.

[13] R. Hathaway, J. Bezdek. Fuzzy c-means clustering of incomplete data. *IEEE Trans. on Systems, Man and Cybernetics*–Part B 31 (2001) 735–744.

[14] R. Hu, L. Hathaway. On efficiency of optimization in fuzzy c-means. *Neural, Parallel & Scientific Computations* 10 (2002) 141–156.

[15] Z. Iscan, M. Kurnaz, Z. Dokur, T. Olmez. Ultrasound image segmentation by using wavelet transform and self-organizing neural network. *Neural Information Processing – Letters and Review* 10 (8–9) (2006) 183–191.

[16] N. Ito, R. Kamekura, Y. Shimazu. The combination of edge detection and region extraction in nonparametric color image segmentation. *Information Science* 92 (1996) 277–294.

[17] M. Kass, A. Witkin, D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision* 1 (1988) 321–331.

[18] D. Kim, K. Lee, D. Lee. A novel initialization scheme for the fuzzy c-means algorithm for color clustering. *Patter. Recog. Let.* 25 (2004) 227–237.

[19] J. Kolen, T. Hutcheson. Reducing the time complexity of the fuzzy c-means algorithm. *Comput. Med. Imag. and Graph.* 10 (2002) 263–267.

[20] J. Leski. Toward a robust fuzzy clustering, *Fuzzy Sets Sys.* 137 (2003) 215–233.

[21] U. Maulik, S. Bandyopadhyay. Genetic algorithm-based clustering technique. *Pattern Recogn.* 33 (9) (2000) 1455–1465.

[22] R. Nevatia. A color edge detector and its use in scene segmentation. *IEEE Trans. on Systems Man, and Cybernetics.* 7 (1977) 820–826.

[23] S. Olabarriaga, A. Smeulders. Anteraction in the segmentation of medical images: A survey. *Med. Imag. Anal.* 5 (2001) 127–142.

[24] N. Pal, S. Pal. A review on image segmentation techniques. *Pattern Recogn.* 26 (1993) 1277–1294.

[25] D. Pham. Fuzzy clustering with spatial constraints. In *International Conf. on Image Processing*, 2002.

[26] P. Sahoo, S. Soltani, A. Wong, Y. Chen. A survey of thresholding techniques. *Comp. Vision, Graphics, and Image Process* 41 (1988) 233–260.

[27] M. Singh, N. Ahuja. Regression based bandwidth selection for segmentation using parzen windows. In *Proc. IEEE Int'l Conference on Computer Vision*, 2003.

[28] L. Szilagyi, Z. Benyo, S. Szilagyii, H. Adam. MR train image segmentation using an enhanced fuzzy c-means algorithm. In *25th Annual Information Conf. of IEEE EMBS*, 2003.

[29] A. Toga, K. Ambach, B. Quinn, M. Hutchin, J. Burton. Postmortem anatomy from cryosectioned human brain. *J. Neurosci. Meth.* 54 (1994) 239–252.

[30] J. Wang, B. Thiesson, Y. Xu, M. Cohen. Image and video segmentation by anisotropic kernel mean shift. In *Proc. European Conference on Computer Vision*, 2004.

[31] S. Wang, D. Fu, M. Xu, D. Hu. Advanced fuzzy cellular neural network: Application to CT liver images. *Artif. Intell. Med.* 39 (1) (2007) 65–77.

[32] S. Wang, M. Xu. A new detection algorithm (NDA) based on fuzzy cellular neural networks for white blood cell detection. *IEEE Trans. Inf. Technol. Biomed.* 10 (1) (2006) 5–10.

[33] A. Wismüller, F. Vietze, J. Behrends, A. Meyer-Baese, M. Reiser, H. Ritter. Fully automated biomedical image segmentation by self-organized model adaptation. *Neural Netw.* 17 (8–9) (2004) 1327–1344.

[34] H. Zhou, T. Liu, F. Lin, Y. Pang, J. Wu, J. Wu. Towards efficient registration of medical images. *Comp. Med. Imag. and Grap.* 31 (6) (2007) 374–382.

Chapter 9

Image Informatics for Clinical and Preclinical Biomedical Analysis

Kenneth W. Tobin, Edward Chaum, Jens Gregor, Thomas P. Karnowski, Jeffery R. Price, and Jonathan Wall

Contents

Biomedical informatics is the study of the application of computational and statistical algorithms, data structures, and methods to improve communication, understanding, and management of biomedical information. Our objective in this chapter is to describe and demonstrate our research in the use of biomedical image databases, in both preclinical and clinical settings, to classify, predict, research, diagnose, and otherwise learn from the informational content encapsulated in historical image repositories. We detail our approach of describing image content in a Bayesian probabilistic framework to achieve learning from retrieved populations of similar images.

We use specific examples from two biomedical applications to describe anatomic segmentation, statistical feature generation and indexing, efficient retrieval architectures, and predictive results.

9.1 Introduction

Imaging performs an important role in the understanding of human disease through its application in the preclinical and clinical environments for both animal and human research. In the preclinical environment, high-resolution, high-throughput, multimodality in vivo imaging provides the ability to carry out nondestructive, longitudinal studies on large populations of animals that previously required animal sacrifice and painstaking dissection to accomplish. In the clinical environment, these imaging modes allow physicians and researchers to visualize the manifestations and consequence of disease. A result of advancing imaging capabilities is the generation of copious amounts of digital image and ancillary data regarding study animals, human patients, physiology, disease, treatments, and so on. We refer to this ancillary data as metadata. Most preclinical researchers today maintain this image and metadata in an ad hoc manner, distributed across a number of computers, on different media types, and in different physical locations. In the clinical environment, picture archiving and communication systems (PACS) are providing improved capabilities for managing image and patient data. For those who have dedicated PACS in their facilities, the largest data component in the system—the imagery—is only marginally indexed or structured for search and retrieval using simple associated text fields. PACS technology for indexing the information content of imagery has not kept pace [1–3]. Over time, these image data lose their informational power simply because they become irretrievable, limiting an investigator's ability to pursue research questions that leverage the historical value of the image repository.

To effectively tap into this historical value, it is necessary that the content of the imagery be described in such a way that it can be queried, referenced, searched, and otherwise accessed in a database environment. For our purposes, image content is defined as the intrinsic characteristics of the array of numbers that encapsulate a recorded scene from an imaging modality. Content represents the spatial and spectral relationships that exist between groups of pixels in a scene that are understood by the human observer to represent structure (e.g., anatomy), morphology (e.g., abnormal bone formation), texture (e.g., tissue type), size and contrast (e.g., benign vs. malignant tumor), and myriad other attributes of human interest and interpretation [4–6]. To describe image content, our goal is to produce an index, that is, a succinct sequence of numbers used to organize image content in a data structure for efficient future access. An index is an alternative representation of an image in an informatics system. It is derived from the representation of statistical features that are

extracted from the image and image regions. It provides a critical foundation on which the structure of association, search, and retrieval is built.

The subject of this chapter is the definition and creation of a structured, searchable environment that leverages the interpretable content of imagery. We refer to this content and its meaning and relationship to the various forms of nonimage data present as *biomedical informatics*. Biomedical informatics is the application of computational and statistical algorithms, data structures, and methods to improve communication, understanding, and management of biomedical information. There are many ways to formulate a biomedical informatics system that do not index image content. These systems and methods store and retrieve imagery, but only secondary to the text-based data associated with, for instance, a preclinical animal study or a clinical trial of human patients [2,7]. To describe and structure image content and its associated metadata, we use the methods and terminology of content-based image retrieval (CBIR). Because we are interested in an environment of information analysis—taking into account prior knowledge and taking advantage of the visual evidence of image content—we formulate our method in a framework of Bayesian probability.

CBIR refers to techniques used to index, retrieve, and manage images from large image repositories based on visual content, as described above. Many researchers today are applying CBIR to the general problem of image retrieval [5,8–10] and to the biological or biomedical fields [11–14]. In a CBIR environment, a user typically queries a data system through the use of an example image or a group of images representing an unknown state. The system ranks images in the archive according to their similarity to the query. When performing retrieval operations, we formulate the returned population of visually similar images as a posterior probability, $P(\omega_i|\boldsymbol{v}) \pm \sigma(\omega_i)$, where P represents the probability of the occurrence of a class ω_i given the visual evidence represented by a vector of statistical features represented by our index, $\boldsymbol{v}$. The second term, $\sigma(\omega_i)$, is a measure of the uncertainty associated with our estimate of the prevalence of these states in our system. The class, ω_i, refers to the "state of nature" [15] of the object of our search, with i referring to the number of states or classes, C (i.e., $i = 1, 2, \ldots, C$). For example, in the small animal study environment, ω_i may refer to the particular 3-D morphology associated with an organ, the appearance of a particular tissue cross-section, or the structure of trabecular bone. In the clinical environment, ω_i may refer to the state of disease present in the image or the level of stratification of that disease. It might also refer to the manifestation of disease such as the presence and type of lesions in retinal tissue or the characteristics of tumors in breast or lung tissue. The query may be formulated to determine answers to questions such as *Do I have other patients in a similar disease state? What are their manifestations? or How did they respond to treatment?* For our discussion, we estimate this uncertainty by assuming that events ω_i occur with a frequency defined by a Poisson distribution [16]; that is, we treat the relative occurrence of the state of disease or disease manifestation

represented in the archive as being discrete with the probability of the phenomenon being constant in time.

The concept of interpreting CBIR response in a probabilistic framework is not new. In essence, the goal is to estimate a posterior probability in a nearest-neighbor-measurement environment where the parameters of the statistical space are unknown; that is, we are working in a nonparametric statistical space without making assumptions on the underlying probability densities. We base our estimate of posterior probability, $P(\omega_i|v)$, on the extraction and analysis of k similar images retrieved from a large image archive in an L-norm metric space [15]. Others use Bayesian methods to estimate probability densities, $p(v|\omega_i)$, to incorporate user relevance feedback, and to indirectly estimate posterior probabilities [17,18]. Still others apply Bayesian methods to use text-based queries to expand on a small population of text-labeled images to access a larger population of visually similar unlabeled images [19] using image content or to apply Bayesian belief networks to integrate low-level and semantic features to achieve higher levels of image understanding [20]. Our method of direct posterior probability estimation provides a framework for gauging the specificity and sensitivity of the estimates we make regarding the state of nature in our archive. This estimate can be shown to improve as the archive increases in size.

In this chapter, we describe six basic elements required to produce a system for archiving and accessing biomedical information. Figure 9.1a–e represents the major design elements that must be considered in the creation of a biomedical informatics system whose goal is to achieve a particular application result supporting scientific research, medical diagnostics, training, teaching, and so on (i.e., Figure 9.1f). In Section 9.2, we review a subset of the many different modes of preclinical and clinical imaging that are used to examine animals and people with particular emphasis on the modes and applications associated with our previous experience. By understanding what is measured in an anatomic

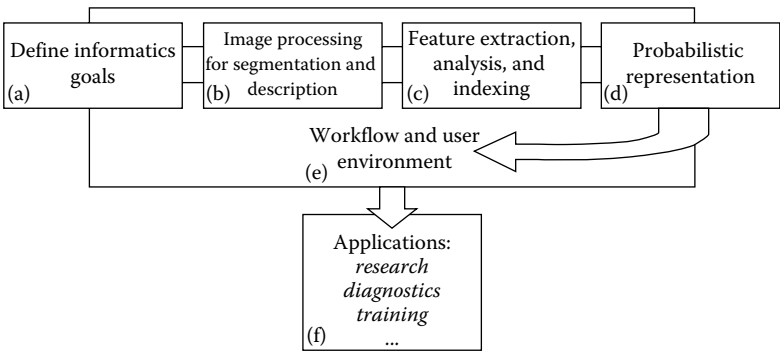


FIGURE 9.1: Simple representation of the five basic elements (a–e) required to design a biomedical informatics system (f) that leverages image content.

or functional imaging mode, we can begin to formulate the goals (Figure 9.1a) that our informatics system will be designed to address. For example, we may be interested in mining data across multiple studies, asking questions such as what frequency or in what strain of animal do we see a specified physiological effect, or we may be interested in discerning information regarding the state of disease that exists in a patient and determining associated manifestations of that disease and previous treatments. In Section 9.3 (Figure 9.1b), we review the process of image analysis for 2-D and 3-D region segmentation from imagery. Our treatment of this subject is fairly specific to the applications in which we have experience, but our current methods and results will shed light on the important considerations of anatomic segmentation. In this section, we also review methods for feature analysis to extract low-level statistical descriptions of image regions in terms of structure, texture, and morphology (Figure 9.1c). We also detail methods for mapping low-level features to a higher-level semantic index that can achieve user-specific responses. In Section 9.4, we detail the Bayesian probabilistic framework (Figure 9.1d). In Section 9.5, we describe high-level concepts related to the integration of the different components representative of Figure 9.1a–e to achieve a user interface and the implementation of workflow processes and to assure the orderly incorporation of data into the informatics system. In Section 9.6, we propose several important research topics that must be pursued to achieve the production of robust biomedical informatics systems. Finally, in Section 9.7, we briefly review some of the limitations and challenges that exist today and that drive future research in this area.

9.2 Background

The ability to noninvasively visualize a pathological process in vivo or to monitor the response to therapy is a fundamental tool in medicine. Functional and anatomic imaging techniques provide methods for monitoring whole-body pathology, including disseminated, subclinical, and preclinical lesions. Monitoring the response to therapy provides the physician with a way to assess the efficacy of treatment regimens, which can further aid in determining the optimal clinical management of the disease. Chemotherapies and immunotherapies are examples in which monitoring the clinical response can provide rapid benefits for the patient.

For this discussion, we present two areas of preclinical and clinical imaging associated with small animal research and human ocular disease, respectively. While these two areas are specific to our previous research, they generally raise the many considerations that must be evaluated and addressed when designing an informatics system—the "goals" shown in Figure 9.1a. In the clinical arena, high-resolution imaging is routine. With the advent of small animal (micro-) imaging platforms, similar capabilities are now available to

the researcher [12]. In the laboratory, in addition to magnetic resonance imaging (MRI), functional MRI (fMRI), computed tomography (CT) [21], functional CT (fCT) [22], ultrasound (US) [23], positron emission tomography (PET) [24], and single photon CT (SPECT) [25], the researcher can add digital optical imaging, optical coherence tomography [26], and scanning laser ophthalmoscopy [27] to the arsenal of available modalities. Each technique has specific advantages and problems associated with its use. Microimaging commonly is applied to the study of mice or rat subjects. For example, only certain SPECT imaging platforms can accommodate animals that are larger than a rat. In general, small animals are important models for genetics and disease studies. For example, Figure 9.2 shows a visualization of detailed skeletal structure in microCT imaging. The spine influences the structure and function of the spinal cord, heart, lungs, kidneys, and reproductive systems and is an important target of study.

Dual-modality imaging (i.e., the acquisition and registration of both high-resolution anatomic images such as CT or MRI as well as functional PET or SPECT image data on the same subject) provides the greatest information content. Not only can the anatomic dataset be used to enhance the quality of the functional image and provide more quantitative output, but as discussed below, the images can be used to analyze, interpret, and segment the generally lower-resolution functional image content.

In the laboratory, preclinical imaging is widely used to evaluate the efficacy of novel diagnostic imaging agents before they transition into clinical use. For monitoring the action of novel therapeutics in animal models of human disease, imaging can provide a quantitative, noninvasive way to assess longitudinal

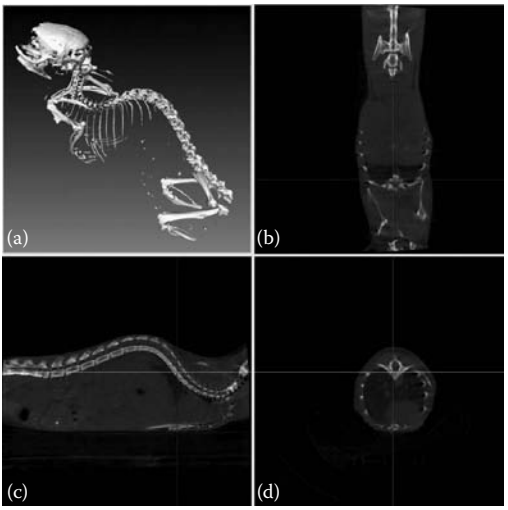


FIGURE 9.2: MicroCT images of mouse spinal and thoracic skeleton. (a) Skeletal visualization derived from volumetric CT data. Selected (b) coronal, (c) sagittal, and (d) axial views (planes) within the CT data volume.

responses in therapeutic trials [28–30]. An important stage in the development of any new pharmaceutical is to ascertain whether any undesirable, nonspecific accumulation occurs in healthy organs or tissues that may compromise the bioavailability, influence dosing, or result in toxicity. By incorporating any one of the positron-emitting isotopes ^{11}C, ^{13}N, or ^{15}O into the reagent of interest it is possible to monitor the uptake kinetics and biodistribution equilibrium in vivo and in real time [31–33].

In addition to the large volumes of image data that accompany a preclinical study, such as the interaction of a novel therapeutic antibody with a tumor antigen, it is prudent if not absolutely necessary to perform supporting studies on each research subject. Although the biodistribution of the radiolabeled tracers can be determined from coregistered functional and anatomic images, it is often difficult to delineate within specific tissues, organs, regions within organs, or vessels. This is particularly true in preclinical small animal CT imaging, which often lacks the functionality to perform the contrast-enhanced CT commonly used in patient care. In addition, with the exception of a few nuclides used in PET imaging, true quantitation of the concentration of the nuclide in the tissues cannot be determined [34–36]. It is therefore necessary to confirm the image-based data. This is achieved by harvesting tissue samples from experimental animals at necropsy and measuring the activity associated with each using a standard counter. The result is a measure of the percentage of injected dose per gram of tissue. In many cases, the number of tissues and organs sampled can be more than 15 or 20 per animal [32,37].

In certain instances, micro-autoradiography can be performed using tissue samples harvested postmortem. This is best performed when gamma-emitting nuclides such as ^{125}I are used for SPECT imaging, although we have used the dual positron- and gamma-emitting isotope ^{124}I with some success. After exposure of micrometer-thick tissue sections to photographic emulsion, the distribution of the radiotracer can be visualized within tissues and even specific cell populations, as evidenced by the presence of punctate silver granules exposed on the surface of the tissue. These data can be easily compared using confocal microscopy on histologically and immunohistochemically stained consecutive tissue sections so that the radiotracer microdistribution can be correlated with the pathology or organ substructure (e.g., vasculature under investigation) [32].

It is clear, therefore, that many options are available to the researcher wishing to perform noninvasive visualization of anatomic structures or the biodistribution of a radiolabeled molecule. Although each modality alone can provide superb image data, there are many advantages to performing dual-modality imaging such as SPECT/CT, PET/CT, and PET/MRI [38–40]. In addition, it is important in preclinical imaging studies that corroborative studies be performed to enhance and confirm interpretation of the findings based on the imagery. Consequently, the amount of data generated in the study of a single mouse can be vast and varied but ultimately is dependent upon imagery. As the use of small animal imaging in basic and biomedical science and drug development increases, it will become ever more important to develop efficient,

secure, and reliable methods to store, retrieve, and analyze the vast volumes of data that accumulate.

In the clinical environment, the application of imaging to the evaluation of ocular, particularly retina, anatomy and physiology, and the use of these methods to diagnose vision-threatening eye diseases is in common practice today. Technological advances in ocular imaging now provide quantifiable, digital data sets about refractive status, corneal and optic nerve topography, retinal thickness, and vitreoretinal anatomy, which permit the development of diagnostic algorithms that can be applied to specific biological and pathological processes. These methods have the potential to improve the sensitivity and specificity of image-based ocular diagnosis for early detection of eye disease and monitoring of therapeutic responses both for preclinical studies and standard medical care. Although the potential of these new technologies is clear, optimization and standardization of the imaging systems, statistical and clinical validation of novel image informatics methods, and network integration protocols for telemedical applications are required to realize their full potential. In particular, the validation of imaging technologies in prospective clinical trials for specific disease-oriented applications is a critical component to broader acceptance and use of these methods.

For example, the World Health Organization estimates that 135 million people have diabetes mellitus worldwide and that the number of people with diabetes will increase to 300 million by the year 2025 [41,42]. Visual disability and blindness have a profound socioeconomic impact on the diabetic population, and diabetic retinopathy (DR) is the leading cause of new blindness in working-age adults in the industrialized world [42]. It is estimated that as much as $167 million and 71,000 to 85,000 sight-years could be saved annually in the United States alone with improved screening methods for DR [43]. By 2025, it will be necessary to examine almost 1 million patients every day for DR screening alone. The implementation of inexpensive image informatics–screening programs for DR would have a significant impact on the economic costs and social consequences of vision loss from this disease.

The limitation of resources and access to health care providers, particularly to eye care specialists, are major obstacles to reducing the incidence of vision loss from diabetes, glaucoma, and other treatable eye diseases. Retinal photography and remote transmission of the examination results is a sensitive and specific method for identifying many retinal lesions, and a number of automated screening approaches have been reported using digitized photographs [44–47]. Some have shown good sensitivity in identifying individual lesions seen in diabetic retinopathy and high-specificity for the absence of disease. These studies have demonstrated that retinal photography is valuable for identifying vision-threatening retinal disease when performed by certified readers [29,48]. However, current reading centers use store-and-forward technology, are labor intensive, and do not provide real-time diagnostic capabilities. To screen large populations for common eye diseases, the use of image informatics that incorporates image content is required.

The challenge is to determine to what degree the examination can be automated to yield the desired accuracy, cost reduction, ease of use, and real-time diagnostic capability. While some small steps have been taken toward automated analysis of the complex features of retinal images, the difficulties of developing computer-based diagnostic algorithms from standard retinal photographs are many. Current unresolved issues in eye disease detection include minimum spatial resolution required to detect clinically significant pathology, minimum retinal area necessary to adequately screen for common blinding diseases, and which methods provide the best sensitivity–specificity ratio, to name just a few.

9.2.1 Preclinical imaging of amyloid

To study the progression and regression of amyloidosis in mice, we used microCT, SPECT, and PET imaging where the term *micro* refers to a small-bore, high-resolution platform specifically designed for small animal imaging. All three types of imaging modalities are volumetric (see Figure 9.3). X-ray CT is an *anatomic* modality that provides a detailed image of the skeletal structure and modest delineation of internal organs and fat tissue within the animal [50,51]. The limits of the visceral organs and vasculature can greatly be enhanced and differentiated by using contrast media such as highly iodinated triglyceride nanoparticles, such as Fenestra (Advanced Research Technologies, Inc., Quebec), which is typically administered intravenously (e.g., see contrast-enhanced vasculature in Figure 9.4). In addition, negative contrast can be achieved by intraperitoneal administration of aqueous iodine-based contrast agents such as those used clinically—Omnipaque™ or Iohexol™—which results in contrast-negative organs delineated by the highly attenuating peritoneal fluid.

SPECT and PET require the use of injected tracer molecules that are labeled with appropriate gamma- or positron-emitting radionuclides, respectively. These imaging techniques are considered *functional* modalities, meaning that the biodistribution and localization of the radiolabeled tracer is dependent on the live subject. To achieve quality images, both methods require concentration of the radiotracer within the target tissue or pathology relative to the background tissue and blood pool; target-to-background ratios of 3:1 are generally considered sufficient to discern specific uptake.

In the case of SPECT, much of our work is centered around ^{125}I-labeled serum amyloid P-component (SAP), which is a human serum protein related to C-reactive protein that binds specifically and essentially irreversibly to the pathogenic fibrils associated with all amyloid diseases [52]. We also use ^{125}I and the positron-emitting isotope ^{124}I-labeled antibodies to study the pathophysiological consequences of amyloid-related diseases [37,40].

Our group has been involved in the development of a dual-modality SPECT/CT microimaging system [32,40] that allows pseudocolored SPECT images to be coregistered and overlaid on an anatomic CT image (generally

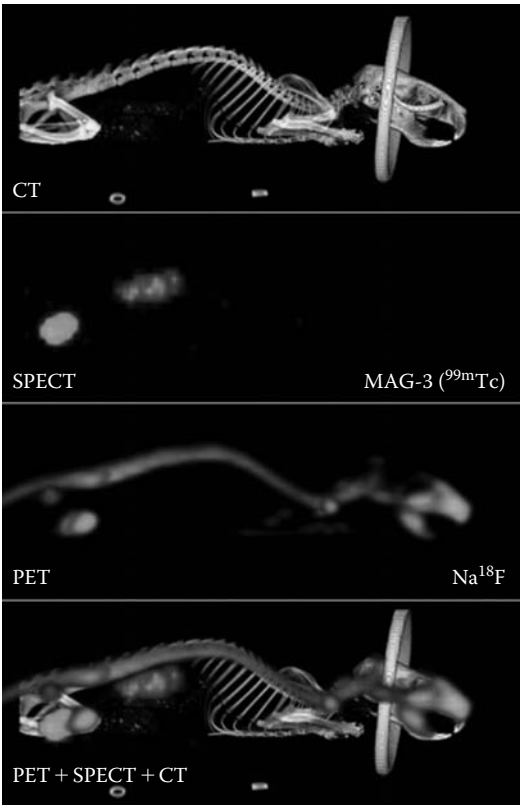


FIGURE 9.3: (See color insert following page 370.) Multimodality imaging using microCT, SPECT, and PET in an individual live mouse. Initially, renal images were acquired by SPECT using 250 μCi of MAG-3 (^{99m}Tc), followed by CT imaging. The mouse then received 100 μCi of Na^{18}F to visualize bone by PET imaging. Data were collected using the microCAT II + SPECT and microPET P4 scanners. The ring visible in the CT is the nose cone used to deliver 2% isoflurane anesthesia. Each modality was coregistered manually using Amira^{TM} software.

with contrast-enhanced vasculature) of the same mouse (see Figure 9.5). This has proven to be an invaluable tool in many ways. First, it allows us to precisely visualize which organs are laden with amyloid. That is, the contrast-enhanced CT data provides readily identifiable anatomic landmarks that enhance interpretation of the biodistribution of the SPECT radiotracer. Second, we can use the CT data to compute attenuation factors needed to make the SPECT and PET data quantitatively more accurate. Lastly, we can use the high-resolution CT data imagery to segment an organ, tissue, vessel, tumor, or other structure and use the resulting binary mask to extract the corresponding volumetric

FIGURE 9.4: Volumetric visualization of two separate mice showing contrast-enhanced CT using intravenously administered Fenestra^{TM} VC.

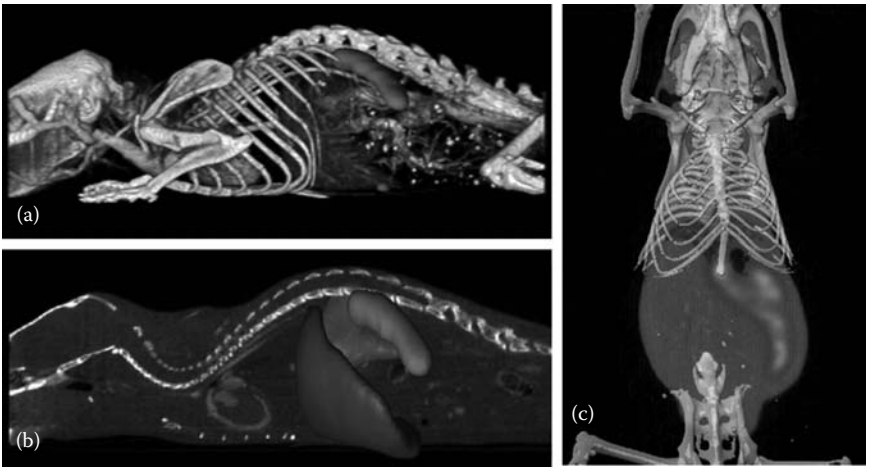


FIGURE 9.5: (See color insert following page 370.) Volume-rendered, contrast-enhanced SPECT/CT visualization of hepatosplenic amyloid imaged using ^{125}I-serum amyloid P-component. (a) Isosurface representation of the SPECT activity in the spleen with volumetric CT data. (b) Sagittal view of the same mouse with isosurface representation of the segmented spleen and liver volume overlaid. (c) A single coronal plane through fused SPECT/CT image of another mouse with splenic uptake of ^{125}I-serum amyloid P-component in which the skeletal structure (visualization) is overlaid for visual interpretation.

region of the SPECT data [53]. When the functional images are correctly calibrated and coregistered, this facilitates computerized organ- or tumor-specific quantitative analysis. That is, we can semiautomatically calculate the percentage of injected dose per unit volume for the segmented organ, which in turn provides a measure of the concentration of associated amyloid. The CT imagery can be similarly combined with a PET image, and if ^{18}F is used, the concentration of isotope can be directly calculated (in nCi/cc) from regions of interest (planar or volumetric) selected by the user. The benefits are the same as for SPECT/CT imaging, although more quantitative in nature.

Generally, we reconstruct high-resolution volumetric CT images with isotropic voxels on the order of 77 μm using a modified version of the Feldkamp algorithm, which applies to circular orbit cone-beam imaging [54]. With respect to SPECT, then we use an in-house developed ordered-subset expectation maximization (OSEM) algorithm [55]. While 1 and 2 mm pinhole collimators allow us to reconstruct pinhole images with isotropic ∼1 mm voxels, submillimeter voxel dimensions can be achieved using a 0.5 mm pinhole aperture [32]. This software has been integrated with the RVA^{TM} software implemented on the Siemens (Siemens Preclinical Solutions, Knoxville, TN) microCAT^{TM} family of imaging platforms from which the SPECT/CT images shown in this chapter were obtained. The microPET data shown, which have millimeter resolution, were obtained using the microPET^{TM} P4 large-axial field of view imaging system, also available from Siemens (Figure 9.3).

To achieve all the benefits mentioned above, accurate registration of the anatomic and functional images is critical. To this end, we use fiducial landmarks that are visible in both functional and anatomic imagery such as ^{57}Co- or ^{68}Ge-sealed sources. We then apply a simple, rigid transformation to achieve the required alignment [32].

9.2.2 Clinical imaging of the human retina

Image extraction algorithms have shown efficacy in automated screening of digital color and monochromatic retinal images for features of nonproliferative diabetic retinopathy such as exudates, hemorrhages, and microaneurysms [56–58] (see Figure 9.6). The use of computerized image extraction algorithms that rely on commonly used statistical methods and trained readers can yield substantial results in automated screening programs; however, the desired objectives of auto-analysis to date have not been adequately met. An ideal imaging method would not only screen for vision-threatening lesions in the central retina, including the macula and optic nerve, but would also detect and quantify the nature, location, and extent of retinal pathology in these critical locations and determine whether significant disease was present and the likely diagnosis. We believe that this is not only the optimal approach but is achievable using the methods outlined below.

The anatomy of the retina concentrates most of the cone photoreceptors that mediate fine visual acuity in a small central region called the fovea. These

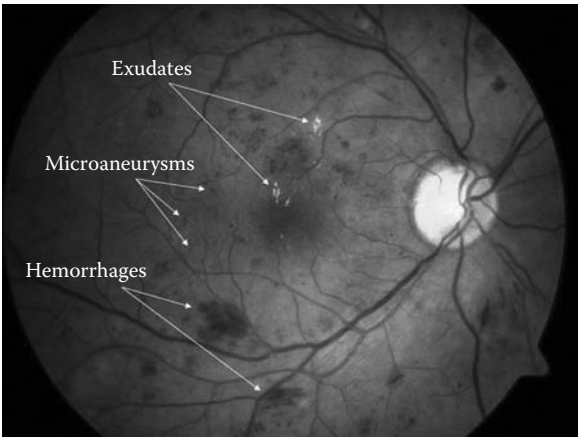


FIGURE 9.6: (See color insert following page 370.) Color fundus image showing manifestations of nonproliferative diabetic retinopathy including exudates, hemorrhages, and microaneurysms as indicated.

cells have high spatial sensitivity because there is very little convergence of input from cones to ganglion cells. Instead, the foveal cones are represented by about equal numbers of bipolar and ganglion cells. This results in a convergence ratio of nearly 1:1 and provides for precise information about the retinal areas being stimulated. It is in the fovea that all of the 20/20 vision of the retina exists. Conversely, spatial and temporal summation from rods in the peripheral retina is very high, permitting vision under conditions of dim illumination but with poor acuity. This unique anatomical organization means that most epidemiologically significant diseases that affect central visual acuity can be detected by changes in the small central region of the posterior retina, encompassing the central macula (about 2500 μm across) and the optic nerve.

Commercially available retinal cameras can take high-quality color and monochromatic images of the macula and posterior pole of the retina. Standardized fields of the posterior pole and periphery (seven fields) have been established for clinical evaluation of retinal diseases [59], but recently, more limited fields centered on the optic nerve and macula (with one peripheral field) have been used to detect retinal pathology in screening studies. Fluorescein angiography of the retinal vasculature generates a high-contrast image of the retinal structures, emphasizing the dual intravascular compartments of the inner and outer retina (see Figure 9.7). The tight junctions between adjacent retinal pigment epithelial cells and healthy, mature retinal capillary endothelial cells block the ingress of fluorescein dye from the choroid to the subretinal space, and from the retinal vessels into the neurosensory retina or vitreous, respectively, in the normal retina. The loss of this selective barrier effect as a result of disease underlies the value of fluorescein angiography as an

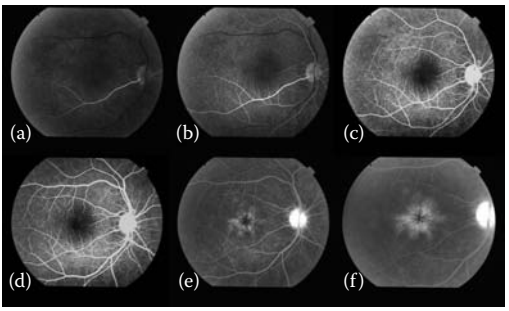


FIGURE 9.7: Fluorescein angiogram showing the choroidal (a) and retinal circulations (b, c, d). Inflammation of the macular vessels and optic nerve causes leakage of fluid with cystic changes in the retina and dye leakage from the optic nerve head (e, f).

imaging modality. Fluorescein (and indocyanine green) are useful diagnostic tools in the evaluation of the state of the blood-ocular barriers and provide data regarding disease severity, mechanisms, and response to treatment. It enhances the ability of the ophthalmologist to distinguish between overlapping tissues and to differentiate retinal, retinal vascular, pigment epithelial, and choroidal structures and diseases.

Optical coherence tomography (OCT) uses an 850 nm diode light source to generate high-resolution cross sections of retinal tissue based on the principle of low-coherence interferometry and spectral reflectance. OCT is particularly suitable for imaging retinal thickness and pathology with a resolution of about 10 μm. In humans, the retina is 150 to 300 μm thick; visually important pathology can therefore be detected by analysis of small cross sections through critical portions of the central retina. The data can be presented either as cross-sectional images or as topographic maps. False-color topographical maps of the macula and representations of low- and high-reflectivity tissues assist the interpretation of the digitized data. OCT can be utilized to analyze retinal and optic nerve structure even through the undilated pupil. These devices generate digital cross sections of the retina and optic nerve, as shown in Figure 9.8.

High-contrast reflectivity among retinal, preretinal, and subretinal structures identifies retinal pathology. Well-defined spectral boundaries permit reliable and reproducible determinations of numerous parameters of diagnostic significance, including [60]

1. Retinal thickness
2. Foveal depression thickness
3. Intraretinal cysts

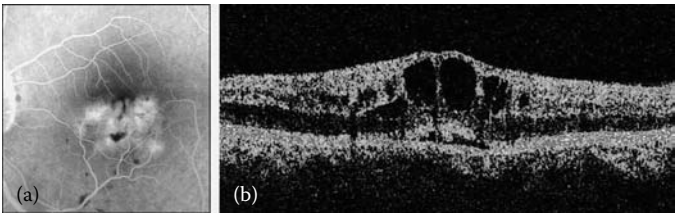


FIGURE 9.8: (See color insert following page 370.) (a) Fluorescein angiogram showing fluid with cystic changes in the retina. (b) Optical coherence tomography image showing fluid with cystic changes in the same retina. Retinal thickness can be quantified accurately using the OCT cross-sectional method.

4. Preretinal traction
5. Subretinal fluid
6. Macular holes
7. Epiretinal membranes and distortion
8. Geographic RPE atrophy (dry macular degeneration)
9. Choroidal neovascular membranes (wet macular degeneration)
10. Retinal pigment epithelial detachments
11. Subretinal fibrosis

Internal algorithms in commercially available units permit accurate measurement of retinal thickness and volume relative to population norms. This cross-sectional approach to retinal analysis, particularly its ability to detect and quantify retinal edema, complements the lesion detection sensitivity of digitized fundus photography (Figure 9.6).

Other cross-sectional methods to analyze the retina and optic nerve include confocal scanning laser microscopy, which acquires and analyzes 3-D images of the optic nerve head and enables the quantitative assessment of retinal and optic nerve topography relative to population norms and from serial data from subsequent patient examinations to identify changes. Glaucoma and other diseases cause optic atrophy and result in the loss of nerve fibers and subsequent loss of visual field. Quantitative measurements of nerve fiber layer thickness can be obtained, making this a useful test in the management and diagnosis of glaucoma. The topographic analysis of the optic nerve head is a quantitative description of its current state and, over time, permits quantification of the progression of the disease.

9.3 Image Processing for Segmentation and Description

Whether analyzing imagery in the preclinical, clinical, small animal, or human environment, the segmentation of images into regions of interest are of paramount importance to achieve a detailed encoding of content in a data system. Statistical descriptions of these subregions representing morphology, texture, tracer activity, and so on, and the relationships among subregions in a population are representations of low-level attributes of imagery that must be adequately comprehended to achieve effective and specific indexing of content. In this section, we provide examples from our experience on a subset of methods that have been critical to our applications. These methods generally demonstrate the goals that segmentation performs in characterizing image regions.

9.3.1 Two- and three-dimensional organ segmentation

Quantitative analysis of small animal imagery requires the segmentation of anatomic structures of interest. The results of such segmentation—the shape and volume of a specific organ, for example—can serve as important features for query and retrieval from an informatics system. The field of medical image segmentation is certainly very rich; in our own work [53,61,62], we have developed semiautomatic methods based on the well-known level set methodology [63,64].

In the organ segmentation problem, we seek the closed contour C that defines the organ and/or anatomic structure(s) of interest, where the organ or structure is contained within the contour. In the level set approach, this contour C is defined implicitly as the zero level set of a function $f(\cdot)$ over the spatial image domain (x, y, z); more formally, $C = \{(x, y, z) : f(x, y, z) = 0\}$, and the organ or structure of interest—inside the contour—is defined by $f(x, y, z) > 0$, while the region outside is defined by $f(x, y, z) \leq 0$. The goal of the segmentation is then to find the optimal function $f(\cdot)$ according to some defined criteria. This optimization problem is posed as a partial differential equation and solved iteratively.

In our work, we have been primarily interested in identifying the mouse spleen and follicular architecture from blood pool contrast-enhanced microCT data, such as that shown in Figure 9.9a. As evident in this figure, the spleen boundaries are weak and the interior grayscale values are inhomogeneous. The darker regions inside the spleen are in fact follicles, which contain no blood; they appear darker because we are employing a blood pool contrast agent. Because of the weak boundaries, we employ a region-based level set model [64]. Because of the inhomogeneity, we first compute a statistically transformed version of the image [62] that is based on a manually marked seed region inside the spleen. An example result from this transformation is

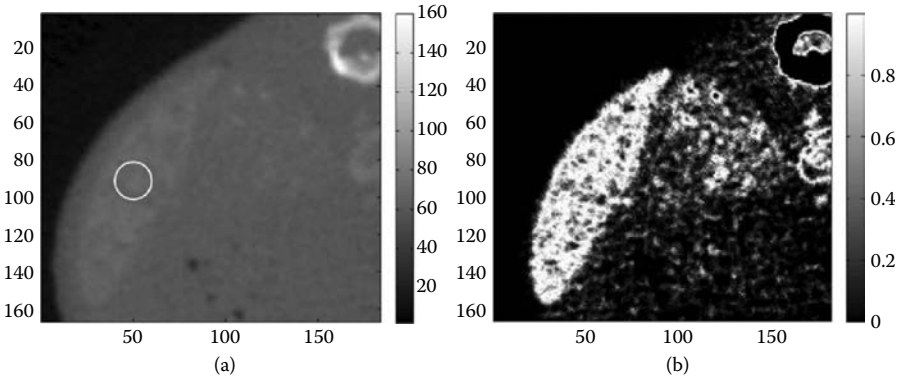


FIGURE 9.9: Example of raw data and corresponding likelihood image. (a) Grayscale image of raw CT data containing a portion of the spleen. The small circle in the spleen region surrounds a manually marked seed point used to initialize our semiautomatic segmentation process. (b) Likelihood image corresponding to raw data from (a). A two-component Gaussian mixture is estimated from the seed region. The resulting higher mean component represents the spleen blood pool and is used to compute the likelihood image.

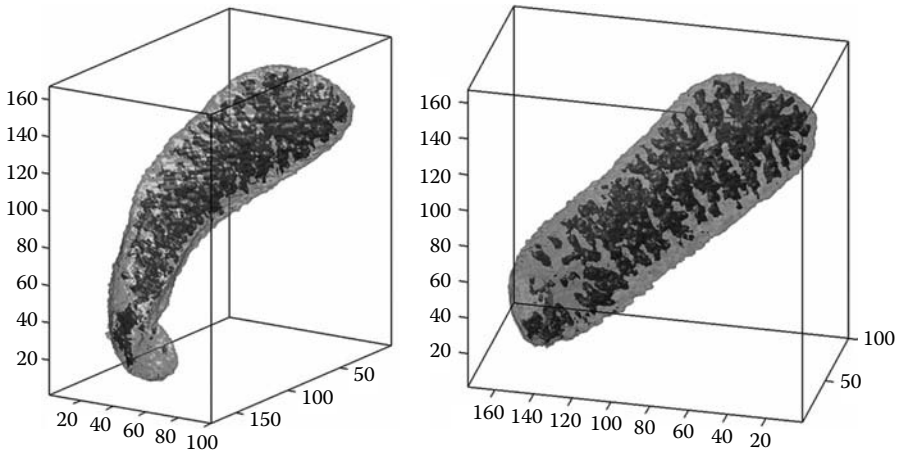


FIGURE 9.10: Renderings of 3-D spleen and follicle segmentation. The follicular architecture is shown darker inside the transparent spleen.

shown in Figure 9.9b. We then apply the region-based level set algorithm to this transformed image. Once the spleen boundary has been found in this manner, we apply a modified [53] region-based method to the spleen interior to segment the follicular architecture. An example result from this approach is shown in Figure 9.10.

Although level set methods have proven their efficacy in a variety of challenging segmentation problems, they are not without limitations and are hence a very active area of research. Papandreou and Maragos [65], for example, improved the computational speed through the use of enhanced finite difference schemes and fast distance transforms. Another avenue of ongoing research involves topology constraint. Since the level set contour is defined implicitly, it is generally not topologically constrained. Although this can be a benefit, such as in follicle segmentation, it can also be detrimental when the topology of the organ or structure is not known ahead of time. A topology-preserving level set algorithm was first proposed by Han, Xu, and Prince [66]. This method, however, often suffers from serious drawbacks [62,67]. A more elegant solution to topology-preserving level sets was proposed by Sundaramoorthi and Yezzi [67], who added a topology-preserving term into the optimization criteria. Finally, we note that despite all the efforts in the field to date, there is still no segmentation approach that generalizes exceptionally well. In other words, user interaction is almost always required when some characteristics of the images change, as they inevitably do in practice. One potential solution is to explicitly incorporate user interaction into the algorithm. Cremers and colleagues [68] proposed and implemented such an approach using a Bayesian framework on a graphics card for real-time performance.

Regardless of the requirements of user interaction, methods such as these facilitate the segmentation of enhanced detail in a manner that is repeatable over multiple animals and organ structures (see Figure 9.11). For example, we are interested in quantifying the volume of the spleen occupied by the "blood pool" red pulp. This blood pool decreases as amyloid deposits accumulate in the spleens of mice with systemic serum amyloid A protein (AA) amyloidosis. Also, variations in the follicular architecture may indicate changes in the extent of the amyloid disease and that of other lymphoproliferative disorders that involve the spleen, such as lymphoma. Quantifying the follicle structure and storing in the CBIR system will allow us to detect changes—statistically across populations and/or in single-subject longitudinal studies—that indicate important biological changes. Thus, indexing the volumetric data by image content becomes extremely relevant to the study of amyloid disease. Methods such as level sets help us achieve adequate segmentation for content description.

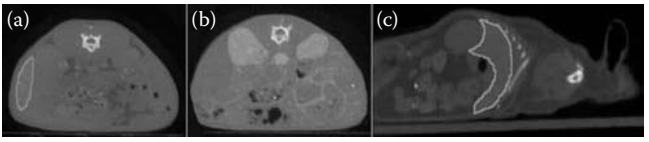


FIGURE 9.11: Application of level sets to 2-D segmentation of a variety of visceral structures such as an axial view of a spleen in (a) and kidneys in (b), plus a sagittal view of the liver in (c).

9.3.2 Segmenting anatomic structure in the retina

The problems of anatomic segmentation in the retina are quite distinct from those in small animal CT. Although inherently two dimensional (i.e., for standard fundus photography), the structures of the retina are generally diffuse and nonspecific. Although the optic nerve head and the vasculature have well-defined boundaries and are readily segmented [69,70], the challenge exists in the segmentation of lesions. There has been considerable research in the development of automated screening for DR. An excellent overview of this area is provided by Teng, Lefley, and Claremont [71], who summarize the image segmentation as a two-step process. The first step identifies the regular or expected physiological features of the retina and includes image preprocessing, optic disk localization and segmentation, and vasculature segmentation. The second step is concerned with identifying the pathology of the retina and includes registration of images when multiple images are recorded, microaneurysm segmentation, and exudate segmentation. Some applications have also investigated drusen detection for age-related macular degeneration (AMD). Overall, this is an active area of research but suffers greatly from a lack of standard image sets, which makes results comparison difficult to assess. This is being remedied in some part by publicly available databases,* although complete ground-truth data are not available for these pioneering efforts.

To address accurate segmentation of anatomic structure, we developed methods that take advantage of a number of visual attributes of the vascular tree in a probabilistic framework [69,70]. Figure 9.12 presents an overview of our method for optic nerve and macula detection. In (a), the original red-free fundus image undergoes analysis to provide a segmentation of the vessels as shown in (b). Using characteristics of the vasculature such as the density, orientation, and change in thickness, we are able to produce a spatial likelihood map of the probable locations of the optic nerve. This map defines the likely coordinate of the optic nerve as in (c). Once the nerve coordinates have been detected, we make use of a parabolic model of the vascular structure

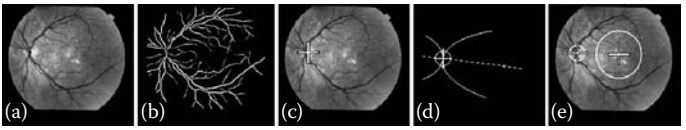


FIGURE 9.12: Overview of results for detection and localization of the optic nerve, vasculature, and macula regions in a red-free fundus image. (a) Original red-free fundus image. (b) Segmentation of the vessels after analysis. (c) Likely coordinate of the optic nerve. (d) Location of the horizontal raphe. (e) Center of the macula positioned near the optic nerve coordinate.

*STARE database, http://www.parl.clemson.edu/stare/.
DRIVE database, http://www.isi.uu.nl/Research/Databases/DRIVE/.

emanating from the optic nerve coordinate to locate the horizontal raphe as in (d). Finally, the center of the macula, our primary reference point in the description of disease manifestation, is positioned 2.5 optic nerve diameters from the temporal edge of the optic nerve coordinate, as shown in (e).

The localization of the optic nerve and macula in retinal imagery provides a reference coordinate system for the description of lesion populations that may exist in the retina. Automated lesion segmentation processing begins with an assessment of image quality [72,73]. Some studies suggest that 10% of retinal images have artifacts that preclude human grading [74–76], with another 10% likely unsuited for automatic processing. Once an image has passed a quality check, subsequent processing is applied to detect and segment the various forms of lesions that occur.

Microaneurysms are important because they are a significant indicator of DR as a single feature, as some studies have indicated [47,77]. DR includes red lesions such as microaneurysms and hemorrhages, yellow lesions such as exudates, and white lesions such as cotton wool spots (i.e., focal retinal infarcts); however, the red lesions are typically the first unequivocal signs of DR. Li and Chutatape [47] identify exudates as one of the most commonly occurring lesions. Strategies used for detection include thresholding, edge detection, and classification. Threshold is straightforward (relatively speaking), but the uneven intensity of exudates and low contrast cause problems. Edge detection has issues with the edges of vessels and other lesions. Classifications have also been attempted, but no method seems to be complete. Exudates, cotton wool spots, and drusen can be distinguished somewhat on the basis of color features [78]. Simple thresholding methods have been applied by numerous authors and include noise reduction methods prior to threshold and removal of the physiological landmarks such as the optic disk and vasculature [79–81]. Although not explicitly mentioned by Teng, Lefley, and Claremont [71] as a separate processing step, drusen segmentation papers receive a large amount of attention in the literature. Exudates and drusen have a rather similar appearance in that both are yellowish lesions, although drusen tend to be darker. Indeed, drusen are more often associated with age-related macular degeneration and thus are considered a confounding factor in exudate detection [82].

In our research, we adapt the method of Ben Sbeh and colleagues [83] for general lesion detection by using their algorithm across multiple scales. The geodesic morphological reconstruction approach is attractive for two reasons. First, the authors state that they attempted several other methods that were not successful. Second, the method seemed to have some degree of flexibility, particularly with regard to the parameters and structuring elements. We extended this parameter flexibility to dark lesions (e.g., microaneurysms and hemorrhages) by simply working with the negative of the original image and filtering out the vasculature with a previously created segmentation of the vessels. We used a ground-truth image set created by one of our authors [84], which is described in detail below. We apply a variety of different scales and determine a simple segmentation quality metric, the mutual overlap between

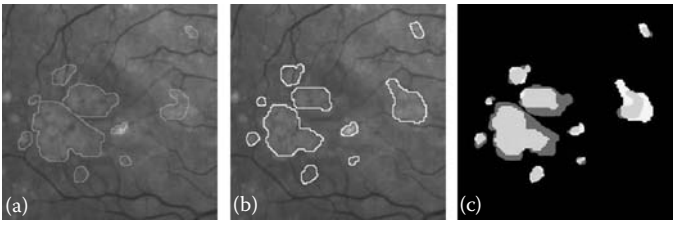


FIGURE 9.13: (a) Hand-segmented lesion example. (b) Machine-segmented lesions selected on the basis of minimum distance to hand-segmented population features. (c) Comparison of hand- and machine-segmented segmentations; dark gray (undersegmented), white (oversegmented), light gray (correct segmentation).

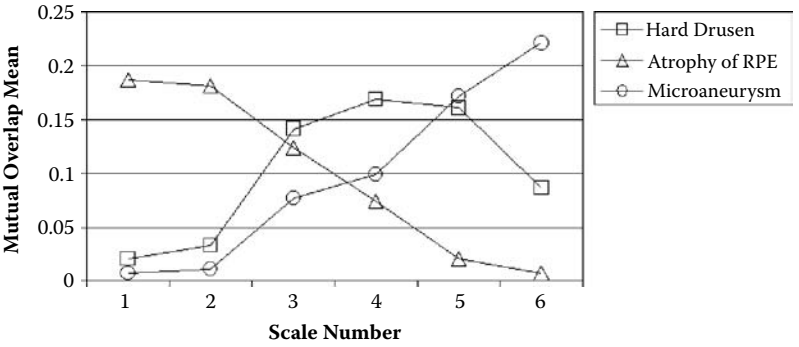


FIGURE 9.14: Plot of scale versus mutual overlap value for three lesion types. This shows, for example, that microaneurysms are segmented best at the higher scale.

the ground-truth segmentation and the machine segmentation. The mutual overlap, a unitless parameter, is the area of the overlap between the hand- and machine-segmented objects normalized by the sum of those areas. For example, Figure 9.13 shows hand-segmented lesions in (a) and the corresponding machine-segmented lesions in (b). The ratio between the number of green pixels in (c) divided by the red and yellow enclosed areas in (a) and (b) represent the mutual overlap. For each lesion category, we then determined the mean of the mutual overlap. An example of this is shown in Figure 9.14 for three representative lesion categories of hard drusen, atrophy of the retinal pigmented epithelial (RPE) layer, and microaneurysm. This is used to define the segmentation scale to optimize performance of each lesion type.

This algorithm tends to oversegment data [83]; therefore, a postprocessing step is applied to separate true lesions from false positives. Ben Sbeh and colleagues [83] performed the postprocessing with thresholds applied to

a small set of numerical descriptors for each lesion. In a similar problem, Niemeijer and colleagues [85] used an extensive set of 68 features to separate blobs from lesions in their algorithm. Among their 68 features, they included size measurements (area and perimeter), shape measurements (aspect ratio, circularity, compactness), and intensity measurements (mean intensity, intensity normalized by standard deviation of intensity, contrast, etc.) with some features measured on different color planes and spaces and at different scales. In our case, we treat the oversegmentations as a two-class problem in which we seek to keep some blobs and eliminate others. We apply a set of features similar to the other research, but the deeper question is how to determine which blobs to keep and which to eliminate. We should note that our overall goal is not specifically accurate segmentation and classification of lesion types; rather, we seek to create a descriptive numeric vector indicating the pathology and disease grade of the image for computer-assisted diagnostics with a human in the loop. Consequently, we are experimenting with a variety of segmentation evaluation algorithms designed to capture as much of the overall pathology as possible. An example of this was presented earlier in Figure 9.13, where we created a machine-segmented image and evaluated its blob structure to determine which objects to keep by generating lesion population features on the ground-truth image. We ranked the segmentations on the basis of mutual overlap and add them one by one until we found the best match to the ground-truth segmentation description. The resulting machine segmentation image (prior to postprocessing) is shown in Figure 9.13c, with the combined image shown for comparison (light gray is areas of overlap, white is machine segmented, and dark gray is ground-truth segmented). Another example is illustrated in Figure 9.15, where we show a hand-segmented image in (a), a machine-segmented image at one scale in (b), and a postprocessed result in (c). A histogram of the mean of the blob interiors was used to design the postprocessing filter.

In general, there is still much research to do in the area of retina lesion detection. The growing body of literature representing research in this area, along with our preliminary results, indicates that these methods will provide

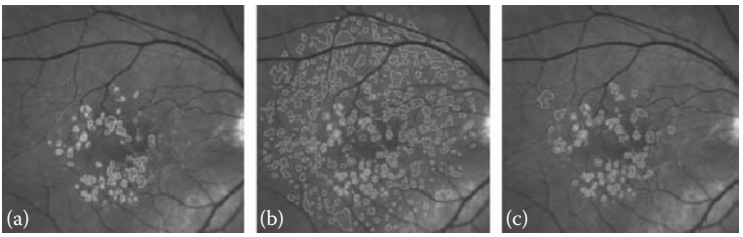


FIGURE 9.15: (a) Hand-segmented lesion example. (b) Machine-segmented lesions prior to postprocessing. (c) Postprocess results. The postprocessing filter was designed using lesion classifications of "keep" or "reject" with the lesion population histogram of lesion mean values as a metric.

reliable segmentation for indexing manifestations of retinal disease in image databases.

9.3.3 Statistical feature analysis

There are many ways to describe electronic imagery for content-based indexing. Low-level features (e.g., overall image color and texture attributes) can be extracted that globally describe a scene, or the scene can be broken up into subregions [86] through rudimentary block processes (e.g., tessellation [87]) or segmented into logical subregions. To achieve a level of specificity that is compatible with the detailed description of anatomic, functional, or disease structure, we chose to decompose critical regions of the imagery, as described above. Once the image is segmented into appropriate regions, we encode the organs and/or regions of interest statistically. That is, we extract a collection of application-specific numerical-valued features. For a mouse organ extracted from a CT image, features may include total volume, surface area, ratio of the two, and variance along the principal axes. Features associated with the corresponding SPECT image of the organ might include the total photon count as well as the mean and the variance computed over the individual voxels. For a retina image, we may measure features related to the size and shape of the optic disc, the scale and shape of the vasculature, and the number, size, and distribution of retinal lesions or microaneurysms.

The generated features are stored in a feature vector, $\boldsymbol{u}$ in $\mathbb{R}^N$, where $\boldsymbol{u} = \{u_1, u_2, \ldots, u_N\}^t$. Under the assumption that similar images produce similar features, we can compare one image to another by comparing the feature vectors. This can be done geometrically by calculating normalized vector distances such as the Euclidean, Manhattan, or others in an L-norm metric space. We can also apply maximum likelihood methods if the feature vectors can be considered to be samples from a high-dimensional distribution whose parameters we estimate from class-labeled learning samples [15]. In either case, the comparison can be carried out in the high-dimensional data space corresponding to all the features or in a subspace thereof depending on the task at hand. We exploit this idea below.

Highly correlated features should be avoided. Given a sample of feature vectors, we can compute the pairwise correlation of the individual features and discard those that provide no additional insight. The remaining features may still exhibit some degree of correlation. Principal component analysis (PCA) [88] may be applied to determine latent information while decorrelating the features at the same time. Probabilistic PCA [89,90] is a related maximum likelihood method for determining latent information that can also be applied.

While PCA and related techniques can provide a useful linear projection and representation of data to reduce dimensionality and decorrelate, there is no guarantee that it will improve discrimination between different classes of data. Recall that our goal is to determine an index that can be used to locate similar imagery by comparison. This can be achieved to some optimal

extent if we can produce an index that provides discrimination between the different categories of interest ω_i in the archive. We addressed this issue by using linear discriminant analysis (LDA) to project our image features $\boldsymbol{u}$ in $\mathbb{R}^N$ onto a subspace $\boldsymbol{v}$ in $\mathbb{R}^M$, where $M < N$ and the dimension of $\boldsymbol{v}$ is $M = C - 1$. The index, $\boldsymbol{v}$, is designed to discriminate between our category classes $\omega_{i,}$, for $i = 1, 2, \ldots, C$, with C the number of user-identified categories. For this discussion, we apply a traditional multiple LDA method based on the well-known Fischer discriminant [15] that maximizes the following criteria function,

$$J(\mathbf{W}) = \frac{|W^t S_B W|}{|W^t S_W W|}$$

where $\mathbf{W}$ is the projection matrix we wish to determine, $\mathbf{S}_B$ is the between-class scatter matrix, and $\mathbf{S}_W$ is the within-class scatter matrix as defined in [91]. The $M \times N$ matrix $\mathbf{W}$ is used to project our image vectors $\boldsymbol{u}$ onto $\boldsymbol{v}$ according to the relationship $\boldsymbol{v} = \mathbf{W}\boldsymbol{u}$. Note that we interpret our projected data to be a semantic approximation to the descriptive intent of the user given that the class categories are defined by the user according to the end-use goals of the system. In general, we are taking a vector of low-level features, which may or may not provide discrimination (relative to user interpretation), and projecting them onto a linear subspace that encodes semantic intent.

To demonstrate the principle of semantic feature mapping, we use an example population of 395 digital retina images from an anonymized data set taken through dilated pupils using a Topcon TRC 50IA retina camera. The images are from the ophthalmic clinical practice of one of the authors [84], a retina specialist at the University of Tennessee Hamilton Eye Institute, and were selected and characterized based on determination of the specific disease diagnosis present in the eye. We tested and selected statistical features that describe the retina in a region immediately surrounding the macula based on lesion shapes, textures, lesion types and counts, and other structural attributes, as shown in Figure 9.16. These features include region morphology

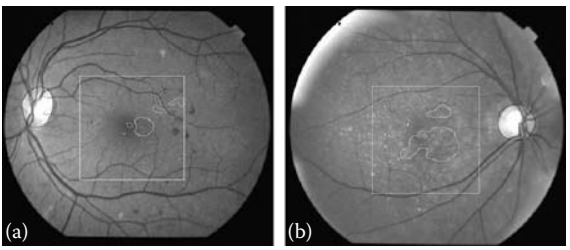


FIGURE 9.16: Examples showing the central macula regions of two retina images indicating the location of various lesion types. (a) Nonproliferative diabetic retinopathy. (b) Age-related macular degeneration.

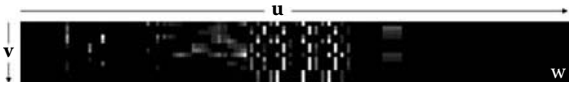


FIGURE 9.17: Visual representation of the projection matrix, **W**, showing the mapping of original features, u, to an image index, v. Note that lighter colors represent higher significance than darker colors.

based on invariant moments, shape descriptors, intensity descriptors, texture derived from statistical region variations, directional textures derived from spatial frequency (occurrence) representations, and proximity features based on relationships between multiple regions in the data set. In total, there are $N = 170$ features representing our initial vector, $\boldsymbol{u}$.

In this data set, we defined one normal class and 14 disease classes for a total of $C = 15$ separate states corresponding to our categories, ω_i. These disease states include stratified grades (levels) of disease for AMD and for nonproliferative and proliferative diabetic retinopathy (NPDR, PDR), and correspond to currently used or common strategies for disease quantification in clinical use. These disease states represent the semantic intent of the user, and our goal is to provide a mapping of the original features as described above. Figure 9.17 shows a visual representation of the projection matrix, **W**, for this dataset. The 170 columns shown represent the original features, $\boldsymbol{u}$, and the rows represent the linear combinations of the relevant features that produce the index $\boldsymbol{v}$ (of length $C - 1 = 14$ features) that is designed to discriminate between the 15 defined categories. It is apparent from the figure that we have greatly overspecified the number of features, $\boldsymbol{u}$, required to produce a semantic mapping to the index, $\boldsymbol{v}$. This is indicated by the large areas of low contribution to the mapping (i.e., dark colors).

The index becomes the primary representation of the imagery in the archive. As described earlier, we perform a comparison of a query index to the indices represented in the archive to produce a ranked listing of images from most to least similar. We produce a similarity value, $S(\boldsymbol{q}, \boldsymbol{v}_j)$, between the query image vector, $\boldsymbol{q}$, and the archive image vectors, $\boldsymbol{v}_j$, for $j = 1, 2, \ldots, n$ images in the archive. Similarity is defined on the basis of the distance between two points, $(\boldsymbol{q}, \boldsymbol{v}_j)$, in the index space and should be in the range [0,1]. Distance, $D(\boldsymbol{q}, \boldsymbol{v}_j)$, is defined for a metric space in which $D(\boldsymbol{q}, \boldsymbol{v}_j) \geq 0; D(\boldsymbol{q}, \boldsymbol{v}_j) = 0$ if and only if $\boldsymbol{q} = \boldsymbol{v}_j; D(\boldsymbol{q}, \boldsymbol{v}_j) = D(\boldsymbol{v}_j, \boldsymbol{q})$; and $D(\boldsymbol{a}, \boldsymbol{b}) + D(\boldsymbol{b}, \boldsymbol{c}) \geq D(\boldsymbol{a}, \boldsymbol{c})$. For this discussion, we use the well-known L-norm metric defined by

$$D(\mathbf{q}, \mathbf{v}_j) = \left(\sum_{i=1}^{M} |q_i - v_{i,j}|^L \right)^{1/L}$$

To produce a similarity value on [0,1], the feature space of the index is normalized using either a z-score or range method [92]. To ensure that our

similarity metric is defined on a unit-normal hypercube, and therefore on [0,1], $S(\boldsymbol{q}, \boldsymbol{v}_j)$ is defined as follows:

$$S(\mathbf{q}, \mathbf{v}_j) = 1 - \frac{D(\mathbf{q}, \mathbf{v}_j)}{M^{1/L}}$$

recalling that M is the dimension of the index. It should be noted that we could use the distance metric, $D(\boldsymbol{q}, \boldsymbol{v}_j)$, to define similarity between two points in our index space, $(\boldsymbol{q}, \boldsymbol{v}_j)$. In the next section, we show that similarity values ranging from 1 for most similar to 0 for least similar are useful for providing a weighting value when estimating the posterior probability from a collection of similar images.

9.4 Information from Image Content

A biomedical informatics archive built on image content assumes that there exists relationships (i.e., cause and effect) among the image records that are stored. For example, we assume that our image content is related by animal type, imaging modes, disease state, study protocol, and so on. We assume that the images are not random, unrelated compilations, and therefore there exists a bias in the predictive characteristics of the system due to the prior distribution of categories of interest. For example, Figure 9.18 shows the prior distribution as probability, $P(\omega_i)$, of the disease states that are present in the

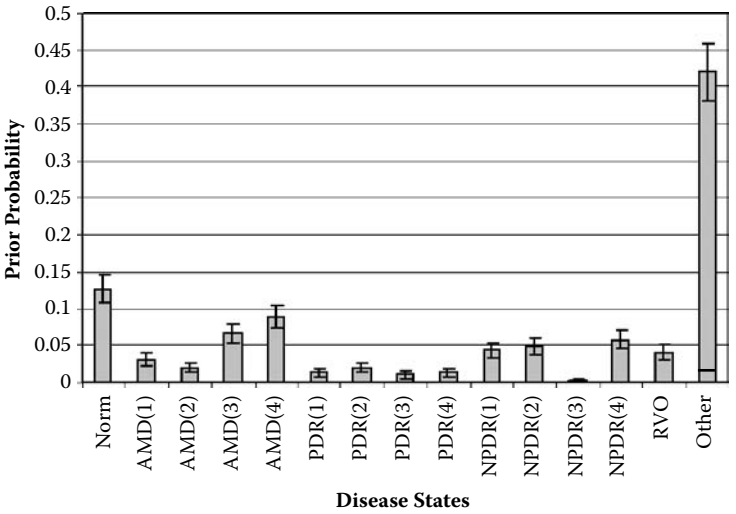


FIGURE 9.18: Prior distribution of the disease state categories (labeled Pathologies) representing the 395 retina images used in our example.

retina data population introduced above. The error bars associated with each bin represent the Poisson uncertainty for the measurement of each disease state. Through our estimation of the Bayesian posterior probability, $P(\omega_i|\boldsymbol{v})$, we are assuming that the bias of the prior distribution of events is inherent because the prior probability, $P(\omega_i)$, is explicit in Bayes theorem:

$$P\left(\omega_i|\mathbf{v}\right) = \frac{p(\mathbf{v}|\omega_i)P(\omega_i)}{p(\mathbf{v})}$$

Therefore, formulating our CBIR methodology in a Bayesian framework will facilitate our ability to modify our predictive response to accommodate changing goals of the system. For example, if the patient database was collected from a practicing ophthalmologist, any predictive estimates of disease will show bias toward the detection of diseases that are most prevalent in that environment. If our goal is to search the diagnostic information of this archive to better understand subjects in a general screening environment, it can be shown that the posterior probability can be corrected through prior probability modification as

$$P'(\omega_i|\mathbf{v}) \sim \frac{P'(\omega_i)}{P(\omega_i)} \cdot P(\omega_i|\mathbf{v})$$

where $\mathrm{P}'(\omega_i)$ represents the prior distribution in the general population (in this case). It is characteristics such as this that provide motivation to use this framework.

9.4.1 Estimating posterior probabilities

Considering the inherent complexity in a data archive that is capable of representing multiple states of the data, we are interested in using nonparametric approaches to the estimation of the posterior probability. The retrieval mechanism introduced earlier can be thought of as a nearest neighbor method. Nearest neighbor classifiers function by locating the population of labeled data points nearest to an unknown point in feature space, either by specifying a number of neighbors, k, or by specifying a volume in the feature space, V. Classification is achieved by performing some form of "voting" according to the labeled population and subsequently assigning a "winning" class to the unknown point. The population of points in the feature space surrounding the unknown (i.e., analogous to the query, $\boldsymbol{q}$, in a CBIR system) represents a representation and direct measurement of the posterior probability. If we specify the retrieval of k-nearest neighbor (k-NN) points surrounding our query index, $\boldsymbol{q}$, we can form an estimate of the posterior probability as

$$P(\omega_i|\mathbf{v}) = \frac{p(\mathbf{v},\omega_i)}{\sum_{j=1}^{c} p(\mathbf{v},\omega_i)} = \frac{k_i}{k}$$

where $p(\boldsymbol{v},\omega_i)$ is the joint probability estimated as $p(\boldsymbol{v},\omega_i) = (k_i/n)/V$, k is the number of samples specified in the neighborhood, k_i is the number of

samples of category ω_i in a volume V surrounding the query point, and n is the number of records in the archive.

Intuitively, an image vector that is nearest to the query vector should provide the strongest evidence of a query's state. One method for incorporating the distance between points into the posterior estimate is to use the similarity ranking as follows:

$$P(\omega_i|\mathbf{v}) = \frac{k_i}{k} = \frac{\sum\limits_{i} S(\mathbf{q},\mathbf{v})^w}{\sum\limits_{k} S(\mathbf{q},\mathbf{v})^w}$$

where the exponent, $w \geq 0$, increases the influence of the points closest to the query point. When $w = 0$, this estimate reverts to the basic k_i/k estimate.

The posterior estimate could also be formulated by specifying a fixed volume, known as a Parzen window, about the query point in the index space as opposed to a fixed number of neighbors. In general, it can be shown [15] that the k-NN produces a better estimate of $P(\omega_i|\boldsymbol{v})$ because the window (i.e., volume, V) dynamically adjusts to accommodate different densities of points in the index space, resulting in a smoother estimate.

It is also possible to show that the estimate of $P(\omega_i|\boldsymbol{v})$ approaches a nearly optimal posterior estimate as the number of records in the system increases—that is, as $n \to \infty$. This provides further support for using a nonparametric k-NN estimator to perform information analysis in a large image archive. In essence, the information-extraction performance of the archive should theoretically improve as the data population grows.

It should also be noted that the notation used to express the posterior probability has been simplified for this discussion. In general, an informatics archive can be thought of as a relational database. Although image content (i.e., $\boldsymbol{v}$) has constituted our primary posterior "evidence," we also have evidence in the form of metadata that can be used to logically bound the extent of our retrievals, and therefore the posterior estimate; that is, we can write our estimate as $P(\omega_i|\boldsymbol{v}, \alpha, \beta, \gamma, \ldots)$, where α, β, γ, and so on represent metadata attributes. For example, in the retina archive, there are data regarding patient age, gender, ethnicity, and medical history. Considering that various disease states are associated with these attributes, this information can be used to modify how a query is performed (e.g., the prevalence of AMD increases with age, the risk of diabetic retinopathy increases with duration of disease and comorbid diseases such as hypertension). In the small animal preclinical environment, the same condition holds true. Metadata related to animal parentage, strain, SPECT or PET tracer, CT contrast agent, route of administration, and so on all add useful evidence to the process of posterior estimation and analysis.

9.4.2 Confidence and performance

Unlike a k-NN classifier, where a class label is assigned to an unclassified event, an information archive provides a broader context for analysis and

decision making (i.e., it is not a classifier). The posterior estimate that results from a retrieval of k similar images from a data population represents a summary of the statistical distribution of categories retrieved from a system. The goal is to infer information about the query based on this retrieved population. Therefore, it is desired that we determine something about the variability or confidence in our estimate of the posterior probability: $P(\omega_i|\boldsymbol{v}) \pm \sigma(\omega_i)$. Knowledge of this uncertainty, $\sigma(\omega_i)$, is useful when determining a level of confidence regarding the retrieved information. For example, $\sigma(\omega_i)$ will vary as the number of data records in the system, n, changes over time (as described in the previous section), and it will vary as the number of examples in any given category, ω_i, changes.

To measure the variability associated with changes in the data population in the system, Poisson statistical analysis is used. Recall that a Poisson process can be applied to phenomena of a discrete nature whenever the probability of the phenomena happening is constant over time—that is, phenomena that may occur $0, 1, 2, \ldots$ times during a given time interval, such as a patient with a given pathology walking into a doctor's office or a mouse exhibiting signs of amyloidosis in a study population. By using the property that the standard deviation in the sampling of the category, ω_i, is the square root of the number of counts in the category, the uncertainty can be expressed as

$$P(\omega_i|\mathbf{v}) \pm \sigma = \frac{k_i}{k} \pm \frac{\sqrt{k_i}}{k}$$

Note in this expression that $\sum_i k_i \equiv k$ is the number of nearest neighbors retrieved in the k-NN method. The uncertainty terms in the numerator and denominator have been propagated through the expression assuming that the k_i are uncorrelated.

Figure 9.19 shows an example of three images used to query the population of retina images using the hold-one-out (HOO) method. By the HOO method, each image is removed from the archive and submitted as a query to generate a posterior estimate and uncertainty, as shown in the right-hand charts of the figure. In all three cases, the disease state prediction is correct even though the error bar in each category is significant. The large uncertainty is a direct result of the small population of data available in this archive. As larger archives are assembled, this uncertainty reduces, thereby improving user confidence in the predictive results. Confidence is lowest when there exists overlap in the estimates for different disease states. The error bars in Figure 9.19 represent one standard deviation (i.e., encompassing ∼68% of the data). For this small data population, it would be impossible to achieve, for example, a 3σ confidence (99.7%) due to the overlap of uncertainty between categories.

Of course, viewing only three examples from the archive is anecdotal. To understand the predictive capability of this method, we can perform HOO testing across the entire population. Also, we can experiment with our ability to use the uncertainty estimate by rejecting data that are defined as having

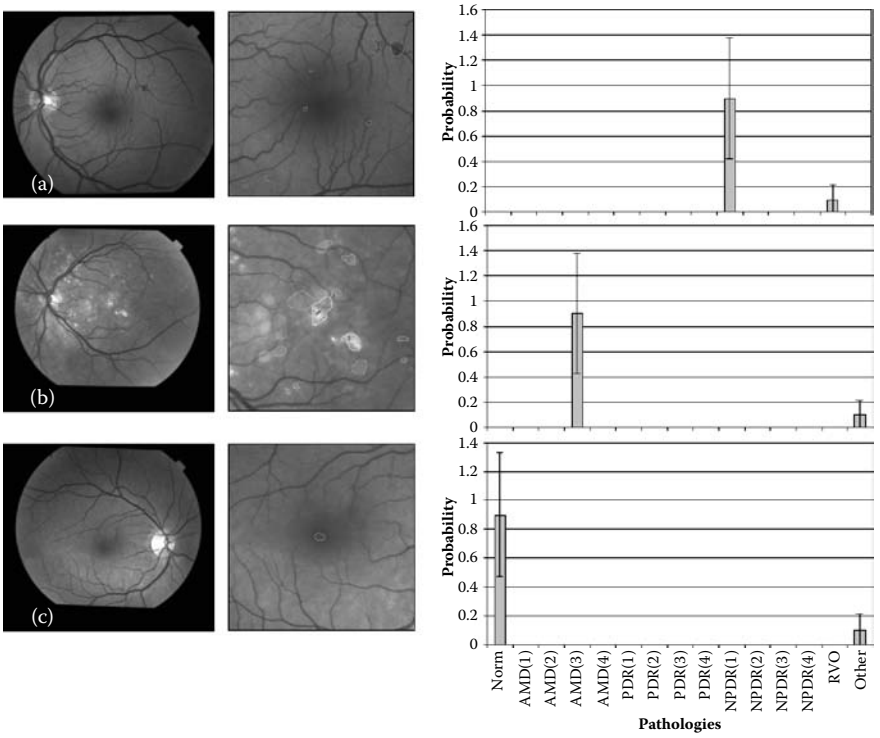


FIGURE 9.19: Examples of the estimation of the posterior probability of disease for (a) NPDR, (b) AMD, and (c) normal fundus. The left column shows the original image, the middle column details the manifestations of the disease, and the right column shows the estimated probability of each disease state.

low confidence due to overlap in uncertainty between categories. Table 9.1 shows a confusion matrix that was developed by testing the system as if it were a classifier. For this example, our class is defined to correspond to the highest probability category in the posterior estimate for a specified σ. The estimated class is compared to the known class, and the confusion matrix is formed. The column labeled "UNKNOWN" collects all data records that are below our confidence level. In Table 9.1, $\sigma = 0.0$; therefore, none of the data are rejected and all records are classified (either correctly or incorrectly). In Table 9.2, the confusion matrix is shown for $\sigma = 1.0$. The efficacy of this method is demonstrated through the overall improvement in performance from 85% for $\sigma = 0.0$ to 91.3% for $\sigma = 1.0$. Note that in the $\sigma = 1.0$ case, 18.8% of the data is rejected (not classified). As σ increases, more data are rejected, but performance continues to improve.

TABLE 9.1: Confusion matrix for the retina data archive. The left-hand column represents hand-labeled (known) disease classification; the top row represents CBIR-estimated classifications based on the posterior estimate. For this case $\sigma = 0.0$ and none of the data is rejected.

σ = 0.0

	Norm	AMD(1)	AMD(2)	AMD(3)	AMD(4)	PDR(1)	PDR(2)	PDR(3)	PDR(4)	NPDR(1)	NPDR(2)	NPDR(3)	NPDR(4)	RVO	Other	UNKNOWN	Performance
Norm	44	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	88.0%
AMD(1)	1	4	0	0	0	0	0	0	0	0	0	0	0	0	7	0	33.3%
AMD(2)	0	1	4	0	0	0	0	0	0	0	0	0	0	0	3	0	50.0%
AMD(3)	0	0	0	22	0	0	0	0	0	0	0	0	0	0	4	0	84.6%
AMD(4)	0	0	0	0	28	0	0	0	0	0	0	0	0	0	7	0	80.0%
PDR(1)	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	100.0%
PDR(2)	0	0	0	0	0	0	7	0	0	0	0	0	0	0	1	0	87.5%
PDR(3)	0	0	0	0	0	0	0	0	0	1	1	0	2	0	0	0	0.0%
PDR(4)	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	80.0%
NPDR(1)	0	0	0	0	0	0	0	0	0	12	0	0	0	0	4	0	75.0%
NPDR(2)	0	0	0	0	0	0	0	0	0	1	12	0	0	0	6	0	63.2%
NPDR(3)	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.0%
NPDR(4)	0	0	0	0	0	0	0	0	0	0	0	0	21	0	2	0	91.3%
RVO	0	0	0	0	0	0	0	0	0	0	0	0	0	14	2	0	87.5%
Other	7	0	0	1	0	0	0	0	0	0	0	0	0	0	158	0	95.2%
Purity	84.6%	80.0%	100.0%	95.7%	100.0%	100.0%	100.0%	ND	100.0%	85.7%	92.3%	ND	84.0%	100.0%	79.0%	0.0%	85.0%

TABLE 9.2: Confusion matrix for the retina data archive with $\sigma = 1.0$. Note in this case that 18.8% of the data has been rejected due to low confidence in the predictive values for those records.

σ = 1.0

	Norm	AMD(1)	AMD(2)	AMD(3)	AMD(4)	PDR(1)	PDR(2)	PDR(3)	PDR(4)	NPDR(1)	NPDR(2)	NPDR(3)	NPDR(4)	RVO	Other	UNKNOWN	Performance
Norm	40	0	0	0	0	0	0	0	0	0	0	0	0	0	4	6	90.9%
AMD(1)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	8	0.0%
AMD(2)	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1	4	75.0%
AMD(3)	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	9	100.0%
AMD(4)	0	0	0	0	25	0	0	0	0	0	0	0	0	0	6	4	80.7%
PDR(1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	ND
PDR(2)	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	5	100.0%
PDR(3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	ND
PDR(4)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	ND
NPDR(1)	0	0	0	0	0	0	0	0	0	11	0	0	0	0	1	4	91.7%
NPDR(2)	0	0	0	0	0	0	0	0	0	0	9	0	0	0	3	7	75.0%
NPDR(3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND
NPDR(4)	0	0	0	0	0	0	0	0	0	0	0	0	18	0	1	4	94.7%
RVO	0	0	0	0	0	0	0	0	0	0	0	0	0	14	1	1	93.3%
Other	7	0	0	0	0	0	0	0	0	0	0	0	0	0	152	7	95.6%
Purity	83.3%	ND	100.0%	100.0%	100.0%	ND	100.0%	ND	ND	100.0%	100.0%	ND	100.0%	100.0%	88.4%	18.8%	91.3%

Performance can be defined in a number of ways. For medical systems, it is common to discuss performance in terms of sensitivity, specificity, and accuracy, defined as follows [13]:

$Sensitivity =$ (images with disease classified as disease)/(all disease images)

$Specificity =$ (normal images classified as normal)/(all normal images)

$Accuracy =$ (images classified correctly)/(all classified images)

By varying σ through a range of values, we generate a set of confusion matrices from which we produce a series of receiver-operator characteristic (ROC) curves [93] detailing performance as a function of the number of data rejected—that is, in terms of specificity, sensitivity, and accuracy. A set of ROC curves for the 395-image database is shown in Figure 9.20. Note the positive and increasing performance of the system as the confusing data are rejected.

Recall that an informatics system is not a classifier but rather a system for locating and analyzing historic data represented in the archive. The results of Tables 9.1 and 9.2 along with Figure 9.20 serve primarily to represent expected predictive performance of such a system. The type of data represented in Figure 9.19 showing the posterior estimate, detected anomalies, and potentially a host of other attributes for user consumption is the true intent and output of the system.

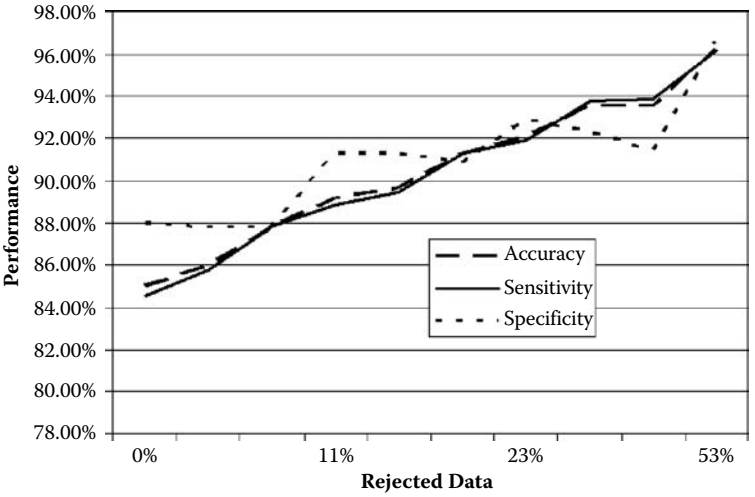


FIGURE 9.20: Predictive performance as a function of data rejected for accuracy, sensitivity, and specificity using the archive of 395 retina images.

9.5 Discussion

The goals of an informatics system will necessarily vary depending on the end-use applications and specification for usability. Basic applications can be defined that provide organization and access to large populations of animal or human data in preclinical or clinical studies. But beyond the fundamental concept of putting things away where they can later be found, the variability of needs to be addressed in the application domain renders it difficult to provide specifics. Software design, network architectures, user interfaces, and workflow protocols are domain specific. Systems designed to support research (e.g., searching for data correlations and data mining) within and across animal studies will have different architectures than systems designed to provide prospective information from a clinical database of patients. Systems designed to provide broader public access to study data for training and teaching purposes will have yet another. Regardless, we discuss in this section some general considerations regarding a system concept for small animal research and a system concept for retinal screening. Before discussing these system concepts, we must first discuss the issue of data structures and indexing in very large archives.

9.5.1 Search structures for large image archives

A goal of indexing that has not yet been discussed relates to the organization of image indices, $\boldsymbol{v}_i$, such that a ranked list of nearest neighbors can be retrieved from an archive without performing an exhaustive comparison with all the records in the database. It is clear that an image archive is expected, at some point, to manage hundred of thousands, if not millions of images. The comparison of a query vector to an entire population in an archive must be accomplished in an efficient manner. One method that can be used to address efficient indexing is the generation of binary decision trees to represent image indices. Figure 9.21a shows a representation of a binary decision tree. A bin, b, is defined as a bottom-level element in a tree structure, sometimes described as a "leaf" or terminal node, that contains a small list of images; a bottom-level bin may contain, for example, a list of image vectors $\{\boldsymbol{v}_a, \boldsymbol{v}_b, \boldsymbol{v}_c, \ldots\}$. The query vector, $\boldsymbol{q}$, is evaluated at the top level of the tree, and a decision is made as to which subtree path to traverse. There are many ways to implement decision trees. We used an approximate nearest neighbor (ANN) indexing and search method that builds on *kd-tree* methods [94,95]. An ANN partition is shown in Figure 9.21b corresponding to the decision tree in (a). The rules for partitioning the index space can vary greatly according to the desired outcome of the partitioning, but in the simplest case partitioning is achieved by simply parsing the data populations along each feature dimension, v_j (denoted by $(\boldsymbol{v}_x, \boldsymbol{v}_y)$ in the figure for simplicity) into equal halves. Whereas an exhaustive nearest neighbor search of the n vectors (i.e., images) in the

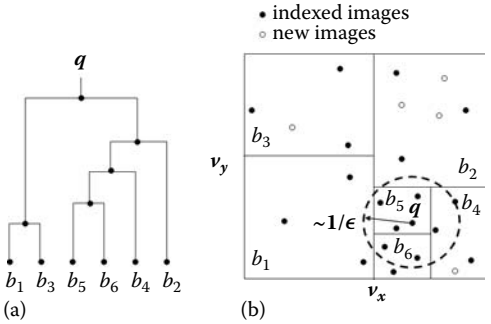


FIGURE 9.21: Schematic representation of a binary decision tree (a) generated from an approximate nearest neighbor partition of the index space represented in (b).

database would be of $O(n)$ computations, searching the kd-tree is of $O(\log(n))$ operations.

The ANN method incorporates a search window that results in the collection of neighboring partitions about the query point. As this window increases in radius, the nearest neighbor error, ε, decreases, but the performance of the system also decreases to $O(n)$. The efficiency of the ANN method is proportional to $O((1/\varepsilon)^{M/2}\log(n))$, where M is the dimension of the index space, n is the number of data points, and ε is the nearest neighbor error [94]. The nearest neighbor error is therefore inversely proportional to the size of the search window. Therefore, the accuracy of retrieval is selectable as a trade-off between nearest neighbor performance and computational search efficiency.

Note also in Figure 9.21b that ANN partitioning results in a data-dependent number of bins that contain a specified number of vectors, $\boldsymbol{v}$. ANN partitioning is required only periodically. As new data are entered into the system through regular use (as shown in the figure), the tree becomes unbalanced. As partitions accumulate large numbers of image vectors, retrieval speed will be impacted. Repartitioning results in rebalancing of the tree and therefore improved retrieval speed. Using this approach, the archive can collect new data continuously, and repartitioning can take place transparently to users during off-hours (e.g., at 2:00 A.M. on Sundays).

The application of ANN or related methods is not required for small data populations (e.g., the archive of 395 retina images described herein) but becomes increasingly important as the image population grows. Based on past experience, retrieval times will be noticeably impacted as archives grow significantly (e.g., over ∼20,000 images). Retrieval time depends, of course, on a number of factors, such as the computing platform, dimension of the index, and quantity of metadata. Retrieval performance speed can also be controlled by judiciously organizing the intrinsic and extrinsic data content using

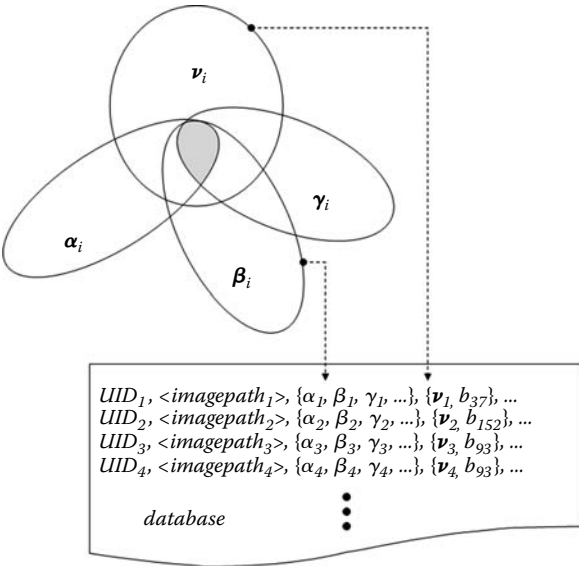


FIGURE 9.22: Simplified representation of database structure to support image and metadata management in an informatics environment.

a commercial relational database product. Commercial components are easy to service, maintain, and upgrade and can be efficiently supported long after technology development cycles are complete. Figure 9.22 represents a simple organizational structure that facilitates the type of image and metadata that has been referenced throughout this discussion. The Venn diagram at the top of the figure represents the permutations of a query that a user may wish to perform. Recall that the goal of an informatics system is to provide flexible query interaction that will support multiple users with multiple search goals. While image content, represented by the index, $\boldsymbol{v}$, provides the key evidence or link to the image files, the metadata represented by $\{\alpha_i, \beta_i, \gamma_i, \ldots\}$ is also used through the Boolean process of structured query language (SQL) processes, for example, to formulate posterior estimates such as $P(\omega_i|\boldsymbol{v}, \alpha, \beta, \gamma, \ldots)$, as described earlier. Note in the pictured elements, $\{\boldsymbol{v}_i, b\}$, that b corresponds to a leaf-node bin from the ANN implementation. Its purpose is to rapidly isolate a population of image indices for nearest neighbor analysis using standard SQL commands.

9.5.2 System concepts

Figure 9.23 provides a conceptual representation of the various elements of a biomedical informatics system to support a small animal research environment. The system is built on a commercial database and PACS [1] components,

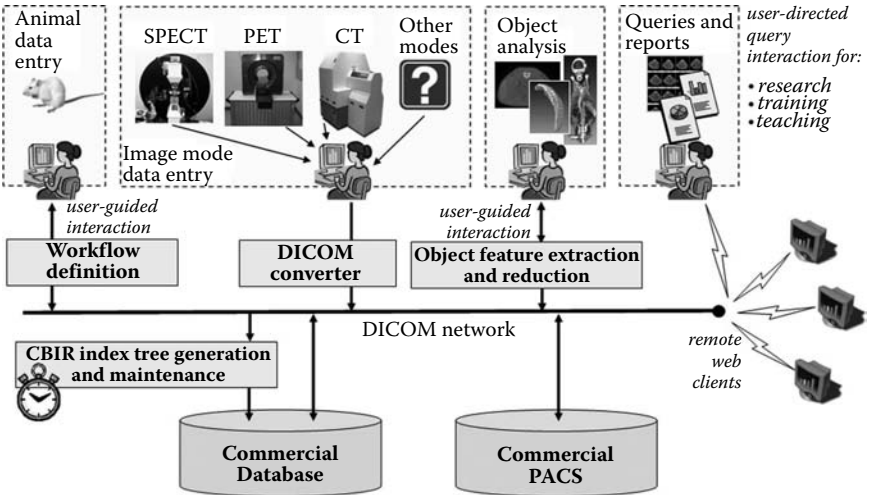


FIGURE 9.23: Conceptual representation of the elements of a biomedical informatics system. Data entry, analysis, and flow processes are built on top of commercial components such as relational database and PACS technology along with accepted standards and communication protocols such as DICOM.

and uses a DICOM*-compliant network and standards [96] to establish a protocol workflow for tracking animals through a study; managing the collection of imagery from the microimaging facility; and postprocessing image data to extract, index, and manage intrinsic content. Note that the system is represented by components that involve human interaction working in conjunction with automatic features such as index tree generation (2:00 A.M. Sundays), DICOM conversion, and feature generation. User interaction is envisioned to assist in the segmentation and merging of anatomic and functional data prior to database submission of this content. As described earlier in our discussion of 3-D organ segmentation using level sets, user-assisted interaction will be required to select appropriate results for subsequent feature analysis and database submission. Figure 9.24 shows an example interface for processing and manually submitting segmentation results of a particular anatomic structure to the data system. Platform-independent Web clients (e.g., constructed using Java and Active Server pages) provide system accessibility by remote users within or outside of the institution to the extent defined and desired to support research, training, and teaching goals of the organization.

*Digital Imaging and Communications in Medicine (DICOM) is a standard defined by the American College of Radiology Manufacturers Association for data communications among medical imaging devices.

The small animal informatics system architecture in Figure 9.23 is proposed for users interested in indexing image data and supporting metadata in an archive to maintain and access information generated through a preclinical study of animal models (e.g., in the study of amyloidosis). The primary goal for this type of system is to archive study data for later analysis of the population or to perform analysis across study populations over time (i.e., data mining). A separate informatics concept is represented in Figure 9.25

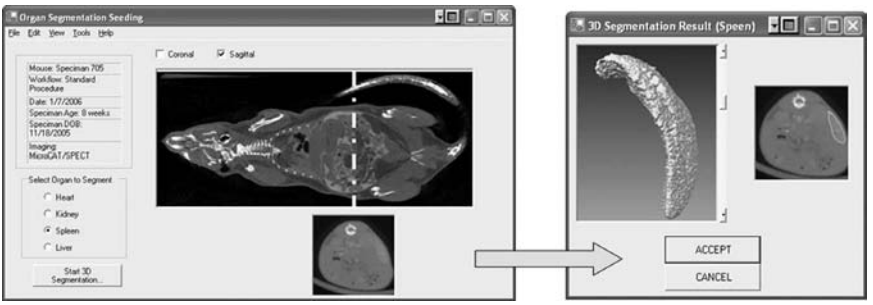


FIGURE 9.24: Example of a user interface component in a small animal informatics system for performing and accepting and submitting anatomic segmentation results.

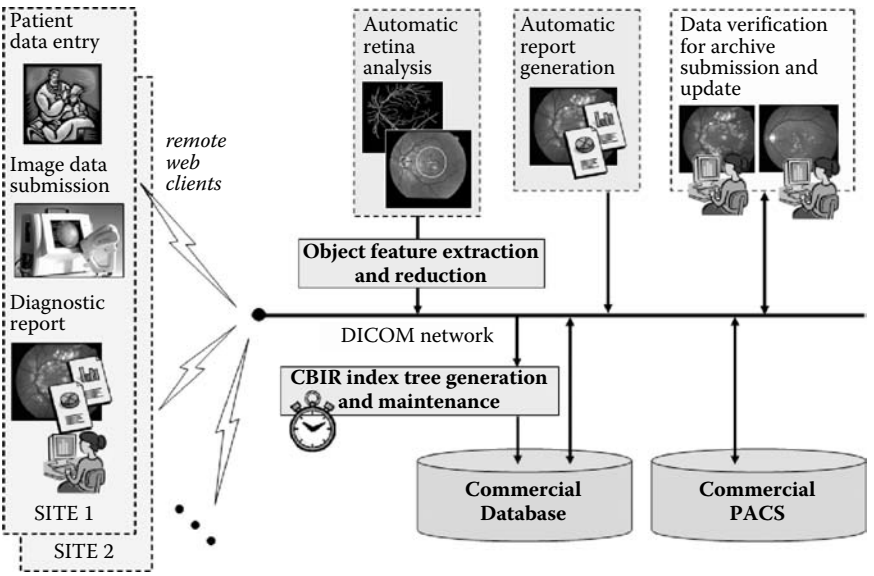


FIGURE 9.25: Conceptual representation of an informatics architecture to support image and metadata management in a medical diagnostics environment.

to support high-throughput, distributed screening of a large population for diabetic retinopathy. Currently, networking technology applied to the remote assessment of retinal disease consists primarily of store-and-forward digital retinal photography with subsequent image analysis by certified technicians at commercial and academic reading centers. The reading center paradigm is an established and validated method, with high sensitivity and specificity for retinal disease detection and quantification in clinical trials [97–100]. However, limited health care resources, access to imaging platforms and eye care specialists, and the absence of real-time diagnostic potential remain major obstacles to achieving the goals of population-based disease detection and management using the reading center method. The application of computer-based informatics to aid disease diagnosis has the potential to facilitate retinal disease detection and management by increasing throughput, reducing costs, and assisting or potentially automating diagnostic capabilities if they can be effectively integrated into a telemedical network delivery system.

Under this operating condition, remote Web clients (e.g., point-of-service physicians, optometrists, etc.) submit retinal images through a Web client to a central service that provides automated analysis and diagnosis of potential retinal pathology. As in Figure 9.23, both automated and assisted components of the system are required on the server side of the system. For the diabetic screening application, the process requires a higher level of automation to accommodate high-throughput Web client submissions. The primary role of users on the server side is to provide periodic manual review to verify and track predictive performance of the archive over time. Unlike the research environment described above, characterized (i.e., manually diagnosed and validated) data sets that are used to augment the archive must be evaluated by trained ophthalmologists and medical technicians before incorporation into the system for future use. Only a small fraction of data submissions need to be incorporated into the diagnostic archive, significantly reducing the amount of manual data manipulation required to maintain the system over time.

9.6 Future Research Directions

Throughout this chapter, the topic of biomedical informatics was presented and demonstrated using the examples of a preclinical informatics system to support small animal research and an informatics system to address high-throughput, low-cost retinal screening to detect those at risk for diabetic retinopathy. Although many of the ongoing issues in biomedical CBIR are touched on through these two disparate examples, there remain many outstanding issues of image segmentation; depiction; semantic description; indexing; data structuring; and efficient, extensible architectures that need to be addressed through future research and development. A number of researchers are looking into these and other issues today. Deserno, Antani, and Long

[101] provide a comprehensive review and extensive categorization of gaps in medical CBIR research through comparison with three medical CBIR implementations: the content-based PACS approach (cbPACS), the medical GNU image-finding tool (medGIFT), and the Image Retrieval in Medical Applications (IRMA) project. Other researchers are describing additional applications for accessibility to visual content. Tahmoush and Samet [102] describe a Web collaboration system, and Deserno, Antani, and Long [103] describe concepts for indexing biomedical images in scientific literature. With the advent of numerous anatomic and functional imaging modes working at multiple resolutions, the field of molecular imaging and medicine, as presented by Poh, Kitney, and Shrestha [104], necessitates future informatics technology capable of concurrently providing access to viscera, tissue, cells, proteins, and genes.

Regarding small animal informatics research, it is anticipated that the needs of this area of development will continue to expand. Since 1998, the U.S. National Institutes of Health (NIH) has funded many small animal imaging resource programs in which instrumentation for all common clinical modalities is either being used and/or being improved in some way to perform research on animals. The NIH continues to fund these programs and to issue new grant requests. As a result, our informal survey has found more than 40 universities that have published plans to build small animal imaging laboratories. In addition, more than 20 of the world's largest pharmaceutical companies either have small animal imaging facilities, are planning to purchase such facilities, or have shown interest through attendance and sponsorship of meetings with a focus on small animal imaging, such as the Academy of Molecular Imaging, and the Society of Molecular Imaging. There is clearly a growth path from the university environment—the early adopters of such technology—to private companies such as pharmaceuticals. As animal imaging techniques become standard assays in the animal research community, further expansion of the technology will be exponential. Along with the increased application of imaging comes ever-increasing volumes of image data that are produced and managed to support research.

In human subjects research and clinical medicine, informatics system designers must also consider protections for the use and dissemination of human patient data. Data encryption protocols are required for the remote use of and access to electronic Protected Health Information (PHI) as dictated by U.S. Department of Health and Human Services (HHS). In 1996, HHS issued patient privacy protections as part of the Health Insurance Portability and Accountability Act (HIPAA). HIPAA included provisions designed to encourage electronic transactions but also required new safeguards to protect the security and confidentiality of PHI, establishing national standards for electronic health care transactions. HHS also addressed the security and privacy of PHI, including provisions for the levying of civil monetary penalties for violation of HIPAA regulations. Therefore, the encryption and protection of PHI in the form of identifiable medical images must remain a top priority in the development of image informatics databases. Ongoing QA,

accessibility, and procedure analyses will be required to protect against HIPAA violations.

There are a number of topics in the field of retinal image informatics that require further investigation and are potential subjects for future research. Networking issues include, for example, maximizing encryption measures to assure HIPAA compliance, optimal image resolution, and file compression standards that minimize bandwidth consumption without corruption or loss of important clinical information. Prospective, blinded clinical trials are required to validate the potential of automated analysis and prescreening throughput. Other applications that may leverage the power of image informatics applied to clinical medicine include the development of searchable genotype/phenotype repositories of retinal and other diseases for which digital image data can be acquired.

9.7 Conclusion

We reviewed and presented many methods and techniques related to the development of bioinformatics systems in the preclinical and clinical environments. Multiple modes of imaging are available to biomedical researchers and physicians alike that provide detailed, high-resolution visualizations of anatomy, pathology, and function. These planar, volumetric, and longitudinal images and sequences provide biomedical information content that can be explored immediately or indexed in a data system to perform research, medical diagnostics, teaching, and training using CBIR methods. The primary philosophy described to establish a searchable archive based on image content begins with specification of end-use goals for the system, followed by subregion image analysis to localize regions of imagery that represent the state of the animal or patient under study. Subsequent to subregion segmentation, these regions are extensively described using statistical features. These features are then mapped to a content-based index using the various categories of anatomy, pathology, and other physiological attributes of the data defined by the end users of the system. This semantic index becomes the primary representation of the image data in the system, providing the visual evidence required to perform queries (e.g., to generate query-based posterior probability estimates that describe query data based on the historical information encoded in the archive population).

Finally, we reviewed and presented both the technologies of CBIR and some of the issues related to the effective and productive use of image content in informatics systems. These systems will continue to evolve to address future needs that are only now beginning to emerge, from the semantic protection of human patient data to the description and integration of imagery across mode, scale, and time. In general, methods that can consistently self-describe

imagery based on visual content will reduce tedious human interaction and make image data much more valuable and accessible to the overall process of biomedical preclinical and clinical research.

Acknowledgments

We would like to acknowledge the Plough Foundation, Memphis, Tennessee, and Research to Prevent Blindness, New York, and the National Eye Institute under grant EY017065 for their support of the retinal components of the presented research. The small animal imaging and amyloid research was supported in part by the National Institutes of Health under grant EB00789. This paper was prepared by the Oak Ridge National Laboratory, Oak Ridge, Tennessee, operated by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

References

[1] J. M. Bueno, F. Chino, A. J. M. Traina, and C. Traina Jr. How to add content-based image retrieval capability in a PACS. In *Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002)*, 2002.

[2] C. P. Schultz, et al. Molecular imaging portal: New development IT platform for imaging, nonimaging and genomics. *Molecular Imaging*, 4: 71–77, 2005.

[3] C. Le Bozec, E. Zapletal, M. Jaulent, D. Heudes, and P. Degoulet. Towards content-based image retrieval in a HIS-integrated PACS. In *Proceedings AIMA Symposium*, 2000, pp. 477–81.

[4] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18, 1996.

[5] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22: 1349–80, 2000.

[6] P. Pala and S. Santini. Image retrieval by shape and texture. *Pattern Recognition*, 32: 517–27, 1999.

[7] L. R. Long, S. K. Antani, and G. R. Thoma. Image informatics at a national research center. *Computerized Medical Imaging and Graphics*, 29: 171–93, 2005.

[8] S. Santini. *Exploratory Image Databases, Content-based Retrieval*. San Diego: Academic Press, 2001.

[9] J. P. Eakins. Towards intelligent image retrieval. *Pattern Recognition*, 35: 3–14, 2002.

[10] Y. Liu, D. Zhang, G. Lu, and W. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40: 262–82, 2007.

[11] C.-R. Shyu, C. Pavlopoulou, A. C. Kak, and C. E. Brodley. Using human perceptual categories for content-based retrieval from a medical image database. *Computer Vision and Image Understanding*, 88: 119–51, 2002.

[12] K. W. Tobin, D. Aykac, V. P. Govindasamy, S. S. Gleason, J. Gregor, T. P. Karnowski, J. R. Price, and J. Wall. Image-based informatics for preclinical biomedical research. In *2nd International Symposium on Visual Computing*, 2006, pp. 1740–50.

[13] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of content-based image retrieval systems in medical applications: clinical benefits and future directions. *International Journal of Medical Informatics*, 73: 1–23, 2004.

[14] M. Rahman, P. Bhattacharya, and B. C. Desai. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. *IEEE Transactions on Information Technology in Biomedicine*, 11: 58–69, 2007.

[15] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

[16] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[17] E. de Ves, J. Domingo, G. Ayala, and P. Zuccarello. A novel Bayesian framework for relevance feedback in image content-based retrieval systems. *Pattern Recognition*, 39: 1622–32, 2006.

[18] L. Duan, W. Gao, W. Zeng, and D. Zhao. Adaptive relevance feedback based on Bayesian inference for image retrieval. *Signal Processing*, 85: 395–99, 2005.

[19] K. A. Heller and Z. Ghahramani. A simple Bayesian framework for content-based image retrieval. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, pp. 2110–17.

[20] J. Luo, A. E. Savakis, and A. Singhal. A Bayesian network-based framework for semantic image understanding. *Pattern Recognition*, 38: 919–34, 2005.

[21] D. W. Holdsworth and M. M. Thornton. Micro-CT in small animal and specimen imaging. *Trends in Biotechnology*, 20: S34–S39, 2002.

[22] T. Lee. Functional CT: Physiological models. *Trends in Biotechnology*, 20: S3–S10, 2002.

[23] J. Stypmann. Doppler ultrasound in mice. *Echocardiography*, 24: 97–112, 2007.

[24] P. K. Marsden. Detector technology challenges for nuclear medicine and PET. *Nuclear Instruments & Methods in Physics Research A*, 513: 1–7, 2003.

[25] S. R. Meikle, P. Kench, M. Kassiou, and R. B. Banati. Small animal SPECT and its place in the matrix of molecular imaging technologies. *Physics in Medicine and Biology*, 50: R45–61, 2005.

[26] I. Voo, E. C. Mavrofrides, and C. A. Puliafito. Clinical applications of optical coherence tomography for the diagnosis and management of macular diseases. *Ophthalmology Clinics of North America*, 17: 21–31, 2004.

[27] P. F. Sharp and A. Manivannan. The scanning laser ophthalmoscope. *Physics in Medicine and Biology*, 42: 951–66, 1997.

[28] H. Barthel, M. C. Cleij, D. R. Collingridge, O. C. Hutchinson, S. Osman, Q. He, S. K. Luthra, F. Brady, P. M. Price, and E. O. Aboagye. 3′-deoxy-3′-[18F]fluorothymidine as a new marker for monitoring tumor response to antiproliferative therapy in vivo with positron emission tomography. *Cancer Research*, 63: 3791–98, 2003.

[29] C. P. Reynolds, B. C. Sun, Y. A. DeClerck, and R. A. Moats. Assessing growth and response to therapy in murine tumor models. *Methods in Molecular Medicine*, 111: 335–50, 2005.

[30] G. Korpanty, J. G. Carbon, P. A. Grayburn, J. B. Fleming, and R. A. Brekken. Monitoring response to anticancer therapy by targeting microbubbles to tumor vasculature. *Clinical Cancer Research*, 13: 323–30, 2007.

[31] W. Zhang, S. Oya, M. P. Kung, C. Hou, D. L. Maier, and H. F. Kung. F-18 stilbenes as PET imaging agents for detecting beta-amyloid plaques in the brain. *Journal of Medicinal Chemistry*, 48: 5980–88, 2005.

[32] J. S. Wall, M. J. Paulus, S. Gleason, J. Gregor, A. Solomon, and S. J. Kennel. Micro-imaging of amyloid in mice. *Methods in Enzymology*, 412: 161–82, 2006.

[33] W. B. Mathews, T. G. Zober, H. T. Ravert, U. Scheffel, J. Hilton, D. Sleep, R. F. Dannals, and Z. Szabo. Synthesis and in vivo evaluation of a PET radioligand for imaging the endothelin-A receptor. *Nuclear Medicine and Biology*, 33: 15–19, 2006.

[34] P. Vaska, D. J. Rubins, D. L. Alexoff, and W. K. Schiffer. Quantitative imaging with the micro-PET small-animal PET tomograph. *International Review of Neurobiology*, 73: 191–218, 2006.

[35] S. G. Nekolla. Dynamic and gated PET: Quantitative imaging of the heart revisited. *Nuklearmedizin*, 44 (Suppl 1): S41–5, 2005.

[36] Y. Du, B. M. Tsui, and E. C. Frey. Model-based compensation for quantitative 123I brain SPECT imaging. *Physics in Medicine and Biology*, 51: 1269–82, 2006.

[37] J. S. Wall, S. J. Kennel, M. Paulus, J. Gregor, T. Richey, J. Avenell, J. Yap, D. Townsend, D. T. Weiss, and A. Solomon. Radioimaging of light chain amyloid with a fibril-reactive monoclonal antibody. *Journal of Nuclear Medicine*, 47: 2016–24, 2006.

[38] T. M. Blodgett, C. C. Meltzer, and D. W. Townsend. PET/CT: form and function. *Radiology*, 242: 360–85, 2007.

[39] J. Gaa, E. J. Rummeny, and M. D. Seemann. Whole-body imaging with PET/MRI. *European Journal of Medical Research*, 9: 309–12, 2004.

[40] J. S. Wall, S. J. Kennel, M. J. Paulus, G.S.S., J. Gregor, J. Baba, M. Schell, T. Richey, B. O'Nuallain, R. Donnell, P. N. Hawkins, D. T. Weiss, and A. Solomon. Quantitative high-resolution microradiographic imaging of amyloid deposits in a novel murine model of AA amyloidosis. *Amyloid*, 12: 149–56, 2005.

[41] A. F. Amos, D. J. McCarty, and P. Zimmet. The rising global burden of diabetes and its complications: Estimates and projections to the year 2010, *Diabetic Medicine*, 14: S57–85, 1997.

[42] National Diabetes Fact Sheet, Centers for Disease Control and Prevention, http://www.cdc.gov 2003.

[43] J. C. Javitt, L. P. Aiello, L. J. Bassi, Y. P. Chiang, and J. K. Canner. Detecting and treating retinopathy in patients with type I diabetes mellitus: Savings associated with improved implementation of current guidelines. *Ophthalmology*, 98: 1565–73, 1991.

[44] J. Staal, M. D. Abrámoff, M. Niemeijer, and B. V. Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23: 501–9, 2004.

[45] A. Hoover and M. Goldbaum. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22: 951–58, 2003.

[46] M. Foracchia, E. Grisan, and A. Ruggeri. Detection of the optic disk in retinal images by means of a geometrical model of vessel structure. *IEEE Transactions on Medical Imaging*, 23: 1189–95, 2004.

[47] H. Li and O. Chutatape. Automated feature extraction in color retinal images by a model based approach. *IEEE Transactions on Biomedical Engineering*, 51: 246–54, 2004.

[48] A. A. Cavallerano, J. D. Cavallerano, P. Katalinic, A. M. Tolson, L. P. Aiello, and L. M. Aiello. Joslin Vision Network Clinical Team. Use of Joslin Vision Network digital-video nonmydriatic retinal imaging to assess diabetic retinopathy in a clinical program. *Retina*, 23: 215–23, 2003.

[49] S. R. Fransen, T. C. Leonard-Martin, W. J. Feuer, and P. L. Hildebrand. Inoveon Health Research Group. Clinical evaluation of patients with diabetic retinopathy: Accuracy of the Inoveon diabetic retinopathy-3DT system. *Ophthalmology*, 109: 595–601, 2002.

[50] M. J. Paulus, S. S. Gleason, M. E. Easterly, and C. J. Foltz. A review of high-resolution X-ray computed tomography and other imaging modalities for small animal research. *Laboratory Animals (NY)*, 30: 36–45, 2001.

[51] M. J. Paulus, S. S. Gleason, S. J. Kennel, P. R. Hunsicker, and D. K. Johnson. High resolution X-ray computed tomography: An emerging tool for small animal cancer research, *Neoplasia*, 2: 62–70, 2000.

[52] P. N. Hawkins and M. B. Pepys. Imaging amyloidosis with radiolabelled SAP. *European Journal of Nuclear Medicine*, 22: 595–99, 1995.

[53] J. Price, D. Aykac, and J. Wall. A 3D level sets method for segmenting the mouse spleen and follicles in volumetric microCT images. In *IEEE Engineering in Medicine and Biology Conference*, 2006, pp. 2332–36.

[54] J. Gregor, S. Gleason, M. Paulus, and J. Cates. Fast Feldkamp reconstruction based on focus of attention and distributed computing. *International Journal of Imaging Systems and Technology*, 12: 229–34, 2002.

[55] J. Gregor, T. Benson, S. Gleason, M. Paulus, S. Figueroa, T. Hoffman, S. Kennel, and J. Wall. Approximate volumetric system models for microSPECT. *IEEE Transactions on Nuclear Science*, 53: 2646–52, 2006.

[56] M. Larsen, J. Godt, and M. Grunkin. Automated detection of diabetic retinopathy in a fundus photographic screening population. *Investigative Ophthalmology & Visual Science*, 44: 767–71, 2003.

[57] S. C. Lee, E. T. Lee, R. M. Kingsley, Y. Wang, D. Russell, R. Klein, and A. Warn. Comparison of diagnosis of early retinal lesions of diabetic retinopathy between a computer system and human experts. *Archives of Ophthalmology*, 119: 509–15, 2001.

[58] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine*, 19: 105–12, 2002.

[59] T. Chanwimaluang, G. Fan, and S. R. Fransen. Hybrid retinal image registration. *IEEE Transactions on Information Technology in Biomedicine*, 10: 129–42, 2006.

[60] J. S. Schuman, C. A. Puliafito, and J. G. Fujimoto. *Optical Coherence Tomography of Ocular Diseases*, 2nd ed. Thorofare, NJ: Slack, 2004.

[61] D. Aykac, J. Price, and J. Wall. 3D segmentation of the mouse spleen in microCT via active contours. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Puerto Rico, 2005, pp. 1542–45.

[62] J. Price, D. Aykac, and J. Wall. Improvements in level set segmentation of 3D small animal imagery. In *SPIE Medical Imaging: Image Processing*, 2007, pp. 651233-1-10.

[63] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79: 12–49, 1988.

[64] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10: 266–77, 2001.

[65] G. Papandreou and P. Maragos. Multigrid geometric active contour models. *IEEE Transactions on Image Processing*, 16: 229–40, 2007.

[66] X. Han, C. Xu, and J. Prince. A topology preserving level set method for geometric deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25: 755–68, 2003.

[67] G. Sundaramoorthi and A. Yezzi. Global regularizing flows with topology preservation for active contours and polygons. *IEEE Transactions on Image Processing*, 16: 803–12, 2007.

[68] D. Cremers, O. Fluck, M. Rousson, and S. Aharon. A probabilistic level set formulation for interactive organ segmentation. In *SPIE Medical Imaging: Image Processing*, 2007, pp. 65120V-1–9.

[69] T. P. Karnowski, V. P. Govindasamy, K. W. Tobin, and E. Chaum. Locating the optic nerve in retinal images: Comparing model-based and Bayesian decision methods. *28th Annual International Conference of the IEEE EMBS*, 2006.

[70] K. W. Tobin, E. Chaum, V. P. Govindasamy, T. P. Karnowski, and O. Sezer. Characterization of the optic disk in retinal imagery using a probabilistic approach. In *SPIE International Symposium on Medical Imaging*, San Diego, CA, 2006.

[71] T. Teng, M. Lefley, and D. Claremont. Progress towards automatic diabetic ocular screening: A review of image analysis and intelligent systems for diabetic retinopathy. *Medical & Biological Engineering & Computing*, 40: 2–13, 2002.

[72] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Investigative Ophthalmology & Visual Science*, 47: 1120–25, 2006.

[73] M. Niemeijer, M. D. Abramoff, and B. van Ginneken. Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. *Medical Image Analysis*, 10: 888–98, 2006.

[74] J. Kristinsson, M. Gottfredsdottir, and E. Stefansson. Retinal vessel dilation and elongation precedes diabetic macular oedema. *British Journal of Ophthalmology*, 81: 274–78, 1997.

[75] R. Klein, S. Moss, B. Klein, M. Davis, and D. Demets. The Wisconsin epidemiologic study of diabetic retinopathy: XI. The incidence of macular edema. *Ophthalmology*, 96: 1501–10, 1989.

[76] B. Liesenfeld, E. Kohner, W. Piehlmeier, S. Kluthe, M. Porta, T. Bek, M. Obermaier, H. Mayer, G. Mann, R. Holle, and K. Hep. A telemedical approach to the screening of diabetic retinopathy: Digital fundus photography. *Diabetes Care*, 23: 345–48, 2000.

[77] J. H. Hipwell, F. Strachan, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester. Automated detection of microaneurysms in digital red-free photographs: A diabetic retinopathy screening tool. *Diabetic Medicine*, 17: 588, 2000.

[78] M. H. Goldbaum, N. P. Katz, M. R. Nelson, et al. The discrimination of similarly colored objects in computer images of the ocular fundus. *Investigative Ophthalmology & Visual Science*, 31: 617–23, 1990.

[79] N. Ward, S. Tomlinson, and C. Taylor. Image analysis of fundus photographs. *Ophthalmology*, 96: 80–86, 1989.

[80] B. Kochner, D. Schulman, M. Obermaier, G. Zahlmann, G. Mann, and K. H. Englmeier. An image processing system for analysing color fundus photographs with regard to diabetic retinopathy. *Klinische Monatsblatter für Augenheilkunde (in German)*, 211: 11, 1997.

[81] G. Zahlmann, B. Kochner, I. Ugi, D. Schulhmann, B. Liesenfeld, A. Wegner, M. Obermaier, and M. Mertz. Hybrid fuzzy image processing for situation assessment: A knowledge-based system for early detection of diabetic retinopathy. In *IEEE Engineering in Medicine and Biology*, 2000, pp. 76–83.

[82] K. A. Goatman, A. D. Whitwam, A. Manivannan, J. A. Olson, and P. F. Sharp. Colour normalisation of retinal images. In *Proceedings of Medical Image Understanding and Analysis*, 2003.

[83] Z. Ben Sbeh, L. D. Cohen, G. Mimoun, and G. Coscas. A new approach of geodesic reconstruction for drusen segmentation in eye fundus images. *IEEE Transactions on Medical Imaging*, 20: 1321–33, 2001.

[84] K. W. Tobin, E. Chaum, V. P. Govindasamy, and T. P. Karnowski. Detection of anatomic structures in human retinal imagery. *IEEE Transactions on Medical Imaging*, 26: 1729–39, 2007.

[85] M. Niemeijer, B. van Ginneken, J. Staal, M. S. Suttorp-Schulten, and M. Abramoff. Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 24: 584–92, 2005.

[86] A. Bosch, X. Munoz, and R. Marti. Which is the best way to organize/classify images by content? *Image and Vision Computing*, 25: 778–91, 2007.

[87] K. Tobin, B. Bhaduri, E. Bright, A. Cheriydat, T. Karnowski, P. Palathingal, T. Potok, and J. Price. Automated feature generation in large-scale geospatial libraries for content-based indexing. *Photogrammetric Engineering & Remote Sensing*, 72: 531–40, 2006.

[88] K. Fukunaga. *Introduction to Statistical Pattern Recognition*, 2nd ed. London: Academic Press, 1990.

[89] N. Lawrence. Probabilistic non-linear principal component analysis with Gaussian process latent variable models. *Journal of Machine Learning Research*, 6: 1783–1816, 2005.

[90] D.-Y. Cho and B.-T. Zhang. Continuous estimation of distribution algorithms with probabilistic principal component analysis. In *Proceedings of the IEEE Conference on Evolutionary Computation*, 2001, pp. 521–26

[91] J. R. Price and T. F. Gee. Face recognition using direct, weighted linear discriminant analysis and modular subspaces. *Pattern Recognition*, 38: 209–19, 2005.

[92] C. H. Chen, L. F. Pau, and P. S. P. Wang. *Handbook of Pattern Recognition and Computer Vision*. Salem, MA: World Scientific, 1993.

[93] A. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30: 1145–59, 1997.

[94] V. N. Gudivada and V. V. Raghavan. Content-based image retrieval systems. *IEEE Computer Magazine*, 8: 18, 1995.

[95] K. W. Tobin, T. P. Karnowski, L. F. Arrowood, R. K. Ferrell, J. S. Goddard, and F. Lakhani. Content-based image retrieval for semiconductor process characterization. *Journal on Applied Signal Processing*, 2002: 704–13, 2002.

[96] Digital Imaging and Communications in Medicine (DICOM) – PS 3.1-2001, National Electrical Manufacturers Association, 1300 N. 17th Street, Rosslyn, VA 22209 USA, 2001.

[97] Risk factors associated with age-related macular degeneration. A case-control study in the age-related eye disease study: Age-related eye disease study report number 3. *Ophthalmology*, 107: 2224–32, 2000.

[98] Photocoagulation for diabetic macular edema: Early treatment diabetic retinopathy study report number 1. *Archives of Ophthalmology*, 103: 1796–1806, 1985.

[99] Preliminary report on effects of photocoagulation therapy. *American Journal of Ophthalmology*, 81: 1–15, 1976.

[100] Baseline and early natural history report. *Archives of Ophthalmology*, 111: 1087–95, 1993.

[101] T. M. Deserno, S. Antani, and R. Long. Gaps in content-based image retrieval. In *Medical Imaging: PACS and Imaging Informatics*, 2007, p. 65160J.

[102] D. Tahmoush and H. Samet. A Web collaboration system for content-based image retrieval of medical images. In *Medical Imaging: PACS and Imaging Informatics*, 2007, p. 65160E.

[103] T. M. Deserno, S. Antani, and R. Long. Exploring access to scientific literature using content-based image retrieval. In *Medical Imaging: PACS and Imaging Informatics*, 2007, p. 65160L.

[104] C.-L. Poh, R. I. Kitney, and R. Shrestha. Addressing the future of clinical information systems: Web-based multilayer visualization. *IEEE Transactions on Information Technology in Biomedicine*, 11: 127–140, 2007.

Chapter 10

Parts-Based Appearance Modeling of Medical Imagery*

Matthew Toews and Tal Arbel

Contents

*Portions reprinted, with permission, from "A Statistical Parts-Based Appearance Model of Anatomical Variability," Matthew Toews and Tal Arbel, *IEEE Transactions on Medical Imaging*, *Special Issue on Computational Neuroanatomy*, Volume 26, Number 4, pages 497–508.

In this chapter, we present a statistical parts-based model (PBM) of image appearance designed to address the problem of modeling natural intersubject anatomical variability in medical imagery. In contrast to global image models, the PBM consists of a collection of localized image regions, called *parts*, whose appearance, geometry, and occurrence frequency are quantified statistically. The parts-based approach explicitly addresses the case in which one-to-one correspondence does not exist between all subjects in a population due to anatomical differences, as model parts are not required to appear in all subjects. The model is constructed through an automatic machine learning algorithm, identifying image patterns that appear with statistical regularity in a large collection of subject images with minimal manual input. Because model parts are represented by generic image features, the model can be applied in a wide variety of contexts to describe the anatomical appearance of a population or to determine intersubject registration. Experimentation based on 2-D magnetic resonance (MR) slices of the human brain shows that a PBM learned from a set of 102 subjects can be robustly fit to 50 new subjects with accuracy comparable to three human raters. Additionally, it is shown that, unlike global models, PBM fitting is stable in the presence of unexpected, local perturbation.

10.1 Introduction

The rapid development of medical imaging devices has led to a proliferation of medical image data. Where human experts once performed a detailed analysis of individual images, the sheer volume of data currently being generated can be better addressed using automated image analysis techniques for routine applications such as image data storage and querying as well as for intelligent applications such as registration, classification, and physiological discovery. The research community has begun responding to this paradigm shift in the way medical image data are analyzed, spawning new research fields such as computational anatomy. A major underlying theme of many medical imaging applications, from intersubject registration to morphometric analysis, is modeling the appearance of the images in question. An appearance model should be able to effectively represent salient features of the population of interest in the presence of noise and intersubject variation, robustly identify if and where these features are present in new subjects, and be efficient to learn from large image databases with minimal manual intervention.

In this chapter, we describe a general modeling technique that addresses these issues, first presented by Toews and Arbel [47]. The technique is parts based, describing image appearance in terms of a set of localized image patches or parts, each of which is quantified probabilistically in terms of geometry, appearance, and occurrence statistics. Model parts are based on invariant image features, localized image patterns that can be reliably extracted from imagery in the presence of a range of image intensity and geometrical

deformations. Such features reduce computational requirements by reducing large data sets into small sets of features that can be automatically and robustly extracted from images. Furthermore, invariant features represent image regions that are informative both in terms of image appearance and geometry and can be used to form a rich description of medical imagery. Because of the geometrical description and distinctive appearance inherent to invariant features, the model can be learned automatically from large image databases with minimal manual supervision.

In Section 10.2, we outline work related to this chapter, including the general problem of intersubject medical image registration and the potential for invariant image feature methods to resolve issues of modeling and computational efficiency. In Section 10.3, we describe probabilistic parts-based modeling of invariant features, which can be generally applied to describing medical imagery from large data sets. In Section 10.4, we present the results of a case study modeling MR brain images of a database of 152 different human subjects. We finish with a discussion in Section 10.5, an outline of future research directions in Section 10.6, and a conclusion in Section 10.7. Although the PBM we present is generally applicable in a wide variety of medical imaging contexts and the underlying theory is independent of image dimension, we provide an illustrative example using 2-D MR brain imagery throughout the chapter.

10.2 Related Work

To automatically analyze a set of images of different subjects of the same anatomical structure or population, they must first be aligned within a common reference frame in order to compare corresponding tissues and structures. This requires addressing the task of intersubject registration, or determining correspondence between different subjects of the same population. Despite decades since the emergence of medical image analysis, intersubject registration remains an open area of research, due primarily to the phenomenon of intersubject appearance variability, or the manner in which images of different subjects vary within a population. Anatomical structure or tissue may exhibit significant appearance variation from one subject to the next or may simply not exist in all subjects, in the case of multiple structural morphologies or pathology, for example. Effectively coping with and quantifying intersubject variability is of great importance to the medical imaging community because it lies at the heart of understanding how anatomical structure varies within a population. In this section, we provide a brief overview of intersubject registration and describe the relevance of recent local invariant feature methods to medical image registration.

10.2.1 Difficulty of intersubject registration

In order to quantify variability between subjects in a population, correspondence must first be established between subjects via intersubject

registration. In general, most registration techniques attempt to determine a one-to-one mapping from one subject to the next [22,39], driven by a measure of image similarity. Low-parameter linear registration techniques are simple and can determine initial coarse alignment between subjects but do not properly account for intersubject variation on a local scale. Deformable registration techniques involving the estimation of highly parameterized deformation fields from one image to the next have been proposed to account for variation on a local scale. Different deformable registration formulations can generally be distinguished by the manner in which the deformation field is constrained or regularized [25]; examples include elastic registration [2], fluid registration [6], finite element models [21], and thin plate splines [5].

In the case of intersubject registration, however, it is not clear that the estimated deformation fields represent meaningful correspondence between underlying tissues [39]. Different regularization approaches will generally result in different deformation field solutions for a given pair of subjects, particularly in regions where correspondence is ambiguous, such as areas of homogenous image intensity. In general, the higher the number of deformation parameters, the more ways a deformation field can be constructed between any two subjects to obtain a near-perfect mapping in terms of pixel intensity error. This could explain the proliferation of novel highly parameterized registration formulations in the medical imaging literature that report low error in terms pixel intensity. Recently, quantitative comparison of six different intersubject registration methods showed no significant difference between high- and low-parameter registration approaches [25], in which case the principle of Occam's razor [17] suggests that the simpler, reduced parameter model would be more plausible.

In general, these difficulties illustrate the challenge of coping with intersubject variability in registration, particularly in the absence of a gold standard for verification of intersubject registration. Given these difficulties, the a priori assumption of the existence of a one-to-one mapping between different subjects is often difficult to justify, considering highly variable regions such as the cortex, the cases of pathology and tissue resection, and similar issues. We propose that intersubject registration relax the requirement for a one-to-one mapping and focus instead on identifying distinct patterns of local image structure that occur with a certain probability within a population. Correspondence via local invariant image features presented in the following section offers an efficient means of achieving this goal.

10.2.2 Generic invariant feature correspondence

Effectively determining correspondence between different subjects requires addressing (a) the situation in which one-to-one correspondence is ambiguous or nonexistent, as highlighted in the previous section, and (b) the issue of computational efficiency for processing large data sets. Feature-based correspondence offers a solution to these issues by focusing on matching a set of

local, informative image regions between images. Because processing is based on a small set of localized features instead of entire images, correspondence can be calculated from a fraction of the total image data and can be avoided in regions where a valid correspondence may not exist. Features used in medical imaging can be domain specific, designed for identifying particular structures of interest, such as particular brain folds or sulci [40], or they can be generic, identifying natural salient patterns in arbitrary imagery. The generic feature approach we describe has the advantage that it can be applied in a variety of different imaging contexts.

Automatic, generic feature detection for image correspondence is based on the notion that a small set of distinct, localizable image patterns can be reliably extracted and matched from one image to the next. In the computer vision literature, early approaches to identifying generic features focused on automatically identifying the image location x of patterns, such as corners (e.g., the Harris detector based on analysis of image derivatives within a local window [24]). The development of scale-space theory, however, showed that the notion on image feature is intimately tied to the resolution or scale σ at which the images are processed [29]. Scale-invariant features were subsequently developed to address the issue of image scale, localizing image patterns in terms of their scale as well as their location [42]. Recent invariant feature detectors are invariant to linear intensity variation, in addition to geometrical deformations such as translation, orientation, scale, and affine transformations [31,35] and have been used in a wide range of contexts (e.g., natural scenes, faces, cars).

The general invariant feature approach is based on detecting image patterns in a manner such that they can be automatically normalized with respect to intensity and geometric variation, at which point they can be efficiently matched between images without requiring an explicit search over deformation parameters. Invariant feature detection requires a search over the geometrical parameters of the image transformation under which features are invariant, which for scale-invariant features are image location x, scale σ, and orientation. Invariant feature detection in R^N image space begins by identifying extrema in an $N+1$-dimensional scale-space function $G(x, \sigma)$ defined over location x and scale σ:

$$\{x_i, \sigma_i\} = \underset{x_i, \sigma_i}{\operatorname{argmax}} \{|G(x, \sigma)|\} \tag{10.1}$$

The process of extrema identification can be efficiently implemented by representing $G(x, \sigma)$ as a scale-space image pyramid [31]. Once parameters of feature location x_i and scale σ_i have been extracted, the local orientation θ_i is determined from histograms of image derivatives within an image window of size proportional to σ_i surrounding location x_i, thereby achieving invariance to orientation changes. Figure 10.1 illustrates scale-invariant features extracted in a sagittal MR image slice.

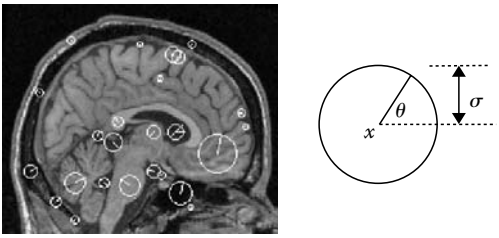


FIGURE 10.1: A subset of scale-invariant features extracted in a sagittal slice of T1-weighted MR brain image. Scale-invariant features, illustrated as white circles inset by radial lines, are oriented regions characterized geometrically by their location x, orientation θ, and scale σ within an image. The geometric characterization results from the detection process and the local image content. The same features can be detected in the presence of image translation, scale, orientation change, and local intensity change. Features here are extracted using the scale-invariant feature transform (SIFT) method [31]. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

A variety of types of invariant features have been proposed in the literature and can be generally differentiated by the scale-space function $G(x, \sigma)$ used, where different functions identify different image characteristics. Often, $G(x, \sigma)$ can be thought of as a blurring kernel applied at image location x with blurring parameter σ. For instance, the scale-invariant feature transform (SIFT) method [31] identifies bloblike image structures as extrema in a difference-of-Gaussian (DOG) scale-space, where the scale of features identified is proportional to the standard deviation of the Gaussian kernel, as illustrated in Figure 10.2. Scale-spaces based on Gaussian spatial derivatives can be used to identify edgelike image structures [33]. Other scale-spaces used have been based on image characteristics such as entropy [27], phase [7], or color moments for multivalued intensities [37]. In general, a variety of invariant feature types can be used simultaneously to provide a rich description of image content. Examples of different invariant feature types identified in the same image can be seen in Figure 10.3.

Once invariant feature geometrical parameters have been identified, they are used to geometrically normalize the image content associated with each feature to a canonical scale and orientation for the purpose of fast correspondence. Feature-to-feature correspondences are based on a measure of image similarity, such as the sum of squared differences or mutual information. Raw image intensities are typically a suboptimal image representation for similarity calculation, however, and most invariant feature methods transform image content into a more discriminative form. For example, the SIFT method [31] represents features in terms of image derivative histograms, thereby both improving the discriminativeness of feature appearance [36] and providing

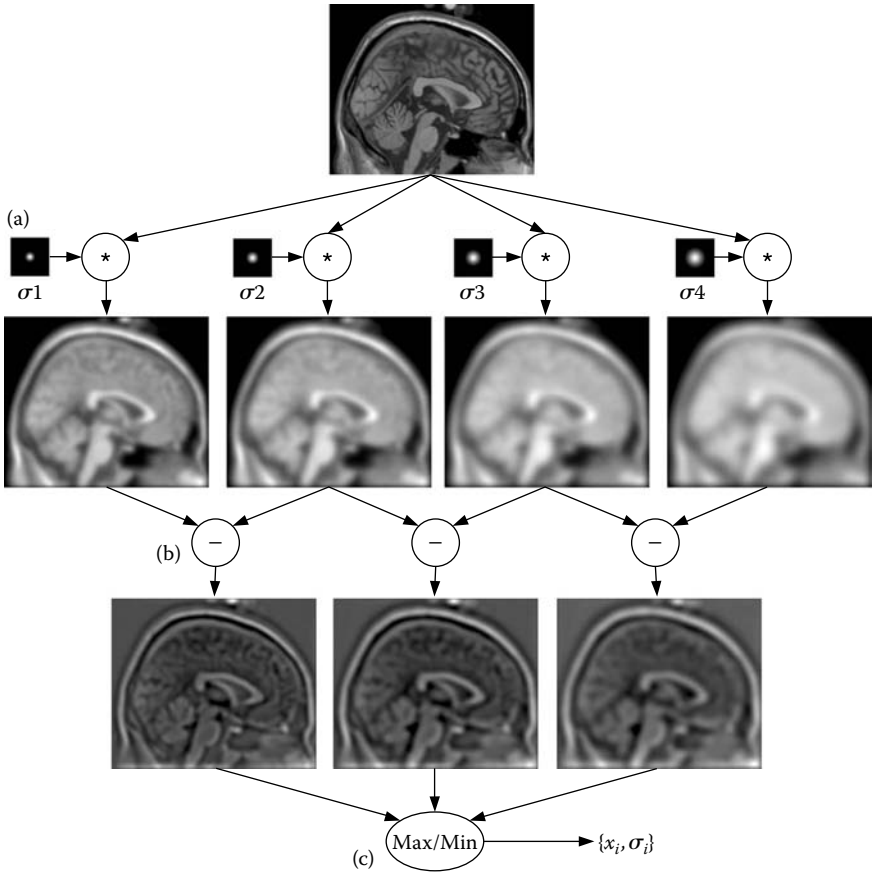


FIGURE 10.2: Illustrating invariant feature extrema detection based on a difference-of-Gaussian scale space. In step (a), the original image is first convolved with Gaussian kernels of varying scale σ, generating Gaussian blurred images. In step (b), the image difference is computed between blurred images of adjacent scales, resulting in difference-of-Gaussian images. Finally, in step (c), maxima and minima are detected between difference-of-Gaussian images in adjacent scales, resulting in a set of invariant feature locations and scales $\{x_i, \sigma_i\}$.

robustness to geometrical deviations arising from the extraction process. In addition to providing fast, search-free correspondence, geometrical parameters can be used with robust techniques such as the Hough transform [31] to achieve occlusion and outlier-resistant correspondence by identifying independent correspondences that agree on the geometrical transform between images.

Despite their attractive properties, scale-invariant features are not widely used for intersubject registration, primarily because of intersubject variability, wherein the variation of features extracted in images of different subjects

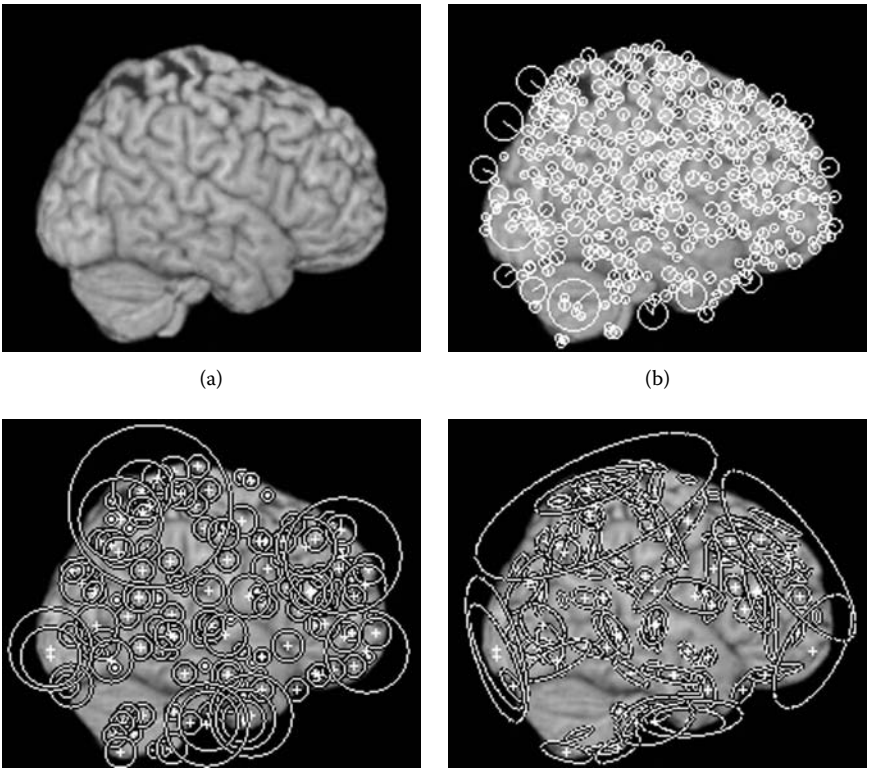


FIGURE 10.3: Various types of invariant features extracted from the same volume rendering of the cortex shown in (a). SIFT features [31], shown in (b), are extracted from a difference-of-Gaussian scale-space and tend to correspond to bloblike features. Scale-invariant Harris features [33], shown in (c), are based on the Harris edge criteria defined by spatial image derivatives and tend to indicate edges or corners. Affine-invariant Harris features [33], shown in (d), are calculated by growing 6-parameter affine regions around scale-invariant Harris features, and they capture elongated regions. Note how large-scale features correspond roughly to larger anatomical structures such as cerebral lobes, while small-scale features arise from smaller structures such as cortical sulci and gyri.

is generally too significant to reliably determine correspondence. As seen in Figure 10.3, many features result from ambiguous or subject-specific image structure and as such are not useful for determining correspondence between subjects. Although this can be considered a shortcoming of generic feature correspondence, we argue that it merely reflects the difficulty of intersubject registration and supports the notion that one-to-one correspondence between different subjects may not generally exist. The PBM presented in the following

section is based on statistically quantifying the occurrence frequency, appearance, and geometry of features over a large set of subjects, thereby learning a set of parts that can be used to determine statistically meaningful correspondence between images of different subjects.

10.3 Probabilistic Appearance Modeling of Invariant Features

In order to evaluate and quantify intersubject feature correspondence in a meaningful way, one must first learn which features can be reliably identified in different subjects of a population, including their appearances, geometries, and occurrence statistics. In this section, we provide an overview of the issues and approaches to modeling medical image appearance and describe a probabilistic PBM derived from invariant image features, which can lead to efficient, robust appearance modeling from large sets of medical imagery.

10.3.1 Appearance modeling: issues and approaches

Given that exact correspondence between subjects of a population is generally difficult and ill posed, research has focused on statistical appearance models to quantify intersubject variability. Statistical models consist of parameters that can be estimated from a set of data in order to quantify variability in a meaningful manner. Parameters of image-based models are typically based on variables of image intensity, image space, and mappings between different images. Because images consist of a large number of data samples (intensities), directly modeling all samples and their interdependencies is generally intractable. Statistical modeling must thus resort to simplifying assumptions regarding parameters and their dependencies, which can be broadly classified according to the image scale at which modeling takes place (i.e., global vs. local models).

Global models refer to those considering the entire spatial extent of the image simultaneously [12,41]. Variation in global models is typically quantified via a linear Gaussian model, consisting of additive "modes" of covariance about a mean. Because all image data are related linearly, the modeling assumption is one of statistical dependence between spatially separated image regions, and the simplification arises from the fact that the majority of data variance can be accounted for by a small number of principal components [50]. Global models can be applied in a variety of ways, such as over combined intensity and shape [12] and over deformation fields between subjects [41]. Local models refer to those that consider variation on a spatially localized scale, typically once images have been brought into alignment, such as morphometry methods [1]. The modeling assumption is that once aligned within a common reference frame, spatially separated image regions can be

treated as statistically independent, and model simplification arises from the fact that only local interdependencies are modeled. Other models, such as Markov models [38], can be considered a hybrid of global and local characteristics, where modeling is based on spatially local interactions that can propagate globally.

A major difficulty in attempting to statistically quantify intersubject variability is that models based either implicitly or explicitly on the assumption of one-to-one correspondence tend to break down when the assumption is invalid. Global models assuming statistical dependency between spatially distinct regions are generally unable to properly account for deformation on a local scale. While this may be less important when considering the shape of simple isolated structures such as the hypothalamus [38], it is problematic when modeling images as a whole where localized variations are commonplace. While local models are able to account for spatially localized phenomena, they are prone to confounding observations of different underlying structure when the assumption of one-to-one correspondence is invalid and are susceptible to error introduced by the method used to obtain alignment, such as bias [26]. To minimize these problems, modeling is often based on subjects free of abnormality [22,49]. It has been observed, however, that even subjects classified as clinically "normal" can vary significantly [41], for example, in cortical folding patterns of the human brain.

An alternative to global modeling or local models assuming the existence of one-to-one correspondence between all subjects is parts-based modeling. Parts-based modeling generally advocates representing a complex image pattern, such as a brain image, in terms of a collection of simpler, independently observable parts [4]. A practical PBM requires a reliable automated means of identifying or detecting model parts in images. A growing body of research in the computer vision community has begun addressing parts-based appearance modeling of general object classes from invariant features, using probability and machine learning theory [15,19,46]. Such modeling has proven to be efficient for learning from large, noisy data sets of natural imagery and robust to complicating factors such as occlusion missing features, although it is only beginning to emerge in the medical imaging field [47,48]. A common probabilistic framework is to consider the conditional posterior probability of an object class of interest o (i.e., a brain in MR imagery) given a set of feature data $\{m_i\}$, which can be expressed as follows:

$$p(o|\{m_i\}) \propto p(o)p(\{m_i\}|o) \tag{10.2}$$

where the proportionality in Equation 10.2 results from Bayes rule, $p(o)$ represents a prior term over the object of interest, and modeling generally focuses on suitably defining the likelihood term $p(\{m_i\}|o)$, based on variables of feature appearance and geometry. While most approaches assume conditional independence between individual feature appearances, they typically differ in the independence assumptions made regarding feature geometries. Independence assumptions include full independence [15,44], naive Bayes dependence

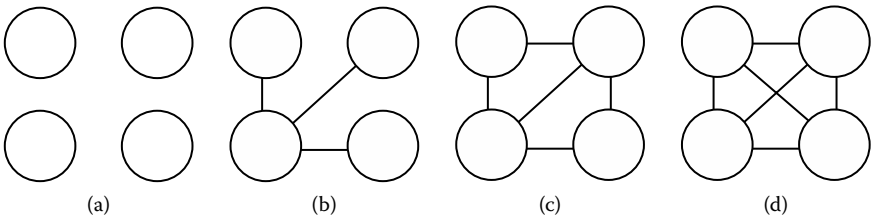


FIGURE 10.4: A graphical illustration of independence assumptions regarding feature geometries. Nodes represent individual feature geometries, and arcs represent statistical dependence. The representational complexity varies from independent models (a) having no interdependencies to fully dependent models (d) having $O(N^2)$ feature-to-feature dependencies. In between are naive Bayes models (b) and Markov models (c).

in which features are conditionally independent given a single parent feature [20,46], Markov dependence in which features are dependent only on those within a neighborhood [8], fully dependent models in which all features are dependent [19,52], and intermediate approaches [14]. These independence assumptions can be visualized graphically in Figure 10.4. While fully independent models are computationally simple, they cannot be used to model a set of features with a consistent global geometrical structure because individual feature geometries are unconstrained. Fully dependent models considering all feature interdependencies are computationally complex and infeasible for rich models consisting of a large number of features. Naive Bayes and Markov models both offer reduced computational complexity, while maintaining the notion of geometrical interdependence.

10.3.2 A parts-based appearance model

In this section, we describe a PBM suitable for learning descriptions of general medical imagery, first published in [47,48]. The model makes the minimal naive Bayes assumption of interfeature geometrical dependence and can thus be learned efficiently from large sets of data. Additionally, the model is well suited to describing the complex nature of anatomy, including cases where one-to-one correspondence is ambiguous or nonexistent, as missing features and multiple distinct modes of appearance are explicitly represented. As mentioned, we carry through an example of modeling brain imagery from 2-D magnetic resonance slices for illustration.

The PBM consists of a set of parts $\{m_i\}$ modeled with respect to a common reference frame o. By reference frame, we mean a geometrical structure with respect to which the geometric variability of m_i can be normalized and quantified. Although the particular reference frame used is application specific and can be chosen in a variety of ways, a well-known example in the context of MR brain imagery is the midplane line defining the Talairach stereotaxic

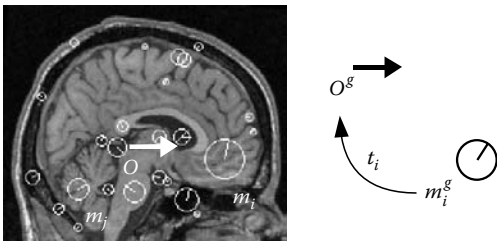


FIGURE 10.5: Illustrating model parts and a reference frame in a sagittal slice of a T1-weighted MR brain image. In the left image, reference frame o is illustrated as a white arrow and represents the projection of the Talairach AC–PC line onto the slice. Reference frame and model part geometry o^g and m_i^g are related via an invertible linear transform $t_i : m_i^g \rightarrow o^g$, as illustrated by the diagram on the right, and thus a single observed part is sufficient to infer the reference frame geometry. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

space [45], which passes from the superior aspect of the anterior commissure (AC) to the inferior aspect of the posterior commissure (PC), as illustrated in Figure 10.5. The significance of the reference frame to statistical modeling is that different parts m_i and m_j can be considered as conditionally independent given o (i.e., the naive Bayes assumption). Specifically, knowing o, parts m_i can be automatically normalized with respect to reference frame scale, rotation and translation, at which point their appearance and remaining geometrical variation can be statistically quantified locally. This is tied closely to the Bookstein definition of shape as the geometric variability remaining in a pattern once global scale, rotation and translation are removed [16].

A model part is denoted as $m_i : \{m_i^b, m_i^g, m_i^a\}$ representing the occurrence, geometry, and appearance, respectively, of a scale-invariant feature within an image. Part occurrence m_i^b is a binary random variable representing the probability of part presence (or absence) in an image. Part geometry $m_i^g : \{x_i, \theta_i, \sigma_i\}$ is an oriented region in $\mathbf{R}^N$ image space, represented by N-parameter location x_i, an $N(N-1)/2$ parameter orientation θ_i, and a scale σ_i. Part appearance m_i^a describes the image content at region m_i^g and can generally be parameterized in a number of ways, such as by using principal components [50].

The reference frame is denoted as $o : \{o^b, o^g\}$ representing the occurrence and geometry of a common structure with respect to which the geometric variability of parts can be quantified. Note that o is identical to a model part with the exception of the appearance component. This is because o cannot be observed directly but must be inferred via model parts m_i. Reference frame occurrence o^b is a binary random variable indicating the presence or absence of the reference frame, the significance of which will be made clear in the discussion of model learning and fitting. Reference frame geometry o^g is parameterized in the same manner as model part geometry m^g, implying that

a single observed part m_i is sufficient to infer o^g in a new image via a learned linear relationship. In the following section, we describe the formulation of the probabilistic model underpinning quantification of part appearance, geometry, and occurrence frequency.

10.3.3 Probabilistic model formulation

The model consists of a set of M parts $\{m_i\}$, which when observed in a new image can be used to infer the reference frame o. Assuming that parts m_i are conditionally independent given o, the posterior probability of o given $\{m_i\}$ can be expressed using Bayes rule as

$$p(o|\{m_i\}) = \frac{p(o)p(\{m_i\}|o)}{p(\{m_i\})} = \frac{p(o)\prod_i^M p(m_i|o)}{p(\{m_i\})} \tag{10.3}$$

where $p(o)$ is a prior over reference frame geometry and occurrence and $p(m_i|o)$ is the likelihood of part m_i given o. The assumption of conditional independence of localized parts generally states that knowing reference frame o, all remaining part variation m_i can be described locally. This assumption represents a significant departure from global modeling approaches, such as the multivariate Gaussian based on principal components, where all image intensities are correlated and statistically dependent. Such global models cannot account for deformation on a local scale, however, where the appearance of one part of the image violates the learned linear model with respect to other parts of the image. We demonstrate this later in experimentation.

Our model focuses principally on the likelihood term $p\,(m_i|o)$, which can be expressed as

$$\begin{aligned} p(m_i|o) &= p(m_i^a, m_i^b|o)p(m_i^g|o) \\ &= p(m_i^a|m_i^b)p(m_i^b|o^b)p(m_i^g|o^b, o^g) \end{aligned} \tag{10.4}$$

under the assumptions that m_i^a and m_i^b are statistically independent of m_i^g given o, and that m_i^a and o are statistically independent given m_i^b. The first of these assumptions generally states that knowing o, the appearance and occurrence of a part offer no further information regarding part geometry. The second assumption generally states that given knowledge of part occurrence, knowing o offers no additional information regarding part appearance. We describe the three terms of Equation 10.4 below.

10.3.3.1 Appearance likelihood $p(m_i^a|m_i^b)$

Appearance likelihood $p(m_i^a|m_i^b)$ represents the appearance of a part after geometry normalization, and can generally be modeled as a multivariate Gaussian distribution in an appearance space with mean and covariance parameters $\mu_i^a, \sum_i^a$. There are two distributions that must be estimated, corresponding to the cases $m_i^{b=1}$ and $m_i^{b=0}$. In the latter case, we take $p(m_i^a|m_i^b)$ to be constant or uniform. We discuss the specific appearance space used in Section 10.4.1.

TABLE 10.1: Occurrence probability $p(m_i^b|o^b)$.

	o^b	m_i^b	**Interpretation**
π_i^0	0	0	Not applicable
π_i^1	0	1	False part occurrence
π_i^2	1	0	Occluded part
π_i^3	1	1	True part occurrence

Source: M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.

10.3.3.2 Occurrence probability $p(m_i^b|o^b)$

Occurrence probability $p(m_i^b|o^b)$ represents part occurrence given reference frame occurrence, and is modeled as a discrete multinomial distribution with event count parameters $\pi_i = \{\pi_i^0, \ldots, \pi_i^3\}$ for the four possible combinations of binary events. Table 10.1 lists the events and their interpretations. Event π_i^0 is not of interest, as neither the part nor the reference frame are present. Event π_i^1 represents the case where part m_i is observed in the absence of reference frame o; this can be viewed as a false part occurrence. Event π_i^2 represents the case where reference frame o is present but part m_i is absent; this is the case where the part is occluded or unobservable due to intersubject variability. Finally, event π_i^3 represents the case that both part m_i and reference frame o are present.

An aspect of the occurrence probability bearing mention at this point is the likelihood ratio of true versus false part occurrences:

$$\frac{p\left(m_i^{b=1}\,|o^{b=1}\right)}{p\left(m_i^{b=1}\,|o^{b=0}\right)} = \frac{\pi_i^3}{\pi_i^1} \tag{10.5}$$

We refer to this ratio as the *distinctiveness* of a part within this chapter, as it provides a measure of how reliably a part can be identified in the presence of noise and false matches [15]. The distinctiveness plays an important role both in automatic model learning and in model fitting, described later.

10.3.3.3 Geometrical likelihood $p(m_i^g|o^b, o^g)$

Geometrical likelihood $p(m_i^g|o^b, o^g)$ models the residual error of a linear transform from part to reference frame geometry $t_i : m_i^g \rightarrow o^g$, and it is represented as a Gaussian distribution with mean and covariance parameters $\mu_i^g, \sum_i^g$. As with the appearance likelihood, there are two distributions corresponding to cases $o^{b=1}$ and $o^{b=0}$. In the latter case, o^g and thus $m_i^g \rightarrow o^g$ are undefined, and we treat the geometrical likelihood as uniform or constant. In order to characterize geometrical error in a scale-invariant manner, the scale dimension is transformed logarithmically, and translation magnitude is normalized by reference frame scale.

10.3.4 Model learning

Learning involves automatically identifying a set of informative model parts $\{m_i\}$ and estimating the parameters of their appearance, geometry, and occurrence frequency distributions based on a set of subject images. Learning begins by labeling a reference frame o^g and automatically extracting features $\{m_i^a, m_i^g\}$ in each training image. For the purpose of the brain modeling example in this chapter, we adopt the AC–PC line defining the Talairach stereotactic reference frame as o^g, as illustrated in Figure 10.5. Other definitions could be used depending on the image context, the sole constraint being that the reference frame represent a stable scale-invariant structure shared by all subjects being modeled. Labeling can be done manually by defining a single line segment corresponding to o^g in each subject image or in an approximate manner via linear registration of MR volumes into the same stereotactic space. We adopt the latter approach, using MR volumes preregistered into the MNI stereotactic space [10].

Parameter estimation is based on a set of data vectors of the form $\{m_i^a, m_i^g, o_i^g\}$, where o_i signifies the reference frame instance associated with feature m_i. A model part corresponds to a cluster of data vectors that are similar in geometry and appearance, and parameter estimation requires identifying these clusters. A variety of clustering techniques could be used for this purpose, but as the data vector space contains a large, unknown number of clusters or modes, approaches such as EM (expectation-maximization) or k-means are ineffective [17] because they require prior knowledge of the number of parts present and tend to either group vectors arising from different clusters or separate vectors arising from the same cluster. We thus adopt a different approach, identifying all isolated clusters in a robust manner similar to the mean shift technique [11]. We do this by treating each feature m_i as a potential cluster center or model part, grouping features into a reduced set of parts, and finally estimating part statistics.

Treating each extracted feature m_i as a potential model part, feature grouping proceeds as follows. The first step is to generate a set G_i of data vectors m_j similar in geometry to m_i, such that the error in predicting the reference frame geometry is less than an empirically determined threshold $Thres^g$:

$$G_i = \{\forall\ m_j : |t_i(m_j^g) - o_j^g| < Thres^g\} \tag{10.6}$$

Recall that $t_i : m_i^g \rightarrow o_i^g$ is a linear transform between feature and reference frame geometries. $t_i(m_j^g)$ thus represents the geometry of reference frame o_j^g as predicted by m_i^g, o_i^g, and m_j^g. $Thres^g = \{T_x, T_\theta, T_\sigma\}$ represents a scale-invariant threshold on the maximum acceptable error permitted in the location, orientation, and scale of the predicted reference frame geometry. Because $Thres^g$ is not related to features themselves but to their ability to predict the reference frame geometry, a single threshold is applicable to all features. We adopt a threshold of $T_x = \frac{\sigma}{2}$ voxels (where σ is the scale of o_j^g),

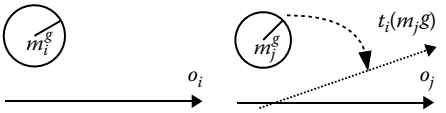


FIGURE 10.6: Illustrating feature geometrical similarity. Here, feature geometries m_i^g and m_j^g differ slightly relative to their respective reference frames o_i^g and o_j^g. Features m_i and m_j are said to be geometrically similar if reference frame o_j^g and its geometry $t_i(m_j^g)$ *(dashed arrow)* as predicted by m_i^g, o_i^g, and m_j^g differ by less than a threshold $Thres^g = \{T_x, T_\theta, T_\sigma\}$ in location, orientation, and scale.

$T_\theta = 15$ degrees, and $T_\sigma = log(1.5)$ (a scaling factor range of 0.66 to 1.5 times). A graphical illustration of geometrical similarity can be seen in Figure 10.6.

The next step is to generate a set A_i of data vectors m_j similar in appearance to m_i, such that the difference between feature appearances m_i^a and m_j^a is less than a threshold $Thres_i^a$:

$$A_i = \left\{ \forall m_j : dist\left(m_i^a, m_j^a\right) < Thres_i^a \right\} \tag{10.7}$$

where $Thres_i^a$ represents a threshold on a function of distance or difference $dist()$ between m_i^a and m_j^a. The particular distance function used can vary depending on the appearance representation used and assumptions regarding the data—for example, Mahalanobis distance [17] or mutual information [51]. Here, $Thres_i^a$ is automatically set to maximize the ratio of features that are similar in appearance and geometry versus features that are similar in appearance but not in geometry:

$$Thres_i^a = \underset{Thres_i^a}{\operatorname{argmax}} \left\{ \frac{|G_i \cap A_i|}{|\bar{G}_i \cap A_i|} \right\} \tag{10.8}$$

Note that the ratio in Equation 10.8 is equivalent to the distinctiveness in Equation 10.5, and $Thres_i^a$ is thus determined to maximize the likelihood of a correct match versus an incorrect match.

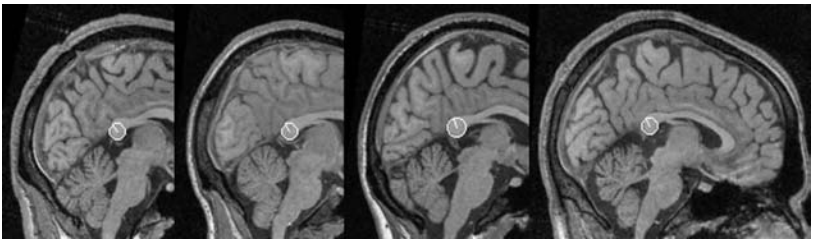
Once G_i and A_i have been determined for each feature m_i, the set of features can be reduced into a small number of representative model parts. There are several mechanisms by which this is possible. Features with arbitrarily low distinctiveness can be removed [46], as they are uninformative. Features with high mutual information with other features can also be removed [3,46], as they are redundant. We generate a set R of features m_j to remove, where m_j are similar to, but occur less frequently than, some other feature m_i:

$$R = \{\forall\ m_j : m_j \in G_i \cap A_i, |G_j \cap A_j| < |G_i \cap A_i|\} \tag{10.9}$$

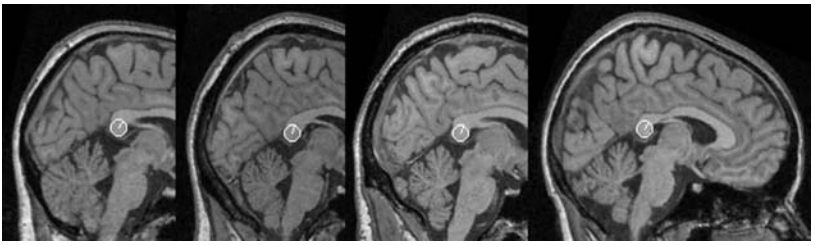
This approach has the effect of discarding redundant features while maintaining those that are most representative as determined by their occurrence

frequency. Feature removal generally reduces the entire feature set by an order of magnitude into a set of model parts. Estimation of part parameters then proceeds as follows. Events o^b and m_i^b are determined by membership in sets G_i and A_i, respectively, allowing the estimation of event count parameters π_i. Geometry and appearance parameters $\{\mu_i^g, \sum_i^g\}$ and $\{\mu_i^a, \sum_i^a\}$ are determined from the geometries and appearances of features in set $G_i \cap A_i$.

It is important to note that this learning process is effective at determining a set of distinct model parts useful for the task of model-to-subject registration, as the ratio of correct to incorrect matches is maximized. Model parts do not necessarily correspond one-to-one with anatomical structures. Features resulting from a single anatomical structure or region, for instance, generally exhibit several stable modes of geometry and appearance due to the interaction of anatomical variability and the particular feature detector used. In such cases, model learning generally represents each mode as a different model part. This is illustrated in Figure 10.7, where the corpus callosum splenium results in



Model part A



Model part B

FIGURE 10.7: Illustrating instances of two different model parts, labeled A and B, arising from the same anatomical structure. In general, the same underlying anatomical structure can give rise to multiple model parts in the learning process, due to the interaction between the feature detector used and the image characteristics of the anatomical structure. Here, the corpus callosum splenium results in features with two significant orientation modes, which are grouped into two distinct model parts by the learning process. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

two different parts with distinct modes of orientation. Relaxing the threshold $Thres^g$ on the permitted geometrical error results in fewer parts, each capable of describing a wider range of variability of the underlying image content, at the cost of decreased part distinctiveness. For the purpose of clarity, we provide a high-level summary of the model learning process in Algorithm 10.1.

Algorithm 10.1 Model learning.

1: Label references frame geometries o^g in training images.
2: Extract invariant features $\{m_i\}$ in training images for each feature m_i do.
3: Identify a set of features G_i similar to m_i in geometry.
4: Identify a set of features A_i similar to m_i in appearance.
5: Discard redundant features.
6: Estimate distribution parameters.

10.3.5 Model fitting

Once the model has been learned, it can be fit to a new image by inferring the geometry of the reference frame o based on features extracted in the new image. Unlike other registration or fitting techniques based on iterative algorithms, which tend to go away when started outside a "capture range" of the optimal solution, the PBM can be fit globally in a robust manner. Fitting begins by first matching features extracted in the new image to the learned model parts. An image feature m is considered to match a model part m_i if $dist(m_i^a, m^a) < Thres_i^a$, as in Equation 10.7. While the reference frame geometry is initially unknown, each matching image feature/model part fixes m_i^a and m_i^g as evidence in the likelihood in Equation 10.4 and infers a hypothesis as to the reference frame geometry o^g in the new image via the learned linear relationship. Dense clusters of similar hypotheses indicate the presence of a reference frame instance, and their corresponding features indicate model-to-subject matches. We are interested in evaluating whether a hypothesis cluster is the result of a true model instance or a random noisy match, that is, $o = \{o^g, o^{b=1}\}$ or $\overline{o} = \{o^g, o^{b=0}\}$. These two hypotheses can be compared via a Bayes decision ratio:

$$\gamma(o) = \frac{p(o|\{m_i\})}{p(\overline{o}|\{m_i\})} = \frac{p(o)}{p(\overline{o})} \prod_{i=1}^{M} \frac{p(m_i|o)}{p(m_i|\overline{o})} \tag{10.10}$$

where high $\gamma(o)$ indicates the presence of a model, and $\frac{p(o)}{p(\overline{o})}$ is a constant representing the expected ratio of true-to-false model instances. Note that the value of $\gamma(o)$ is in large part determined by the part distinctiveness defined in Equation 10.5, as highly distinctive parts will carry greater weight in determining the model fit. Figure 10.8 illustrates the processing of fitting the model to a new image. Note that due to the nature of invariant features, the model

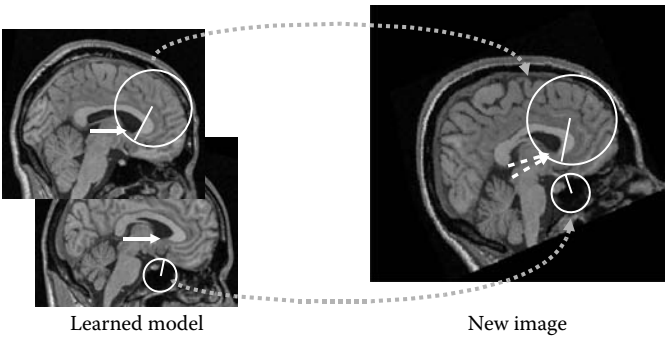


FIGURE 10.8: Illustrating fitting of the parts-based model to a new image based on sagittal slices of MR brain imagery. Fitting begins by matching features in a learned model (left) to those extracted in a new image (right). Each model-to-image correspondence infers a hypothesis as to the reference frame o^g in the new image (dashed white arrows). Clusters of geometrically similar hypotheses indicate the presence of a valid reference frame instance. Note that the new image has been rotated and magnified to illustrate the invariance of model fitting to location, orientation, and scale changes.

can be automatically fit in the presence of global image translation, orientation, and scale changes, in addition to intersubject variability. For the purpose of clarity, we provide a high-level summary of the model fitting process in Algorithm 10.2.

Algorithm 10.2 Model fitting.

1: Extract invariant features m in new image.
2: Match image features to model parts m_i; determine reference frame hypotheses o^g.
3: Cluster reference frame hypotheses.
4: Evaluate the Bayes decision ratio of hypothesis clusters.

10.4 Case Study: MR Brain Image Modeling

The PBM technique we introduced in the previous section can be applied to a wide variety of medical imaging contents, in general, any population that can be described in terms of a collection of distinct, spatially localized anatomical parts. Here, we outline the results of a study modeling the appearance of 2-D sagittal slices of MR brain images of 152 different subjects from the International Consortium on Brain Modeling (ICBM) 152 data set [32], originally published by Toews and Arbel [47].

10.4.1 Learning brain image appearance

For the purpose of experimentation, we learn a PBM from 102 sagittal slices of the ICBM 152 data set using the fully automatic procedure described in Section 10.3.4. As mentioned in Section 10.2.2, a variety of techniques can be used to identify scale-invariant features from a number of image characteristics. For the purpose of this study, we use the SIFT technique [31]. The SIFT feature detector and appearance representation have been shown to perform well in comparisons with other approaches in terms of detection repeatability [43] and appearance distinctiveness [36] and are currently widely used in the computer vision literature. Briefly, SIFT features are extracted as maxima/minima in a DOG scale-space image pyramid. The pyramid location at which features are extracted determines the location and scale components x_i and σ_i of the feature geometry m_i^g. The orientation component θ_i is then determined from peaks in an orientation histogram generated from local image derivatives. Features extracted in the DOG scale-space tend to correspond to bloblike image structures. The SIFT appearance representation m_i^a is a 128-value vector corresponding to bins of a histogram of image first derivatives quantized into $8 \times 4 \times 4 = 128$ bins over orientation and 2-D spatial location, respectively. The distance function $dist()$ used is the Euclidean distance of normalized appearance vectors, which is directly related to the normalized cross-correlation measure and to the Mahalanobis distance in a feature space of independent and identically distributed features [17].

The result of learning is a set of spatially localized parts, including their occurrence frequency, geometrical variability, and appearance variability, a natural and intuitive representation for describing anatomical variability. Figure 10.9 illustrates a graph of the likelihood of model part occurrence in the population, that is, $p(m_i^{b=1}|o^{b=1})$, sorted in order of descending likelihood. Note that a small number of model parts occur in a relatively high percentage of subjects, indicating structure that is stable and repeatably identifiable in many members of the population. Conversely, a relatively large number of model parts are related to subject-specific characteristics or noise, occurring in only a small number of subjects. Such parts are of limited use in describing the anatomy of the population and may be disregarded. Note that, in general, no model parts are detected in all 102 subject brains. This is not to say that common anatomical structures such as the pons or corpus callosum are not present in all brains but that scale-invariant features arising from such structures tend to cluster into several distinct modes of appearance and geometry and are identified as distinct model parts, as illustrated in Figure 10.7. This is a characteristic of our learning approach, which seeks a set of maximally distinctive image patterns that arise from regularity in underlying brain anatomy for the task of intersubject registration.

An efficient SIFT implementation is publicly available [30]. Model parts can be potentially useful in a number of ways. In the following sections, we demonstrate that model parts serve as a basis for robust, stable model-to-subject registration and that a subset of common parts can be used to quantify

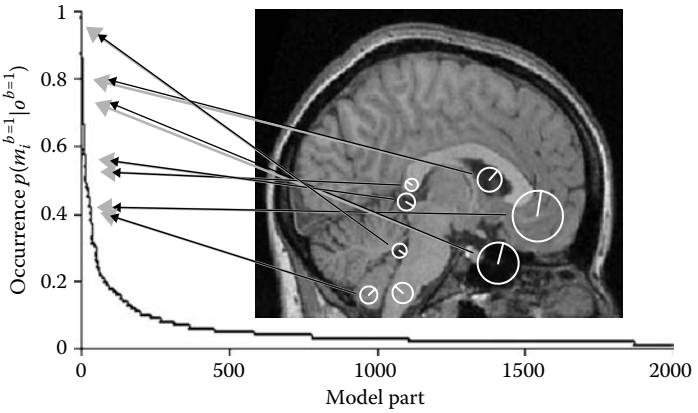


FIGURE 10.9: A graph of learned model parts sorted by descending occurrence frequency $\pi_i^3 = p\left(m_i^{b=1} \left| o^{b=1} \right.\right)$. Features (white circles) overlaying the image illustrate instances of parts that occur at the indicated frequency within the population. Note that part occurrence drops off sharply, indicating that a relatively small number of model parts are common to all brains, whereas a large number are specific to a small number of brains.

model fitting accuracy. Model part statistics can be used in order to interpret the result of model fitting in a meaningful, quantitative manner. For example, the distinctiveness in Equation 10.5 represents the certainty with which a model part can be identified in a new image, and the geometrical likelihood in Section 10.3.3.3 represents the variability that can be expected in localizing a model part in space, orientation, and scale. For the purpose of anatomical study, model parts could be grouped and assigned semantic labels according to their underlying anatomical structures and tissues, after which point model-to-subject registration could be used to propagate these labels in an automatic, robust manner to hundreds of new subjects. It is possible that distributions of part appearance and geometry could serve as useful indicators of disease or abnormality. For example, the geometry of features lying on the corpus callosum could potentially serve as robust indicators of schizophrenia, as in the work of Toga and colleagues [49]. Part geometry and occurrence variables could potentially be used to improve morphometric analysis [1] by improving intersubject alignment and indicating where intersubject alignment may not be valid. Although many of these possibilities fall outside the scope of this chapter, they represent future directions of application for parts-based modeling.

10.4.2 Model-to-subject registration

The goal of model-to-subject registration is to determine how a new subject relates to its population. The PBM does this by identifying the features in

the subject image that are representative of the population, as determined through learning. In this section, we describe model-to-subject registration trials in which the model learned from 102 subjects described in Section 10.4.1 is automatically fitted to the remaining 50 new test subjects not used in learning, via the algorithm described in Section 10.3.5. No gold standard exists for evaluating the ground truth accuracy of intersubject registration [39], and we thus compare the automatic part registration with respect to manual part registration, as established by three different human raters. Since there are many model parts, none of which is generally identified in all subjects, we focus on a set of 25 test parts that occur frequently and throughout the brain during model learning. Fitting trials identify a subset of these 25 parts in each test subject, which serves as the basis for fitting evaluation. The number of test parts identified per subject image is 10 on average and ranges from 4 to 16. The number of instances of each test part identified over all 50 trials is 20 on average and ranges from 4 to 47. In total, 516 part instances are considered. Figure 10.10 illustrates a set of four test subject images, which together contain at least one instance of each test model part.

Since model parts are defined via a fully automatic learning procedure and may not necessarily match obvious anatomical structures, human raters themselves must first be taught the appearances and geometries of the parts before attempting to localize them in new images. This is done by showing the raters images of model part instances identified in different subjects during the model learning process. Specifically, 10 images such as those in Figure 10.7 of a single model part are shown in a looping video sequence. The rater is asked to watch the videos and then determine the model part locations in all test subject images within which the part was identified during model fitting. Note that model parts contain a much richer description than simple spatial location, including orientation and scale information in addition to a measure of distinctiveness. These aspects could also be established and verified by human raters; however, to do so is difficult and labor intensive, and we thus restrict our evaluation to part location in this study. As a measure of fitting quality, we calculate the target registration error (TRE) [23] for each part, between locations identified by the model and human raters (model-to-rater) and between raters (interrater). The TRE generally measures the discrepancy between different locations of a target identified by different methods in an image.

The TRE for each test model part averaged over all 50 test images is illustrated in Figure 10.10. Overall, the interrater and model-to-rater TREs are similar, indicating that individual model parts can be automatically fit with similar accuracy to human raters on a part-by-part basis. Localization is more precise for certain parts than for others, for both interrater and model-to-rater cases. This is primarily due to part scale, as large-scale parts such as those arising from cerebral lobes are intrinsically more difficult to localize with precision than are small-scale features. For part W, the interrater TRE is somewhat lower than the model-to-rater TRE, indicating some disagreement between human raters and automatic model fitting for this particular

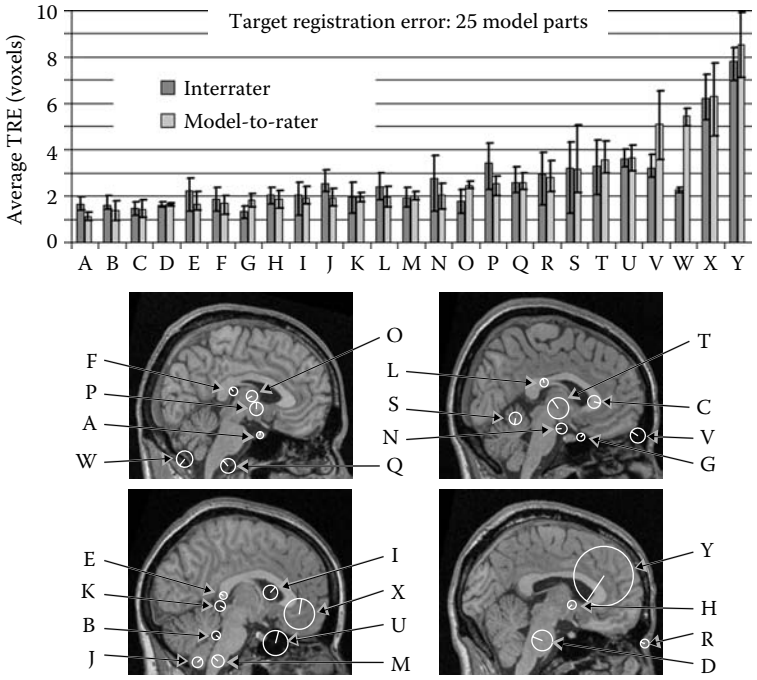


FIGURE 10.10: The interrater and model-to-rater TRE for 25 test model parts, averaged over 50 test subject images and sorted in order of ascending model-to-rater TRE. The height of the bars represents the mean TRE, and error bars indicate the minimum and maximum TRE. The four images below illustrate instances of the indicated test parts. Agreement between interrater and model-to-rater TRE indicates that individual model parts can be automatically localized with similar precision to human raters, validating model fitting on a part-by-part basis. The TRE varies from one part to the next, primarily due to the scale of the part in question, where the larger the part scale, the greater the error associated with its localization. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

model part. Subsequent investigation revealed that model part W had a relatively high intrinsic geometrical variability, as reflected in term $p(m_i^g|o^b, o^g)$, whereas human raters tended to agree as to where they felt the part should occur. Note that model-to-rater agreement could be forced by tightening the geometrical consistency constraint in model learning, as mentioned in Section 10.3.4, but as is, the high geometrical uncertainty has already been quantified by $p(m_i^g|o^b, o^g)$ and accounted for in model fitting.

It is also of interest to know the error with which the model can be registered to individual test images. The TRE for each of the 50 test images

averaged over identified test model parts is illustrated in Figure 10.11. Here, agreement between interrater and model-to-rater TRE indicates that automatic model fitting is similar in accuracy to human raters on a per-image basis. The average interrater and model-to-rater TREs over all images are similar: 2.3 and 2.4 voxels, respectively. The two images in Figure 10.11 illustrate two test subjects for which the TRE is somewhat higher than average for both interrater and model-to-rater cases, indicating increased fitting difficulty for both man and machine. On the left, the higher TRE occurs because the feature detector has fused model parts X and Y into a single region with two dominant orientations. On the right, model part W has been mismatched to

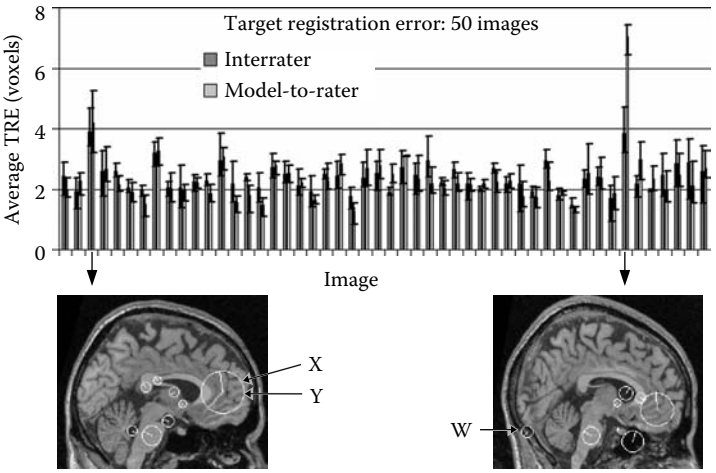


FIGURE 10.11: The interrater and model-to-rater TRE for each test subject image averaged over identified test model parts. The height of the bars represents the mean TRE, and error bars indicate the minimum and maximum TRE. General agreement between interrater and model-to-rater TRE indicates that the model can be automatically fit with similar precision to human raters, validating PBM fitting on an image-by-image basis. The images below illustrate two test subjects for which the TRE is noticeably higher than average for both interrater and model-to-rater cases. On the left, the higher TRE occurs because the feature detector has fused parts X and Y into a single region with two dominant orientations. On the right, the part W has been mismatched to a similar-looking yet incorrect feature, in the absence of a feature detected at its correct location just beneath the cerebellum. Note that part W is known to have relatively high intrinsic geometrical uncertainty, as quantified by model term $p(m_i^g|o^b, o^g)$, although this uncertainty is not accounted for by the average TRE measure. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

a similar-looking yet incorrect feature in the absence of a feature detected at the correct location just beneath the cerebellum. As previously mentioned, the high geometrical uncertainty associated with part W is quantified in term $p(m_i^g|o^b, o^g)$, and its influence on fitting accuracy is already discounted by the model. The TRE measure in Figure 10.11 does not reflect this uncertainty, however, because the errors of all identified model parts are weighted equally.

10.4.3 Parts-based modeling versus global modeling

In this section, we compare parts-based modeling to the global modeling approach common in the literature. The PBM describes a set of brain images as a collection of spatially localized, conditionally independent parts. In contrast, global models such as the active appearance model (AAM) [12] assume one-to-one correspondence between all subject images and represent the population terms of global modes of covariance about a mean, in which spatially distinct regions are coupled in linear relationship and are thus statistically dependent. Forcing such a global model to fit in locations where correspondence is invalid can adversely affect the entire fit, including locations where valid correspondence exists. We hypothesize that the PBM fitting is therefore more robust than the AAM to unexpected intersubject variability on a local scale because the PBM specifically accounts for such variability and avoids forcing the assumption of one-to-one correspondence in locations where it is invalid. To test this hypothesis, we compare AAM and PBM fitting, where an AAM and a PBM are trained on the same set of 102 subjects and fit to the same independent test set of 50 subjects.

The AAM and the PBM differ significantly in both training and fitting. AAM training requires manually determining a set of point correspondences in all 102 training images, after which a linear Gaussian model over image intensity and point location is estimated. Establishing manual point correspondence is tedious, requires a human to decide which and how many points to use, and is subject to interrater variability. PBM learning is fully automatic and requires no manual intervention, as features are determined by an automatic detector. AAM fitting is an iterative process, starting from an initial guess and occasionally falling into suboptimal local maxima when outside of a particular capture range. During experimentation, we performed multiple restarts in order to obtain the best possible AAM fitting solutions. In contrast, the PBM fitting produces a robust, globally optimal solution, even in the presence of image translation, rotation, and scale change.

Directly comparing AAM and PBM fitting to new subjects is difficult for several reasons. First, the fitting solution output of the two approaches is fundamentally different: the AAM produces a smooth mapping from model to subject, whereas the PBM identifies the presence and local geometries of a discrete set of modeled parts. Second, there is no gold standard for evaluating the

ground-truth accuracy of intersubject registration [39]. Third, little guidance exists in selecting a set of manual AAM landmarks leading to an optimal model. We thus compare the two models in terms of their stability in the presence of artificial perturbation. To this end, we perform two different sets of fitting trials: the first is based on 50 normal test subjects. The second is based on the same subjects with the addition of a localized, artificial intensity perturbation. Model fitting stability can then be evaluated in terms of the per-image TRE averaged over all model locations identified before and after perturbation. For the PBM, locations are based on identified test part locations, as described in the previous section, and for the AAM, locations are based on the landmarks defining the model.

The first step in comparing the two models is to establish a baseline in terms of AAM model fitting accuracy. To do this, we construct an AAM based on the set of 102 training subject images, which we then fit automatically to the remaining 50 test subjects. The AAM is defined by six manually chosen landmark points, as illustrated in Figure 10.12. The results of automatic AAM fitting trials are compared to manually labeled solutions established by a single human rater in terms of the TRE, and illustrated in Figure 10.12. Note that in 2 of 50 trials, the TRE is exceptionally high, indicating that the AAM has converged to suboptimal, incorrect solutions. We determine an average TRE threshold of 10 voxels as our definition of a successful/unsuccessful model fitting trial and exclude all subjects for which the TRE of model fitting is greater than this threshold. Note that as illustrated in Figure 10.11, all PBM fitting trials are successful by this definition. The average per-image TRE for successful fitting trials before perturbation is 2.3 voxels for the PBM and 3.8 voxels for the AAM fitting trials. The AAM implementation used is publicly available [13].

Having established that both the AAM and the PBM can be successfully fit to a set of 48 subjects, where a successful fit is defined by an average TRE of less than 10 voxels, we can evaluate the stability of fitting in the presence of perturbation. Given that an accurate, stable solution was obtained in fitting to the 48 normal test subjects, how does this solution vary when a localized artificial perturbation is introduced in each of the test subjects? The perturbation we consider consists of a single black circle inserted at random locations in each test subject image, thereby modifying the image appearance locally in a manner reminiscent of a tumor or a resection. The circle is of radius 16 voxels, occupying approximately 2% of the slice of size 217×181 voxels. The image intensity at the circle center is 0, with a border blended smoothly into the original image in order to simulate a more natural image boundary. Intuitively, since the perturbation has only affected the image intensity in a local image region, the fitting solution should not change except perhaps in a neighborhood surrounding the perturbed area. As a measure of model fitting stability, we consider the per-image TRE between fitting solutions obtained before and after the perturbation, which we refer to as the original-to-perturbed TRE.

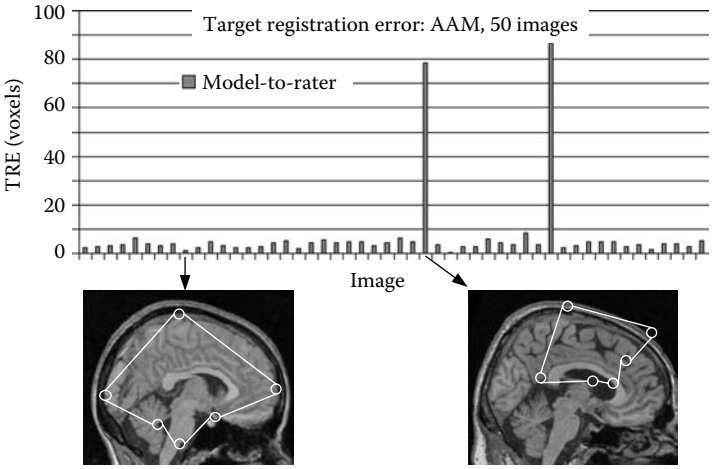


FIGURE 10.12: The model-to-rater TRE of AAM fitting for 50 test subject images, averaged over all landmark locations identified in each image. As illustrated in the lower left image, the AAM created for the purpose of experimentation is defined by six manually selected points, three lying on distinct subcortical structures and three on cortical extremities. The modeled image content lies within the convex hull of the six points and includes both cortical and subcortical structure. The TRE is calculated from the six landmark point locations determined both via automatic AAM fitting and manual identification by a single human rater, and the height of the bars represents the average TRE of all six points. Note that in 2 of the 50 subjects, the model incorrectly converges to suboptimal incorrect solutions with extraordinarily high TRE. The lower left and right images illustrate examples of successful and unsuccessful AAM fitting trials, respectively, where a successful fitting trial is deemed to be one in which the average TRE is less than 10 voxels. The unsuccessful trials are due to the inability of the AAM to cope with normal, unexpected intersubject variability, such as the large, diagonally oriented ventricles of the subject in the lower right image. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

The original-to-perturbed TRE for both the AAM and the PBM is illustrated in Figure 10.13. A large number of AAM fitting trials failed to converge to similar solutions before and after perturbation, as evidenced by exceptionally high original-to-perturbed TRE. As hypothesized, all AAM fitting solutions undergo a global change after perturbation, extending to image regions obviously unaffected by the perturbation; see Figure 10.14. This contrasts sharply with the PBM fitting solutions. While the *minimum* original-to-perturbed TRE for the AAM is 2.9 voxels, the *maximum* for the PBM is 0.5 voxels and is barely visible in the graph in Figure 10.13, indicating that

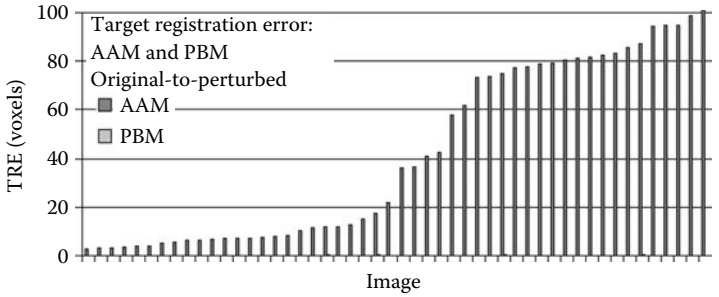


FIGURE 10.13: The original-to-perturbed TRE for AAM and PBM fitting, averaged over AAM landmarks and identified PBM test parts in each of 50 test images, and sorted according to increasing AAM TRE. Note that AAM fitting is generally unstable because a local image perturbation generally induces a global perturbation in the fitting solution, as evidenced by nonzero AAM TRE. In many cases, AAM solutions become completely invalid, resulting in exceptionally high TRE. While the original-to-perturbed TRE for the AAM ranges from 2.9–110 voxels, the PBM TRE ranges from 0.01–0.5 voxels and is barely visible on the graph. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

PBM fitting is stable in the presence of unexpected local variation. As seen in Figure 10.14, PBM fitting solutions are virtually identical before and after perturbation, with the exception of fewer matches in a local neighborhood around the perturbation. This is because the perturbation is recognized as a new and unmodeled image structure by the PBM and can be safely ignored. The size of the neighborhood affected by the perturbation is defined by both the scale of the perturbation and the extent of the filter used in feature detection, which in the case of SIFT features is the width of a truncated Gaussian filter.

10.5 Discussion

In the previous section, we presented the results of experimentation establishing that the PBM can be (a) learned from a database of MR brain imagery consisting of 102 subjects, (b) successfully fit to 50 new subjects in the presence of intersubject variability, and (c) fit in a manner that is robust to artificial local perturbation, where the global active appearance model fails. Model fitting results were quantified using the target registration error measure, evaluating the discrepancy between manually labeled ground-truth solutions and

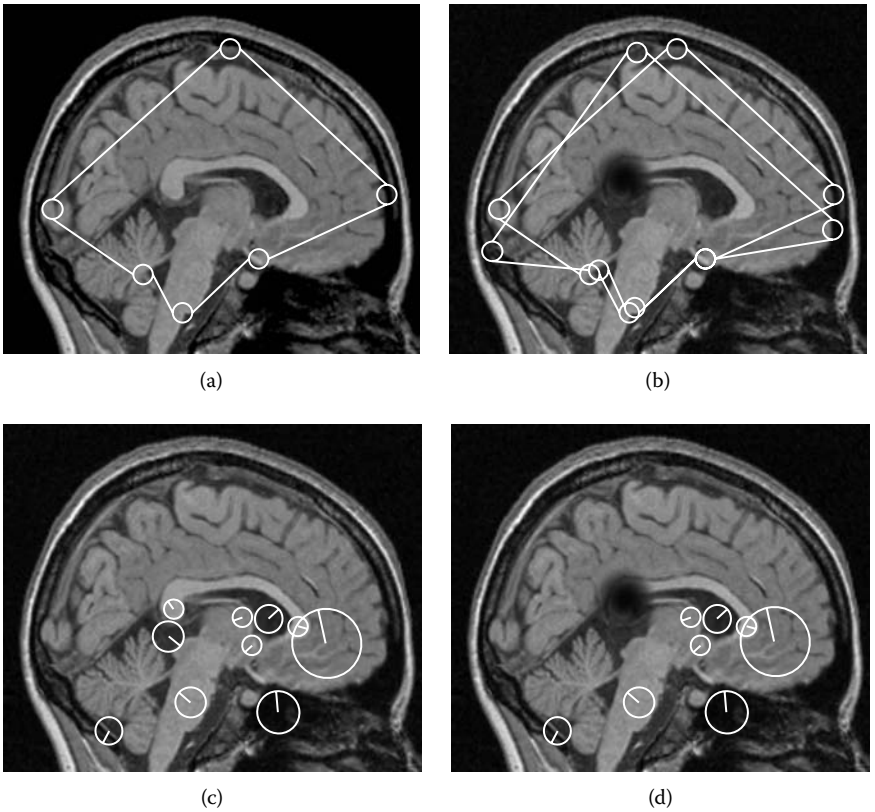


FIGURE 10.14: Illustrating the stability of AAM and PBM fitting in the presence of a local perturbation. (a) illustrates the original AAM fit to a new subject. (b) illustrates the original and the perturbed AAM fits to the same subject after the introduction of a local, artificial perturbation in image content akin to the effect of a tumor or a resection. The local perturbation gives rise to a global change in the AAM solution, extending to brain structure unaffected by the change. (c) and (d) illustrate PBM fitting for the same two images. (c) is the original PBM while (d) is the perturbed PBM. The PBM fit remains virtually unaffected except in a neighborhood of the perturbation, where two model test parts disappear due to the appearance of the novel, unmodeled image structure. The original-to-perturbed TRE for this test subject is 8.4 voxels for the AAM and 0.03 voxels for the PBM. (From M. Toews and T. Arbel, *IEEE Transactions on Medical Imaging—Special Issue on Computational Neuroanatomy*, 26(4):497–508, 2007. With permission.)

automatic model fitting solutions. The effect of perturbation was measured using original-to-perturbed target registration error between fitting solutions before and after perturbation.

The experimentation on learning demonstrates that the PBM can be used to identify local image patterns characteristic of the image population and to quantify them probabilistically in terms of their geometry, appearance, and occurrence frequency. These patterns, or model parts, provide an interpretation of the underlying data, which is fundamentally different from that offered by typical global modeling approaches. Here, individual subjects are described as a collage of distinct model parts instead of as smooth mappings between a model template and a subject. Consequently, this type of interpretation necessarily implies discontinuities between adjacent or overlapping features, an aspect that does not arise in models based on smooth mappings. It is precisely such discontinuities, however, that make it possible to account for occlusion or multiple modes of appearance, phenomena that are, by definition, discontinuous and difficult to describe using smooth, global mappings.

The success of model fitting is due to the fact that image appearance is modeled in terms of statistically independent, spatially localized image parts that are not expected to occur in all subjects of a population. In image regions where valid model-to-subject correspondence exists, multiple independent model-to-subject matches that agree on the referenced frame geometry will be found. Regions containing subject-specific or noisy image characteristics are unlikely to either match to model parts or offer geometrical agreement and are therefore disregarded. Fitting of global models, which often treat all image regions as statistically dependent, is generally poor in the presence of local noise or subject-specific image content because locally poor fitting adversely affects the entire fitting solution, including regions where valid model-to-subject correspondence exists.

10.6 Future Research Directions

The PBM offers a means of automatically learning and describing consistent structures within a set of medical images, where consistency is defined in terms of local image patterns that occur with statistical regularity in a population. In this chapter, we applied the technique to a database of brain images; however, the framework is generally applicable to any imaging context in which the population of interest can be described as a collection of spatially distinct parts in images. Furthermore, in this study, we present results with human brain images from approximately 100 subjects; however, invariant feature-based models have been successfully applied using thousands of images in nonmedical, computer vision contexts, such as in object and face detection [18]. It is possible that model learning on larger medical imaging databases may prove useful in identifying subtle structural characteristics that become observable only in large numbers of images. The goal is that the framework

will then complement traditional anatomical descriptions generated by human experts. Future work will explore how model parts can serve as indicators of anatomical subpopulations, such as normal and pathological subjects. The hope is that this will lead to a robust and objective framework for automatic classification of new images into one of a predefined set of categories.

There remain many open research issues for investigation that result from this work. As mentioned, a gap exists between local parts-based modeling and the more traditional approach of considering smooth, global mappings from model to subject. Further work in bridging this gap is required in order combine the benefits of both modeling approaches (i.e., modeling local appearance changes and providing a smooth interpretation of all image data). It is possible that different modeling assumptions regarding interfeature dependencies may prove useful for this purpose; for instance, the Markov dependency model offers a mechanism for dealing with interfeature boundaries.

Investigation of the invariant features underlying the PBM will likely become an active area of research. As mentioned above, a variety of invariant feature types exist and can be used simultaneously, thereby capturing and modeling different and possibly complementary image characteristics. Certain feature types may prove more effective than others depending on the particular imaging context. Application of the PBM to higher image dimensionalities, such as 3-D volumetric or 4-D temporal sequences, will follow the development of robust invariant feature detectors in higher dimensions. Recent work has shown the feasibility of feature extraction in 3-D MR brain imagery [9]. While identifying the location and scale of invariant features in N-dimensional imagery is straightforward using $N + 1$-dimensional image pyramids, the more difficult aspect is the estimation of local feature orientation, which has a representational complexity of $N(N - 1)/2$. Estimating additional orientation parameters in higher dimensions can be accomplished using higher dimensional histograms of local image derivatives. It may be that estimating all orientations locally is unnecessary, however, as features can be matched using rotationally invariant image appearance representations, and orientation information not available locally can be estimated by considering relationships between multiple, different features.

Another important issue to address in medical imagery is the registration and analysis of images of different modalities, such as magnetic resonance and ultrasound imagery. In general, invariant features corresponding to the same underlying anatomical image structures cannot be extracted in images of different modalities. Image data of different image modalities can, however, be fused during the learning process via a common geometrical reference frame. At this point, techniques for learning intermodal intensity relationships [28] could be applied at a local scale for individual features, and the resulting model could be used as a basis for multimodal registration. This should prove to be an interesting area of research to explore.

10.7 Conclusions

In this chapter, we described a general parts-based approach to modeling medical imagery from local invariant features, which offers a solution to two major challenges facing current computational approaches to medical image analysis: (1) efficient learning from large sets of data with minimal manual supervision and (2) coping with the situation where one-to-one correspondence between different subjects of a population is ambiguous or nonexistent due to natural intersubject variability. The PBM represents several important advancements with respect to statistical appearance models typically applied to quantifying variability in medical imagery:

1. The model can be constructed via a fully automatic machine learning procedure capable of dealing with a large image set containing significant intersubject variation.
2. The model can be robustly fit to new subjects to obtain a globally optimal solution in the presence of intersubject variability in addition to global image translation, rotation, and scale changes.
3. Model fitting is stable in the sense that a localized image deformation results in a localized change in the fitting solution.
4. All subjects of a population can be modeled simultaneously without making a priori classifications as to which subjects are "normal."
5. The spatially localized model parts identified by the model offer an intuitive means of describing and communicating anatomical variability within a population.

We outlined the results of a case study involving parts-based brain modeling in 2-D slices of MR images, demonstrating (a) the feasibility of model learning from a large image set and (b) the robustness of model fitting in the presence of intersubject variability and unexpected local perturbation, in comparison with the well-established active appearance modeling technique. Finally, we discussed issues and offered pointers to future research directions relating to parts-based modeling of medical imagery.

Acknowledgment

We would like to thank Dr. D. Louis Collins of the McConnell Brain Imaging Centre, Montreal Neurological Institute, for his helpful comments and insight regarding the experimental portion of this work and for providing the images used throughout.

References

[1] John Ashburner and Karl J. Frison. Voxel-based morphometry: The methods. *NeuroImage* 11 (2000), no. 23, 805–821.

[2] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing* 46 (1989), 1–21.

[3] E. Bart, E. Byvatov, and S. Ullman. View-invariant recognition using corresponding object fragments. *European Conference on Computer Vision*, 2004, pp. 152–165.

[4] I. Beiderman. Recognition-by-components: A theory of human image understanding. *Psychological Review* 94 (1987), no. 2, 115–147.

[5] F. L. Bookstein. Thin-plate splines and the atlas problem for biomedical images. *Information Processing in Medical Imaging*, 1991, pp. 326–342.

[6] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. *Visualization in Biomedical Computing*, 1996, pp. 267–276.

[7] G. Carneiro and A. D. Jepson. Multi-scale phase-based local features. *Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. 736–743.

[8] G. Carneiro and D. G. Lowe. Sparse flexible models of local features. *European Conference on Computer Vision*, vol. III, 2006, pp. 29–43.

[9] W. Cheung and G. Hamarneh. N-sift: N-dimensional scale invariant feature transform for matching medical images. *International Symposium on Biomedical Imaging*, 2007.

[10] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3D inter-subject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography* 18 (1994), no. 2, 192–205.

[11] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence* 24 (2002), no. 5.

[12] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *Pattern Analysis and Machine Intelligence* 23 (2001), no. 6, 681–684.

[13] T. Cootes. AM Tools. http://www.isbe.man.ac.uk/bim/software.

[14] D. Crandal, P. Felzenswalb, and D. Huttenlocher. Spatial priors for part-based recognition using statistical models. *Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 10–17.

[15] G. Dorko and C. Schmid. Selection of scale-invariant parts for object class recognition. *International Conference on Computer Vision*, 2003, pp. 634–640.

[16] I. L. Dryden and K. V. Mardia. *Statistical shape analysis.* New York: Wiley, 1998.

[17] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*, 2nd ed. New York: Wiley, 2001.

[18] M. Everingham, A. Zisserman, C. K. I. Williams, and L. Van Gool. The PASCAL Visual Object Classes Challenge 2006 (VOC2006) Results. www.pascal-network.org/challenges/VOC/voc2006/results.pdf.

[19] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. *Computer Vision and Pattern Recognition*, 2003, pp. 264–271.

[20] R. Fergus, P. Perona, and A. Zisserman, Weakly supervised scale-invariant learning of models for visual recognition. *International Journal of Computer Vision* 71 (2006), no. 3, 273–303.

[21] J. C. Gee, D. R. Haynor, M. Reikvich, and R. Bajcsy. Finite element approach to warping of brain images. *SPIE Medical Imaging 1994: Image Processing*, vol. 2167, 1994, pp. 18–27.

[22] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics* LVI (1998), no. 4, 617–693.

[23] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes. *Medical image registration.* Boca Raton: CRC Press, 2003.

[24] C. Harris and M. Stephens. A combined corner and edge detector. *Proceedings of the 4th Alvey Vision Conference*, 1988, pp. 147–151.

[25] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen, and H. J. Johnson. Retrospective evaluation of intersubject brain registration. *IEEE Transactions on Medical Imaging* 22 (2003), no. 9, 1120–1130.

[26] S. Joshi, B. David, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage* LVI (2004), no. 23, 151–160.

[27] T. Kadir and M. Brady. Saliency, scale and image description. *International Journal of Computer Vision* 45 (2001), no. 2, 83–105.

[28] M. E. Leventon and W. E. L. Grimson. Multi-modal volume registration using joint intensity distributions. *Lecture Notes in Computer Science* 1496 (1998), 1057–1066.

[29] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision* 30 (1998), no. 2, 79–116.

[30] D. Lowe. Demo software: Sift keypoint detector. http://www.cs.ubc.ca/lowe/keypoints/.

[31] D. G. Lowe. Distinctive image features from scale-invariant key-points. *International Journal of Computer Vision* 60 (2004), no. 2, 91–110.

[32] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London B: Biological Science* 356 (2001), no. 1412, 1293–1322.

[33] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. *International Conference on Computer Vision*, 2001, pp. 525–531.

[34] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. *European Conference on Computer Vision*, 2002, p. 128.

[35] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision* 60 (2004), no. 1, 63–86.

[36] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence* 27 (2005), no. 10, 1615.

[37] F. Mindru, T. Tuytelaars, Luc Van Gool, and Theo Moons. Moment invariants for recognition under changing viewpoint and illumination. *Computer Vision and Image Understanding* 94 (2004), no. 1–3, 3–27.

[38] S. M. Pizer, P. T. Fletcher, Sarang Joshi, Andrew Thall, James Z. Chen, Yonatan Fridman, Daniel S. Fritsch, A. Graham Gash, John M. Glotzer, Michael Jiroutek, Conglin Lu, Keith E. Muller, Gregg Tracton, Paul Yushkevich, and Edward L. Chaney. Deformable m-reps for 3D medical image segmentation. *International Journal of Computer Vision* 55 (2003), no. 2–3, 85–106.

[39] Daniel Reuckert. *Nonrigid registration: Concepts, algorithms and applications*, ch. 13, pp. 281–301. Boca Raton: CRC Press, 2003.

[40] D. Riviere, Jean-Francois Mangin, Dimitri Papadopoulos-Orfanos, Jean-Marc Martinez, Vincent Frouin, and Jean Regis. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis* 6 (2002), 77–92.

[41] D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3-D statistical deformation models of the brain using non-rigid registration. *IEEE Transactions on Medical Imaging* 22 (2003), no. 8, 1014–1025.

[42] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *Pattern Analysis and Machine Intelligence* 19 (1997), no. 5, 530–535.

[43] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision* 37 (2000), no. 2, 151–172.

[44] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and William T. Freeman. Discovering objects and their localization in images. *International Conference on Computer Vision*, 2005, pp. 370–377.

[45] J. Talairach and P. Tournoux. *Co-planar stereotactic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. Stuttgart: Georg Thieme Verlag, 1988.

[46] M. Toews and T. Arbel. Detection over viewpoint via the object class invariant. *International Conference on Pattern Recognition*, vol. 1, 2006, pp. 765–768.

[47] M. Toews and T. Arbel. A statistical parts-based appearance model of anatomical variability. *IEEE Transactions on Medical Imaging – Special Issue on Computational Neuroanatomy* 26 (2007), no. 4.

[48] M. Toews, D. Louis Collins, and T. Arbel. A statistical parts-based appearance model of inter-subject variability. *Medical Image Computing and Computer Assisted Intervention*, vol. I, 2006, pp. 232–240.

[49] A. W. Toga, P. M. Thompson, M. S. Mega, K. L. Narr, and R. E. Blanton. Probabilistic approaches for atlasing normal and disease-specific brain variability. *Anatomy & Embryology* 204 (2001), 267–282.

[50] M. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3 (1991), no. 1, 71–96.

[51] P. A. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision* 24 (1997), no. 2, 137–154.

[52] M. Weber, M. Welling, and P. Perona. Unsupervised learning of models for recognition. *European Conference on Computer Vision*, 2000, pp. 18–32.

Chapter 11

Reinforced Medical Image Segmentation

Farhang Sahba, Hamid R. Tizhoosh, and Magdy M. A. Salama

Contents

In many medical imaging applications, we need to segment one object of interest. The techniques used for segmentation vary depending on the particular situation and the specifications of the problem at hand. This chapter introduces a new multistage image segmentation system based on reinforcement learning (RL). In this system, the RL agent takes specific actions, such as changing the tasks parameters, to modify the quality of the segmented image. The approach starts with a limited number of training samples and improves its performance in the course of time. It contains an offline mode, where

the reinforcement learning agent uses some images and manually segmented versions of these images to provide the segmentation agent with basic information about the application domain. The reinforcement agent is provided with reward and punishment to explore and exploit the solution space. After using this mode, the agent can choose the appropriate parameter values for different processing tasks on the basis of its accumulated knowledge. The online mode consequently guarantees that the system is continuously training. By using these two learning modes, the RL agent allows us to recognize the parameters for the entire segmentation process. The results on transrectal ultrasound (TRUS) images demonstrate the potential of this approach in the field of medical image segmentation.

11.1 Introduction

Medical image segmentation plays an invaluable role in medicine by exhibiting anatomical structures. In many applications, an object of interest is extracted according to measurable features such as intensity, texture, and gradient. Segmentation methods are assessed in terms of the degree of automation, robustness, efficiency, and so on. Techniques used in these methods are highly dependent on the particular situation and the specifications of the problem at hand. Often, segmentation methods suffer from factors such as poor image contrast, noise, and missing or diffuse boundaries. Further complications arise when the quality of the image is influenced by the type and settings of equipment. Usually, there is no universal method that yields acceptable results for all cases. However, there are several methods specialized for particular applications. The resultant performance is usually improved by taking into account expertise or prior knowledge. The amount of required knowledge for user intervention is another issue. Typically, the segmentation tasks are not easy to resolve, and most algorithms must be optimized for specific applications where the parameters are set for the entire process. Ultrasound imaging is a widely used technology for diagnosis and treatment. These images are the result of reflection, refraction, and deflection of ultrasound beams from different types of tissues with different acoustic impedances [2]. However, factors such as poor contrast, speckle, and weak edges result in images that are challenging cases for segmentation.

The prostate segmentation of transrectal ultrasound (TRUS) images is a renowned case study [7,9]. The detection of the prostate boundary in such images is crucial for automatic cancer diagnosis and classification. However, due to a very low signal-to-noise ratio, it is difficult to extract all of the correct boundaries, and any improvements in their segmentation process is desirable [10].

In this chapter, an agent-based segmentation system is introduced, and to validate the findings, the system is applied to segment the prostate in TRUS images.

11.2 Background

Many methods to facilitate more accurate automatic or semiautomatic segmentation of the prostate boundaries in ultrasound images are introduced in the literature [1,3,4,8,10–12,19,25]. By studying the existing methods, we can observe that they may require many training samples if they rely on learning techniques, or that some user interactions are necessary to determine the initial values. Also, many of them cannot improve their performances over time. Considering these factors, a new algorithm based on reinforcement learning (RL) is proposed. In this system, the goal is to construct a framework that has the following characteristics:

- Requires a limited amount of training data
- Improves performance through continuous online feedbacks

Some approaches showing the application of RL for image-based problems appear in the literature [15–18,20,23,24]. Such an intelligent system gains an advantage in training by employing a limited number of samples and gaining more knowledge during the online evaluation. Due to the nature of RL, in terms of the state, action, and reward definitions and their interactions with each other, this system can acquire knowledge and adapt to new input images. This system learns from interactions in two modes: offline and online. It recognizes the parameters for all the processing stages through exploratory learning in the offline mode. This information is then exploited during the online mode, where the system modifies its knowledge simultaneously. The structure used in this system can also incorporate subjective evaluation as a critical feedback. The goal of this system is to identify the object of interest in an image. It is important to note that the proposed approach is not designed to compete with the existing segmentation approaches. The aim of this work is the proof of concept for a new segmentation method based on reinforcement learning.

RL is derived from the idea that an agent learns behavior through trial-and-error interactions in a dynamic environment [22]. The history of RL can be traced to the solution of an optimal control problem by using value functions and dynamic programming [21]. The agent automatically determines the ideal behavior within a specific context that maximizes performance with respect to predefined measures. Figure 11.1 shows the components constituting the concept behind RL. An RL agent, the decision maker of the process, observes the state of the environment. It then takes an action that influences the environment. This action is based on previous experience associated with current observation and accumulated reinforcement. Reward or punishment is determined from the environment, depending on the action taken. The RL agents discover the optimal policy for decision making in two stages: *exploration* and *exploitation*. In the exploration stage, the agent attempts to discover which actions yield the maximum reward by taking different actions

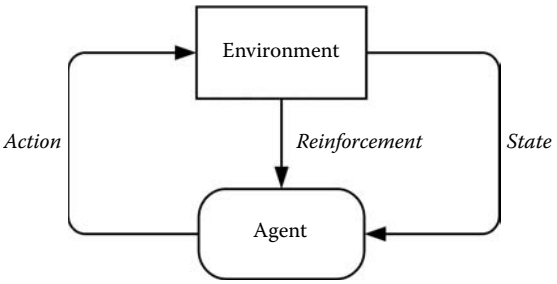


FIGURE 11.1: A general model for a reinforcement learning agent.

repeatedly; in the exploitation stage, it selects the actions that yield more rewards. The agent also receives information concerning the state of the environment. At the beginning of the learning process, the RL agent does not have any knowledge of the result of choosing different actions. The agent takes various actions and observes the results. After a while, the agent explores many actions that bring the highest reward and gradually exploits them. In fact, it acquires knowledge from the actions and eventually learns to perform those that are the most rewarding. During this process, the agent tries to meet a certain goal with respect to the optimal state. The reward and punishment are defined objectively when they are based on desired properties of the results, or they are gained subjectively when the agent receives them directly from the interactive user.

The action policy, π, in RL is the strategy used by the agent to select an action to change the current state. The agent must find a trade-off between an immediate and a long-term return. It must explore the unseen states as well as the states that maximize the return by choosing what the agent already knows. Therefore, a balance between the exploration of unseen states and the exploitation of familiar (rewarding) states is crucial. Watkins developed Q-learning, a well-established online learning algorithm, as a practical RL [26, 27]. In this algorithm, the agent maintains a numerical value for each state-action pair, representing a prediction of the worthiness of taking an action in a state. It can be given as follows [22]:

1. Visit state s_t and select an action a_t.

2. Receive reward $r_t = r(s_t, a_t)$ and observe the next state s_{t+1}.

3. Update the action value $Q_t(s_t, a_t)$ according to:
   $Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha[r_t + \gamma \hat{V}_t(s_{t+1}) - Q_t(s_t, a_t)]$,
   where $\hat{V}_t(s_{t+1}) = \max_a [Q_t(s_{t+1}, a)]$ is the current estimated of the optimal expected cost.

4. Repeat steps 1 to 3 for new state until stopping criterion is satisfied.

In these equations, a is the learning rate, and γ is the discounting factor that controls the degree in future rewards that affect the total value of a policy [14,21,22].

The Boltzman policy is used to estimate the probability of taking each action a given state s. Also, there are other policies for Q-learning, such as ϵ-greedy and greedy. In the greedy policy, not all actions are explored, whereas the ϵ-greedy selects the action with the highest Q-value in a given state with a probability of $1-\epsilon$ and other ones with a probability of ϵ [14,21]. The reward $r(s,a)$ is defined according to each state-action pair (s,a). The goal is to find a policy to maximize the discounted sum of rewards received over time. The principal concern in RL is the cases in which the optimal solutions cannot be found but can be approximated in some way. The online nature of RL makes it possible by focusing more on learning with frequently encountered states and less for infrequently encountered states [22]. Ideally, the RL agent does not require a set of training samples. Instead, it can continuously learn and adapt while performing the required task. Although this is not supervised learning, a reward function is nevertheless employed wherein weak supervision is assumed. This behavior is applicable in cases in which a sufficient number of precise learning samples is difficult or impossible to obtain.

Now we introduce a segmentation system based on RL that attains the highest reward by finding the parameters for the different processing tasks.

11.3 Reinforced Image Segmentation

A framework for adjusting the parameters of a multistage segmentation system by using an RL agent is introduced in this section. The framework is depicted in Figure 11.2. As shown, an intelligent agent is employed to find the appropriate parameters for image processing tasks and enable us to segment the input images. This system is designed as a general framework such that a range of segmentation problems can be integrated with it. The system contains a series of image processing tasks with parameters that must be adjusted to manipulate the input images in some way. The goal is to choose a final set of parameters for various tasks, such that an object of interest is extracted.

A layout is required where the states are the features that describe the image in various stages (after each processing task). Also, the actions have the capability to change the task parameters. The reward can be evaluated subjectively or objectively.

11.3.1 Design aspects

There is a series of N image processing tasks $T_1, T_2, \ldots, T_N$ with specific parameters that must be adjusted. The system adaptively computes the set of parameters that optimize its performance for a given segmentation problem.

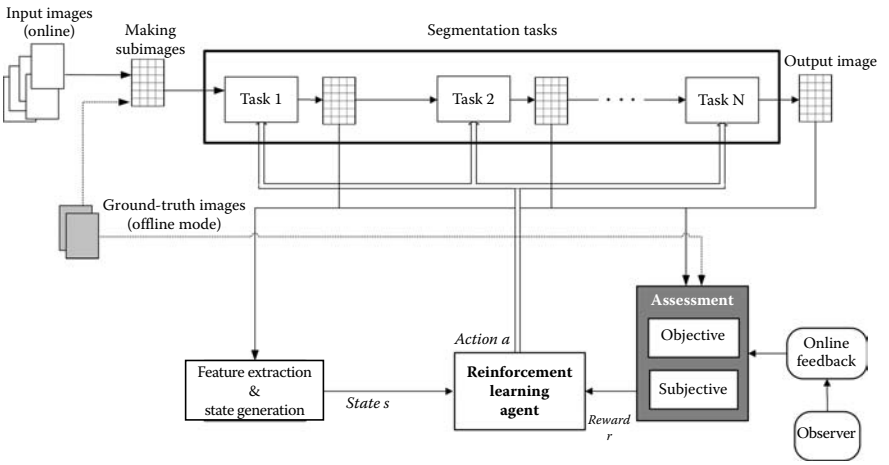


FIGURE 11.2: The general model used in the proposed system for reinforced image segmentation.

Limiting ourselves to the model described in Section 11.2, we must first define the states s, actions a, and reward r. To generate the state, features of the input and output images for each processing task are used. These features must describe the status of these images. It is desirable to use the features that do not lead to extremely large state spaces.

To define the actions, the concept that the actions are modifications of the parameters of the processing tasks is used. After the agent adjusts these parameters, it receives a reward. The reward is an external signal provided by the environment and must correctly reflect the goal of the agent. When the agent has modified the parameters of the processing tasks, an evaluation metric is used to assess the result and produce a reward. The RL method for object segmentation has two modes: offline and online. Offline training is conducted by using the manually segmented samples represented as ground-truth images. In this mode, the agent is adopted in a simulation environment and interacts with training samples to provide the fundamental information about the required segmentation parameters. Once the agent's behavior is acceptable in terms of specific criteria, the agent switches to the online mode. In this mode, the RL agent is used directly in a real-time case with new images. With the agent learning continuously in this mode, it can adapt to changes in the input images. Also, subjective feedback is provided for the system in this mode.

In the proposed system, input image I is divided into $R_I \times C_I$ subimages (for R_I rows and C_I columns). The RL agent works on each of them separately. Local processing on subimages is carried out to find the best segmentation parameters for each of them. To construct the processing task chain, the subimages are thresholded using local values. Due to disturbing factors such as poor contrast or noise, artifacts exist after thresholding. Therefore,

morphological operators are applied to postprocess each thresholded subimage. The RL agent determines the local threshold value and the size of the structuring elements for each individual subimage.

During the offline training where the desired output image is available, the agent works on each subimage and explores the solution space until a sufficient number of actions are taken. In this mode, the agent tries different actions in different states via an exploratory procedure. After the RL agent changes the parameters for each subimage, the agent receives a corresponding reward for that state–action pair and updates the corresponding value in the Q-matrix. After the offline mode, the agent explores many actions and tries to exploit the most rewarding ones.

In the online mode, the system operates episodically. During each iteration, the agent works on each subimage and completes the iteration by using the knowledge previously gained. The rewards are not calculated immediately, and the agent waits until the whole image is scanned. The reward is then provided for each subimage objectively and/or subjectively. For each subimage, it is calculated on the basis of the quality of the segmented object before and after the action is taken. Figure 11.3 summarizes the algorithm in the system. The state transition diagram for a sample subimage is shown in Figure 11.4. In this diagram, $s_{k,l}$ indicates the state number in iteration k. The action a_i is one of the possible actions $\{a_1, \ldots, a_M\}$ that changes the current state $s_{k,l}$ to the next state $s_{k+1,l'}$ and receives the reward $r^{s_l, s_{l'}}$ (s_l and $s_{l'}$ represent two arbitrary states). For each subimage the diagram starts from initial state $s_{0,\text{Initial}}$, and after K iterations reaches to a final desired state $s_{K,\text{Final}}$. The goal of the system is to find an optimal path between initial and final state for each subimage by trial-and-error interactions. For each input image, the knowledge obtained from previous input images is used. Specifically, this method is useful where input images have widely similar object and background characteristics. Components required to construct the RL agent, states, actions and rewards, are defined in the following sections.

11.3.2 State definition

The idea of the state containing some features that represent the status of the image at each step of the processing task is considered,

$$s = [\xi_1, \xi_2, \ldots, \xi_q] \tag{11.1}$$

where $\{\xi_1, \xi_2, \ldots, \xi_q\}$ are the selected features. Each must reflect the local information after different processing tasks. In other words, the states must represent the features of the input gray-level image and the object in the binary image for each subimage. This is the fundamental concept of our proposed system. Because of applying the method to each subimage, employing features that result in an extremely large state space or that significantly

Offline mode
Select a training image
  Create the subimages
  Set for more exploration policy (e.g., using ϵ)
  Repeat for all subimages
    Extract the subimage
    Initialize s
    Repeat
      Calculate the state s using features $[\xi_1, \xi_2, \ldots, \xi_q]$
      Take the action a_i for processing tasks 1, ... , N based on the policy π derived from Q
      Observe the objective reward r using ground-truth and calculate the next state s'
      Update the Q-matrix
    Until a maximum number of iterations k is reached or a specific criterion is met
  Until all subimages are visited
Until all training images are learned

Online mode
Select a test image
  Repeat
    Set for more exploitation (e.g., using ϵ)
    Repeat for each subimage
      Extract the subimage
      Calculate the state s using features $[\xi_1, \xi_2, \ldots, \xi_q]$
      Take the action a_i for processing tasks 1, ... , N based on the policy π
    Until all subimages are visited
    Evaluate the result objectively
    Calculate the online reward r objectively
    Evaluate the result subjectively
    Calculate the online reward r subjectively
    Update the Q-matrix
  Until the user confirms the result (no manual change is made)
Until all test images are segmented

FIGURE 11.3: Summary of the algorithm used for reinforced image segmentation.

increase computational complexity is not applicable. The solution for choosing these features is highly dependent on the specifications of the problem at hand and the accuracy expected. The potential features can be explained as follows:

- The features reflect the local information in the gray-level input image, such as the first, second, or a higher order of spatial distribution. Local histogram or texture information are included in this group.

- The features represent the properties after each processing task. Shape or boundary properties such as area, Euler number (total number of

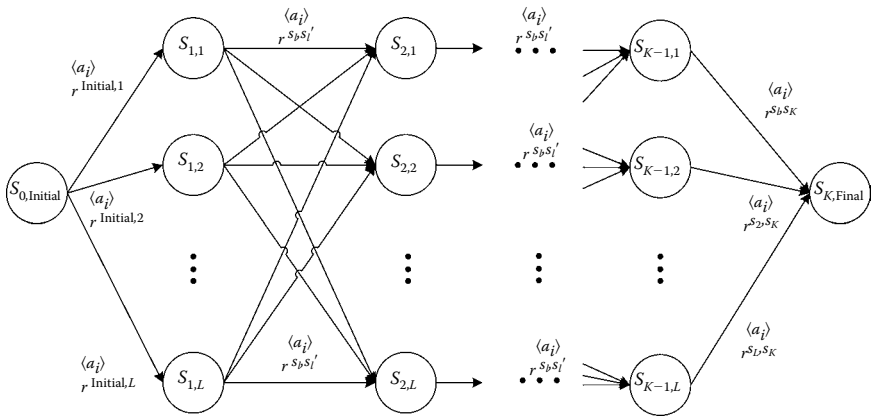


FIGURE 11.4: State transition model for a subimage.

objects minus the number of holes in those objects), compactness, convexity, slope density function, curvature, and edge properties can be used if there is a binary image after the task.

- The feature that shows the geometrical situation of each subimage, such as relative position with respect to particular landmarks.

These features show the local situation in the input, output, and intermediate images. For the application of prostate segmentation in TRUS images, considering the above conditions (representing the local quality with low computational complexity), Figure 11.3 shows the features in each subimage that are extracted to define the states.

11.3.2.1 Area

In each subimage A_{object}, the area of the largest object is chosen as a feature to define the state. The normalized area A_Δ is calculated with respect to the whole area of the subimage as follows:

$$A_\Delta = \frac{A_{\text{subimage}} - A_{\text{object}}}{A_{\text{subimage}}} \tag{11.2}$$

11.3.2.2 Compactness

The compactness Ψ of the largest object after thresholding, defined as

$$\Psi = \frac{4\pi A}{P^2} \tag{11.3}$$

is adopted as another feature to represent the state, where A and P are the area and perimeter of the object in the subimage, respectively [5].

11.3.2.3 Relative position to the center point

We suppose that the geometric center of the prostate in the original image, (x_c, y_c), is given by the user. The relative distance and angle of the subimage, ρ and ϕ, with respect to the geometric center, is used as a state parameter:

$$\rho = ((x - x_c)^2 + (y - y_c)^2)^{1/2} \tag{11.4}$$

$$\phi = \tan^{-1}\left(\frac{y - y_c}{x - x_c}\right) \tag{11.5}$$

where x and y are the coordinates of the center of the current subimage.

11.3.2.4 Gray-level information

A feature based on spatial gray-level information is extracted for the state definition. By using the ground-truth image, a gray-level histogram prototype on the edge of the prostate can be made. The histogram distance between each subimage and the prototype is then calculated and used to define a feature. One of the most popular histogram distances, χ^2 distance, is selected for this purpose such that

$$D_{\chi^2}(h_1, h_2) = \sum_{b=1}^{M_h} \frac{(h_1(b) - h_2(b))^2}{(h_1(b) + h_2(b))} \tag{11.6}$$

For h_1 and h_2, the normalized gray-level histograms (gray-level probability density function) is used. To incorporate the spatial gray-level information, texture—a feature generally used in ultrasound image analysis—can also be employed. Because of the increase in the computational expense, however, texture is not used in this work.

11.3.2.5 Number of the objects

The last feature, used in state definition, is the number of revealed objects, N_O, after the morphological operation.

11.3.3 Action definition

To extract the object of interest, the actions are defined according to changing of the parameters for each subimage. The assigned values are increased or decreased or are chosen from the predefined values to control the effects of each processing task. As a result, there exist a set of actions, $\langle a_i \rangle$ as follows:

$$\langle a_i \rangle = \{a_1, \ldots, a_M\} \tag{11.7}$$

where $a_1, \ldots, a_M$ represent the changes of the parameters of the processing tasks, $T_1, T_2, \ldots, T_N$. In order to extract the prostate, the subimages are thresholded using two local levels to make the primary binary image. Due to disturbances such as speckle or poor contrast, artifacts exist after thresholding. Therefore, opening and closing, as two morphological operators, are employed to postprocess each thresholded subimage. These make three processing tasks, $\{T_1, T_2, T_3\}$, as illustrated in Figure 11.2. The RL actions are defined as changes in the values of the threshold and the size of the structuring elements for each subimage. The agent changes the local threshold values by a specific amount Δ. Using this definition, the values τ_{g_1} and τ_{g_2} among a predefined set $(\tau_1, \tau_2, \ldots, \tau_n)$, equally spaced between the local maximum gray level $g_{l\max}$ and local minimum gray level $g_{l\min}$, are selected. For opening and closing morphological operators, the sizes of the structuring elements are chosen from the predefined values to control the effects of these operators. Thus, the set of all actions a can be presented as follows:

$$\langle a_i \rangle = \{\tau_{g_1} \pm \Delta, \tau_{g_2} \pm \Delta\} \cup \{v_j\} \cup \{v_u\} \tag{11.8}$$

where v_j and v_u are the sizes of the structuring elements for the morphological opening and closing, $g_1, g_2 \in \{1, \ldots, n\}, \Delta = \tau_{g_1} - \tau_{g_1-1} = \tau_{g_2} - \tau_{g_2-1}, j \in \{1, \ldots, J\}$, and $u \in \{1, \ldots, U\}$.

11.3.4 Reward definition

The rewards and punishments are defined according to the quality criterion that represents how well the object is segmented in each subimage. A straightforward method is to compare the results with the ground-truth image after each action. To measure this value for each subimage, we note how much the quality has changed after the action. In each subimage, for improving the quality of the segmented object, the agent receives a reward; otherwise, it is punished. A general form for the reward function is represented by

$$r = \begin{cases} \kappa_1 & D_\Delta \geq 0 \\ \kappa_2 & D_\Delta < 0 \end{cases} \tag{11.9}$$

where D_Δ is a measure indicating the difference between the quality after and before taking the action. D_Δ can be calculated based on the normalized number of misclassified pixels in the segmented subimages. In Equation 11.9, κ_1 and κ_2 are constant values.

11.3.5 Offline learning procedure

In the offline mode, the ground-truth images are available as manually segmented versions. The RL agent works on each subimage individually and continues until the maximum number of iterations is attained. Based on the

segmentation quality after the action, the agent receives an objective reward or punishment r_{OFF} for the subimage and updates its knowledge. For the quality measure of each subimage, it is compared with the ground-truth image. To measure the similarity, the percentage of pixels that are the same in the ground-truth and the output image generated by the system is calculated. During this procedure, the agent must explore the solution space. The agent attempts different actions in different states by using an exploratory policy. After a number of iterations, the Q-matrix is filled with the appropriate values and the agent can estimate the best action for each given state.

11.3.6 Online evaluation

For new samples where no ground-truth images are available, the agent takes action according to the knowledge it has previously gained. The RL agent finds the appropriate local parameters such that the object of interest can be correctly segmented. In this mode, in each iteration the RL agent finds the parameters for each subimage and then moves to the next subimage. This continues until the whole image is scanned in one episode. After each episode, the bounding contour is considered in a global manner to identify the image regions that correspond to the object of interest. It is the guidance from the higher level of perception that helps the system update its information. The RL agent then uses this feedback as online rewards r_{ON} for the action taken on each subimage. If a subimage is segmented correctly, in terms of the global feedback (objective or subjective), the parameters remain unchanged and its process is terminated. The online learning procedure is performed by the following evaluations.

11.3.6.1 Objective evaluation

Objective evaluation is applied as a primary assessment. There are several methods for implementing this evaluation. One way is to use the techniques that compute a measure of distance between two shapes. Typically, these techniques compute the best global match between two shapes by pairing each element of shape Γ_1 with each element of shape Γ_1 to give the minimum difference. The transformations in these algorithms are generally invariant to translation, scaling, and rotation, and they preserve the information contained in the curve. The technique introduced by Rodriguez and colleagues [13] is adopted in a 2-D plane to measure the similarity. In this technique, there is an unknown curve

$$C^1 = \left\{p_1^1\left(x_1^1, y_1^1\right), p_2^1\left(x_2^1, y_2^1\right), \ldots, p_n^1\left(x_n^1, y_n^1\right)\right\}$$

and a prototype curve

$$C^2 = \left\{p_1^2\left(x_1^2, y_1^2\right), p_2^2\left(x_2^2, y_2^2\right), \ldots, p_m^2\left(x_m^2, y_m^2\right)\right\}$$

corresponding to shape Γ_1 and shape Γ_2, respectively, and $p_{i_p}^{\{1,2\}}(x_{i_p}^{\{1,2\}}, y_{i_p}^{\{1,2\}})$ are the points located on these curves. If the characteristics of the points located on the two curves are represented as strings $\tilde{x}$ and $\tilde{y}$, the edit distance between the two strings is defined as the shortest way (minimum cost) to transform $\tilde{x}$ into $\tilde{y}$. The idea is to find a minimum cost path as a sequence of edit operations that transform the string corresponding to the curve C^1 into curve C^2. To achieve this transformation, the technique allows changing one element of a string into another element, deleting one element from the string, or inserting an element into the string [13]. With these operations, a measure of the similarity between two strings is computed. The changes performed in this measurement correspond to the insertion, substitution, and deletion operations. The algorithm gives not only the distance $\Delta(\tilde{x}, \tilde{y})$ between two strings $\tilde{x}$ and $\tilde{y}$ but also the way in which the symbols of $\tilde{x}$ correspond to those of $\tilde{y}$. The signature of a shape is used to generate its string. The shape signature is defined as a 1-D function representing a 2-D area or boundary of the shape. Complex coordinates, central distance, curvature, and cumulative angles are among the most used signatures in the literature and are shown to be useful for shape representation [6]. In the proposed system, the central distance is calculated as the distance from the points on the boundary to the central point of the object. It is represented as a 2π periodic function. This method is especially useful when a reliable central point of the object can be calculated (e.g., represented by the expert). In this case, angle θ is assigned to a distance d, which is the nearest corresponding contour point. It is represented as a function $d = f(\theta)$ to generate the central distance signature. If a reliable central point is not available, other methods such as curvature can be used. Because of the sensitivity to noise, however, the boundary of the object must be smooth. The signature of the extracted object is then compared with the standard signature of the object for which we are looking. The standard signature is extracted by using the ground-truth images. Finally, the significant deviations are estimated and the objective reward (or punishment) r_{ON}^{O} for the corresponding subimage is provided.

11.3.6.2 Subjective evaluation

Another alternative is subjective evaluation from the online feedback. The user considers the results of the segmentation system for each image. If not satisfied, he or she can change the results manually. These changes are evaluated as subjective punishment for the agent, where the manually corrected result is used as a new ground-truth image to improve the agent's knowledge. The agent then proceeds to the next image with its updated information. By adopting this method, the agent can be further trained online by a subjective evaluation as r_{ON}^{S}. The Q-matrix is updated, and the segmentation system can follow the changes in the new input image. During all previous stages, the object of interest (in our case, the prostate) is extracted by using the position of the central point.

11.4 Results and Discussions

To evaluate the performance of the segmentation system, validation experiments are performed. The results of the segmentation system on the image set, containing 80 TRUS image slices from five patients, are represented in this section. To train the agent, 12 manually segmented images from the data set are chosen. A ϵ-greedy policy is implemented to explore and exploit the solution space when the RL agent is in the offline mode. The size of the TRUS images are 460×350 pixels. The sizes of R_I and C_I are empirically set to 11 and 14, respectively. The threshold action is defined by changing the two values between the local maximum and minimum gray levels in each subimage by $\Delta = 1/10$. For the postprocessing tasks, the size of the disk shape structuring element for the opening is increased or decreased by 5 in the set $\{0, \ldots, 5\}$ and for the closing by 1 in the set $\{0, \ldots, 5\}$. The criterion to terminate the process for each subimage is to reach a pixel difference of less than 5% by comparing it with the ground-truth image. The reward is calculated according to $\kappa_1 = 10$, $\kappa_2 = 0$. The average time for training was 61 seconds per patient set, and for test images, 7.9 seconds per image. After the offline learning, the Q-matrix is filled with the appropriate values, and the agent gains enough knowledge to recognize the appropriate values for the subimages. The method is applied on the remaining sample images in the online mode (objectively and subjectively). Figure 11.5 displays the results of the proposed approach for nine sample images. The choice of criteria to measure the accuracy is dependent on the application and can be derived from the region or boundary information [9]. For all images, the accuracy in the final segmented object is defined as the *area overlap* of the segmented images and the ground-truth images, a commonly used metric. Table 11.1 shows the results for training images, and Table 11.2, for test images. (Because in offline mode there are few training samples for each patient, the standard deviations are not shown in Table 11.1).

The results in terms of visual appearance and accuracy can be employed as suitable coarse-level estimations to serve a fine-tuning segmentation algorithm [9]. For instance, these results can be used as the initial points for the well-known method introduced by Ladak and colleagues [8]. In some cases, the results can even be regarded as the final segmentation.

As mentioned previously, this work offers the proof of concept by presenting a new learning-based segmentation method. Our proposed method (with its preliminary results) introduces a new class of knowledge-based method that requires a limited number of samples for training and can improve its performance in an online learning procedure. To understand the advantages of the proposed method, we must know that static methods do not have any ability to be trained. On the other hand, in the methods that use learning algorithms, we usually need a large number of training samples. They do not have online training unless we retrain them. To the best of our knowledge,

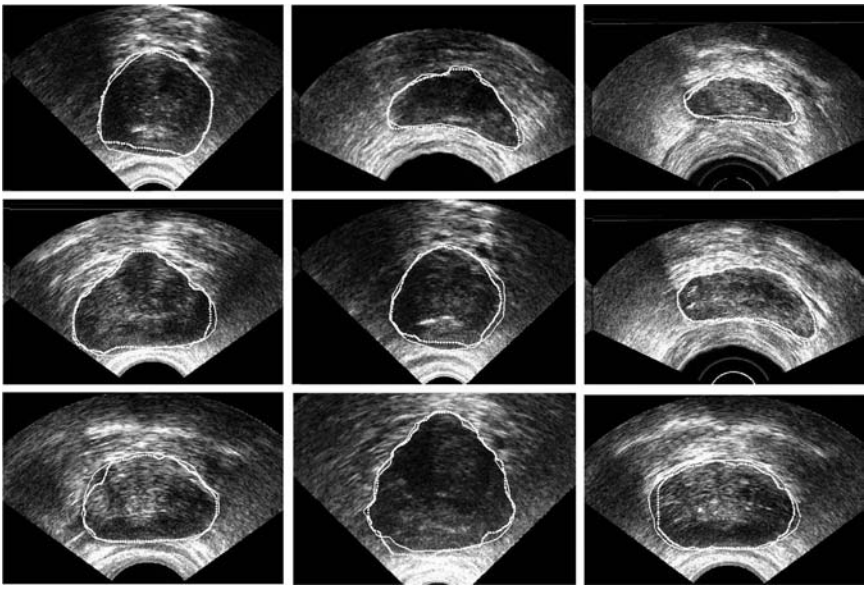


FIGURE 11.5: Demonstration of the segmentation of the prostate in TRUS image. In each image, the result of the proposed system is indicated with the dashed contour, and for the manually segmented, with the solid contour.

TABLE 11.1: The average area overlap AO between the segmented and the ground-truth images for the training images.

	$AO(\%)$ **Patient 1**	$AO(\%)$ **Patient 2**	$AO(\%)$ **Patient 3**	$AO(\%)$ **Patient 4**	$AO(\%)$ **Patient 5**
μ	93.56	95.30	94.07	94.58	93.83

TABLE 11.2: The means μ and standard deviations σ of area overlap AO between the segmented and the ground-truth images for test images.

	$AO(\%)$ **Patient 1**	$AO(\%)$ **Patient 2**	$AO(\%)$ **Patient 3**	$AO(\%)$ **Patient 4**	$AO(\%)$ **Patient 5**
μ	90.74	91.05	89.07	92.26	91.12
σ	1.04	1.12	1.06	0.91	0.72

none of the existing segmentation methods can start their work using a limited number of training samples and improve their performance using online training.

11.5 Future Research Directions

The work presented in this chapter is an attempt to use an RL agent for segmentation of images. To improve this version and achieve more accurate results, the following investigations can be concluded:

- State definition is a crucial step in this system. For both input gray level and output binary images, more information from different classes of features can be added. It is expected that this additional information will increase the computational complexity by forming a larger state space. An appropriate approximation method may be employed to reduce this problem. An extension using the more sophisticated actions can also be considered.

- The proposed system could be the basis for other parameter control applications such as contrast enhancement or noise filtering.

- As a preliminary version, we chose the parameters empirically to introduce the concept. Automatic selection of most parameters must be investigated for the next versions.

- From an application point of view, the results of this work lead to the application of the proposed method to other medical image modalities.

11.6 Conclusions

This chapter presented the concept of reinforcement learning in the field of medical image segmentation by showing some results. In this work, a reinforcement learning agent for ultrasound image segmentation is proposed. First, the image is divided into several subimages. Then, in an offline procedure, the agent takes some actions (i.e., changing the threshold value and the size of the structuring elements) to change its environment (the quality of the segmented parts) and accumulates the local information for each subimage. After the Q-matrix is filled, the agent takes actions with maximum reward for each possible state. The agent, on the basis of its accumulated knowledge, can

choose the appropriate values for those input images with similar characteristics. This method is used for one object segmentation in ultrasound images. The proposed system can be adopted as the main segmentation approach or as an interim stage to serve other segmentation methods. This segmentation system is applied to TRUS ultrasound images as a case study to confirm its effectiveness.

References

[1] N. Betrounia, M. Vermandela, D. Pasquierc, S. Maoucheb, and J. Rousseaua. Segmentation of abdominal ultrasound images of the prostate using a priori information and an adapted noise filter. *Computerized Medical Imaging and Graphics*, 29, 2005, 43–51.

[2] Cancer Facts and Figures 2002. American Cancer Society. http://www.cancer.org.

[3] B. Chiu, G. H. Freeman, M. M. A. Salama, and A. Fenster. Prostate segmentation algorithm using dyadic wavelet transform and discrete dynamic contour. *Physics in Medicine and Biology*, 49, 2004, 4943–4960.

[4] L. Gong, S. Pathak, D. Haynor, P. Cho, and Y. Kim. Parametric shape modeling using deformable superellipses for prostate segmentation. *IEEE Transactions on Medical Imaging*, 23, 2004, 340–349.

[5] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[6] R. J. Holt and A. N. Netravali. Using line correspondences in invariant signatures for curve recognition. *Image and Vision Computing*, 11, 1993, 440–446.

[7] M. F. Insana and D. G. Brown. *Acoustic Scattering Theory Applied to Soft Biological Tissues: Ultrasonic Scattering in Biological Tissues*. Boca Raton: CRC Press, 1993, 76–124.

[8] H. M. Ladak, F. Mao, Y. Wang, D. B. Downey, D. A. Steinman, and A. Fenster. Prostate boundary segmentation from 2D ultrasound images. *Medical Physics*, 27, 2000, 1777–1788.

[9] N. D. Nanayakkaral, J. Samarabandul, and A. Fenster. Prostate segmentation by feature enhancement using domain knowledge and adaptive region based operations. *Physics in Medicine and Biology*, 51, 2006, 1831–1848.

[10] J. A. Noble and D. Boukerroui. Ultrasound image segmentation: a survey. *IEEE Transactions on Medical Imaging*, 25, 2006, 987–1010.

[11] S. D. V. Pathak, D. R. Chalana, D. Haynor, and Y. Kim. Edge-guided boundary delineation in prostate ultrasound images. *IEEE Transactions on Medical Imaging*, 19, 2000, 1211–1219.

[12] J. S. Prater and W. D. Richard. Segmenting ultrasound images of the prostate using neural networks. *Ultrasound Imaging*, 14, 1992, 159–185.

[13] W. Rodriguez, M. Lastb, A. Kandel, and H. Bunked. 3-Dimensional curve similarity using string matching. *Robotics and Autonomous Systems*, 49, 2004, 165–172.

[14] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[15] F. Sahba and H. R. Tizhoosh. Filter fusion for image enhancement using reinforcement learning. *Canadian Conference on Electrical and Computer Engineering (CCECE03)*, Montreal, May 2003, pp. 847–850.

[16] F. Sahba, H. R. Tizhoosh, and M. M. A. Salama, Using reinforcement learning for filter fusion in image enhancement. *Fourth IASTED International Conference on Computational Intelligence*, July 2005, Calgary, Canada, pp. 262–266.

[17] F. Sahba, H. R. Tizhoosh, and M. M. A. Salama. A reinforcement learning framework for medical image segmentation. *IEEE World Congress on Computational Intelligence (WCCI)*, July 2006, Vancouver, Canada, pp. 1238–1244.

[18] F. Sahba, H. R. Tizhoosh, and M. M. A. Salama. Increasing object recognition rate using reinforced segmentation. *IEEE International Conference on Image Processing (ICIP)*, October 2006, Atlanta.

[19] D. Shen, Y. Zhan, and C. Davatzikos. Segmentation of prostate boundaries from ultrasound images using statistical shape model. *IEEE Transactions on Medical Imaging*, 22, 2003, 539–551.

[20] M. Shokri and H. R. Tizhoosh. Q(λ)-based image thresholding. *Canadian Conference on Computer and Robot Vision*, 2004.

[21] S. Singh, P. Norving, and D. Cohn. *Introduction to Reinforcement Learning*. New York: Harlequin, 1996.

[22] R. S. Sutton and A. G. Barton. *Reinforcement Learning*. Cambridge, MA: MIT Press, 1998.

[23] G. A. Taylor. A reinforcement learning framework for parameter control in computer vision applications. *Proceedings of the First Canadian Conference on Computer and Robot Vision*, 2006.

[24] H. R. Tizhoosh and G. W. Taylor. Reinforced contrast adaptation. *International Journal of Image and Graphic*, 6, 2006, 377–392.

[25] Y. Wang, H. Cardinal, D. Downey, and A. Fenster. Semiautomatic three-dimensional segmentation of the prostate using two-dimensional ultrasound images. *Medical Physics*, 30, 2003, 887–897.

[26] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8, 1992, 279–292.

[27] C. J. C. H. Watkins. *Learning from Delayed Rewards*, Cambridge: Cambridge University Press, 1989.

Chapter 12

Image Segmentation and Parameterization for Automatic Diagnostics of Whole-Body Scintigrams: Basic Concepts

Luka Šajn and Igor Kononenko

Contents

Bone scintigraphy or whole-body bone scan is one of the most common diagnostic procedures in nuclear medicine. Since expert physicians evaluate images manually, some automated procedure for pathology detection is desired. Scintigraphy segmentation into the main skeletal regions is briefly presented. The algorithm is simultaneously applied on anterior and posterior whole-body bone scintigrams. The expert's knowledge is represented as a set of parameterized rules, used to support image processing algorithms. The segmented bone regions are parameterized with algorithms for classifying patterns so the pathologies can be classified with machine learning algorithms. This approach enables automatic scintigraphy evaluation of pathological changes; thus, in addition to detection of pointlike high-uptake lesions, other types can be discovered. We extend the parameterization of the bone regions with a multiresolutional approach and present an algorithm for image parameterization using the association rules.

Our study includes 467 consecutive, nonselected scintigrams. Automatic analysis of whole-body bone scans using our segmentation algorithm gives more accurate and reliable results than previous studies. Preliminary experiments show that our expert system based on machine learning closely mimics the results of expert physicians.

12.1 Introduction

The whole-body scan or the whole-body bone scintigraphy is a well-known clinical routine investigation and one of the most frequent diagnostic procedures in nuclear medicine [1]. Indications for bone scintigraphy include benign and malignant diseases, infections, degenerative changes, and other clinical entities [2]. Bone scintigraphy has high sensitivity, and the changes of the bone metabolism are seen earlier than the changes in bone structure detected on skeletal radiograms [1].

The investigator's role is to evaluate the image, which is of technically poor resolution due to the physical limitations of gamma camera. There are approximately 158 bones visible on anterior and posterior whole-body scans [3]. Poor image resolution and the number of bones to recognize make the evaluation of images difficult. Some research on automating the process of counting the bone lesions has been done, but only a few studies attempted to automatically segment individual bones prior to the computerized evaluation of bone scans [4–6].

12.1.1 Related work

First attempts to automate scintigraphy in diagnostics for thyroid structure and function were made in 1973 [7]. Most of the research on automatic localization of bones was performed at the former Institute of Medical Information Science at the University of Hildesheim in Germany from 1994 to 1996. The main contribution was made by Bernauer [4] and Berning [5], who developed semantic representation of the skeleton and evaluation of the images. Benneke [6] realized their ideas in 1996.

Yin and Chiu [8] tried to find lesions using a fuzzy system. Their preprocessing of scintigrams includes rough segmentation of six fixed-size regions, regardless of individual image properties. Those parts are rigid and not specific enough to localize a specific bone. Their approach for locating abnormalities in bone scintigraphy is limited to pointlike lesions with high uptake.

When dealing with lesion detection, other authors, such as Noguchi [3], have used intensity thresholding and manual lesion counting or manual bone region of interest (ROI) labeling. Those procedures are sufficient only for more evident pathologies, whereas new emerging pathological regions could be overlooked.

12.1.2 Aim and approach

The aim of our study is to develop a robust procedure for diagnosing whole-body bone scans. Some possible methods for individual bone extraction are also presented. Segmented scans allow further development of automated procedures for recognition of pathological condition in specific bone regions. The experience with segmentation and pathology classification is presented. The steps of segmentation and diagnosing are shown in Figure 12.1.

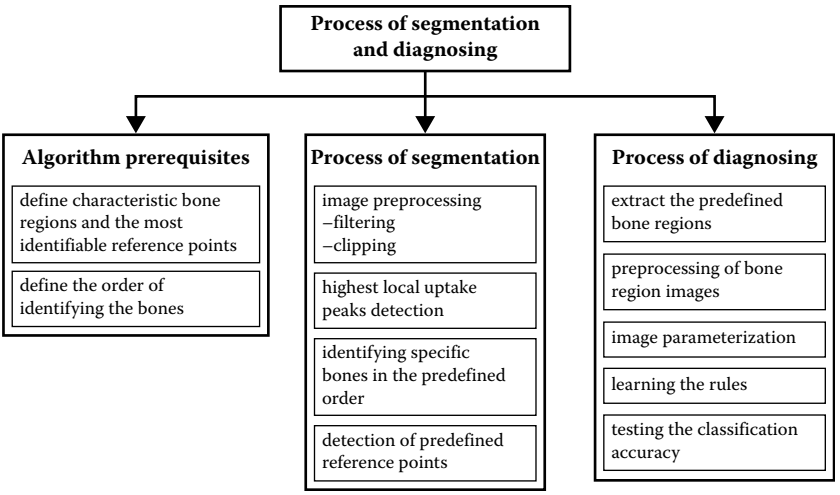


FIGURE 12.1: Algorithm steps.

12.1.2.1 Segmentation

When a scintigraphy is observed by an expert physician, each bone region is diagnosed according to several possible pathologies (lesions, malignom, metastasis, degenerative changes, inflammation, other pathologies, or no pathologies). The process of detecting the lesions can be aided by some advanced machine learning classifier [9] that produces independent diagnoses. The implementation of such a system was also the goal of the study. It can be used as a tool to remind a physician of some possibly overlooked spots or even to give some additional insight into the problem. It also enables further studies of individual bone regions with other algorithms (e.g., pattern classification or custom tailored algorithms).

In order to achieve a robust segmentation algorithm, we defined the most characteristic bone regions and the most identifiable points in those regions (reference points, Figure 12.2). Those points are chosen so that they can be uniformly identified over all distinct images in as many cases as possible, which is necessary since the images and skeletons vary considerably.

There are some algorithms for detecting image features (e.g., scale-invariant feature transform, or SIFT [10]). Since the scintigrams are so variable, we cannot rely solely on features detected in images alone but must use some existing background knowledge. In our case, this background knowledge is the human anatomy.

Because scintigrams are represented by relatively small images, many algorithms (e.g., PCA [11], image correlation) are not appropriate or directly applicable. The idea is to use simple and easy-to-control algorithms. In our study, we used several image processing algorithms, such as binarization, dilatation, skeletonization, Hough's transform, Gaussian filtering [12], beam search, circle fitting, and ellipse fitting with least square method (LSM), in combination with background knowledge of anatomy and scintigraphic patterns.

12.1.2.2 Image parameterization

In order for images to be automatically analyzed, they must be described in terms of a set of (numerical) parameters. There exist many different approaches to characterize textures of images. Most texture features are based on structural, statistical, or spectral properties of the image. For the purpose of diagnosis from medical images, it seems that structural description is most important. For that purpose, we use algorithm ArTeX for textural attributes that are based on association rules. The association rules algorithms can be used for describing textures if an appropriate texture representation formalism is used. This representation has several good properties, such as invariance to global lightness and invariance to rotation. Association rules capture structural and statistical information and are very convenient to identify the structures that occur most frequently and have the most discriminative power.

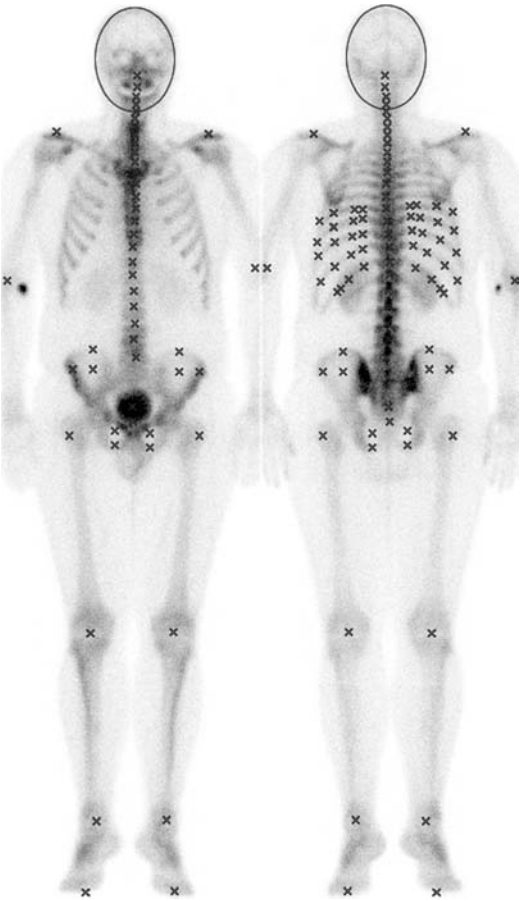


FIGURE 12.2: Defined characteristic reference points on anterior (left) and posterior (right) whole-body bone scintigrams.

Another issue is the selection of the appropriate resolution for extracting most informative textural features. Image parameterization algorithms use descriptors of some local relations between image pixels where this search perimeter is bounded to a certain size. This means that they can give different results at different resolutions. The resolution used for extracting parameters is important and depends on the properties of the observed domain.

12.1.2.3 Image classification

The ultimate goal of the medical image analysis is a decision about the diagnosis. When images are described with (hopefully) informative numerical attributes, we can use various machine learning algorithms [13] for generating a classification system. For that purpose, we can use decision trees, the naive Bayesian (NB) classifier, neural networks, k-nearest neighbors, support vector

machines, and others. Our experience in a similar problem of diagnosing the ischemic heart disease from scintigraphic images of the heart muscle [14] is that the NB classifier gives the best results. It turned out that also for diagnosing (parts of) bone scintigrams, as described in this chapter, the NB classifier performs the best. Therefore, we describe results obtained by the NB. It is also well known that, in other problems of medical diagnosis, the NB usually outperforms other classifiers [15].

This chapter is organized as follows. First, the medical domain and the algorithm for segmenting whole-body bone scans are presented. Then, the algorithm ArTeX for pattern parameterization using association rules with an extended parameter set is described. The algorithm ARes for selecting the appropriate resolutions for better pattern description follows. Finally, the segmentation and pathology diagnosing results are presented.

12.2 Materials and Methods

12.2.1 Patients and images

Retrospective review of 467 consecutive, nonselected scintigraphic images from 461 different patients investigated at the Nuclear Medicine Department, University Medical Centre in Ljubljana, Slovenia, from October 2003 to June 2004, was performed. Images were not preselected, so the study included a representative distribution of patients coming to examinations over the 9-month period. Images contained some artifacts and nonosseous uptake, such as urine contamination and medical accessories (i.e., urinary catheters) [16]. In addition, the site of radiopharmaceutical injection is frequently visible (i.e., obstructs the image).

Twenty-one percent of the images were diagnosed as normal with no artifacts, meaning that no pathologies and no artifacts were present. Sixty percent of the images were diagnosed with slight pathology regardless of artifacts, 18% with strong pathology, and 1% of the images were classified as super-scans. A super-scan is obtained when the patient has very strong pathologies, which absorb most of the radiopharmaceutical so the other bone regions are vaguely visible.

In 18% of the images, partial scans (missing a part of the head or upper/lower extremities in the picture) were acquired, which complicates the segmentation process. There were also children/adolescents with growth zones (5% of the images), manifested as increased osteoblastic activity in well-delineated areas with very high tracer uptake.

12.2.2 Bone scintigraphy

All patients were scanned with gamma camera model Siemens Multi-SPECT with two detectors, equipped with LEHR (low-energy high-resolution)

collimators. Scan speed was 8 cm per minute with no pixel zooming. Technetium-labeled phosphonate (99 m-Tc-DPD, Techneos®) was used. Bone scintigraphy was obtained 2 to 3 hours after intravenous injection of 750 MBq of radiopharmaceutical. The whole-body field was used to record anterior and posterior views digitally with a resolution of 1024×256 pixels (approx. $205\,\text{cm} \times 61\,\text{cm}$). Images represent counts of detected gamma decays in each spatial unit with 16-bit grayscale depth.

12.2.3 Detection of reference points

Bone scans vary considerably in their size, contrast, brightness, and skeletal structure. In practice, many scans are only partial because only a determined part of the body is observed or because of the scanning limitations when the patient is too obese to fit in the screening area.

The idea of detecting the reference points is to find an image region that can easily be found in most cases. This region's reference point is then used as a starting location for further reference point detection. Further reference point search is guided with the background knowledge of the spatial relations between bone regions and specific bones. The following image processing algorithms aid the search within some predefined boundaries (regions of interest, or ROIs): beam search, dynamic binarization, dilatation, skeletonization, and circles and ellipses fitting using LSM. Afterwards, all detected reference points are shifted vertically and horizontally to the neighboring regions with the highest uptake (intensity).

In our study, we observed that on only 2 images out of 467, the shoulders were not visible. Many other possible starting points are missing on the images more often (i.e., head, arms, one or both legs). Therefore, we chose the shoulders as the main bone region to start with. The last assumption is the upward orientation of the image. This assumption is not limiting since all scintigraphies are made with the same upward orientation.

In order to make the detection of reference points faster (linear time complexity) and more reliable, we tried to automatically detect peaks that would roughly cover the reference points and build some structural skeleton for reference point search guidance. The idea is similar to the SIFT [10] algorithm, where image features are also represented by the most outstanding pixels over different resolutions.

12.2.3.1 Detection of the reference point candidates (peaks)

There are numerous methods for detecting image peaks. We tried several. The SIFT [10] method returned too few peaks to guarantee a robust region detection on a vast variety of skeletons. More classic image filters, such as Canny edge detection, returned too many peaks. Peaks with such filters can be acquired by connecting the peaks to polylines, which are then reduced to vertices. We tried another customized approach based on orthogonal two-way Gaussian and linear filtering that mimics the SIFT algorithm.

The algorithm smooths the image with the averaging window, whose size was experimentally determined. The algorithm works as shown in the pseudocode (see Algorithm 12.1).

Algorithm 12.1 Detection of relevant peaks Ω.

Require: Original image $I[x_i, y_i], i \in [0, m-1], j \in [0, n-1]$; Gaussian filter $G_{3\times 3} = \left(\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \right)$; size of averaging window $W(w \times h$ where w and h are odd); $d =$ minimal distance between detected peaks;

$$W_{w\times h} = \left(\frac{2}{w+h} \begin{bmatrix} 1 & 2 & \cdots & \frac{w+1}{2} & \cdots & 2 & 1 \\ 2 & 3 & \cdots & \frac{w+1}{2}+1 & \cdots & 3 & 2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ \frac{h+1}{2} & \frac{h+1}{2}+1 & \cdots & \frac{w+h}{2} & \cdots & \frac{h+1}{2}+1 & \frac{h+1}{2} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 2 & 3 & \cdots & \frac{w+1}{2}+1 & \cdots & 3 & 2 \\ 1 & 2 & \cdots & \frac{w+1}{2} & \cdots & 2 & 1 \end{bmatrix} \right)$$

Ensure: Relevant set of peaks Ω

1: $I_1 \Leftarrow (I \times G$ {Gaussian filter applied}).
2: $I_2 \Leftarrow (I_1 \times W_{w\times h}$ {bigger linear averaging filter applied}).
3: Add all elements from I_2 to Ω_1 where both left and right-side elements in I_2 have lower intensities {horizontal pass}.
4: Add all elements from I_2 to Ω_1 where both upper and lower elements in I_2 have lower intensities {vertical pass}.
5: Sort Ω_1 by element intensities.
6: Add consecutively all elements I_i from Ω_1 to Ω where all elements in Ω have distance to $I_i > d$.

The limited level of radioactivity injected due to radiation protection regulation in typical studies causes low intensities (count levels of gamma rays), which causes vague images of bone scans. Bone edges are more visible after images are filtered with some averaging algorithm (i.e., wavelet-based, median, or Gaussian filter) [12]. A Gaussian filter is used in order to enhance the detection of peaks. The image is smoothed with the Gaussian filter where the pixels of higher intensities (>90 gamma rays per pixel) are set to the predefined upper limit (90). Images smoothed in such a way are less prone to be obstructed by high-intensity lesions or artifacts since at this stage we are interested only in finding anatomical reference points and not possible lesions.

In the case of scintigraphic images, we may introduce scale-variant methods, since the nature of this image modality guarantees a fixed scale (100 pixels $\approx$ 24 cm). The size of the averaging window $W_{w\times h}$ was experimentally

set to 11×5 ($2.6\,\text{cm} \times 1.2\,\text{cm}$) because it gave best results by the means of segmentation accuracy.

The reference points 1.1 are searched for using the detected peaks (Figure 12.3). Both images, anterior and posterior, are simultaneously processed in the same detection order. Detected points from the anterior image are mirrored to the posterior, and vice versa. Mirroring is possible because both images are taken at the same time and are therefore aligned and have the same size. Some point $pt_{(x,y)}$ on one image is mirrored to the other one as $pt'_{(x',y')}$, where $x' = image_{width} - x, y' = y$ on the other image. Some bones are better visible on the anterior and some on the posterior images due to the varying distances from the collimators (gamma-ray sensors). In each step, the detected reference points visible in both images are compared. The algorithm

FIGURE 12.3: Example of detected image peaks on anterior and posterior image using Algorithm 12.1 ($d = 1, w \times h = 11 \times 5$) (all peaks are mirrored to the counter image).

chooses the one that is estimated to better represent the desired reference point. For each reference point type (knee, ankle, ilium, pubis, shoulders, etc.), we determined the rules that decide better choice.

When deciding the better reference point of the anterior and posterior scan, two aspects are observed. The expected neighboring uptake and relations to other detected reference points on the basis of expected skeletal ratios (e.g., the length of the upper arm humerus is expected to be approximately 0.67% of the spine length, spanning from the neck nape to the sacrum bone). More detailed algorithms (source code) can be found in Šajn [17].

Bigger bone regions have the same shape and position in both images, but usually the edge of one side of the bone is better expressed on one image due to the distance to collimator, whereas the other side is better expressed on the other image. With combining peaks from both anterior and posterior images, both sides are clearly expressed (i.e., ilium bone). Some bone regions (skull, ilium, pubic bones, etc.) can be represented by some basic geometric shapes (circles, lines, and ellipses), which can be determined by using the LSM method. The fitting of the geometric shapes is improved also using the mirrored points from anterior to posterior image, and vice versa.

The order in which the reference points were detected was determined using knowledge of the human anatomy as well as the physicians' recommendations. The anatomical knowledge is represented as a list of parameterized rules specific to each bone region. Rule parameters (e.g., thresholds, spatial and intensity ratios) were initially set by physicians and further refined on a separate tuning set.

12.2.3.2 Scintigraphy segmentation

Respective skeletal regions are processed in the following order: shoulders, pelvis, head, thorax, and extremities. Here we briefly describe the idea of locating the bone regions. More detailed procedures are described in our previous work [18].

The two shoulder reference points as the main starting points are detected with the first peaks (see Figure 12.3) found from diagonal directions (Figure 12.4).

The pelvis is located at the end of the spine and has approximately the same width as the shoulders. In order to find the pelvis, the estimation of the spine position is required. This is performed with a beam search. The most identifiable bone in pelvic region is the ilium bone, which has higher uptake values than its neighboring soft tissue. The ilium bone has a circular shape in the upper part and is therefore convenient for circle fitting with LSM.

The pubic bone is detected by estimating the pubis ROI using the detected ilium location, the distance between detected ilium circle centers, and the angle between them. The experimentally determined ROI's size is narrowed with the binarization algorithm. Circles representing the two ilium and pubic bones are detected with LSM using the detected peaks.

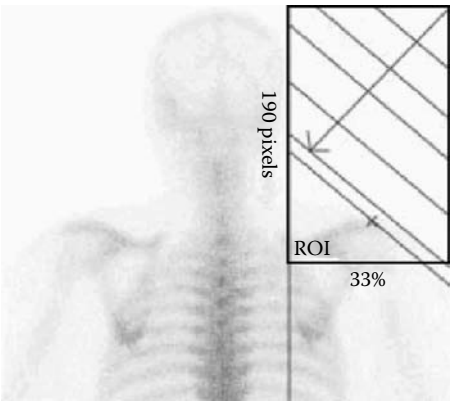


FIGURE 12.4: Locating shoulders.

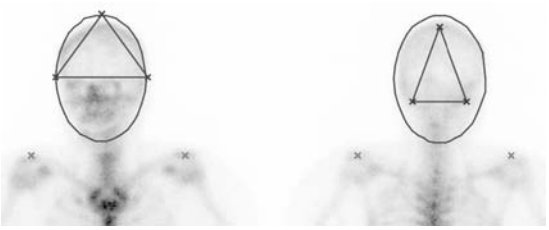


FIGURE 12.5: Head detection with ellipse fitting sizes.

When at least the image orientation and the location of the shoulders are known, some part of the neck or the head is visible between the shoulders. Finding the head is not difficult, but determining its orientation is, especially in cases where a part of the head is not visible in the scan. The most reliable method for determining the head orientation and position is ellipse fitting of the head contour, as shown in Figure 12.5.

Vertebrae have more or less specific intervertebral spaces [19]; the only problem is that on a bone scintigram, only a planar projection of the spine is visible. Since the spine is longitudinally curved, the vertebral spatial relations vary due to different axial orientation of the patients. The average intervertebral disc spaces have been experimentally determined from normal skeletons. After the approximate spine ROI is determined, it is divided into 10 vertical parts according to the predefined intervertebral disc.

Ribs are the most difficult skeletal region to detect because they are quite unexpressive on bone scans, their formation can vary considerably, and their contours can be disconnected in the case of stronger pathology. Using the detected vertebrae, the ribs' ROIs are defined. For the rib contour detection, we use the morphology-based image operations, particularly three well-known

algorithms: dynamic binarization, dilation/skeletonization [20], and Hough transform [21].

Extremities are often partly absent from the whole-body scan because of the limited gamma camera detector width. The regions of humerus, ulna, and radius as well as femur, tibia, and fibula bones are located with the use of beam search. The detection is designed in such a way that a part or the entire extremity or head may not be visible in the scan.

12.3 ArTeX Algorithm

12.3.1 Related work

Researchers have tried to characterize textures in many different ways. Most texture features are based on structural, statistical, or spectral properties of the image. Some methods use textural features that include several of these properties. Well-known statistical features are based on gray-level cooccurrence statistics [22], which are used in the "Image Processor" program [23]. Examples of structural features are features of Voronoi tesselation [24], representations using graphs [25], representations using grammars [26], and representations using association rules [27]. The spectral features are calculated in space, which is closely related to textural features—for example, frequency and amplitude. The most frequently used space transformations are the Fourier, Laws [28], Gabor [29], and wavelet transforms.

12.3.2 Association rules

Association rules were introduced by Agrawal, Imielinski, and Swami [30] in 1993. The following is a formal statement of the problem: Let $\mathcal{I} = \{i_1, i_2, \ldots, i_m\}$ be a set of literals, called items. Let $\mathcal{D}$ be a set of transactions, where each transaction T is a set of items such that $T \subseteq \mathcal{I}$. We say that a transaction T contains X if $X \subseteq T$. An *association rule* is an implication of the form $X \Longrightarrow Y$, where $X \subset \mathcal{I}, Y \subset \mathcal{I}$ and $X \cap Y = 0$. The rule $X \Longrightarrow Y$ holds in the transaction set $\mathcal{D}$ with confidence c if $c\%$ of transactions in $\mathcal{D}$ that contain X also contain Y. The rule $X \Longrightarrow Y$ has support s in the transaction set $\mathcal{D}$ if $s\%$ of transactions in $\mathcal{D}$ contain $X \cup Y$. The problem of discovering association rules says: Find all association rules in transaction set $\mathcal{D}$ with confidence of at least *minconf* and support of at least *minsup*, where *minconf* and *minsup* represent the lower boundary for confidence and support of association rules.

12.3.3 Texture representation

The use of association rules for texture description was independently introduced by Rushing and others [27]. Here we present a slightly different

approach, which uses different texture representation and a different algorithm for association rules induction and which we developed before we become aware of the work by Rushing and colleagues.

Association rules are most widely used for data mining of very large relational databases. In this section, we give a representation of texture, which is suitable for processing with the association rules algorithms. To apply the association rules algorithms on textures, we must first define the terms used for association rules in the context of textures.

Pixel $\vec{A}$ of a texture P is a vector $\vec{A} = (X, Y, I) \in P$, where X and Y represent the absolute coordinates and I represents the intensity of pixel A.

Root pixel $\vec{K}$ is the current pixel of a texture $\vec{K} = (X_K, Y_K, I_K)$.

R neighborhood $N_{R,\vec{K}}$ is a set of pixels located in the circular area of radius R with root pixel $\vec{K}$ at the center. Root pixel $\vec{K}$ itself is not a member of its neighborhood.

$$N_{R,\vec{K}} = \{(X, Y, I) | \delta \leq R\} \setminus \vec{K}$$

$$\delta = \left[\sqrt{(X_K - X)^2 + (Y_K - Y)^2} + 0.5 \right] \tag{12.1}$$

Transaction $T_{R,\vec{K}}$ is a set of elements based on its corresponding neighborhood. The elements of transaction are represented with Euclidean distance and intensity difference from the root pixel.

$$T_{R,\vec{K}} = \left\{ (\delta, I_K - I) | (X, Y, I) \in N_{R,\vec{K}} \right\}$$

$$\delta = \left[\sqrt{(X_K - X)^2 + (Y_K - Y)^2} + 0.5 \right] \tag{12.2}$$

Transaction element is a two-dimensional $(r, i) \in T_{R,\vec{K}}$, where the first component represents the Euclidean distance from the root pixel and the second component represents the intensity difference from the root pixel.

Association rule is composed of transaction elements; therefore, it looks like this:

$$(r_1, i_1) \wedge \cdots \wedge (r_m, i_m) \Longrightarrow$$
$$(r_{m+1}, i_{m+1}) \wedge \cdots \wedge (r_{m+n}, i_{m+n})$$

Transaction set $D_{P,R}$ is composed of transactions that are derived from all possible root pixels of a texture P at a certain neighborhood size R:

$$D_{P,R} = \left\{ T_{R,\vec{K}} | \forall \vec{K} : \vec{K} \in P \right\}$$

This representation of a texture replaces the exact information of location and intensity of the neighboring pixels with more indecisive information of the distance and the relative intensity of neighboring pixels. This description is also rotation invariant.

Figure 12.6 illustrates the association rule $(1,1) \wedge (2,10) \Longrightarrow (1,15) \wedge (3,5)$, which can be read as follows: If a pixel of intensity 1 is found on distance 1 and a pixel of intensity 10 is found on distance 2, then there is also a pixel of intensity 15 on distance 1 and a pixel of intensity 5 on distance 3.

This representation is almost suitable for processing with general association rule algorithms. What is still to be considered is the form of a transaction element. Association rule algorithms expect scalar values for transaction elements, whereas our representation produces a two-dimensional vector for a transaction element. Luckily, this issue can be easily solved. Let us say that the intensity of each texture point can have values from interval $[0..(Q-1)]$ and that the neighborhood size is R. Take some transaction element (r,i), where i has a value from $[-(Q-1)..+(Q-1)]$ and r has a value from $[1..R]$. What is needed here is a bijective mapping that transforms each vector to its scalar representation. This can be achieved in many ways. A possible and quite straightforward solution is:

$$s = (2Q-1)(r-1) + i + (Q-1)$$

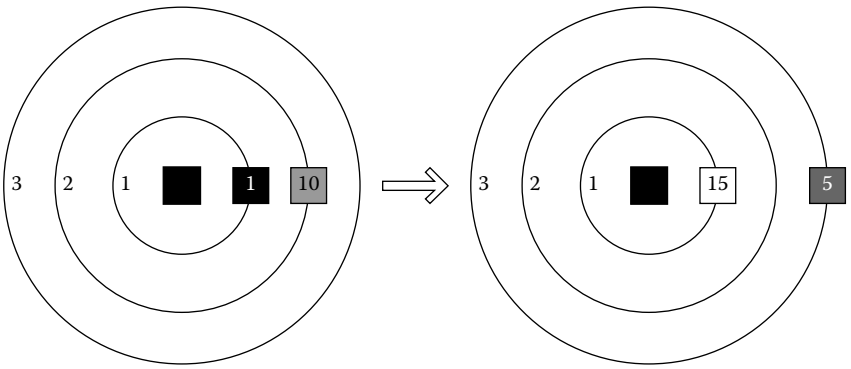


FIGURE 12.6: An illustration of association rule $(1,1) \wedge (2.10) \Longrightarrow (1,15) \wedge (3,5)$.

The transformation is also reversible:

$$r = 1 + s \operatorname{div}(2Q - 1)$$
$$i = s \bmod (2Q - 1) - (Q - 1)$$

Now it is possible to define a transaction that suits the general association rule algorithms:

$$T_{R,Q,\vec{K}} = \left\{ s \middle| \begin{matrix} (r,i) \in T_{R,\vec{K}} \\ s = (2Q-1)(r-1) + i + (Q-1) \end{matrix} \right\}$$

And finally, we obtain the appropriate transaction set definition:

$$\mathcal{D}_{P,R,Q} = \left\{ T_{R,Q,\vec{K}} \middle| \forall \vec{K} : \vec{K} \in P \right\}$$

12.3.4 From association rules to feature description

Using association rules on textures allows us to extract a set of features (attributes) for a particular domain of textures. The ArTeX algorithm is defined with the following steps:

1. *Select a (small) subset of images F for feature extraction.* The subset F can be considerably small. Use at least one example of each typical image in the domain—that is, at least one sample per class, or more if the class consists of subclasses.

2. *Preprocessing of images in F.* Preprocessing involves the transformation of images to grayscale if necessary, the quantization of gray levels, and the selection of proper neighborhood size R. The initial number of gray levels per pixel is usually 256. The quantization process downscales it to, say, 32 levels per pixel. Typical neighborhood sizes are 3, 4, 5.

3. *Generate association rules from images in F.* Because of the representation of texture, it is possible to use any algorithm for association rules extraction. We use *Apriori* and *GenRules*, as described in [30].

4. *Use generated association rules to extract a set of features.* There are two features associated with each association rule: support and confidence. Use these two attributes of all association rules to construct a feature set. The number of extracted features is twice the number of association rules (which could be quite a lot).

To clarify what we said, we also provide a formal algorithm (see Algorithm 12.2). The algorithm takes five input parameters: a set of images I, neighborhood size R, texture quantization Q, minimum support *minsup*, and minimum confidence *minconf*. Functions φ_{sup} and φ_{conf} are used to calculate support and confidence given an image and an association rule. The output of the algorithm is feature set matrix d, where $d_{i,j}$ represents the jth feature of image i.

Algorithm 12.2 ArTeX (image i, radius R, quantization Q, *minsup*, *minconf*).

Select F so that $F \subset I$
Preprocess(F, R, Q)
for all $f \in F$ **do**
  $\mathcal{D} =$ transactionModel(f, R, Q) STATE $r_1 =$ apriori$(\mathcal{D}, \text{minsup})$
  $r_2 =$ genRules$(r_1, \mathcal{D}, \text{minconf})$
  $\rho_{sup} = \rho_{sup} \cup r_1$ {itemsets with support > minsup}
  $\rho_{conf} = \rho_{conf} \cup r_2$ {rules with confidence > minconf}
end for
$i = 0$
for all $f \in (I \setminus F)$ **do**
  $j = 0$
  for all $\varrho \in \rho_{sup}$ **do**
    $d_{i,j} = \varphi_{sup}(f, \varrho)$ {jth attribute of ith image}
    $j = j + 1$
  end for
  for all $\varrho \in \rho_{conf}$ **do**
    $d_{i,j} = \varphi_{conf}(f, \varrho)$
    $j = j + 1$
  end for
  $i = i + 1$
end for
return d {d is a matrix of attribute values}

12.3.5 Extending the parameter set

Our model of texture is such that the structure of the association rule also describes some aspects of the textural structure. Since we are interested in the parametric description of a texture, this structure has to be represented with one or more parameters. Until now, we have presented the basic algorithm, which uses only basic interestingness measures, support, and confidence, which were defined together with association rules [30]. They are still most widely used today, but there are some concerns, especially with confidence measure, which can be misleading in many practical situations, as shown by Brin and others [31]. They also offered an alternative to evaluate association rules using the χ^2 test. Contrary to confidence measure, the χ^2 test could be used to find both positively and negatively correlated association patterns. However, the χ^2 test alone may not be the ultimate solution because it does not indicate the strength of correlation between items of association pattern. It only decides whether items of association pattern are independent of each other; thus, it cannot be used for ranking purposes.

We use the χ^2 test just to select interesting association patterns, which are later described by the Pearson correlation coefficient (ϕ-coefficient) as advised in Tan and Kumar [32]. ArTeX also uses an additional interestingness

measure, which was selected with thorough experiments on various domains from a subset of collections made by Tan and colleagues [33]. From all tested measures, the J-measure gave best results [34]. The J-measure for rule $A \rightarrow B$ is defined as follows:

$$P(A,B)\log\left(\frac{P(B|A}{P(B)}\right) + P(A,\overline{B})\log\left(\frac{P(\overline{B}|A}{P(\overline{B})}\right)$$

12.4 Multiresolutional Parameterization

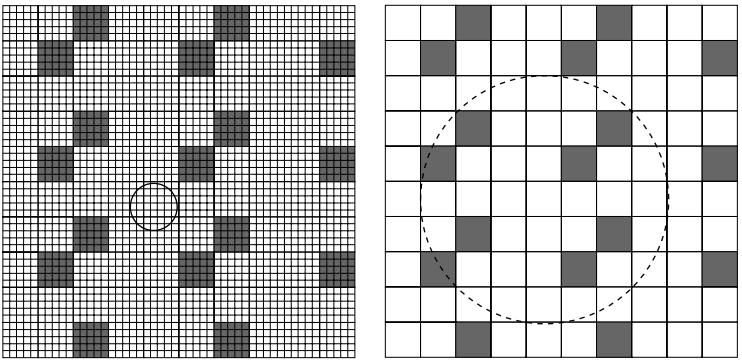
Why use more resolutions? Digital images are stored in matrix form, and algorithms for pattern parameterization basically use some relations between image pixels (usually first- or second-order statistics). By using only a single resolution, we may miss the big picture and proverbially not see the forest for the trees. Since it is too computationally complex to observe all possible relations between at least any two pixels in the image, we have to limit the search to some predefined neighborhood. These limitations make relations vary considerably over different resolutions. This means that we may get completely different image parameterization attributes for the same image at different scales.

12.4.1 Parameters from many resolutions

Many authors in their multiresolutional approaches use only mor resolutions that are independent of the image content [35–37]. Usually two or three resolutions are used. Authors report better classification results when using more resolutions and also observe that when using more than three resolutions, the classification accuracy starts to deteriorate. We have observed that in many cases authors use a set of resolutions by exponentially decreasing the resolution size ($\frac{100}{2^i}, i = 0..n-1$). However, we noticed that in many cases, equidistant selection of resolutions ($\frac{100i}{n}, i = 1..n$, where n is the number of resolutions used) gives better results. When using exponential forms of resolutions, less pattern content is examined and consecutively less significant attributes are derived. Another frequently used multiresolutional approach is the wavelet transform [38], which describes textures with measures calculated with iterative image division. None of the procedures mentioned above observes the contents of images.

Another extension of parameters for texture parameterization comes from the issue of pattern's scale. Not every combination of scale and neighborhood size can guarantee that the pattern is detected. The problem is illustrated in Figure 12.7.

To increase the possibility that the pattern will be detected, we propose a framework in which the extraction of attributes is repeated at different texture resolutions and combined in one feature vector.



(a) A very inadequate resolution (b) An adequate resolution

FIGURE 12.7: Detecting patterns at different scales.

12.4.2 Automatic selection of a small subset of relevant resolutions

The idea for the algorithm for automatic selection of a small subset of relevant resolutions is derived from the well-known SIFT algorithm [10]. The SIFT algorithm is designed as a stable local feature detector represented as a fundamental component of many image registration and object recognition algorithms. Since we are not interested in detecting stable image key points but in detecting resolutions at which the observed image has the most extremes, we devised a new algorithm, ARes (see Algorithm 12.3) for determining the resolutions for which more informative features can be obtained. The algorithm was designed especially for the ArTeX parameterization algorithm (see Algorithm 12.2) but also usually improves the results with other parameterization algorithms, as can be seen in Section 12.5. Resolutions for the bone scintigraphy are also determined using the ARes algorithm.

The ARes algorithm consequently resizes the image from 100% down to some predefined lowest threshold at some fixed step when detecting the appropriate resolutions. Both the lowest threshold and the resolution step are determined using the observed image data set. At each resize step, the peaks are counted. Peaks are represented by pixels that differ from their neighborhood either as highest or lowest intensity. This algorithm can also be implemented with the difference-of-Gaussian (DOG) [10] method, which improves the time complexity with a lower number of resizes required to search the entire resolution space.

The detected peak counts are recorded over all resolutions as a histogram. From the histogram, the best resolutions are detected as the highest counts. The number of resolutions we want to use in our parameterization is predefined by the user. When there are several equal counts, we choose resolutions as diverse as possible.

Algorithm 12.3 Algorithm ARes for detecting a small subset of relevant resolutions.

Require: Set of input images Θ with known classes, number of desired resolutions η, number of images to inspect in each class γ, radius ϕ used by the parameterization algorithm later in the process

Ensure: Subset of resolutions Π

1: $$W_{max} = \max_{i=1}^{|\Theta|}(\Theta_{i(width)}), H_{max} = \max_{i=1}^{|\Theta|}(\Theta_{i(height)})$$
{find the biggest image height and width}

2: extend the image sizes $\Theta_i \in \Theta$ to $W_{max} \times H_{max}$ by adding a frame of intensity equal to the average intensity of the original image Θ_i. New resized images are saved in the set Θ' {image sizes must be unified in order to be able to compare resolutions over different images}

3: $\delta = \frac{2*\phi}{3} \cdot \frac{1}{\min\{W_{max}, H_{max}\}}$ {set the resize step}

4: for each class, add γ randomly selected images from the set Θ' into the set Θ_1

5: $\Omega = \{\}$

6: **for** $(\forall\theta \in \Theta_1)$ **do**

7: $\nu = 1.0$ {start with 100% resolution}

8: **while** $(\min\{W_{max}, H_{max}\} \cdot \nu > 3 \cdot \phi)$ **do**

9: $\theta_1 = resize(\theta, \nu)$ {change the observed image's size}

10: Find local peaks in θ_1 by comparing each pixel's neighborhood inside $[3 \times 3]$ window

11: Add the pair {ν, number of peaks} into the set Ω

12: $\nu = \nu - \delta$

13: **end while**

14: **end for**

15: order the set Ω by the number of descending peaks and resolutions

16: add first η resolutions from the ordered set Ω into the final set Π

12.5 Results in Diagnostics of the Whole-Body Scintigrams

12.5.1 Segmentation

Approximately half of the scans were used for tuning the parameters required for reference point detection and the other half for testing the segmentation process. All 246 patients examined between October 2003 and March 2004 were used as the tuning set, and the 221 patients examined between April and June 2004 were used as the test set. In the tuning set, 38.9% of the images showed various nonosseous uptakes, 47.5% images contained visible injection points, and 6.8% were of children/adolescents with visible growth zones.

Similar distribution was found in the test set (34.5% nonosseous uptakes, 41.0% visible injection points, and 2.85% children/adolescents). Most of the artifacts were minor radioactivity points from urine contamination in the pelvic region or other parts (81.4% of all artifacts), whereas only a few other types were observed (urinary catheters 13%, artificial hips 4%, and lead accessories 1.6%). We observed that there were no incorrectly detected reference points in children/adolescents with the visible growth zones since all the bones are homogenous, have good visibility, and are clearly divided with growth zones. The segmentation algorithm works for adult and child/adolescent patients without any user intervention. Results of detecting the reference points on the test set are shown in Table 12.1.

Results are presented for different types of pathologies because we expected the degree of pathology to affect the quality of the detection process. The results show that there are no major differences in the reference point detection between different pathological groups.

The algorithm was implemented in a system called Skeleton 1.2 [17] (written in Java 2 SE, version 1.4.2) (see Figure 12.8). The current system includes image editing, filtering with different linear filters, conversions to other formats, storing images in the database with batch procedures, XML exports of reference points, manual correction of reference points, image annotation, and region localization using the scalable correlation. We also tested the accuracy of the bone region localization using correlation with generalized bone region images. In addition to bigger time complexity, the latter approach gave worse results compared to our segmentation algorithm. Therefore, all high-complexity algorithms were avoided, and applied algorithms were optimized in the sense of computational complexity. The detection of reference points on both anterior and posterior images takes approximately 3 seconds on a Pentium 4 PC with 2.8 GHz, 1 GB RAM.

12.5.2 Diagnosing pathologies with machine learning

The obtained reference points can be used in two ways to segment a scan. One way is to segment the scintigram only by extracting bones along the detected reference points with some predefined offset width. This type of segmentation was used in the classification part of our study. Since there is only a limited number of reference points, we can extract only 26 bones or bone regions (i.e., extremities, lumbal, thoracic, and cervical spine; 10 ribs; bones in pelvic region; and the head). Some extracted bones are shown in Figure 12.9c.

Another possibility of segmenting the scintigram is to map some standard, predefined skeletal mask over the scan observed using the detected reference points, as shown in Figure 12.9b. With this approach, we also need some virtual reference points, which are defined with the detected reference points; see Figure 12.9a. The skeletal mask represented with polygons can be defined by the radiologist on some scans. The algorithm triangulates the scan using the detected reference points. A new scan is triangulated in the same way

TABLE 12.1: Number of correctly detected reference points on the test set in different degrees of pathologies. Both frequencies and percentages are given.

Bone	**No Pathology**		**Slight Pathology**		**Strong Pathology**		**Super-Scan**		**All**	
Females	**22**		**78**		**15**		**1**		**116**	
Ilium	22	100%	76	97%	13	87%	0	0%	111	96%
Pubis	22	100%	76	97%	14	93%	1	100%	113	97%
Trochanter	22	100%	77	99%	15	100%	1	100%	115	99%
Shoulder	22	100%	78	100%	14	93%	1	100%	115	99%
Extremities	21	96%	74	95%	15	100%	1	100%	111	96%
Spine	22	100%	78	100%	14	93%	1	100%	115	99%
Ribs	16	73%	70	90%	12	80%	1	100%	99	85%
Neck	22	100%	76	97%	15	100%	1	100%	114	98%
Average	96%		97%		93%		88%		96%	
Males	24		55		24		2		105	
Ilium	24	100%	55	100%	20	83%	2	100%	101	96%
Pubis	22	92%	54	98%	23	96%	2	100%	101	96%
Trochanter	24	100%	55	100%	24	100%	2	100%	105	100%
Shoulder	24	100%	55	100%	24	100%	2	100%	105	100%
Extremities	20	83%	48	87%	24	100%	2	100%	94	90%
Spine	24	100%	53	96%	24	100%	2	100%	103	98%
Ribs	19	79%	46	84%	24	100%	2	100%	91	87%
Neck	22	92%	53	96%	24	100%	2	100%	101	96%
Average	93%		95%		97%		100%		95%	
Sum	46		133		39		3		221	
Ilium	46	100%	131	98%	33	85%	2	67%	212	96%
Pubis	44	96%	130	98%	37	95%	3	100%	214	97%
Trochanter	46	100%	132	99%	39	100%	3	100%	220	100%
Shoulder	46	100%	133	100%	38	97%	3	100%	220	100%
Extremities	41	89%	122	92%	39	100%	3	100%	205	93%
Spine	46	100%	131	98%	38	97%	3	100%	218	99%
Ribs	35	76%	116	87%	36	92%	3	100%	190	86%
Neck	44	96%	129	97%	39	100%	3	100%	215	97%
Average	95%		94%		96%		96%		96%	

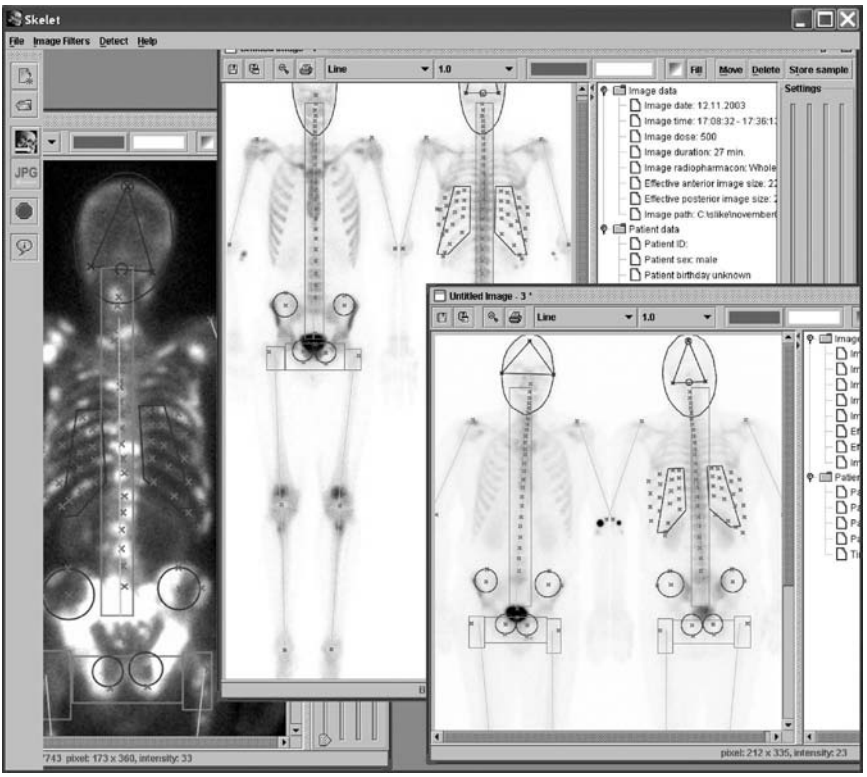


FIGURE 12.8: (See color insert following page 370.) View of the program skeleton.

so the defined skeletal mask can be mapped. Mapping is carried out with the linear transform for each triangle. The defined skeletal mask polygons are transformed with a linear operation that multiplies the polygon vertices $p_{im}(x_{im}, y_{im})$ with the matrix A calculated from the belonging triangle coordinates in the original image $v_{1o}(x_{1o}, y_{1o})$, $v_{2o}(x_{2o}, y_{2o})$, $v_{3o}(x_{3o}, y_{3o})$ and in the new image $v_{1m}(x_{1m}, y_{1m})$, $v_{2m}(x_{2m}, y_{2m})$, $v_{3m}(x_{3m}, y_{3m})$.

$$A = \begin{bmatrix} x_{1m} & x_{2m} & x_{3m} \\ y_{1m} & y_{2m} & y_{3m} \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{1o} & x_{2o} & x_{3o} \\ y_{1o} & y_{2o} & y_{3o} \\ 1 & 1 & 1 \end{bmatrix}^{-1}, p_{im} = A \cdot p_{io}$$

When all reference points are obtained, every bone is assigned a portion of original scintigraphic image according to relevant reference points.

The results of our parameterization with our ArTeX algorithm are compared with the results of four other image parameterization algorithms (Haar wavelets [38], Laws filters [39], Gabor filters [40], image processor [23] (implements many parameters of the first- and second-order statistics [22]), and Laws texture measures [39]).

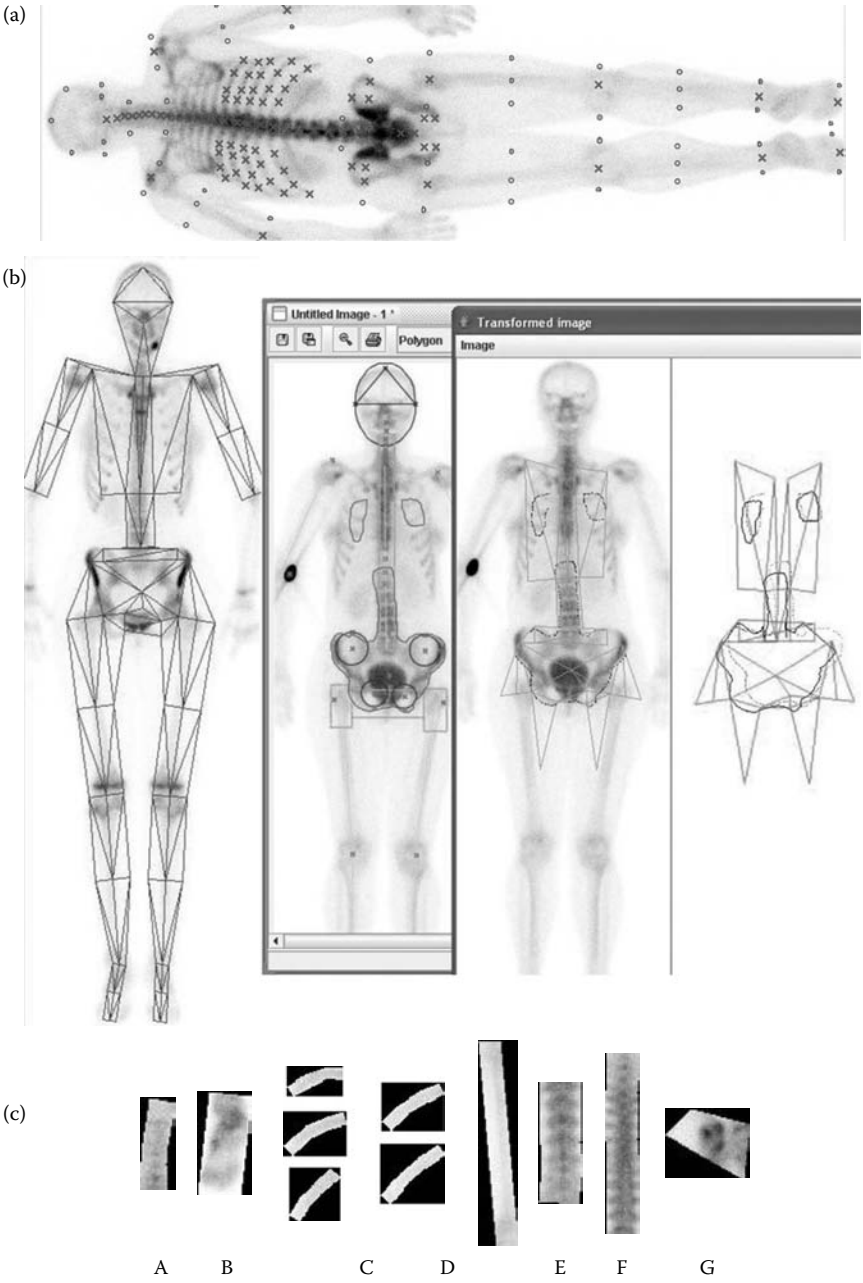


FIGURE 12.9: Using reference points for scintigraphy segmentation. (a) Additional virtual reference points (b) Triangulation mask and an example of a mapped annotation between different scans (c) Examples of extracted bones (A, cervical spine; B, foot; C, ribs; D, femur; E, lumbal spine; F, thoracic spine; G, sacroiliac joint).

The images of bones were described with several hundred automatically generated attributes with the ArTeX algorithm. Attributes are invariant to rotation and illumination changes. Rotation invariance is very important in this case because it compensates for different patients' positions inside the camera, whereas the illumination compensates for different absorptions of radiopharmaceutical throughout the body. Attributes were used for training the naive Bayesian classifier implemented in Weka [41]. In our preliminary experiments, various pathologies were not discriminated; that is, bones were labeled with only two possible diagnoses: no pathology or pathology.

From our complete set of 467 patients, pathologies were thoroughly evaluated by physicians only for 268 patients. These 268 patients were used for the experiments with machine learning. In 21% of scans, no pathology or other artifacts were detected by the expert physicians. In the remaining 79% of the scans, at least one pathology or artifact was observed. All bones were classified as normal or pathological and grouped into 10 regions. For each region, we have a binary classification problem. For machine learning, we used all pathological bones and randomly sampled healthy bones so that in all 10 problems, 30% of the bones were pathological, and 70% were healthy. Results were evaluated with tenfold cross validation and are shown in Tables 12.2 and 12.3. Table 12.2 gives the classification accuracy with the ArTeX algorithm, and Table 12.3 compares the other algorithms. The best results were achieved when ArTeX with J-measure and resolutions determined with ARes were used for parameterization of images. The Friedman rank test showed that ArTeX was significantly better ($\alpha = 0.05$) than the Laws and Haar algorithms; however, the difference compared with the Gabor and image processor algorithms using resolutions determined with ARes is not significant.

TABLE 12.2: Experimental results (classification accuracy [%]) with machine learning on two-class problem at different resolutions. Parameterized with ArTeX using the J-measure.

Bone Region	**100% Resolution**	**4 Equidistant Resolutions**	**4 Resolutions with ARes**
Cervical spine	76.00	72.00	75.00
Feet	84.44	84.44	85.56
Skull posterior	70.00	80.00	80.00
Ilium bone	88.48	90.30	90.30
Lumbal spine	69.10	74.42	71.41
Femur and tibia	84.15	84.15	86.41
Pelvic region	88.57	91.43	94.29
Ribs	94.31	95.42	95.42
Scapula	95.00	95.00	95.00
Thoracic spine	71.67	71.67	81.12
Average	82.17	83.88	85.45

TABLE 12.3: Experimental results with other algorithms.

Algorithm	100% Resolution	Equidistant	ARes
ArTeX (no J-measure)	83.25	83.00	84.91
Gabor	83.61	84.41	84.12
Haar	80.16	79.80	81.82
Laws	82.60	83.00	83.39
Image processor	81.47	83.59	84.46

TABLE 12.4: Specificity and sensitivity of the classification.

Bone Region	100% Resolution		Equidistant		ARes	
	Sp. %	Sen. %	Sp. %	Sen. %	Sp. %	Sen. %
Cervical spine	88.89%	42.86%	88.89%	28.57%	86.10%	56.08%
Feet	93.75%	61.54%	93.75%	61.54%	93.75%	65.38%
Skull posterior	84.62%	0.00%	100.00%	0.00%	100.00%	0.00%
Ilium bone	91.36%	81.82%	93.83%	81.82%	93.83%	81.82%
Lumbal spine	80.68%	43.90%	88.64%	43.90%	81.82%	48.78%
Femur and tibia	91.20%	67.31%	95.20%	57.69%	94.40%	73.08%
Pelvic region	94.00%	75.00%	96.00%	80.00%	98.00%	85.00%
Ribs	95.16%	92.00%	98.39%	88.00%	98.39%	88.00%
Scapula	100.00%	81.82%	100.00%	81.82%	100.00%	81.82%
Thoracic spine	82.93%	43.75%	85.37%	37.50%	85.37%	61.20%
Average	90.26%	59.00%	94.01%	56.08%	93.16%	64.12%

Table 12.4 gives sensitivity and specificity of the classification results. ArTeX achieved the best specificity. The Friedman rank test showed that in the specificity of the bone classification, ArTeX was significantly better ($\alpha = 0.05$) than the Haar and Gabor algorithms; however, the difference compared with the image processor and Laws algorithms is not significant. Sensitivity was not significantly different in any comparison of ArTeX to other algorithms.

12.6 Discussion

The detection of the reference points gave excellent results for all bone regions except for the ribs and the extremities, where some parts were missing in the scan. The extremities were mainly misdetected in the cases where the humerus bone was partially missing.

As expected, the detection of ribs was the most difficult. The results show that in 14% to 20% of scans, there were difficulties in detecting ribs. The ribs in the thoracic areas were hard to follow due to a vague expression. Generally, one rib was missed or not followed to the very end. We intend to improve

this problem in the future. In the present system (Figure 12.8), such reference points can be manually repositioned by the expert physician.

Since a robust segmentation algorithm should not fail on partial skeletal images, which is often the case in clinical routine (18% of the scans in our study), special attention was paid to such cases (e.g., amputees and skeletal parts entirely invisible in the scan). In our results, such cases do not stand out from the normal scans.

The automatically detected reference points can be used for mapping a standard skeletal reference mask, which we believe is the best way to find individual bones on scintigrams, since individual bones are often not expressive enough that their contours can be followed. Examples of such mask mapping and extracted bones are shown in Figure 12.9b and 12.9c.

Segmented bone images can be treated as textures because they have similar structure over a certain bone region. Bones as textures can be parameterized with some algorithms for pattern parameterization. Pathology represents some unusual texture patterns that can be distinguished from healthy cases. The ArTeX algorithm for pattern parameterization alone performs just as well as other algorithms in terms of classification accuracy. When we extend the parameter set with the J-measure and use the multiresolutional approach, the performance greatly improves. The resolutions for the multiresolutional approach must be carefully selected, as shown in the results. Resolution selection is performed by the algorithm ARes, which is also described in the chapter and improves ArTeX as well as other parameterization algorithms (e.g., Haar, Laws, and image processor). The same performance is achieved in other medical domains (e.g., ischemic heart disease), which are not discussed in this chapter.

12.6.1 Future research directions

The algorithms for pattern parameterization presented here open a new research area of multiresolution image parameterization and enable many applications in medical, industrial, and other domains where textures or texturelike surfaces are classified. The ARes algorithm can be improved with additional resolution search refinements, which would be more domain oriented. In our case, we plan to study malignant pathologies and to seek different criteria for resolution quality evaluation.

While our experimental results with machine learning are quite satisfactory, they were obtained for a simplified (two-class) problem only. Simply extending a problem to a multiclass paradigm is not acceptable since the bone may be assigned several different pathologies at the same time. We are currently developing a new approach in which the problem is rephrased to the multilabel learning problem, and each bone will be labeled with a nonempty subset of all possible labels [42,43].

12.7 Conclusions

The computer-aided system for bone scintigraphy presented here is a step toward automating the routine medical procedures. Some standard image processing algorithms were tailored and used in combination to achieve the best reference point detection accuracy on scintigraphic images, which have technically very low resolution. Because of poorer image resolution compared to radiography, the presence of artifacts and pathologies necessitates that algorithms use as much background knowledge of anatomy and spatial relations of bones as possible in order to work satisfactorily. This combination gives quite good results, and we expect that further studies on automatic scintigram diagnosis using reference points for image segmentation will give more accurate and reliable results than presented in previous studies that did not use segmentation.

For improving classification accuracy in medical image domains, we encourage the multiresolutional parameterization approach. As we have observed, the resolutions should be determined according to the properties of the observed domain.

This approach offers a new view of automatic bone scintigraphy evaluation since in addition to detection of pointlike high-uptake lesions, it also offers:

- More accurate and reliable evaluation of bone symmetry when looking for skeletal abnormalities
- Detection of a greater number of abnormalities, since many abnormalities can be spotted only when the symmetry is observed (differences in length, girth, curvature, etc.)
- Detection of lesions with low uptake or lower activity due to metallic implants (e.g., artificial hip)
- Possibility of comparing uptake ratios among different bones
- More complex pathology detection by combining pathologies of more bones (e.g., arthritis in joints)
- Possibility of automatic reporting of bone pathologies in written language for educational purposes

The machine learning approach described in this work is in an early stage of development. However, the preliminary results are encouraging, and the multiresolutional approach will make them more useful for clinical applications.

References

[1] V. Mäuller, J. Steinhagen, M. deWit, and H. K. Bohuslavizki. Bone scintigraphy in clinical routine. *Radiology & Oncology*, 35(1):21–30, 2001.

[2] A. Hendler and M. Hershkop. When to use bone scintigraphy. It can reveal things other studies cannot. *Postgraduate Medicine*, 104(5):54–66, 1998.

[3] M. Noguchi, H. Kikuchi, M. Ishibashi, and S. Noda. Percentage of the positive area of bone metastasis is an independent predictor of disease death in advanced prostate cancer. *British Journal of Cancer*, 88:195–201, 2003.

[4] J. Bernauer. *Zur semantischen Rekonstruktion medizinischer Begriffssysteme. Habilitationsschrift.* Institut für Medizinische Informatik, Universität Hildesheim, 1995.

[5] K. C. Berning. *Zur Automatischen Befundung und Interpretation von Ganzkörper-Skelettszintigrammen.* PhD thesis, Institut für Medizinische Informatik, Universität Hildesheim, 1996.

[6] A. Benneke. Konzeption und Realisierung eines semi-automatischen Befundungssystems in Java und Anbindung an ein formalisiertes Begriffssystem am Beispiel der Skelettszintigraphie. Diplom Arbeit, Institut für Medizinische Informatik, Universität Hildesheim, mentor Prof. Dr. D. P. Pretschner, 1997.

[7] M. N. Maisey, T. K. Natarajan, P. J. Hurley, and H. N. Jr. Wagner. Validation of a rapid computerized method of measuring 99mTc pertechnetate uptake for routine assessment of thyroid structure and function. *Journal of Clinical Endocrinology & Metabolism*, 36:317–322, 1973.

[8] T. K. Yin and N. T. Chiu. A computer-aided diagnosis for locating abnormalities in bone scintigraphy by a fuzzy system with a three-step minimization approach. *IEEE Transactions on Medical Imaging*, 23(5):639–654, 2004.

[9] M. Kukar, I. Kononenko, C. Grošeli, K. Kralj, and J. Fettich. Analysing and improving the diagnosis of ischaemic heart disease with machine learning. *Artificial Intelligence in Medicine*, 16:25–50, 1999.

[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[11] H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1 1995.

[12] G. Jammal and A. Bijaoui. DeQuant: A flexible multiresolution restoration framework. *Signal Processing*, 84(7):1049–1069, 2004.

[13] I. Kononenko and M. Kukar. *Machine Learning and Data Mining: Introduction to Principles and Algorithms*. Chichester, UK: Horwood, 2007.

[14] M. Kukar, L. Šajn, C. Grošelj, and J. Grošelj. Multi-resolution image parameterization in sequential diagnostics of coronary artery disease. In R. Bellazzi, A. Abu-Hanna, and J. Hunter, Eds., *Artificial Intelligence Image Segmentation and Parameterization for Automatic Diagnostics of Whole-Body Scintigrams in Medicine*, pp. 119–129. Berlin: Springer, 2007.

[15] I. Kononenko. Inductive and Bayesian learning in medical diagnosis. *Applied Artificial Intelligence*, 7:317–337, 1993.

[16] M. G. Weiner, L. Jenicke, V. Müller, and H. K. Bohuslavizki. Artifacts and non-osseous uptake in bone scintigraphy. Imaging reports of 20 cases. *Radiology & Oncology*, 35(3):185–91, 2001.

[17] L. Šajn. Program source code, 2007. lkm.fri.uni-lj.si/skelet/skelet.zip.

[18] L. Šajn, M. Kukar, I. Kononenko, and M. Milčinski. Computerized segmentation of whole-body bone scintigrams and its use in automated diagnostics. *Computer Methods & Programs in Biomedicine*, 80(1):47–55, 2005.

[19] G. Holmaas, D. Frederiksen, A. Ulvik, S. O. Vingsnes, G. Ostgaard, and H. Nordli. Identification of thoracic intervertebral spaces by means of surface anatomy: A magnetic resonance imaging study. *Acta Anaesthesiologica Scandinavica*, 50(3):368–373, 2006.

[20] H. Blum. *Models for the Perception of Speech and Visual Form*. Cambridge, MA: MIT Press, 1967.

[21] P. V. C. Hough. Machine analysis of bubble chamber pictures. International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.

[22] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(11):610–621, 1973.

[23] M. Bevk and I. Kononenko. A statistical approach to texture description of medical images: A preliminary study. In *The 19th International Conference on Machine Learning ICML'02 Workshops*, Sydney, 2002.

[24] M. Tuceryan and A. K. Jain. Texture segmentation using Voronoi polygons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 211–216, 1990.

[25] M. Tuceryan and A. K. Jain. Texture analysis. In C. H. Chen, L. F. Pau, and P. S. P. Wang, Eds., *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed., pp. 207–248. River Edge, NJ: World Scientific, 1998.

[26] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*, 2nd ed., vol. 2. New York: Academic Press, 1982.

[27] J. A. Rushing, H. S. Ranganath, T. H. Hinke, and S. J. Graves. Using association rules as texture features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):845–858, 2001.

[28] K. I. Laws. *Textured Image Segmentation*. PhD thesis, Dept. Electrical Engineering, University of Southern California, 1980.

[29] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.

[30] R. Agrawal, T. Imielinski, and A. N. Swami. Mining association rules between sets of items in large databases. In Peter Buneman and Sushil Jajodia, Eds., *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, pp. 207–216, New York: ACM Press: 1993.

[31] S. Brin, R. Motwani, and C. Silverstein. Beyond market baskets: Generalizing association rules to correlations. In Joan Peckham, Ed., *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, May 13–15, 1997, Tucson, Arizona, USA*, pp. 265–276. New York: ACM Press, 1997.

[32] P. Tan and V. Kumar. Interestingness measures for association patterns: A perspective. Technical Report TR00–036, Department of Computer Science, University of Minnesota, 2000.

[33] P. Tan, V. Kumar, and J. Srivastava. Selecting the right interestingness measure for association patterns. In *Proceedings of the Eight ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002.

[34] M. Bevk. *Derivation of texture features using association rules*. PhD thesis, Faculty of Computer and Information Science, University of Ljubljana, 2005. In Slovene.

[35] C. Bastos Rocha Ferreira and D. Leandro Borges. Automated mammogram classification using a multiresolution pattern recognition approach. *SIBGRAPI01*, 00:76, 2001.

[36] M. L. Comer and E. J. Delp. Segmentation of textured images using a multiresolution Gaussian autoregressive model. *Image Processing, IEEE Transactions on Image Processing*, 8(3):408–420, 3, 1999.

[37] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution grayscale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[38] C. K. Chui. *An Introduction to Wavelets*. San Diego: Academic Press, 1992.

[39] K. Laws. *Textured image segmentation*. PhD thesis, University of Southern California, 1, 1980.

[40] S. E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on Gabor filters. *IEEE Transactions on Image Processing*, 11(10):1160–1167, 2002.

[41] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Francisco: Morgan Kaufmann, 1999.

[42] X. Shen, M. Boutell, J. Luo, and C. Brown. Multi-label machine learning and its application to semantic scene classification. In *Proceedings of the 2004 International Symposium on Electronic Imaging (EI 2004)*, San Jose, California, 2004.

[43] A. McCallum. Multi-label text classification with a mixture model trained by EM. In *Proceedings of the AAAI'99 Workshop on Text Learning*, 1999.

Chapter 13

Distributed 3-D Medical Image Registration Using Intelligent Agents

Roger J. Tait, Gerald Schaefer, and Adrian A. Hopgood

Contents

A major drawback of 3-D medical image registration techniques is the performance bottleneck associated with similarity calculation. Such bottlenecks limit registration applications in clinical situations where fast execution times are required and become particularly apparent in the case of volumetric data sets. In this chapter, a framework for high-performance intensity-based volume registration is presented. Geometric alignment of both reference and sensed volumes is achieved through a combination of translation, rotation, and similarity evaluation. Crucially, similarity estimation is performed intelligently

by agents. The agents work in parallel and communicate with one another by means of a distributed blackboard architecture. Partitioning of the blackboard is used to balance communication and processing workloads. The framework described demonstrates the flexibility of coarse-grained parallelism and shows how high-performance registration can be achieved with nonspecialized architectures.

13.1 Introduction

The ability to visualize hidden structures in detail using 3-D data has become a valuable resource in medical imaging. Importantly, the alignment of volumes makes possible the combining of different structural and functional information for diagnosis and planning purposes. According to Zitova and Flusser (2003), registration can be classified as either landmark or intensity based because a universal registration method is impractical due to the wide variety of noise and geometric deformations contained within captured volumes. During landmark-based registration, corresponding subsets of samples are used to estimate alignment accuracy; this method is in contrast with intensity-based registration, which employs the complete set of voxels from each volume (Temkin, Vaidyanath, and Acosta, 2005).

Transform optimization, resampling, and similarity calculation form the basic stages of an intensity-based registration process. During transform optimization, translation and rotation parameters that geometrically map voxels in the reference (fixed) volume to voxels in the sensed (moving) volume are estimated. Once estimated, intensities that are mapped into noninteger coordinates are interpolated in the resampling stage. After resampling, a metric is used for similarity calculation in which the degree of likeness between corresponding volumes is evaluated. Optimization of the similarity measure is the goal of the registration process and is achieved by seeking the best transform. Transform parameters are in general defined as the search space. Other search parameters, however, may be defined depending on the type of registration algorithm employed (Cordon, Damas, and Santamaria, 2006). Due to the iterative nature of registration algorithms, similarity calculation represents a considerable performance bottleneck, which limits the speed of time critical applications.

Inputs to and output from a similarity metric are illustrated in Figure 13.1. In general, a metric works by examining corresponding voxels in both fixed and moving volumes and then formulating a measure of similarity based on the relationship between these intensities (Penney et al., 1998). The metric assumes that the relationship changes with variations in the spatial transform used to map between voxels, and a maximum similarity is achieved when volumes are in close alignment. Intensity equality, which is high when voxels are similar, is one such relationship employed as a similarity metric when

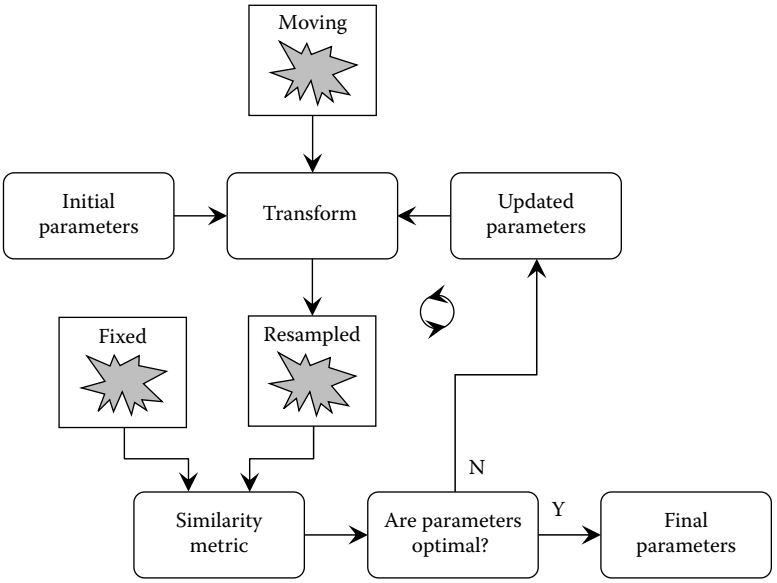


FIGURE 13.1: Inputs to the similarity metric. The output is a single value used to determine suitability of the transform parameters.

volumes are captured using the same sensor type. Total equality is, however, seldom reached due to noise and volume acquisition inconsistencies. Additional robustness is therefore achieved by assessing the ratio of intensities and minimizing the variance of such ratios (Jeongtae and Fessler, 2004). When volumes are acquired with different sensor types, an extension of the ratio method that maximizes the weighted collection of variances can be employed. Examples include the mutual information–based similarity metrics described by Pluim, Maintz, and Viergever (2003).

Although a number of different approaches to overcome the similarity calculation bottleneck exist (Wang and Feng, 2005; Zhang and Blum, 2001), the use of parallel computing to overcome the time constraints associated with volume registration applications is growing in popularity. Conveniently, many of the similarity calculation strategies employed are inherently parallel and therefore well suited to distribution. An important consideration when adopting a parallel processing approach is the architecture of the host system (Seinstra, Koelma, and Geusebroek, 2002). In a computer constructed of multiple processors with shared memory, data distribution is not required. These systems are viewed as tightly coupled architectures. A loosely coupled architecture, in contrast, consists of multiple computers in different locations. Loosely coupled architectures require data distribution, communication, and accumulation mechanisms. Importantly, the most effective distribution scheme will depend on the architecture of the host system. Contrasting architectures of host systems are illustrated in Figure 13.2.

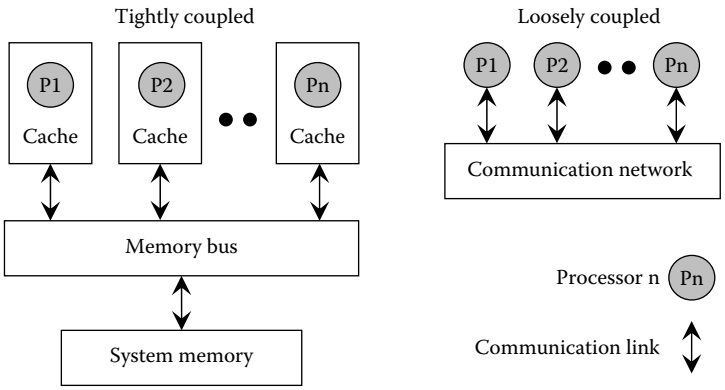


FIGURE 13.2: Tightly and loosely coupled architectures. Data are either fetched from main memory or transferred over a communications network.

13.2 Background

Understandably, parallel architectures have been employed by a number of researchers. Clinically compatible speeds have been achieved by Warfield, Jolesz, and Kikinis (1998), who introduced a parallel nonrigid algorithm based on the work-pile paradigm. Their goal was to develop an interpatient registration algorithm that can be applied without operator intervention to a database of several hundred scans. In an initial step, each scan is segmented using a statistical classification method. This preprocessing stage is used to identify different tissue types, including skin, white matter, gray matter, and bone structure. Once segmented, a transform that brings these features into alignment is estimated. In the system described, a message-passing interface and cluster of symmetric multiprocessors execute parallel similarity calculation operations using multiple threads. Crucially, work is dynamically load-balanced across a cluster of multiprocessor computers. Results published by the group show that successful registration of $256 \times 256 \times 52$ volume brain scans has been achieved in minutes rather than hours.

Christensen (1998) compares two nonthreaded architectures, Multiple Instruction Multiple Data (MIMD) and Single Instruction Multiple Data (SIMD). The work presented raises implementation issues and timing analysis for the registration of $32 \times 32 \times 25, 64 \times 64 \times 50$, and $128 \times 128 \times 100$ volume data sets. During each clock cycle, the SIMD implementation performs calculations in which all processors are performing the same operation. The MIMD implementation, in contrast, breaks an algorithm into independent parts, which are solved simultaneously by separate processors. The movement of data in both shared-memory systems is unrestricted, and during execution, each

processor has access to the whole memory. The main performance bottleneck associated with both approaches was reported as scalability of hardware with increasing numbers of processors. Crucially, the MIMD implementation is recorded as being approximately four times faster than its SIMD counterpart. Reduced performance of the SIMD implementation is reportedly caused by overheads during serial portions of the registration algorithm.

More recently, the demands placed on registration algorithms when aligning deformable structures in 3-D space have been discussed. Salomon et al. (2005) introduced the deformable registration of volumes, which involves optimization of several thousand parameters and requires several hours of processing on a standard workstation. Based on simulation of stochastic differential equations and using simulated annealing, a parallel approach that yields processing times compatible with clinical routines is presented. The approach represents a hierarchical displacement vector field that is estimated by means of an energy function. The energy function is scaled in relation to the similarity measure and is reevaluated at the end of each transform parameter optimization cycle. The algorithm is reportedly suited to massively parallel implementation and has been successfully applied to the registration of $256 \times 256 \times 256$ volumes. Again, the results published show how alignment can be achieved in minutes rather than hours.

Although these research projects demonstrate the performance benefits achievable with such architectures, they lack data distribution capabilities. This limitation strictly bounds the size of volumes to the smallest memory size within the architecture. In order to simultaneously overcome the limitations of memory and performance, a data-distributed parallel algorithm was developed by Ino, Ooyama, and Hagihara (2005). Based on Schnabel's implementation, the algorithm performs multimodal volume registration using adaptive mesh refinement, which requires no user interaction or preprocessing stage. The data-distribution scheme employed allows for increased volume size and is achieved by assigning partitioned volume segments to all available processors. Efficiency of the algorithm is improved through the inclusion of load balancing, which manages the computational cost associated with each volume segment. Experimental results obtained on a 128-processor cluster are based on volumes $1024 \times 1024 \times 590$ voxels in size.

While good for maximizing speedup, the fine-grained parallelism employed complicates distribution of the registration process and reduces flexibility of the approach because the basic alignment steps are distributed between all processors. In this research, we describe a coarse-grained approach to parallelism that increases flexibility and allows the issues of fine-grained parallelism to be ignored. Building on a distributed blackboard architecture, the framework is designed to achieve high-performance intensity-based volume registration. The approach supports multiple distributed intelligent agents organized in a worker/manager model. Crucially, the basic alignment steps are allocated to individual processors, the most computationally intensive of which are performed concurrently.

In this chapter, an intensity-based registration algorithm is introduced, and its mathematical derivation is given. Mapping of algorithm components to individual agents, plus development of information strings to control the transform optimization process, are discussed in detail. Functionality and agent behavior implemented for similarity calculation between volume segments and the updating of transform parameters at the end of each optimization cycle are also presented. The chapter concludes with a discussion of experimental testing, conclusions, and future trends.

13.3 Intensity-Based Image Registration

When described formally, the inputs to a registration process can be defined as the fixed volume F, the moving volume M, and the transform T used to map voxel coordinates. As previously stated, the goal of the registration process is recovery of a spatial mapping that brings the two volumes into alignment. To achieve this, the metric $S(F, M, T)$ is employed to generate a measure of similarity based on how well aligned the transformed moving volume is with the fixed volume. The measure of similarity produced forms a quantitative criterion that can be optimized in a search space defined by transform parameters. Importantly, by employing a gradient-descent optimization technique, the metric can be used to produce derivatives of the similarity measure with respect to each transform parameter. Using the resulting derivatives, updated transform parameters can be estimated and evaluated. To overcome the problems of local minima, common in gradient-based optimization, other techniques can be employed. Genetic algorithms employed in two resolution phases (Chalermwat, El-Ghazawi, and LeMoigne, 2001) and evolutionary algorithms for nonliner optimization (Cordon, Damas, and Santamaria, 2006) are two such examples.

Given the fixed volume F and moving volume M, the calculation of similarity metric derivatives for updating the set of transform parameters P, as described by Yoo (2004), is defined as

$$\frac{\partial S(P|F,M,T)}{\partial p_i} = \sum_{j=1}^{Q} \frac{\partial S(P|F,M,T)}{\partial x'_j} \frac{\partial x'_j}{\partial p_i} \tag{13.1}$$

where Q is the number of valid voxels, identified by j, mapped between volumes, and $\partial x'_j / \partial p_i$ is a matrix called the transform Jacobian. Typically, transform T works on the set of transform parameters P indexed by p_i.

Conveniently, the transform T can be implemented as a Quaternion (Conway and Smith, 2003), which expresses the rotational relationship between two vectors. The Quaternion makes possible the retrieval of one vector by operating with the other. To achieve this, the orientation of the first vector in relation to a second vector is described using a versor. The

change in magnitude between the two vectors, in contrast, is encapsulated as a tensor. Since the versor represents an orientation change, it provides a convenient description of rotations in 3-D space. Also, when the versor is coupled with a translation vector, a rotation and translation in 3-D space can be described. Using the versor type, the voxel location x in the fixed volume is mapped to the new position x' in the moving volume using

$$x' = V^* x + t \tag{13.2}$$

where V is a versor, * stands for the multiplication of the versor with a vector, and t is a translation vector. In general, the versor consists of three components: direction, angle, and norm. The direction of a versor is taken as being parallel to the axis around which the first vector is rotated, in order to be mapped to the second vector, and orthogonal to the plane defined by the two vectors. The angle, in contrast, is the measure of rotation between the two vectors for which the versor is a Quaternion. The norm of the versor is defined as a function of its rotation angle.

The transform Jacobian $\partial x'_j/\partial p_i$ identified in Equation 13.1 is used to determine how mapped voxel location x' moves as a function of variations in the transform parameters. Using p_i as a parameter from the set of transform parameters P, the transform Jacobian is defined as

$$\frac{\partial x'}{\partial p_i} = \begin{bmatrix} \frac{\partial x'_1}{\partial p_1} & \frac{\partial x'_1}{\partial p_2} & \cdots & \frac{\partial x'_1}{\partial p_i} \\ \frac{\partial x'_2}{\partial p_1} & \frac{\partial x'_2}{\partial p_2} & \cdots & \frac{\partial x'_2}{\partial p_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x'_j}{\partial p_1} & \frac{\partial x'_j}{\partial p_2} & \cdots & \frac{\partial x'_j}{\partial p_i} \end{bmatrix} \tag{13.3}$$

To update transform T identified in Equation 13.1, rotational changes are made to the current versor through addition with a versor that represents the change in angle determined by optimization of the transform parameters. When described formally, at the end of each optimization cycle, updates to versor V are computed using

$$V' = dV + V \tag{13.4}$$

where V' is the updated versor to be used in the next optimization cycle, dV is the variation determined by optimization, and $+$ stands for the addition of versors. Importantly, derivative dV is computed using

$$dV_i = \left[\frac{\partial S(V_i)}{\partial V_i}\right]\lambda \tag{13.5}$$

where $\partial S(V_i)/\partial V_i$ is a derivative of the similarity measure with respect to versor component V_i, weighted by step length λ. Understandably, the derivative

expresses how much similarity S changes with variations in V. The transform vector t identified in Equation 13.2 is simply updated using

$$t = C - V'^{*}C \tag{13.6}$$

where C is a fixed point used for reference throughout the optimization process, V' is the updated versor identified in Equation 13.4, and * stands for the multiplication of the versor with a vector.

13.4 Distributed Approach to Medical Image Registration

To distribute the registration process, both fixed and moving volumes require division into segments and distribution. Similarity between volume segments can then be estimated concurrently by propagating transform parameters to all processing nodes in the distributed processing network. On receiving transform parameters, each node computes derivatives of the similarity measure for the segments allocated to it. The local derivatives computed are then accumulated and summed into a global derivative. This allows the transform parameters to be updated on the basis of the similarity of complete volumes. Convergence testing is then performed using the newly updated transform parameters. Depending on the success or failure of convergence testing, propagation of updated transform parameters and hence evaluation of new parameters can occur.

13.4.1 iDARBS framework

Regarded as a computational intelligence technique, DARBS (Distributed Algorithmic and Rule-Based Blackboard System) is a distributed blackboard architecture based on a client/server model (Nolle, Wong, and Hopgood, 2001). In DARBS, the server functions as a blackboard and client modules as agents. In general, each agent represents a structure in which rules and algorithms can be embodied. Building on previous work within the field of automated visual inspection for manufacturing, iDARBS (imaging DARBS) (Tait, Schaefer, and Hopgood, 2005), the underlying framework on which this chapter is based, consists of Distributor, Worker, and Manager Agent types.

- The Distributor Agent splits selected volumes into segments, which are placed on the blackboard.
- Worker Agents take segments from the blackboard and perform local processing.
- The Manager Agent supervises Worker Agent activities.

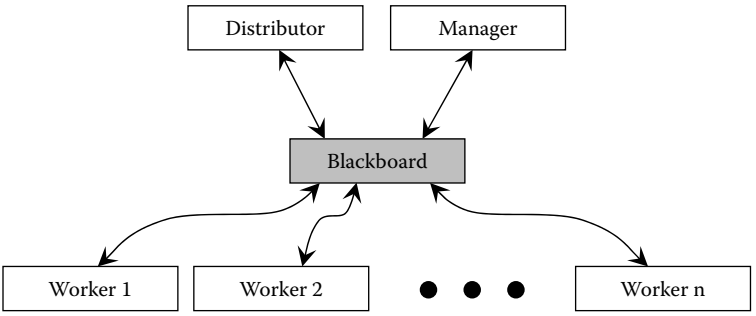


FIGURE 13.3: The worker/manager model and agent types on which the iDARBS framework is based.

As illustrated in Figure 13.3, the implementation of Distributor, Worker, and Manager Agents as client components means equal access to segments stored on the blackboard is possible. As a consequence, the concurrent processing of volume segments can be achieved and allows the speed benefits of parallel processing to be realized. Importantly, the coordination of Distributor, Worker, and Manager Agent activities is achieved by means of reactive behavior and communication via the blackboard. Such behavior removes the need for a dedicated control module and associated overheads.

13.4.2 Partitioning of information on the blackboard

Partitioning of information on the blackboard aids design of the framework by introducing structure to an area of shared memory. Whenever a partition is changed, all agents working with information contained within the partition are restarted. As a consequence, the logical organization of information is used to control the number of partitions from which an agent draws. Due to the exhaustive search required, a drop in performance can be expected with a single-partition implementation. Similar inefficiencies are encountered through management and processing of excess partitions. As shown in Figure 13.4, the iDARBS blackboard is divided into specialized partitions.

- A *Distributor control* partition controls division of selected volumes into segments.
- *Worker n control* partitions are used to manage processing of volume segments.
- Supervision of Worker Agent activities is achieved by means of the *Manager control* partition.
- System variables are maintained in the *Parameters* partition.
- The *Image container* partition holds partitioned volume segments.

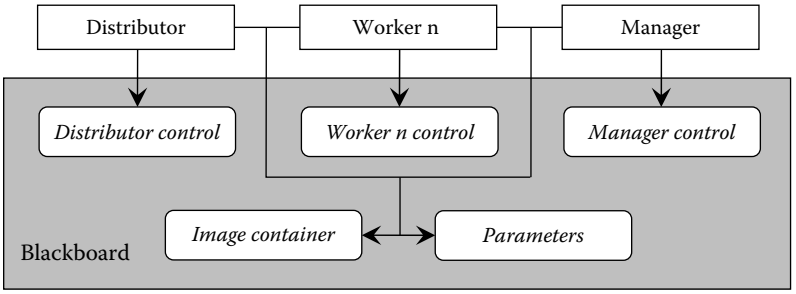


FIGURE 13.4: Blackboard partitions and agent interdependence used to balance communication and processing workloads.

To promote the efficient handling of agent queries and hence increase performance, careful consideration has been given to the format of information stored on the blackboard. As a result, short, concise strings have been adopted. The use of human-readable strings also aids in the debugging of errors and permits easy modifications in the future.

13.4.3 Information strings found on the blackboard

A range of information strings were developed to control the behavior of Distributor, Worker, and Manager Agent types and allow the movement of transform and derivatives parameters between framework components. Example strings are illustrated in Figure 13.5.

- The region of interest string is used to hold the size of a segment without borders. The string is generated by the Distributor Agent and placed in all Worker control partitions.
- A previous transform string is created by the Distributor Agent and used to hold transform parameters from the previous iteration. Updated by the Manager Agent, the string is maintained in the Parameters partition and used in the calculation of updated transform parameters.

```
[ROI 0_0_0_90_109_90]
[Previous 1.34982342_1.42314742 . . . _11.85115785]
[Current 1.98437218_1.56237292 . . . _16.30191341]
[Derivative -203.68349139_68.62940029 . . . _549]
[Final 1.64514585_0.01234546 . . . _15.79934632]
```

FIGURE 13.5: Example information strings for controlling the registration process. Each string consists of an identifying tag followed by an underscore-delimited list of numbers.

- Generated and updated by the Manager Agent, the current transform string is used for propagation of transform parameters to all Worker Agents. Crucially, the string coordinates the start of Worker Agent activities.

- The accumulation of local derivatives is achieved with derivative strings. The creation of a derivative string marks the temporary suspension of a Worker Agent's activities.

- The final transform string contains optimum transform parameters. Once created and propagated to all Worker Agents, the string causes the permanent suspension of Worker Agent activities.

13.4.4 Agent behavior

Framework initialization and volume selection are the responsibility of the Distributor Agent. The Distributor Agent also divides selected volumes vertically and horizontally into segments. Before being placed on the blackboard, the edges of each segment are assigned a border. The border is designed to counter inconsistencies caused by a lack of voxels at a segment's boundary. Upon activation, the Worker Agents retrieve individual segments and compute local derivatives of the similarity measure with respect to each transform parameter. The accumulation and summation of local derivatives is performed by the Manager Agent. The Manager Agent also updates current transform parameters and propagates them to the Worker Agents. Understandably, the calculation of local derivatives and the updating of transform parameters are repeated until predefined thresholds are exceeded. On convergence of optimization, resampled volume segments are retrieved from the blackboard and a registered volume is constructed.

During the alignment process, the mapping of intensities between fixed and moving segments is achieved using a versor rigid transform component. The transform represents a rotation and translation, where the rotation is specified by a (unit) Quaternion and the translation is implemented as a vector. B-spline interpolation is employed by Worker Agents to estimate moving-volume intensities at nongrid coordinates. Understandably, the intensities produced by the interpolator map to non-discrete coordinates in 3-D space. Once similarity between volumes has been calculated, versor rigid 3-D transform optimization is employed by the Manager Agent to search for new transform parameters. An adaptation of gradient-descent optimization, the optimizer combines the current rotation with the computed gradient to produce a new rotation versor. Translation parameters are simply updated in vector space.

13.4.4.1 Initialization

Because a gradient-descent optimization scheme is employed, the Distributor Agent places predefined parameters associated with the optimization

process on the blackboard. The parameters include an initial step length, a minimum step length, and the maximum number of iterations to be performed. Logically, the initial step length is used to initialize the optimization process. Convergence of optimization and hence the selection of final transform parameters is controlled by the minimum step length. The maximum number of iterations prevents the optimization process from entering into an endless loop.

To start the alignment process, the Distributor Agent calculates an initial center of rotation and a translation. Employing both fixed and moving volumes, initial parameters are estimated using centers of mass computed from intensity levels. Once estimated, the fixed-volume center is set as the rotational center of the initial transform. The translation component, in contrast, is set as the vector between the fixed- and moving-volume centers. To simplify calculation, no rotation is specified in the initial transform. Crucially, the use of centers of mass over geometrical volume centers results in a robust estimate of the initial transform because the subject of interest is not always in the geometric center of a volume.

13.4.4.2 Calculating local derivatives of the similarity measure

For similarity between volume segments to be calculated, the Worker n Agent employs region of interest and current transform parameters retrieved from the Worker n control partition. First, the region of interest is used to identify the fixed segment without borders. Then, for all voxel coordinates within the fixed segment, corresponding moving-segment coordinates are computed using the current transform parameters. If the transformation of fixed-segment voxel coordinates results in a corresponding location that falls inside the moving segment, the number of valid voxels is incremented and a contribution to the local derivative is made. Otherwise the voxel is considered invalid and the next fixed-segment voxel coordinates are processed. By evaluating all voxels, local derivatives of the similarity measure with respect to each transform parameter are generated. Crucially, moving-segment voxel coordinates are calculated using scaled fixed-segment voxel coordinates. First, fixed-segment voxel coordinates are mapped into the physical space of the fixed segment and then into the physical space of the moving segment. In a final step, the physical coordinates of the fixed-segment voxel in moving-segment space are mapped to moving-segment voxel coordinates. By employing a scale factor, the mapping of voxel coordinates through physical space allows segments of different size to be registered.

The contribution to a local derivative represents a summation of intensities, from a moving-segment gradient volume, around the mapped voxel coordinates. Using a recursive Gaussian filter, a gradient volume is created from the moving segment. The gradient volume represents a vector field in which every vector points in the direction of its nearest edge, an edge being a rapid increase or decrease in neighboring intensities. Each vector has a

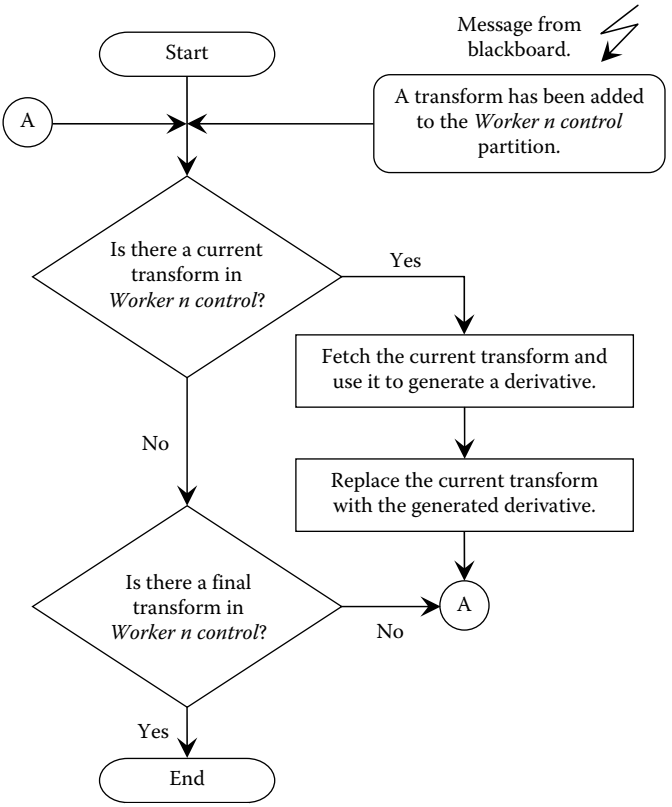


FIGURE 13.6: The Worker n Agent flow diagram. The diagram illustrates the iterative retrieval of transform parameters and the generation of local derivatives.

magnitude proportional to the second derivative of the intensity in the direction of the vector. Created once after retrieval of the moving segment, the moving-segment gradient volume is used for all iterations of the optimization process. Figure 13.6 illustrates the iterative nature of the Worker n Agent engaged in similarity calculation.

13.4.4.3 Advancing transform parameters

To update the current transform, the gradient-descent optimization scheme employed advances the current transform parameters in the direction of the global derivative. If the direction of the global derivative abruptly changes, it is assumed that an optimum has been encountered and the step length is reduced by a half. After repeated iterations, the step length is reduced further and the selection of transform parameters is restricted to a small area of

search space. Once the step length becomes smaller than the predefined minimum, the optimization process is considered as having converged. This allows the precision of the final transform parameters to be specified. If optimization of the transform parameters fails to reach the desired precision, the maximum number of iterations is used to halt the optimization process. Understandably, large numbers of iterations and long computational times result when the initial step length chosen is small. Large step lengths, in contrast, can result in the optimum transform parameters being missed. Figure 13.7

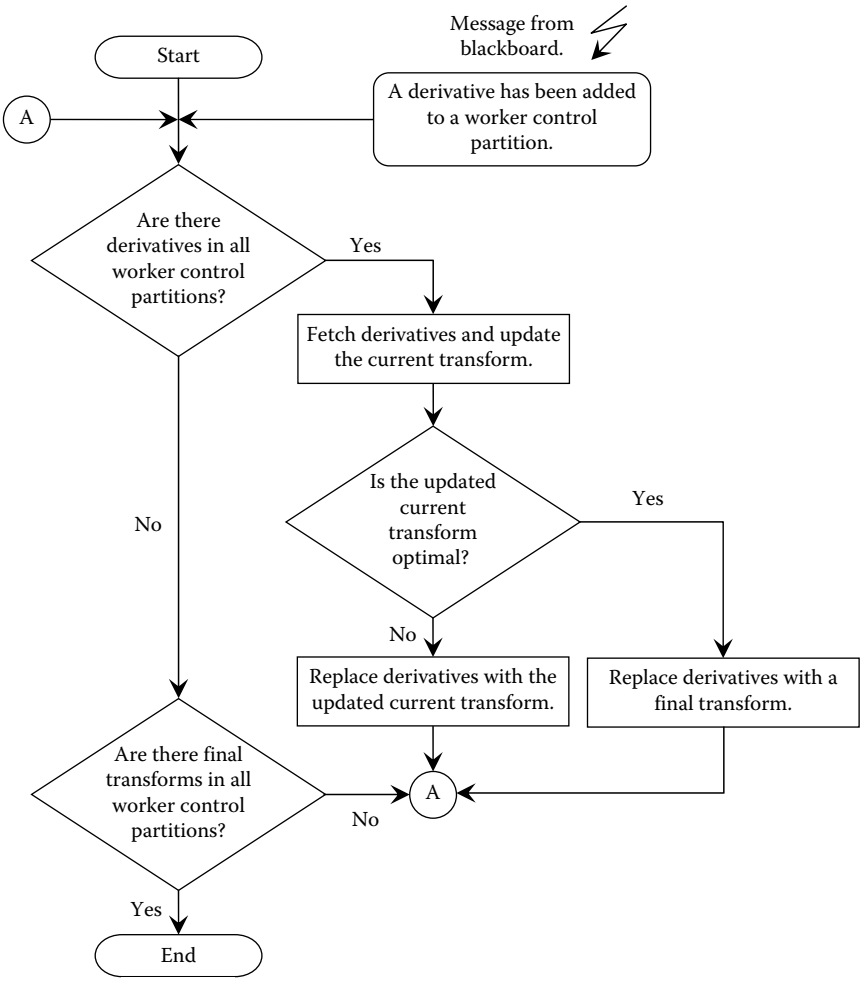


FIGURE 13.7: The Manager Agent flow diagram. The diagram illustrates the iterative updating of transform parameters and their propagation to Worker Agents.

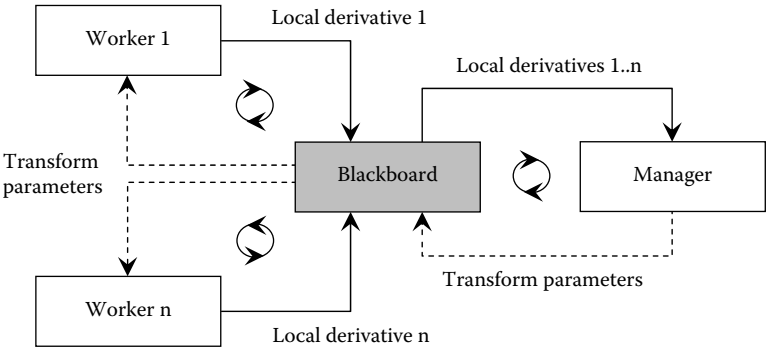


FIGURE 13.8: The movement of information between framework components. Derivatives are accumulated by the Manager Agent, and transform parameters are propagated to all Worker Agents.

illustrates the iterative updating of transform parameters by the Manager Agent.

The replacing of derivatives with current transform parameters ensures that the local derivatives computed by the Worker Agents are based on the same transform. Also, to ensure that the transform parameters updated by the Manager Agent are based on local derivatives from the same iteration, derivatives are replaced with current transform parameters. Unsurprisingly, if the Manager and Worker Agents operate with transform and derivative parameters from different iterations, a corrupt path through transform search space occurs. Similarly, if the Manager and Worker Agents become unsynchronized in the number of iterations performed, a state of deadlock occurs as both wait for either transform or derivative parameters. The accumulation of local derivatives by the Manager Agent and the propagation of updated transform parameters to Worker Agents are illustrated in Figure 13.8.

13.4.5 Volume visualization

In order for volumes to be visually assessed, a 3-D viewer was constructed using the Qt library (Blanchette and Summerfield, 2004). The viewer is created during initialization of an agent and can be shown or hidden at any time. Volume data can be loaded into the viewer by means of a standard open file dialog box, or it can be functionality embedded in agent rule files. The viewer consists of perspective and orthographic projections in the x, y, and z planes. Conveniently, volume data can be displayed as either a complete object or as slices. When viewed as a complete object, only the surface of a structure is visible. When viewed as slices, the internal detail of a structure is presented. By extracting individual slices, assessment of alignment accuracy can be achieved using checkerboard, squared difference, and weighted

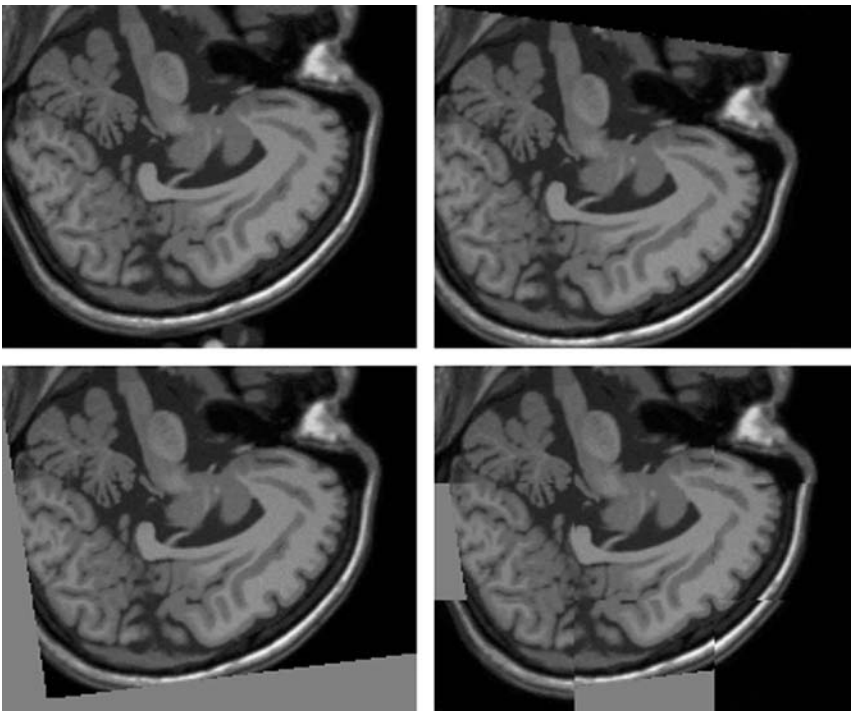


FIGURE 13.9: From top left to bottom right: fixed, moving, registered, and checkerboard images. The checkerboard represents a combination of fixed and registered slices.

overlay images. Volume overlay, whereby each data set is assigned a different color channel, is also provided. Figure 13.9 shows extracted slices rendered by the viewer.

13.5 Experimental Testing

Speed tests are presented and demonstrate the performance increase of volume registration in nondistributed and distributed processing environments. To determine the accuracy of alignment between volumes, mean square error and normalized correlation similarity metrics were selected for testing purposes. Timing of an experiment started when the Manager Agent propagated current transform parameters to all worker control partitions. Timing stopped when the Manager Agent placed final transform parameters in the

Parameters partition. The time required to register volumes was calculated by subtracting the start time from the stop time.

Magnetic resonance imaging (MRI) is a diagnostic imaging modality that has the ability to derive contrast from a number of physical parameters (Stark and Bradley, 1999). Conveniently, MRI volumes specifically designed for the testing of new registration algorithms can be freely obtained and hence were deemed suitable for testing purposes. Spin-lattice relaxation (T1) is an MRI imaging modality used to highlight gray matter contained within the body. For testing of the distributed algorithm, two T1 volumes were obtained from the McConnell Brain Imaging Centre (Cocosco et al., 1997). The McConnell Brain Imaging Centre Web site is an open-access simulated brain database that contains realistic MRI volumes produced by an MRI simulator. Selection of three acquisition parameters allows realistic MRI volumes of the brain to be acquired. The parameters include modality, slice thickness, and noise content. Both volumes obtained are based on an anatomically normal human brain and have a slice thickness of 1 mm and noise content of 3%. Both volumes are also $181 \times 217 \times 181$ cubic voxels in size. The moving volume represents a data set that has been rotated 10 degrees clockwise about the origin and translated to 15 voxels in the x axis. The origin is the voxel located at the bottom left corner of the volume.

Using a loosely coupled architecture, testing was performed three times; results were then combined and an average calculated. In all cases, selected volumes were divided by the Distributor Agent into 1 to 14 segments, and a 20-voxel wide border was allocated. In order to maintain consistency, the experiments represented an ideal case (i.e., one processor for the blackboard, one processor for the Distributor Agent, one processor for the Manager Agent, and one processor for each Worker Agent). Connected via an Ethernet 100 Mbps switch, the network contained approximately 30 computers. Running the Linux Ubuntu operating system, all computers contained AMD Athlon 2 GHz processors with 1 GB of RAM. The memory was of sufficient size to allow all computations to occur in RAM. By employing artificially created volumes, ground-truth parameters were known and could be compared with final transform parameters, thus providing a numerical assessment of alignment accuracy. A sequential algorithm, hosted on a single-processor architecture, was used as a performance benchmark for comparison. As a consequence, the results obtained allow the speedup and efficiencies, achieved through distribution, to be highlighted.

13.5.1 Mean square error

Mean square error is a similarity measure computed over all voxels in both volumes. Importantly, calculation of mean square error is suited to volumes of the same modality, and as a consequence, intensities at corresponding locations need to be similar. The metric is attractive because it is simple to compute and produces a relatively smooth search space. When alignment

between volumes is poor, large values are produced by the metric. Small values, in contrast, occur near optimum alignment. The distributed mean square error measure of similarity is defined as

$$S(F, M, T) = \frac{\sum_{i=1}^{R} \left[\sum_{j=1}^{Q_i} (F(x_{ij}) - M(T(x_{ij})))^2 \right]}{\sum_{i=1}^{R} Q_i} \tag{13.7}$$

where F and M are fixed- and moving-segment intensity functions, respectively, T is the spatial transform, and x_{ij} is the jth voxel of segment i from the fixed volume. R is the number of segments into which a volume is divided, and Q_i is the number of valid voxels mapped between segments identified by i. The derivative of the distributed similarity metric with respect to transform parameter p is computed as

$$\begin{aligned} \frac{\partial S}{\partial p} &= \frac{2 \sum_{i=1}^{R} \left[\sum_{j=1}^{Q_i} [F(x_{ij}) - M(T(x_{ij}, p))] \frac{\partial M(T(x_{ij}, p))}{\partial p} \right]}{\sum_{i=1}^{R} Q_i} \\ &= \frac{2 \sum_{i=1}^{R} \left[\sum_{j=1}^{Q_i} [F(x_{ij}) - M(T(x_{ij}, p))] \frac{\partial M(y)}{\partial y} \Big|_{y=T(x_{ij}, p)} \frac{\partial T(x_{ij}, p)}{\partial p} \right]}{\sum_{i=1}^{R} Q_i} \end{aligned} \tag{13.8}$$

where $M(T(x_{ij}, p))$ represents a discrete input that has been interpolated using a B-spline interpolation scheme, and where $\partial T(x_{ij}, p)/\partial p$ is the transform Jacobian used to estimate variations in the mapped voxel coordinates with respect to transform parameter p. Importantly, when a voxel location that maps outside of the moving segment is encountered, the contribution to the local derivative is discarded.

Figure 13.10 shows the speed of sequential and distributed volume registration, plotted as time against number of Worker Agents, with mean square error as a similarity metric. As can be seen, the average execution time of registration was reduced from 27 minutes to approximately 5 minutes when 12 Worker Agents were employed. For each test, the distributed metric converged after 19 iterations with transform parameters, which matched those computed by the sequential algorithm.

13.5.2 Normalized correlation

Computed using all voxels, normalized correlation calculates the cross-correlation of the volumes to be registered. Once calculated, the cross-correlation is normalized by the square root of the autocorrelation of each

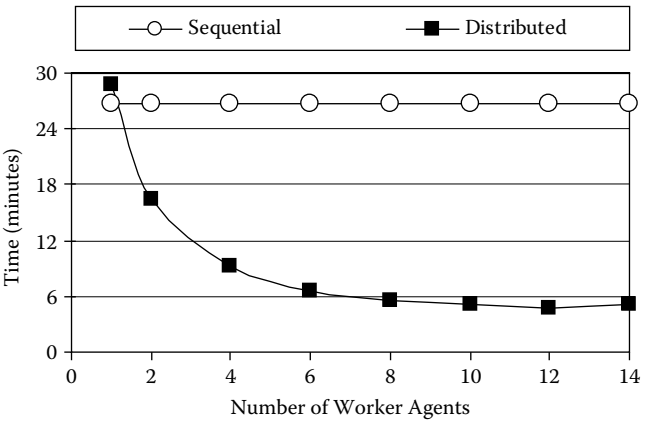


FIGURE 13.10: The sequential and distributed processing speed of volume registration using mean square error as a similarity metric. Increasing numbers of Worker Agents are shown.

volume. Suited to volumes of the same modality, appealing properties of the metric include the production of a search space containing sharp peaks and well-defined throughs. The accurate alignment of volumes results in values near 1 being produced by the metric. Misalignment, in contrast, produces values of less than 1. The distributed normalized correlation measure of similarity is defined as

$$S(F,M,T) = -1 \frac{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i}(F(x_{ij}) - M(T(x_{ij})))\right]}{\sqrt{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i}F(x_{ij})^2 \sum_{j=1}^{Q_i}M(T(x_{ij}))^2\right]}} \tag{13.9}$$

where F, M, T, R, and Q are the intensity functions, spatial transform, number of segments, and number of valid voxels previously defined. The derivative of the distributed normalized correlation similarity metric with respect to transform parameter p is computed as

$$\frac{\partial S}{\partial p} = -\frac{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i}\left(F(x_{ij})\frac{\partial M(T(x_{ij},p))}{\partial p}\right) - b\sum_{j=1}^{Q_i}\left(M(T(x_{ij},p))\frac{\partial M(T(x_{ij},p))}{\partial p}\right)\right]}{a}$$

$$= -\frac{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i}\left(F(x_{ij})\frac{\partial M(T(x_{ij},p))}{\partial p}\right) - b\sum_{j=1}^{Q_i}\left(M(T(x_{ij},p))\frac{\partial M(y)}{\partial y}\Big|_{y=T(x_{ij},p)}\frac{\partial T(x_{ij},p)}{\partial p}\right)\right]}{a} \tag{13.10}$$

with

$$a = \sqrt{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i} F^2(x_{ij}) \sum_{j=1}^{Q_i} M^2(T(x_{ij},p))\right]}$$

$$b = \frac{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i} F(x_{ij})\, M(T(x_{ij},p))\right]}{\sum_{i=1}^{R}\left[\sum_{j=1}^{Q_i} M^2(T(x_{ij},p))\right]}$$

where $M(T(x_{ij},p))$ and $\partial T(x_{ij},p)/\partial p$ are the discrete inputs interpolated using B-spline interpolation and transform Jacobian previously defined. Again, mapped voxel locations that lie outside of a moving segment do not contribute to either the similarity measure or its derivatives.

Figure 13.11 shows results plotted as time against number of Worker Agents, obtained while performing volume registration with normalized correlation as a similarity metric. As can be seen, average processing time of registration was reduced from 68 minutes to approximately 10 minutes when 10 Worker Agents were employed. The distributed algorithm was observed to converge after 54 iterations with transform parameters, which matched those computed by the sequential algorithm.

13.5.3 Speedup and efficiencies achieved

In general, the overall processing time of a registration algorithm will depend on initial alignment and the size of the volumes being registered. As a

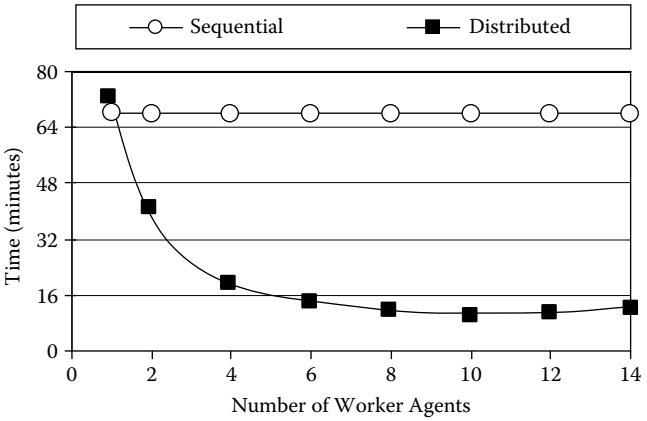


FIGURE 13.11: The sequential and distributed processing speed of volume registration using normalized correlation as a similarity metric. Increasing numbers of Worker Agents are shown.

consequence, the time required to register volumes can be seen as a function of the number of processing nodes employed. Crucially, the increase in processing speed due to parallelization of an algorithm with multiple processing nodes is governed by Amdahl's law (Amdahl, 1967; Gustafsom, 1988). The law states that the distribution of an algorithm can be analyzed in terms of sequential and parallel portions. The law also suggests that the maximum speedup achievable through distribution of an algorithm with N processing nodes is an N times speedup. The maximum speedup, however, is always constrained by the sequential portion of the algorithm and can be calculated as

$$Speedup = \frac{t_s + t_p}{t_s + t_p/N} = \frac{1}{t_s + t_p/N} \tag{13.11}$$

where t_s is the sequential or nonparallelizable portion of the algorithm, and t_p is the parallelizable portion. When considering the distributed similarity metrics described, t_p is the time spent by the Worker Agents computing local derivatives; t_s, in contrast, represents the accumulation of local derivatives, followed by the updating of transform parameters, and their propagation to Worker Agents by the Manager Agent. Crucially, the amount of speedup achieved by each Worker Agent represents efficiency and is calculated using

$$Efficiency = \frac{Speedup}{N} \tag{13.12}$$

where *Speedup* is the speed improvement calculated in Equation 13.11, and N is the number of processing nodes previously defined. In this chapter, both speedup and efficiency are calculated using averaged processing times.

The speedup and efficiency, plotted against number of Worker Agents, achieved during distributed volume registration with mean square error as a similarity metric are shown in Figure 13.12. Ideal speedup and efficiency rates are also shown. As can be seen, a peak speedup factor of 5.5 and an efficiency of 44% were achieved by 12 Worker Agents. The results obtained show that as the number of Worker Agents grows, speedup steadily increases to a peak, and a deterioration in performance then occurs. The initial speed increase is due to the small number of Worker Agents involved in similarity calculation. The deterioration in performance is, however, caused by increasing Worker Agent overheads.

In Figure 13.13, the speedup and efficiency, plotted against number of Worker Agents, achieved during distributed volume registration with normalized correlation as a similarity metric are illustrated. Again ideal speedup and efficiency rates are shown. As can be seen, a peak speedup factor of 7 and an efficiency of 67% were achieved by 10 Worker Agents. The results obtained show that efficiency steadily deteriorates as the number of Worker Agents increases. This is an indication that the Worker Agents are only partly utilized during the alignment process. The cause of underutilization is the wait by Worker Agents until local derivatives have been accumulated and updated transform parameters have been estimated by the Manager Agent.

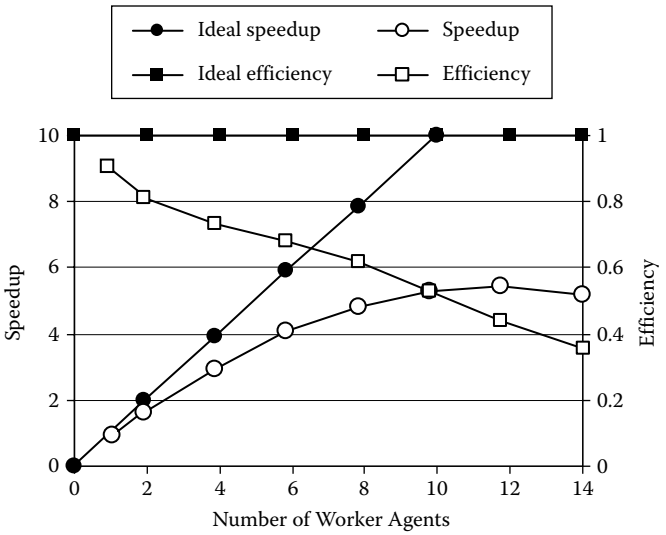


FIGURE 13.12: The speedup and efficiency of volume registration using mean square error as a similarity metric. Increasing numbers of Worker Agents are shown.

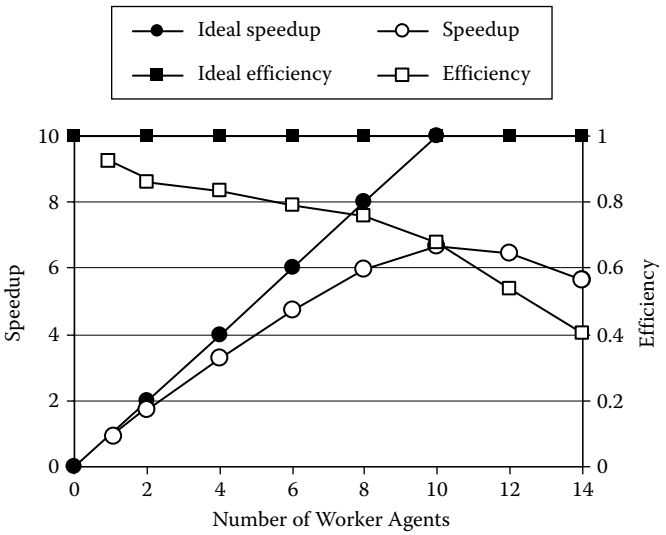


FIGURE 13.13: The speedup and efficiency of volume registration using normalized correlation as a similarity metric. Increasing numbers of Worker Agents are shown.

13.6 Discussion

Because the underlying blackboard architecture on which the iDARBS framework is based does not have a control module, true opportunism and independence can be achieved. The use of reactive behavior also reduces the time required to align distributed volume segments. For example, the autonomous behavior employed by Distributor, Worker, and Manager Agents requires only Agent-to-blackboard communications. The simplicity of Agent implementation increases efficiency and makes behavioral modification a straightforward task. During registration, transform and derivative parameters are placed in worker control partitions. As a consequence, each Worker Agent focuses its attention toward an associated blackboard partition, thus simplifying the search for required information. Crucially, the storage of transforms and derivatives on the blackboard means registration can be stopped and parameters adjusted if required. In contrast with other applications (Ino, Ooyama, and Hagihara, 2005; Salomon et al., 2005), distributed registration on the iDARBS framework means alignment can be restarted in the event of failure. To achieve this, Worker and Manager Agents are simply restarted and supplied with previous transform parameters found in the *Parameters* partition.

The results obtained from iDARBS clearly show that the speed of distributed registration is better than sequential versions of the same algorithms. By hosting each Worker Agent on an individual processor, context switching between processes is avoided. Naturally, the context switching between multiple processes on a single processor introduces additional overhead, resulting in longer execution times. The processing speeds accomplished during testing are, however, less than ideal, mainly due to the overhead caused by the serial processing of Agent requests by the blackboard (Tait, Schaefer, and Hopgood, 2005). These overheads are demonstrated in each experiment by means of the single Worker Agent implementation. An important consideration highlighted during testing is that the limited coordination of Worker Agent activities represents a performance bottleneck in communications to and from the blackboard. Once triggered, Worker Agents try to obtain segments from the blackboard simultaneously, causing an initial overloading of communications. A second overload occurs when the Worker Agents complete their assigned tasks and try to return processed segments to the blackboard.

13.7 Future Research Directions

In general, the use of intelligent agents represents a significant step toward high-performance, intensity-based registration using non-specialized architectures. Despite the success of distributed registration applications, it is clear

that research can be extended in a number of ways. Areas of research currently under investigation include robustness of the transform parameter optimization process and the distribution of deformable registration algorithms.

Conveniently, the transforms used for volume registration are not restricted to a single optimization scheme. To increase robustness of the alignment process, the evaluation of transform parameters can be performed concurrently using multiple optimization techniques. Once assessed, a voting strategy or averaging is employed to select optimum transform parameters for propagation. This has been achieved, for example, through simultaneous global and multidirectional local search methods adapted for image registration purposes (Wachoviack and Peters, 2004). In general, the use of multiple optimization techniques is aimed at shortening the path through transform parameter search space, thus resulting in faster processing speeds.

Deformable registration is based on the assumption that an evenly spaced mesh of control points can be placed within fixed and moving volumes. A transform is then used to estimate the displacement necessary to map between corresponding control point locations. In order for registration to be performed, each control point plus underlying intensities are transformed and compared with their counterpart until an acceptable level of similarity is achieved. Such algorithms can be hosted within distributed architectures by assigning subsets of control points to each volume segments. Conveniently, the use of free-form deformations (Loeckx et al., 2004) permits the independent movement of control points and removes the need for interprocessor communication.

13.8 Conclusions

Registration is an important step in medical analysis tasks in which information is extracted from a combination of sources. Traditionally, volume registration algorithms have been implemented using single-processor architectures, which are limited by memory and speed constraints. Such limitations have a negative impact in clinical applications where real-time processing allows physicians to monitor the progress of treatment while a patient is present. The surveyed literature suggests that the most successful registration algorithms employ intensity-based correlation as a measure of similarity. Although a variety of similarity metrics have been developed, in practice they represent a considerable computational burden during the alignment process, mainly because of the high cost associated with multiple evaluations of a complex transform and the interpolation of nondiscrete intensity coordinates. The literature also makes clear that concurrent similarity calculation can be achieved and provides better performance than nonparallel approaches (Christensen, 1998; Ino, Ooyama, and Hagihara, 2005; Salomon et al., 2005; Warfield, Jolesz, and Kikinis, 1998). The large speedups reported are, however,

difficult to obtain, and only specialized hardware is capable of maintaining such efficiencies when scaled. As a consequence, the applications developed to address the problem of slow volume registration speeds are restricted to high-cost specialized architectures found predominantly in the research environment.

While various approaches to distribution have been employed, it is fine-grained parallelism that achieves the best results (Hastings et al., 2003; Ourselin, Stefanescu, and Pennec, 2002; Rohlfing and Maurer, 2003). These methods are based on the low-level decomposition of an algorithm within a tightly coupled architecture. Understandably, such algorithms are difficult to implement and minimize computational expense by eliminating the exchange of data between processors.

In this chapter, a novel approach to high-performance intensity-based volume registration was presented. The approach has been shown to significantly improve processing speed, clearly outperforming sequential versions of the same algorithm. By employing an intensity-based algorithm, we avoid the complicated segmentation of features fundamental to landmark-based registration methods. As a consequence, the algorithm does not require user intervention. Unlike other distributed registration applications (Christensen, 1998; Warfield, Jolesz, and Kikinis, 1998), the decoupling of algorithm components allows different similarity calculation strategies to be employed with only minor modifications to the existing framework. This was demonstrated through the distribution of mean square error and normalized correlation similarity metrics. At an implementation level, the transferral of transform and derivative parameters as short, concise strings minimizes communication overheads. The method is advantageous because the overhead generated remains approximately the same for all iterations of the optimization process. Crucially, as no specialized hardware is required, the framework can reside on any network connected by the TCP/IP communications protocol, thus making high-performance registration accessible within a clinical environment.

References

Amdahl, G. M. 1967. Validity of the single processor approach to achieving large-scale computing capabilities. *Proceedings of the AFIPS Conference*, Atlantic City, 30: 483–485.

Blanchette, J., and Summerfield, M. 2004. *C++ GUI Programming with Qt 3*. Upper Saddle River, NJ: Prentice Hall.

Chalermwat, P., El-Ghazawi, T., and LeMoigne, J. 2001. 2-phase GA-based image registration on parallel clusters. *Future Generation Computer Systems*, 17: 467–476.

Christensen, G. E. 1998. MIMD vs SIMD parallel processing: a case study in 3-D medical image registration. *Parallel Computing*, 24(9): 1369–1383.

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., and Evans, A. C. 1997. Brainweb: online interface to a 3D MRI simulated brain database. *NeuroImage*, 5(4): 425.

Conway, J. H., and Smith, D. A. 2003. *On Quaternions and Octonions: Their Geometry, Arithmetic, and Symmetry*. Natick, MA: A.K. Peters.

Cordon, O., Damas, S., and Santamaria, J. 2006. Feature-based image registration by means of the CHC evolutionary algorithm. *Image and Vision Computing*, 24: 525–533.

Cordon, O., Damas, S., and Santamaria, J. 2006. A fast and accurate approach for 3D image registration using the scatter search evolutionary algorithm. *Pattern Recognition Letters*, 27: 1191–1200.

Gustafsom, J. L. 1988. The scaled-sized model: a revision of Amdahl's law. *Proceedings of the 3rd International Conference on Supercomputing*, pp. 130–133.

Hastings, S., Kurc, T., Langella, S., Catalyurek, U., Pan, T., and Saltz, J. 2003. Image processing for the grid: a toolkit for building grid-enabled image processing applications. *Proceedings of the 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid*, pp. 36–43.

Ino, F., Ooyama, K., and Hagihara, K. 2005. A data distributed parallel algorithm for non-rigid image registration. *Parallel Computing*, 31: 19–43.

Jeongtae, K., and Fessler, J. A. 2004. Intensity-based image registration using robust correlation coefficients. *IEEE Transactions on Medical Imaging*, 23: 1430–1444.

Loeckx, D., Maes, F., Vandermeulen, D., and Suetens, P. 2004. Non-rigid image registration using free-form deformations with a local rigidity constraint. *Lecture Notes in Computer Science*, 3216: 639–646.

Nolle, L., Wong, K. C. P., and Hopgood, A. A. 2001. DARBS: a distributed blackboard system. *Research and Development in Intelligent Systems*, 18: 161–170.

Ourselin, S., Stefanescu, R., and Pennec, X. 2002. Robust registration of multimodal images: towards real-time clinical applications. *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 140–147.

Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., and Hawkes, D. J. 1998. A comparison of similarity measures for use in 2D–3D medical image registration. *IEEE Transactions on Medical Imaging*, 17: 586–595.

Pluim, J. P., Maintz, J. B. A., and Viergever, M. A. 2003. Mutual information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22: 986–1004.

Rohlfing, T., and Maurer, C. R. 2003. Non-rigid image registration in shared-memory multi-processor environments with application to brains, breasts, and bees. *IEEE Transactions on Information Technology in Biomedicine*, 7: 16–25.

Salomon, M., Heitz, F., Perrin, G. R., and Armspach, J. P. 2005. A massively parallel approach to deformable matching of 3-D medical images via stochastic differential equations. *Parallel Computing*, 31: 45–71.

Seinstra, F. J., Koelma, D., and Geusebroek, J. M. 2002. A software architecture for user transparent parallel image processing. *Parallel Computing*, 28: 967–993.

Stark, D. D., and Bradley, W. G. 1999. *Magnetic Resonance Imaging*, 3rd ed. St. Louis: Mosby.

Tait, R. J., Schaefer, G., and Hopgood, A. A. 2005. Registration and fusion of multi-sensor data using multiple agents. *Proceedings of SPIE: Opto-Ireland Imaging and Vision*, Dublin, Ireland, pp. 104–112.

Temkin, B., Vaidyanath, S., and Acosta, E. 2005. A high accuracy, landmark-based, sub-pixel level image registration method. *International Congress Series*, 1281: 254–259.

Wachoviack, M. P., and Peters, T. M. 2004. Parallel optimization approaches for medial image registration. *Lecture Notes in Computer Science*, 3216: 781–788.

Wang, X., and Feng, D. D. 2005. Hybrid registration for two-dimensional gel protein images. *Proceedings of the 3rd Asia-Pacific Bioinformatics Conference*, Singapore, pp. 201–210.

Warfield, S. K., Jolesz, F., and Kikinis, R. 1998. A high-performance approach to the registration of medical imaging data. *Parallel Computing*, 24(9): 1345–1368.

Yoo, T. S. 2004. *Insight into Images: Principles and Practices for Segmentation, Registration, and Image Analysis*. Natick, MA: A.K. Peters.

Zhang, Z., and Blum, R. S. 2001. A hybrid image registration technique for a digital camera image fusion application. *Information Fusion*, 2: 135–149.

Zitova, B., and Flusser, J. 2003. Image registration methods: a survey. *Image and Vision Computing*, 21: 977–1000.

Chapter 14

Monte Carlo–Based Image Reconstruction in Emission Tomography

Steven Staelens and Ignace Lemahieu

Contents

Nuclear medicine uses radioactively labeled pharmaceuticals to diagnose and assess disease in the human body by single photon emission computed tomography (SPECT) and positron emission tomography (PET). The image reconstruction consists in retrieving the 3-D spatial distribution of the radiopharmaceutical from the projection data acquired under different angles. Monte Carlo methods are statistical simulations using a sequence of random numbers to model different physical processes. The Monte Carlo method has been applied in nuclear medicine for detector design, analyzing quantification issues, correction methods for image degradations, and detection tasks. With the market of affordable miniclusters rapidly gaining importance, there has been a renewed interest in using Monte Carlo simulations for image reconstruction. This method allows for a patient-specific simulation of the data acquisition, resulting in a more accurate quantification when incorporated in an iterative reconstruction algorithm. Quantitative image reconstruction ultimately enhances diagnosis, dosimetry, and treatment planning.

14.1 Introduction

Nuclear medicine imaging involves the use of radioactively labeled pharmaceuticals to diagnose and assess disease in the human body. The radiopharmaceutical is intravenously injected, inhaled, or ingested: the choice of pharmaceutical and route of administration depends on the disease under investigation. Position-sensitive detectors are used to detect the radiation emitted from the pharmaceutical, and so, in principle, both the temporal and spatial distribution of the pharmaceutical within the body can be determined. Historically, nuclear medicine imaging was the first modality for functional imaging (i.e., giving information regarding the way organs work). Today, although functional imaging is also achievable using magnetic resonance imaging, ultrasound, or even x-ray computed tomography, nuclear medicine still plays a major role given its exquisite sensitivity, which makes it possible to target molecular processes. Modern nuclear medicine imaging primarily consists of two main branches: single photon imaging and positron annihilation photon imaging.

14.1.1 Single photon emission tomography

Single photon imaging requires at minimum one detector fixed in one position to obtain a two-dimensional (2-D) projection of a three-dimensional (3-D) radiopharmaceutical distribution. A typical imaging system suitable for this task is the gamma camera. A typical gamma camera consists of a collimator that limits the angle of incidence of the detected photons on the detector to a specific direction, for instance around 90° for a parallel hole collimator. The deposited energy is converted by a scintillating crystal to visible light, which travels through the crystal and the light guide toward a set of photomultiplier

tubes containing a photocathode, which converts the optical photons into electrons that are transported over dynodes toward an Anger logic readout for position and energy signal calculation.

Single projection imaging can be improved by simply rotating the gamma camera around the patient, thus obtaining a series of 2-D projections. These can be used to retrieve depth information (e.g., the 3-D distribution of the radiopharmaceutical), which is known as single photon emission computed tomography (SPECT). The radionuclide most widely used in SPECT is Technetium-99m (^{99m}Tc), which has a half life of 6.03 hours and which emits photons of 140.5 keV. SPECT application fields, in order of importance, are cardiac imaging, bone scanning, neurology, and renal and liver function imaging.

14.1.2 Positron emission tomography

Pharmaceuticals can also be labeled with neutron-deficient isotopes, which are positron emitters. A number of elements that are fundamentally used by the human body can be positron emitters, allowing more scope for radiopharmaceutical design than is possible with single photon emitters. However, production costs are higher because a cyclotron is needed to generate these unstable radionuclides. The emitted positron annihilates with an electron, giving rise to two virtually antiparallel 511 keV photons corresponding to the conversion of the rest mass of the two particles into energy. The primary advantage of detecting both annihilation photons in time coincidence is the electronic collimation. Detection of two photons is sufficient to determine the line on which the annihilation took place, so no physical collimation is required, allowing for higher sensitivity than in single photon imaging. Tomographs appropriate for positron emission tomography (PET) typically consist of a series of ring detectors, having a large number of separate blocks coupled to photomultiplier tubes. Fluorine 18 (^{18}F) is mostly used to label the fluorodeoxyglucose (FDG) compound, which is a glucose analog and thus suitable for the visualization of the glucose metabolism. PET application fields in order of importance are oncology, neurology, and cardiology.

14.2 Background

A key ingredient for SPECT and PET is image reconstruction, which consists of retrieving the 3-D spatial distribution of the radiopharmaceutical from the projection data acquired under different angles. Tomographic reconstruction can be performed in several ways:

- Analytical reconstruction
  - Filtered back-projection: analytical inversion of the x-ray transform based on the central slice theorem

- Deterministic iterative reconstruction
  - ART: algebraic reconstruction technique
  - SIRT: simultaneous iterative reconstruction technique
- Statistical iterative reconstruction
  - ML-EM: maximum likelihood–expectation maximization
  - MAP: maximum a posteriori

A deterministic model is obtained when the system characteristics are incorporated. Accordingly, the reconstruction problem can be solved analytically. If we assume the acquisition process to be statistical, then a statistical reconstruction technique is mandatory. Often used statistics are the Poisson model for the radioactive decay process and the Gaussian model for the intrinsic features (detector response) of the acquisition. The main advantage of a statistical model is that a smaller variance results in the reconstructed image if the appropriate models are used. Iterative reconstruction techniques can incorporate image-degrading effects in the reconstruction algorithm in order to improve the reconstructed image quality. MLEM [60,68], which is an algorithm that incorporates the Poisson statistics of the radioactive decay in the reconstruction, is hereafter discussed as a basic statistical iterative reconstruction technique.

Let us consider a parameterized family of probability density functions:

$$F \rightarrow f(F|P) \tag{14.1}$$

with F the reconstructed image and P the measurement. When $f(F|P)$ is viewed as a function of F with P fixed, it is a probability density function, and when viewed as a function of P with F fixed, it is a likelihood function. When we write down the Bayes rule, we find

$$Pr(F|P) = \frac{Pr(P|F)Pr(F)}{Pr(P)} \tag{14.2}$$

wherein $Pr(P|F)$ is the likelihood (true forward probability) that the true tracer distribution F gives rise to a particular measurement P. This is the well-studied forward problem of the acquisition setup, commonly referred to as the detector response function. $Pr(F|P)$, however, is the probability of an image, given the measurements. This is the a posteriori probability, which is not known. For a posteriori knowledge of the true tracer distribution, $Pr(F|P)$, the likelihood $Pr(P|F)$ is needed. Assuming that the acquisition model (system matrix A) is known, the following expression for the expected value r in the projection bin i is

$$r_i = \sum_{j=1}^{N} a_{ij} f_j \tag{14.3}$$

with N the number of pixels in the image. In reality, p_i rather than r_i is measured. p_i is Poisson distributed with the expectation value r_i as mean:

$$Pr(\,p_i|r_i) = \exp(-r_i)\frac{{r_i}^{p_i}}{p_i!} \tag{14.4}$$

All measured photons are independent, and the total likelihood is the multiplication of the individual likelihoods:

$$Pr(P|F) = \prod_i \exp(-r_i)\frac{{r_i}^{p_i}}{p_i!} \tag{14.5}$$

Maximizing this term is equivalent to maximizing the logarithm, which is less complicated. This function is referred to as the log-likelihood function:

$$L(P|F) = \sum_i \left(-r_i + p_i \ln r_i - \ln p_i!\right) \tag{14.6}$$

The projection data cannot be altered, so the last term is omitted. Maximization of this expression can be done in various ways, but the expectation-maximization (EM) algorithm guarantees convergence. The derivation is not discussed in detail [60,61,68]. The iterative EM algorithm is given by the following expression:

$$f_j^{(k+1)} = \frac{f_j^{(k)}}{\sum_{i=1} a_{ij}} \sum_{i=1}^{M} a_{ij} \frac{p_i}{\sum_{j'=1}^{N} a_{ij'} f_{j'}^{(k)}} \tag{14.7}$$

where M is the number of detection bins and $\sum_{j'=1}^{N} a_{ij'} f_{j'}^{k}$ is the forward-projection of the previous estimation f^k in the projection bin i. The ratio of the measured value p_i and this estimation is back-projected for each projection bin. These values are normalized and then multiplied with the current image estimate to obtain an updated estimate of the true tracer distribution. In a MAP reconstruction, a priori anatomical or analytical information is included in the reconstruction algorithm, which accordingly includes a penalization factor $Pr(F)$ next to the data-fit term $Pr(P|F)$ [43,45,46]. This is illustrated using the formula for the one-step-late algorithm:

$$f_j^{(k+1)} = \frac{f_j^{(k)}}{\sum_{i=1} a_{ij} + \beta \dfrac{\partial U}{\partial f_j^{(k)}}} \sum_{i=1}^{M} a_{ij} \frac{p_i}{\sum_{j'=1}^{N} a_{ij'} f_{j'}^{(k)}} \tag{14.8}$$

where the function U is called the total energy function. U can, for instance, be expressed as a weighted sum of potential functions given by quadratic differences between pixel intensities. The influence of this regularization is tuned by an additional parameter β, which expresses the relative importance of the data-fit term and of the smoothing regularization. If β is small, then the reconstructions are forced to fit the data, resulting in noisier images.

14.3 Monte Carlo Method

Monte Carlo methods are statistical simulation methods wherein a statistical simulation is any simulation that uses a sequence of random numbers. In order to do Monte Carlo calculations, some a priori information about the occurring physics processes in the simulations is needed; this information is expressed in the form of probability density functions (pdfs). When simulating photon interactions, the partial and total cross-section data (based on the material constitution) represent such information used to calculate the path length and the type of interaction. Afterwards, these pdfs are sampled by predefined rules using randomly generated numbers. The energy of a photon can be dissipated along its path, or the photon can penetrate all scattering and attenuating media to reach the detector where a new pdf-sampling decides whether it should be accounted for in the scoring region or whether it should be discarded [74].

14.3.1 Random numbers

Random numbers are important when modeling a physical system by a statistical model. Every random-number generator has to deliver uncorrelated, uniform, and reproducible sequences with a very long period in a short amount of time. Some algorithms are able to generate repeating sequences of C integers, which are, to a fairly good approximation, randomly distributed in the range 0 to $C-1$, where C is a large integer, for instance, $\mathrm{mod}(2^k)$, with k the integer word size of the computer. An example of such an algorithm is the linear congruential algorithm. The formula linking the nth and $(n+1)$th integers in the sequence is

$$I_{n+1} = (AI_n + B)\mathrm{mod}(C) \tag{14.9}$$

where A and B are constants. The first number of the series, the seed value, must be randomly changed, for instance, by triggering a value from a call to the system clock.

14.3.2 Sampling methods

To obtain a stochastic variable that follows a particular pdf, $f(x)$, three different sampling methods can be used.

14.3.2.1 Distribution function method

A cumulative distribution function $F(x) = P(X \leq x)$ is constructed from the integral of $f(x)$ over the interval $[a, x]$ according to

$$F(x) = \int_a^x f(x')dx' \tag{14.10}$$

The variable x is then sampled by replacing $F(x)$ in Equation 14.10 with a uniformly distributed random number in the range of $[0, 1]$ and solving for x. This sampling method is used if the inverse of $F(x)$ can be easily calculated.

14.3.2.2 Rejection method

Possibly, there can arise some mathematical difficulties in calculating the inverse of the cumulative distribution function F. In such case, the rejection method can offer a solution. Let $f(x)$ be defined in the range $[a, b]$, and let $r(x)$ be the rejection function $\forall x \in [a, b] \Rightarrow 0 < r(x) \leq 1$. Here, $r(x)$ is given as $f^*(x)$, being the pdf after normalization to unity. Then, a uniform distributed value of x within the range $[a, b]$ can be sampled from the relation

$$x = a + R_1(b - a) \tag{14.11}$$

where R_1 is a random number in $[0, 1]$. A second random number R_2 in $[0, 1]$ is then drawn to decide whether the sampled x should be accepted by checking whether

$$R_2 < r(x) = f^*(x) \tag{14.12}$$

If this relation is fulfilled, then x is accepted as a properly distributed stochastic value; otherwise, x needs to be resampled.

14.3.2.3 Mixed methods

When the previous two methods are not applicable, a combination of the two methods is used. The $f(x)$ is written as the product of two probability distribution functions $m(x).r(x)$, with $r(x)$ acting as the rejection function defined in the previous section. One determines an x value using the distribution function method on $m(x)$ and applies the rejection method with that x to $r(x)$ [48].

14.3.2.4 Nonanalog sampling

A Monte Carlo simulation involving the exact pdfs may be impractical from a computational point of view, requiring unrealistic computation time to achieve reliable results. One remedy consists of biasing the sampling. When simulating photon transport, this can be done using stratification (i.e., by sampling more from regions that highly contribute to the results), or using techniques such as splitting or Russian roulette in which particles are weighted, or by implementing forced detection (i.e., by forcing particles to interact so that they contribute to the simulation result) [30].

14.3.3 Photon transport modeling

Compton incoherent scattering and photoelectric effect are the two most important interaction mechanisms in nuclear medicine simulations. The total photoelectric cross section for a given energy, E, is calculated using discretized

libraries by means of interpolation. The incident photon is absorbed, and an electron is emitted in the same direction as the incident photon. The kinetic energy of the electron is defined as the difference between the energy of the original photon and the binding energy of the electron. The subshell of emission is again randomly sampled, thereby using cross-section data for all subshells. When simulating the Compton scattering of a photon from an atomic electron, an empirical cross-section formula is used, and sampling of the final state is done by a variant of the mixed Monte Carlo method, as noted in Section 14.3.2.3. The following illustrates how this is implemented for GEANT4 [1], a particle-tracking code that forms the basis of the Monte Carlo simulator GATE [36]. The quantum mechanical Klein-Nishina differential cross section per atom is [27]

$$f(x) = \pi r_e^2 \frac{m_e c^2}{E_0} Z \left[\frac{1}{x} + x\right] \left[\frac{1 - x\sin^2\theta}{1 + x^2}\right] \tag{14.13}$$

where

r_e = classical electron radius

$m_e c^2$ = electron mass

E_0 = energy of the incident photon

E_1 = energy of the scattered photon

$x = \dfrac{E_1}{E_0}$

Assuming an elastic collision, the scattering angle θ is defined by the Compton formula:

$$E_1 = E_0 \frac{m_e c^2}{m_e c^2 + E_0(1 - \cos\theta)} \tag{14.14}$$

The value of x corresponding to the minimum photon energy (backscatter) is given by:

$$a = \frac{m_e c^2}{m_e c^2 + 2E_0} \tag{14.15}$$

Hence, $x \in [a, 1]$. One may state that

$$\begin{aligned} f(x) &\approx \left[\frac{1}{x} + x\right] \left[\frac{1 - x\sin^2\theta}{1 + x^2}\right] \\ &= m(x).\, r(x) \\ &= \left[\alpha_1 m_1(x) + \alpha_2 m_2(x)\right].\, r(x) \end{aligned} \tag{14.16}$$

where

$$\begin{aligned} \alpha_1 &= \ln\left(\frac{1}{a}\right) \\ m_1(x) &= \frac{1}{\alpha_1 x} \\ \alpha_2 &= \frac{\left(1-a^2\right)}{2} \\ m_2(x) &= \frac{x}{\alpha_2} \end{aligned} \tag{14.17}$$

m_1 and m_2 are probability density functions defined on the interval $[a, 1]$, and $r(x)$ is set to

$$r(x) = \left[\frac{1 - x\sin^2\theta}{1+x^2}\right] \tag{14.18}$$

being the rejection function $\forall x \in [a,1] \Rightarrow 0 < r(x) \leq 1$. Given a set of three random numbers R_0, R_1, R_2 uniformly distributed on the interval [0, 1], the sampling procedure for x is the following:

1. Decide whether to sample from $m_1(x)$ or $m_2(x)$: if $R_1 < \dfrac{\alpha_1}{(\alpha_1+\alpha_2)}$ select $m_1(x)$; otherwise, select $m_2(x)$.
2. Sample x from the distributions to m_1 or m_2:

   For $m_1 : x = a^{R_1}$
   For $m_2 : x^2 = a^2 + \left(1-a^2\right) R_1$

3. Calculate $\sin^2\theta = t\,(t-1)$, where $t \equiv (1-\cos\theta) = \dfrac{m_e c^2\,(1-x)}{E_0 x}$.
4. Test the rejection function: if $r(x) \geq R_2$, accept x; otherwise, go to step 1.

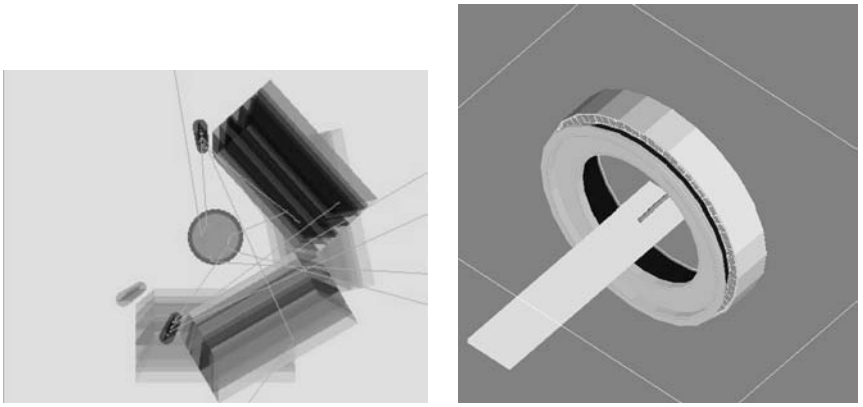
The polar angle θ is deduced from the sampled x value, and in the azimuthal direction, the angular distributions of both the scattered photon and the recoil electron are considered to be isotropic [1,27].

14.3.4 Scoring

A selection of relevant physical quantities such as energy, detection location, and photon interactions must be accumulated into tallies or scores. Also, an estimate of the statistical error as a function of the number of trials (and other quantities) must be determined.

14.3.5 State of the art in emission tomography simulators

Some Monte Carlo simulation codes dedicated to the simulation of SPECT and PET have been developed, such as SimSET [29] enabling SPECT and



(a) (b)

FIGURE 14.1: Example scanner types modeled in GATE: (a) SPECT system (IRIX-Beacon), and (b) PET system (ECAT HR+).

PET modeling; SIMIND [44] and SimSPECT [73] for SPECT; and PET-SIM [66], Eidolon [75], and SORTEO [59] for PET. To ensure high flexibility in simulation design, general-purpose nuclear physics codes such as EGS4 [9], GEANT4 [1], and MCNP [10] have also been used for SPECT and PET simulations. More specifically, GATE [36] has recently been designed as an upper layer of GEANT4 tuned for simulating SPECT and PET acquisitions. Examples of a modeled SPECT and PET scanner using GATE are shown in Figure 14.1.

14.4 Monte Carlo–Based Image Reconstruction

In Section 14.2, we assumed the system matrix (each element a_{ij} denotes the possibility that a decay in voxel i will be detected in bin j) to be known exactly. The success of a reconstruction algorithm is, however, mainly dependent on the accuracy of modeling this system matrix. Using the Monte Carlo method, as in Section 14.3, to simulate this system matrix may be the optimal approach for accurate quantitative image reconstruction, as discussed hereafter both for SPECT and PET.

14.4.1 SPECT

SPECT images are degraded by photon flux attenuation, by detector and collimator blurring, and by scattered photons and photons that penetrate the collimator septa. Such image-degrading factors can have a large

impact on quantitative accuracy and on clinical diagnosis, and all have to be corrected for in order to obtain optimal SPECT images. Iterative reconstruction algorithms such as ML-EM (Section 14.2), ordered subset expectation-maximization (OSEM) [35], and dual matrix ordered subsets (DM-OS) [37,39] can correct for these image degradations if their effects are simulated during the reconstruction. The simulators can be based on analytical models, on Monte Carlo methods, or on a combination of these.

These SPECT simulators face a considerable challenge; they have to balance accuracy against efficiency. The simulator has to be fast enough to serve as an on-the-fly forward-projector in order to avoid the massive storage of the transition matrix, which can hinder efficient 3-D image reconstruction [22,24,42]. A large variety of simulators have been proposed [3,5,25,49,70]. Reconstruction methods based on these simulators have been shown to improve contrast-to-noise [4,6,38,71,72], lesion detectability [23,26,54], and more recently, quantification [11,18,21,33,34,50].

Among these, Beekman and colleagues have developed a computationally efficient approach involving the use of a dual matrix block–iterative expectation–maximization algorithm [8]. The dual matrix reconstruction uses a different transition matrix for forward-projection (calculating the projection data given the 3-D isotope distribution) and back-projection (calculating the 3-D isotope distribution given the projections). The forward-projection step models attenuation, detector blurring, and scatter, whereas the back-projection models only attenuation and detector blurring, which significantly reduces calculation times. For forward-projection, scatter is modeled by using a fast Monte Carlo simulator based on convolution-based forced detection (CFD) [7,15], which combines stochastic photon transport calculation in the patient with an analytic (instead of a stochastic) detector modeling. For low-energy isotopes, the detector response is often simplified to a Gaussian.

Using such a fast Monte Carlo simulator for the projection step makes it possible to perform a Monte Carlo simulation at each iteration of the reconstruction algorithm, thus avoiding the need for massive system matrix storage. These hybrid simulations (i.e., mixing Monte Carlo and analytical modeling) involved in the dual matrix–CFD reconstruction approach converge several orders of magnitude faster to low noise projections than those carried out with brute force Monte Carlo simulators, making the reconstruction computationally practical. Typical reconstruction times of less than 5 minutes have been reported for a cardiac perfusion SPECT study [16].

However, the Gaussian modeling of the collimator response function involved in conventional CFD is a valid approach for all low-energy SPECT applications where the emitted photons do not have enough energy to penetrate the septa of the collimator. Figure 14.2a shows the direct photons, which result in the typical Gaussian after blurring with the intrinsic resolution of the gamma camera and after binning to realistic grid dimensions, as in Figure 14.2b. When medium energy isotopes are considered, a Gaussian function is no longer accurate enough to model the point spread function, since there is

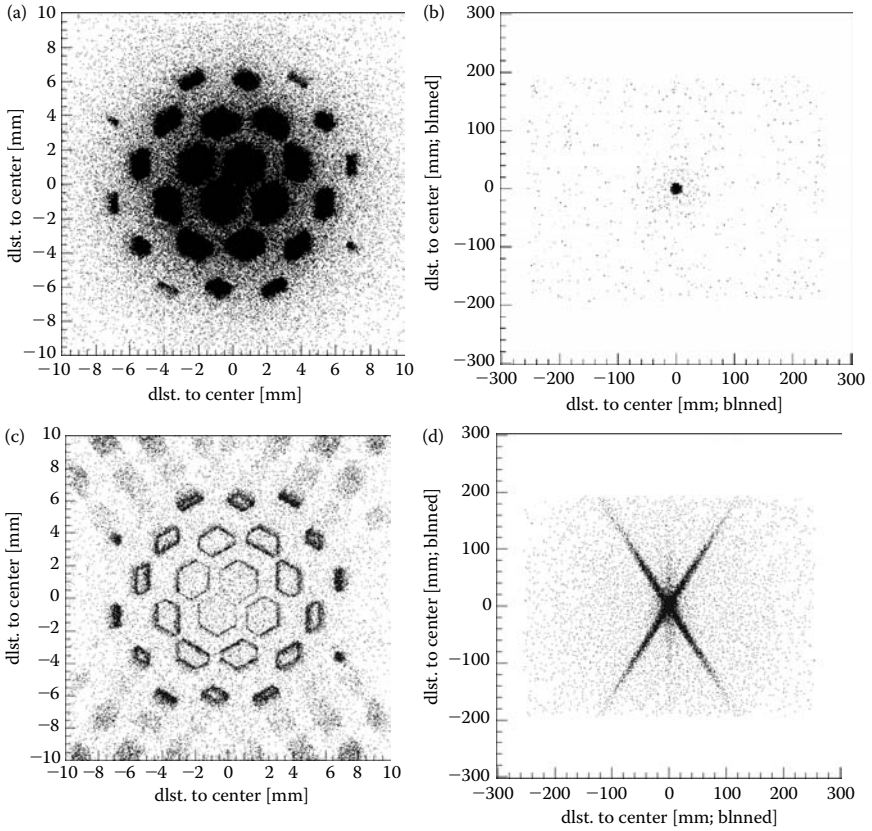


FIGURE 14.2: GATE simulated SPF scatter plots for (a) direct zoom, no intrinsic blur; (b) direct photons binned and with intrinsic resolution blurring; (c) same as (a) but for penetrated photons; and (d) same as (c) but for penetrated photons. (From Staelens, S. et al., *Phys. Med. Biol.* 52(11):3027–3043, June 7, 2007. With permission.)

a significant contribution from photons that penetrate one or more collimator lamella, as shown on Figure 14.2c. The individual hole positions are once again averaged out due to intrinsic resolution and binning, as in Figure 14.2d, which demonstrates the typical four- or six-streak artifacts due to septal penetration causing the point spread functions to deviate from a true Gaussian.

Several groups have proposed ways to correct for image degradations stemming from septal penetration. Spectral-based approaches [17,32,41] are calculation-wise inexpensive, but they increase the noise due to count subtraction. Recently proposed model-based [12–14,19,20,52,63] methods are highly accurate, but they are computationally very demanding. When incorporated in iterative fully 3-D reconstruction, such methods may require many hours of computation time, which may exceed clinically acceptable limits.

We believe that septal penetration for isotopes that suffer from contaminations originating outside the photopeak can often be corrected for on a spectral basis. These sufficiently accurate methods are easy to implement in the clinic using multiwindow acquisition modes. However, these spectral methods cannot compensate for photopeak photons that penetrate the collimator without undergoing any scatter interaction because such events spectrally coincide with the primaries, so next to energy spectra-based compensation, it is important to minimize septal penetration of (high-energy) direct photopeak photons. Therefore, we developed an efficient simulation method for calculating the effects of septal penetration in SPECT projections. We incorporated the latter method in the aforementioned dual matrix–CFD framework, accordingly resulting in a framework that is now generally applicable for a wide range of isotopes and collimators.

14.4.1.1 Methodology

The left part of Figure 14.3 represents the CFD part of the simulator. CFD starts with the stochastic simulation of the photon transport in the object,

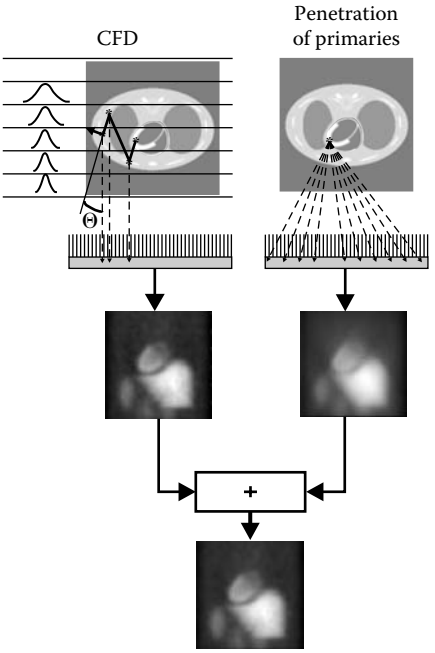


FIGURE 14.3: Methodology for any collimator: separated stochastic photon transport followed by convolution-based forced detection (left) and septal penetration calculation of primary photons (right). (From Staelens, S. et al., *Phys. Med. Biol.* 52(11):3027–3043, June 7, 2007. With permission.)

followed by Gaussian modeling of the gamma camera response. A copy of the original photon is created at the interaction site. However, instead of moving along a path that is randomly chosen by sampling the probability density function, the photon copies are forced to move along the path that runs along the axis of the collimator hole. In order to facilitate convolution with the appropriate blurring kernel, the weight of the copies is first stored in a 3-D subprojection map. When all copies are generated, the layers in the subprojection map, which are oriented parallel to the camera face, are convolved layer by layer using the distance-dependent camera response function. Finally, the projection is formed by adding together all convolved layers of the subprojection map. More details about this implementation are available in Beekman, de Jong, and Slijpen [7] and de Jong, Slijpen, and Beekman [15].

The right side of Figure 14.3 illustrates how the contribution of primary photon penetration is added. Rays are traced from each object voxel to each of the detector pixels. The attenuation along these rays is calculated by incremental Siddon [31]. The value assigned to each detector bin is the product of the source value and the line integral of attenuation through the object, multiplied by the appropriate septal penetration point spread function (SP-PSF) value. This SP-PSF table is presimulated with GATE (see Section 14.3.5) using a point source in air at a single (arbitrary) distance from the collimator front. Only septal penetrating photons that did not scatter elsewhere are recorded, and the resulting SP-PSF projection is stored on a fine grid.

If the SP-PSF were to be incorporated in collimator modeling during CFD, this would need intensive MC simulation of the detector response because the SP-PSF is energy dependent. It would also slow the CFD method by several orders of magnitude since the convolution kernels would no longer be linearly separable and would have a much larger spatial extent. In addition, significant errors would occur in the scatter angle and the attenuation path if the CFD central ray approach were used.

Our method completely decouples the SP-PSF-based simulation from the simulation of primaries and scatter photons. The only approximation in our new approach consists of neglecting the septal penetration of photons that scattered in the object. We have shown in previous work [65] the validity of this approach. Moreover, neglecting septal penetration after object scatter implies that the SP-PSF table has to be simulated using only the primary energy. The effects of object attenuation are calculated separately during the ray tracing. Accordingly, the distance-dependent SP-PSF is object independent and it is linearly scalable. The SP-PSF has to be simulated only once per collimator and per isotope for a single reference distance.

14.4.1.2 Results

Because the septal penetration distribution is spatially smooth for extended objects, the source and attenuation map as well as the detector can be coarsely binned without significant accuracy loss (further details in Staelens, De Wit, and Beekman [65]). In addition, by using a threshold mask,

TABLE 14.1: Calculation times for septal penetration simulation and corresponding accuracies for ^{201}Tl without and with SP-PSF thresholding on a 2.4 GHz dual CPU system.

Source Map	Att Map	Proj Size	No Threshold		2% Threshold	
			Time [s]	Accuracy [%]	Time [s]	Accuracy [%]
64	64	64	293 s	2.1	70.5 s	4.7
32	32	32	5.21 s	3.2	1.55 s	5.6
16	16	16	0.14 s	14.2	0.05 s	15.0
32	32	16	1.37 s	5.6	0.41 s	7.0
32	32	8	0.35 s	10.4	0.11 s	10.7

Source: Staelens, S. et al., *Phys. Med. Biol.* 52(11):3027–3043, June 7, 2007. With permission.

TABLE 14.2: Calculation times for septal penetration simulation and corresponding accuracies for ^{111}In without and with SP-PSF thresholding on a 2.4 GHz dual CPU system.

Source Map	Att Map	Proj Size	No Threshold		2% Threshold	
			Time [s]	Accuracy [%]	Time [s]	Accuracy [%]
64	64	64	296 s	4.3	37.9 s	6.5
32	32	32	5.18 s	6.8	1.06 s	8.6
16	16	16	0.13 s	14.2	0.05 s	15.3
32	32	16	1.38 s	7.9	0.28 s	8.0
32	32	8	0.36 s	18	0.08 s	18

Source: Staelens, S. et al., *Phys. Med. Biol.* 52(11):3027–3043, June 7, 2007. With permission.

we were able to exploit the intrinsic streak pattern of the SP-PSF in order to speed up the calculations. These three factors resulted in a reduction of the calculation cost by four orders of magnitude with only a small loss of accuracy.

When all three acceleration techniques are efficiently combined (source and attenuation map 32^3 and projection 16^2 with a 2% threshold), an accuracy of 7% is obtained for ^{201}Tl at a computational cost of 0.41 seconds per projection angle as shown in Table 14.1. This is equivalent to a 10.6% increase in calculation time when added to the CFD part of the simulator, which takes approximately 4 seconds per projection angle. For ^{111}In, the accuracy is 8.0% versus an increase in computation time of 7.2%, as illustrated by Table 14.2.

14.4.2 PET

As mentioned earlier, one of the major advantages of 3-D iterative reconstruction methods is that the physical degradations that limit spatial

resolution can be modeled into the system response matrix used by the reconstruction algorithm. Several groups use the Monte Carlo method to simulate factors of their PET system response [2,51,55]; others use an on-the-fly simulated system matrix [47]; and finally, a lot of researchers simulate the full system matrix offline [28,53,56–58,62,69]. However, due to the large number of scintillation crystals present in current PET scanners, the system matrix can be very large. For instance, Rafecas and colleagues use a database management system (DB2 from IBM) to handle, sort, store, and access the system matrix for their small animal PET scanner (MADPET-II) [57]. The system matrix was generated using realistic Monte Carlo simulations based on Geant3. Their approach allowed users to store the 3-D system matrix for this scanner and to reconstruct a larger FOV without compromising the resolution, in spite of the high statistical noise affecting a Monte Carlo–generated system matrix. Moreover, they showed that a small reduction in statistical noise within the matrix, achieved by applying the matrix symmetries, improved detectability, contrast, and spatial resolution [56].

14.4.2.1 Methodology

We have been investigating alternative basis functions to parameterize the object with the goal of further exploiting the inherent symmetries present in ring PET scanners [67]. In our approach, the object is represented by basis functions naturally provided by the scanner: natural pixels, as in Figure 14.4. This gives a block-circulant system matrix and provides a framework for

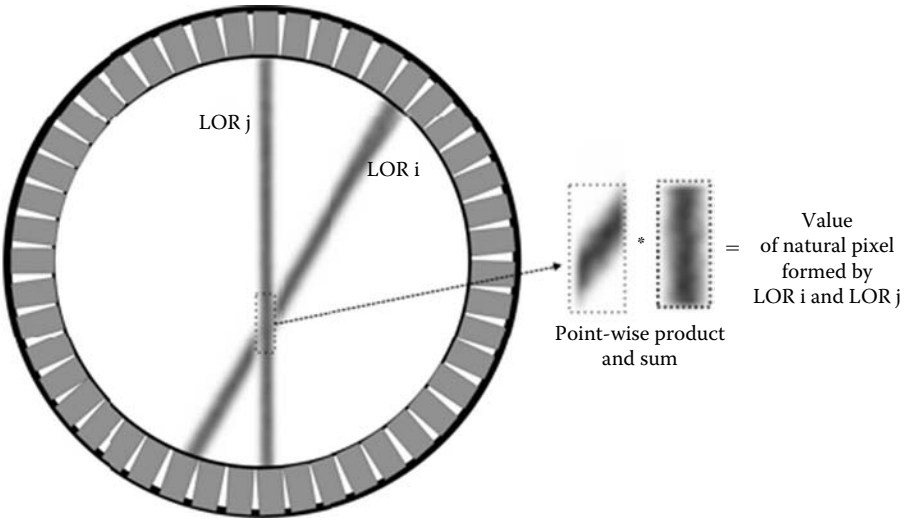


FIGURE 14.4: The calculation of the natural pixel system matrix element formed by two lines of response. (From Vandenberghe, S. et al., *Phys. Med. Biol.* 51(12):3105–3125, June 21, 2006. With permission.)

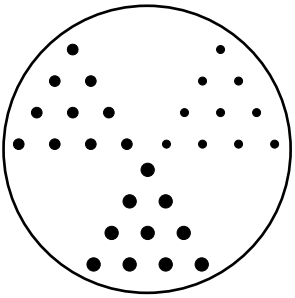


FIGURE 14.5: Derenzo-like phantom used for the evaluation.

the accurate modeling of the detection process. This block-circulant property allows for a substantial reduction in storage requirements for the system matrix. We simulate the system matrix for the natural pixel parameterization using GATE (see Section 14.3.5). Our results are illustrated hereafter by the contrast-to-noise ratio for a Derenzo-like phantom (see Figure 14.5).

The GATE Monte Carlo simulator modeled a PET scanner of 256 BGO crystals, each with size 7.35 mm × 7.35 mm. The thickness of the crystals was 20 mm. A circular phantom was used (Figure 14.5), which consisted of hot spheres on a warm background. The phantom is divided in six segments, and the spheres are positioned in three of the six segments. Spheres of sizes 12 mm, 8 mm, and 4 mm were used with the contrast between activity uptake within the sphere and within the background equal to 33. Three of the segments did not contain any spheres and were used to evaluate the noise in the background.

14.4.2.2 Results

Contrast versus noise for a 12 mm inner and outer sphere is shown in Figure 14.6 [67]. It is observed that the natural pixel-based method, which accurately models the detector response, provides higher contrast (approximately equal to the true sphere contrast of 33:1) than the traditional reconstruction method (simple line integral as system model) for the same noise level. For the traditional method, which does not explicitly model the system response, there is a clear difference between the recovered contrast for the inner and outer spheres. For the natural pixel method, both spheres reach the same contrast, meaning the contrast recovery is manifest even in regions toward the outside FOV where crystal penetration is significant.

14.5 Discussion

14.5.1 SPECT

The SP-PSF-based simulation together with the CFD result in a SPECT framework that is now generally applicable for a wide range of isotopes and

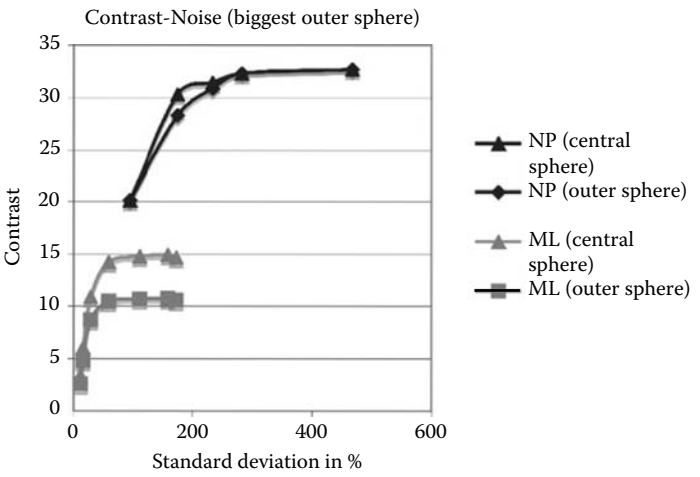


FIGURE 14.6: The contrast of the 12 mm sphere versus noise defined as the standard deviation in a background ROI. Results are shown for reconstructions using the natural pixel method and the traditional ML-EM method for both the centered and off-center sphere. (From Vandenberghe, S. et al., *Phys. Med. Biol.* 51(12):3105–3125, June 21, 2006. With permission.)

collimators and is fast enough to act as a projector in a fully 3-D iterative reconstruction algorithm. Attenuation, scatter, and distance-dependent spatial resolution are simulated using CFD, while the SP-PSF part of the simulator adds septal penetration of high-energetic photopeak photons. A similar approach is recommendable for collimator scatter. On the other hand, septal penetration correction for isotopes suffering from high-energy contaminations outside the photopeak could be excluded from this model-based forward-projector. Since those high-energy contaminants need to lose a vast amount of energy to reach the photopeaks of interest, other processes dominate. Multiple downscatter in the patient, septal penetration followed by Compton effect in the crystal, and backscatter from the camera end parts all result in an approximately constant bias in the energy spectrum, which accordingly obviates spectral-based corrections.

14.5.2 PET

The use of Monte Carlo to determine the natural-pixel PET system matrix elements provides an accurate model for the imaging system that needs to be simulated only once. The results discussed above ultimately prove that using accurate Monte Carlo simulations for detector modeling can compensate for degradation of resolution with radial distance. This does not include the influence of the object. It might be possible to simulate the patient separately

in order to preserve the rotational symmetry of the scanner component of the system matrix. This will result in attenuation and scatter sinograms, which can be used to construct a combined object–system matrix on the fly.

14.6 Future Research Directions

An important issue of Monte Carlo simulations is the calculation burden. In the late 1990s, Monte Carlo simulations in emission tomography were therefore used only by research laboratories to assess correction techniques and to study specific aspects of the PET or SPECT system response. Several approaches are investigated to solve this computational issue both for SPECT and PET. First, variance reduction techniques such as stratification, importance sampling, and forced detection can be used. Also, hybrid methods (as described in Section 14.4.1) can be proposed that combine Monte Carlo simulations in the object coupled to analytical or tabulated models for the detector. Lastly, efforts are conducted to submit simulations to a cluster of computers by distribution software or by intrinsically parallelizing the code.

In the meantime, Monte Carlo simulations have found their way to the clinic. Simulations contribute to the SPECT or PET imaging, for instance, through Monte Carlo–based scatter-correction techniques in the reconstruction process. This evolution is mainly caused by the rise of affordable miniclusters, which are readily incorporated in the acquisition consoles of recent emission tomographs.

Eventually, fully Monte Carlo–Based image reconstruction may lead to improved patient-specific quantification, diagnosis, dosimetry, and treatment planning, resulting in enhanced patient comfort.

14.7 Conclusions

We described the use of Monte Carlo simulations for emission tomography in nuclear medicine. We introduced the two main modalities in nuclear medicine imaging: SPECT and PET. In Section 14.2, we described several approaches for reconstructing the acquired data into a 3-D image. We highlighted statistical iterative reconstruction as the method of excellence because it allows incorporation of corrections for image-degrading effects through the system matrix. In Section 14.3, we described the Monte Carlo technique, which is a computational intelligent method to accurately simulate the aforementioned system matrix. Other noncomputational techniques, such as analytical system matrix descriptions or explicit measurements, do not allow correction of all image degradations at the detector and/or patient levels. Monte Carlo

simulations can correct for the realism of both the patient and the detector simultaneously, resulting in high contrast-to-noise performance for the reconstructed images. Sections 14.4.1 and 14.4.2 demonstrate how this is accomplished for SPECT and PET, respectively. In the case of SPECT, a hybrid simulation method is proposed, resulting in a framework that is now generally applicable for a wide range of isotopes and collimators and is fast enough to act as a projector in a fully 3-D iterative reconstruction algorithm. For PET, an alternative object parameterization through natural pixels is proposed, thereby proving that accurate Monte Carlo simulations for detector modeling can compensate for degradation of resolution with radial distance. All these results indicate that the Monte Carlo method has its place as a computational intelligent technique in nuclear medicine imaging and may eventually lead to improved patient-specific quantitative image reconstruction.

References

[1] Agnostelli, S., and GEANT4. 2003. Geant4: a simulation toolkit. *Nucl. Instrum. Meth. A*. 506: 250–303.

[2] Alessio, A., Kinahan, P., and Lewellen, T. 2006. Modeling and incorporation of system response functions in 3-D whole body PET. *IEEE T. Med. Im.* 25: 828–37.

[3] Beekman, F., and Viergever, M. 1995. Fast SPECT simulation including object shape dependent scatter. *IEEE T. Med. Im.* 14: 271–82.

[4] Beekman, F., Kamphuis, C., and Viergever, M. 1996. Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation. *IEEE T. Med. Im.* 15: 491–99.

[5] Beekman, F., den Harder, J., Viergever, M., and van Rijk, P. 1997. SPECT scatter modeling in non-uniform attenuating objects. *Phys. Med. Biol.* 42: 1133–42.

[6] Beekman, F., Kamphuis, C., and Frey, E. 1997. Scatter compensation methods in 3-D iterative SPECT reconstruction: a simulation study. *Phys. Med. Biol.* 42: 1619–32.

[7] Beekman, F., de Jong, H., and Slijpen, E. 1999. Efficient SPECT scatter calculation in non-uniform media using correlated Monte Carlo simulation. *Phys. Med. Biol.* 44: N183–92.

[8] Beekman, F., de Jong, H., and van Geloven, S. 2002. Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo–based scatter compensation. *IEEE T. Med. Im.* 21: 867–77.

[9] Bielajew, A., Hirayama, H., Nelson, W., and Rogers, D. 1994. History, overview and recent improvements of EGS4. Technical report, National Research Council, Canada.

[10] Briesmeister, J. 1993. MCNP: A general Monte Carlo N-particle transport code. Technical report, Los Alamos National Laboratory.

[11] Cot, A., Jane, E., Sempau, J., Falcon, C., Bullich, S., Pavia, J., Calvino, F., and Ros, D. 2006. Modeling of high energy contamination in SPECT imaging using Monte Carlo simulation. *IEEE T. Nucl. Sci.* 53: 198–203.

[12] Cot, A., Falcon, C., Crespo, C., Sempau, J., Pareto, D., Bullich, S., Lomena, F., Calvino, F., Pavia, J., and Ros, D. 2006. Absolute quantification in dopaminergic neurotransmission SPECT using a Monte Carlo-based scatter correction and fully 3-dimensional reconstruction. *J. Nucl. Med.* 46: 1497–1504.

[13] Chen, J., Garcia, E., Galt, J., Folks, R., and Carrio, I. 2006. Improved quantification in I-123 SPECT imaging with deconvolution of septal penetration. *Nucl. Med. Comm.* 27: 551–58.

[14] Chen, J., Garcia, E., Galt, J., Folks, R., and Carrio, I. 2006. Optimized acquisition and processing protocols for I-123 cardiac SPECT imaging. *J. Nucl. Card.* 13: 251–60.

[15] de Jong, H., Slijpen, E., and Beekman, F. 2001. Acceleration of Monte Carlo SPECT simulation using convolution based forced detection. *IEEE T. Nucl. Sci.* 48: 58–64.

[16] de Wit, T., Xiao, J., and Beekman, F. 2005. Monte Carlo-based statistical SPECT reconstruction: influence of number of photon tracks. *IEEE T. Nucl. Sci.* 52: 1365–69.

[17] Dewaraja, Y., and Koral, K. 1998. Quantitative ^{131}I SPECT with triple energy window Compton scatter correction. *IEEE T. Nucl. Sci.* 45: 3109–14.

[18] Dewaraja, Y., Ljungberg, M., and Koral, K. 2000. Characterization of scatter and penetration using Monte Carlo simulation in ^{131}I imaging. *J. Nucl. Med.* 41: 123–30.

[19] Dewaraja, Y., Ljungberg, M., and Fessler, J. 2006. 3-D Monte Carlo-based scatter compensation in quantitative ^{131}I SPECT reconstruction *IEEE T. Nucl. Sci.* 53: 181–88.

[20] Du, Y., Frey, C., Wang, W. T., Tocharoenchai, C., Baird, W., and Tsui, B. 2002. Combination of MCNP and SimSET for Monte Carlo simulation of SPECT with medium- and high-energy photons. *IEEE T. Nucl. Sci.* 49: 668–74.

[21] Du, Y., Tsui, B., and Frey, C. 2006. Model-based compensation for quantitative I-123 brain SPECT imaging. *Phys. Med. Biol.* 51:1269–82.

[22] El Bitar, Z., Lazaro, D., Coello, C., Breton, V., Hill, D., and Buvat, I. 2006. Fully 3-D Monte Carlo image reconstruction in SPECT using functional regions. *Nucl. Instr. Methods A* 569: 399–403.

[23] Farncombe, T., Gifford, H., Narayanan, M., Pretorius, H., Frey, E., and King, M. 2004. Assessment of scatter compensation strategies for ^{67}Ga SPECT using numerical observers and human LROC studies. *J. Nucl. Med.* 45: 802–12.

[24] Floyd, C., Jaszczak, R., Greer, K., and Coleman, C. 1986. Inverse Monte Carlo as a unified reconstruction algorithm for ECT. *J. Nucl. Med.* 27: 1577–85.

[25] Frey, E., and Tsui, B. 1996. A new method for modeling the spatially variant object dependent scatter response function in SPECT. *Proc. IEEE Med. Im.* (Anaheim, CA) 2, pp. 1082–86.

[26] Frey, E., Gilland, K., and Tsui, B. 2002. Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE T. Med. Im.* 21: 1040–50.

[27] Geant4. http://geant4.web.cern.ch/geant4/UserDocumentation/.

[28] Gimenez, E., Nacher, E., Gimenez, M., Benlloch, J., and Rafecas, M. 2006. Comparison of different approaches based on Monte Carlo methods to calculate the system matrix for small animal PET. *Nucl. Instrum. Meth. A.* 569: 346–49.

[29] Harrison, R., Vannoy, S., Haynor, D., Gillipsie, S., Kaplan, M., and Lewellen, T. 1993. Preliminary experience with the photon generator module of a public-domain simulation system for emission tomography. *Proc. IEEE Med. Im.* (San Francisco, CA), pp. 1154–58.

[30] Haynor, D., Harrison, R., and Lewellen, T. 1991. The use of importance sampling techniques to improve the efficiency of photon tracking in emission tomography simulations. *Med. Phys.* 18: 990–1001.

[31] Han, G., Liang, Z., and You, J. 1999. A fast ray-tracing technique for TCT and ECT studies. *Proc. IEEE Med. Im.* (Seattle, WA) 3, pp. 1515–18.

[32] Hapdey, S., Soret, M., and Buvat, I. 2006. Quantification in simultaneous Tc-99m/i-123 brain SPECT using generalized spectral factor analysis: a Monte Carlo study. *Phys. Med. Biol.* 51: 6157–71.

[33] He, B., and Frey, E. 2006. Comparison of conventional, model-based quantitative planar, and quantitative SPECT image processing methods for organ activity estimation using In-111 agents. *Phys. Med. Biol.* 51: 3967–81.

[34] He, B., Du, Y., Song, X., Segars, W., and Frey, E. 2005. A Monte Carlo and physical phantom evaluation of quantitative ^{111}In SPECT. *Phys. Med. Biol.* 50: 4169–85.

[35] Hudson, H., and Larkin, R. 1994. Accelerated image reconstruction using ordered subsets of projection data. *IEEE T. Med. Im.* 13: 601–09.

[36] Jan, S., Santin, G., Strul, D., Staelens, S., and OpenGATE. 2004. GATE: a simulation toolkit for PET and SPECT. *Phys. Med. Biol.* 49: 4543–61.

[37] Kadrmas D., Frey E., Karimi S., and Tsui B. 1998. Fast implementations of reconstruction-based scatter compensation in fully 3D SPECT image reconstruction. *Phys. Med. Biol.* 43: 857–73.

[38] Kadrmas D., Frey E., and Tsui B. 1998b. Application of reconstruction-based scatter compensation to 201Tl SPECT: Implementations for reduced reconstructed noise. *IEEE T. Med. Im.* 17: 325–33.

[39] Kamphuis, C., Beekman, F., van Rijk, P., and Viergever, M. 1998. Dual matrix ordered subsets reconstruction for accelerated 3-D scatter compensation in single-photon emission tomography. *Eur. J. Nucl. Med.* 25: 8–18.

[40] Lange, K., and Carson, R. 1984. Reconstruction algorithms for emission and transmission tomography. *J. Comput. Assist. Tomogr.* 8: 306–16.

[41] Larsson, A., Ljungberg, M., Jakobson Mo, S., Riklund, K., and Johansson, L. 2006. Correction for scatter and septal penetration using convolution subtraction methods and model-based compensation in I-123 brain SPECT imaging: a Monte Carlo study. *Phys. Med. Biol.* 51: 5753–67.

[42] Lazaro, D., El Bitar, Z., Breton, V., Hill, D., and Buvat, I. 2005. Fully 3-D Monte Carlo reconstruction in SPECT: a feasibility study. *Phys. Med. Biol.* 50: 3739–54.

[43] Leahy, R., and Yan, X. 1991. Incorporation of anatomical MR data for improved functional imaging with PET. In A. C. F. Colchester and D. J. Hawkes, Eds., *Information Processing in Medical Imaging.* New York: Springer, pp. 105–20.

[44] Ljungberg, M., and Strand, S. 1989. A Monte Carlo program for the simulation of scintillation camera characteristics. *Comp. Meth. and Progr. in Biomedicine* 29: 257–72.

[45] Gindi, G., Lee, M., Rangarajan, A., and Zubal, I. 1991. Bayesian reconstruction of functional images using anatomical information as priors. In A. C. F. Colchester and D. J. Hawkes, Eds., *Information Processing in Medical Imaging.* New York: Springer.

[46] Gindi, G., Lee, M., Rangarajan, A., and Zubal, I. 1993. Bayesian reconstruction of functional images using anatomical information as priors. *IEEE T. Med. Im.* 12: 351–59.

[47] Kudrolli, H., Worstell, W., and Zavarzin, V. 2002. SS3D-fast fully 3-D PET iterative reconstruction using stochastic sampling. *IEEE T. Nucl. Sci.* 49: 124–30.

[48] Ljungberg, M., Strand, S., and King, M. 1998. *Monte Carlo Simulations in Nuclear Medicine: Applications in Diagnostic Imaging*. Bristol, UK: Institute of Physics.

[49] Meikle, S., Hutton, B., and Bailey, D. 1994. A transmission-dependent method for scatter correction in SPECT. *J. Nucl. Med.* 23: 360–67.

[50] Moore, S., Ouyng, J., Park, M., and El fahkri, G. Monte Carlo-based compensation for patient scatter, detector scatter, and crosstalk contamination in In-111 SPECT imaging. *Nucl. Instr. and Methods.* 569: 472–476.

[51] Mumcuoglu, E., Leahy, R., Cherry, S., and Hoffman, E. 1996. Accurate geometric and physical response modeling for statistical image reconstruction in high resolution PET. *Proc. IEEE Med. Im.* (Anaheim, CA), pp. 1569–74.

[52] Ouyang, J., Moore, S., Park, M. A., and El Fahkri, G. 2005. Fast Monte Carlo estimation of patient and detector scatter and crosstalk contamination in SPECT imaging. *Proc. IEEE Med. Im.* (Puerto Rico), pp. M03–106.

[53] Ortuno, J., Guerra-Gutierrez, P., Rubio, J., Kontaxakis, G., and Santos. A. 2006. 3D-OSEM iterative image reconstruction for high-resolution PET using precalculated system matrix. *Nucl. Instrum. Meth. A.* 569: 440–44.

[54] Narayanan, M., Pretorius, H., Dahlberg, S., Leppo, J., Botkin, N., Krasnow, J., Berndt, W., Frey, E., and King, M. 2003. Evaluation of scatter compensation strategies and their impact on human detection performance Tc-99m myocardial perfusion imaging. *IEEE T. Nucl. Sci.* 50: 1522–27.

[55] Qi, J., Leahy, R., Cherry, S., Chatziioannou, A., and Farquhar, T. 1998. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Phys. Med. Biol.* 43: 1001–13.

[56] Rafecas, M., Boning, G., Pichler, B., Lorenz, E., Schwaiger, M., and Ziegler, S. 2004. Effect of noise in the probability matrix used for statistical reconstruction of PET. *IEEE T. Nucl. Sci.* 51: 149–56.

[57] Rafecas, M., Mosler, B., Dietz, M., Poegl, M., Stamatakis, A., McElroy, D., and Ziegler, S. 2004. Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data. *IEEE T. Nucl. Sci.* 51: 2597–605.

[58] Rehfeld, N., and Alber, M. 2006. The influence of noise in full Monte Carlo ML-EM and dual matrix reconstructions in positron emission tomography. *Med. Phys.* 33: 3498–507.

[59] Reilhac, A., Lartizien, C., Costes, N., Sans, S., Comtat, C., Gunn, R., and Evans, A. 2004. PET-SORTEO: a Monte Carlo-based simulator with high count rate capabilities. *IEEE T. Nucl. Sci.* 51: 46–52.

[60] Shepp, L., and Vardi, Y. 1982. Maximum likelihood reconstruction for emission tomography. *IEEE T. Med. Im.* 1: 113–22.

[61] Shepp, L., Vardi, Y., Ra, J., Hilal, S., and Cho, Z. 1984. Maximum likelihood PET with real data. *IEEE T. Nucl. Sci.* 31: 910–12.

[62] Shokouhi, S., Vaska, P., Southekal, S., Schlyer, D., Purschke, M., Dzordzhadze, V., Woody, C., Stoll, S., Alexoff, D., Rubins, D., Villaneuve, A., and Krishnamoorthy, S. 2004. Statistical 3D image reconstruction for the RatCAP PET tomograph using a physically accurate, Monte Carlo based system matrix. *IEEE T. Nucl. Sci. Proc. IEEE Med. Im.* (Rome, Italy), pp. 3901–05.

[63] Song, X., Segars, W., Du, Y., Tsui, B., and Frey, E. 2005. Fast modeling of the collimator-detector response in Monte Carlo simulation of SPECT imaging using the angular response function. *Phys. Med. Biol.* 50: 1791–804.

[64] Staelens, S., De Wit, T., and Beekman, F. 2006. Ultra-fast SPECT simulation with combined convolution based forced detection and efficient septal penetration modeling. *J. Nucl. Med.* 47: 198.

[65] Staelens, S., De Wit, T., and Beekman, F. 2007. Fast hybrid SPECT simulation including efficient septal penetration modelling (SP-PSF). *Phys. Med. Biol.* 52: 3027–43.

[66] Thomson, C., Cantu, J., and Picard, Y. 1992. PETSIM: Monte Carlo program simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems. *Phys. Med. Biol.* 37: 731–49.

[67] Vandenberghe, S., Staelens, S., Byrne, C., Soares, E., Lemahieu, I., and Glick, S. 2006. Reconstruction of 2D PET data with Monte Carlo generated natural pixels. *Phys. Med. Biol.* 51: 3105–25.

[68] Vardi, Y., Shepp, L., and Kaufman, L. 1985. A statistical model for positron emission tomography. *J. Am. Statist. Assoc.* 80: 8–20.

[69] Veklerov, E., and Llacer, J. 1988. MLE reconstruction of a brain phantom using a Monte Carlo transition matrix and a statistical stopping rule. *IEEE T. Nucl. Sci.* 35: 603–07.

[70] Walrand, S., Elmbt, L., and Pauwels, S. 1994. Quantification in SPECT using an effective model of the scattering. *Phys. Med. Biol.* 39: 719–34.

[71] Xiao, J., De Wit, T., Staelens, S., and Beekman, F. 2006. Evaluation of 3D Monte-Carlo-based scatter correction for Tc-99m cardiac perfusion SPECT. *J. Nucl. Med.* 47: 1662–69.

[72] Xiao, J., De Wit, T., Zbijewski, W., Staelens, S., and Beekman, F. 2007. Evaluation of 3D Monte-Carlo-based scatter correction for Tl-201 cardiac perfusion SPECT. *J. Nucl. Med.* 48: 637–44.

[73] Yanch, J., Dobrzeniecki, A., Ramanathan, C., and Behrman, R. 1992. Physically realistic Monte Carlo simulation of source, collimator and tomographic data acquisition for emission computed tomography. *Phys. Med. Biol.* 37: 853–70.

[74] Zaidi, H. 1999. Relevance of accurate Monte Carlo modeling in nuclear medical imaging. *Med. Phys.* 26: 574–608.

[75] Zaidi, H., Hermann Scheurer, A., and Morel, C. 1999. An object-oriented Monte-Carlo simulator for 3D cylindrical positron tomographs. *Comp. Meth. and Progr. in Biomedicine* 58: 133–45.

Chapter 15

Deformable Organisms: An Artificial Life Framework for Automated Medical Image Analysis

Ghassan Hamarneh, Chris McIntosh, Tim McInerney, and Demetri Terzopoulos

Contents

We present an emerging, artificial life framework for medical image analysis. It was originally introduced as an extension to the established physics-based image segmentation approach known as *deformable models*. Although capable of extracting coherent, smooth boundaries from low-level image information, classical deformable models rely on human guidance to produce satisfactory segmentation results. Our artificial life approach augments the bottom-up, data-driven deformable model methodologies with top-down, intelligent deformation control mechanisms, yielding intelligent deformable models that we call *deformable organisms*. This is achieved by adding behavioral and cognitive modeling layers atop the physical and geometrical layers of the classical models. The resulting organisms evolve according to behaviors driven by decisions based on their perception of the image data as well as their internally encoded anatomical knowledge. Thus, deformable organisms are autonomous agents that can automatically segment, label, and quantitatively analyze anatomical structures of interest in medical images. This chapter motivates and overviews our novel framework, describes its fundamental principles, demonstrates prototype instances of deformable organisms, and summarizes recent advances.

15.1 Introduction

Medical imaging has become essential to the practice of medicine, but accurate, fully automatic medical image analysis (MIA) continues to be an elusive ideal. A substantial amount of knowledge is often available about anatomical structures of interest—characteristic shape, position, orientation, symmetry, associated landmarks, relationship to neighboring structures, and so on—in addition to plausible image intensity characteristics, subject to natural biological variability or pathological conditions. Even so, MIA researchers have not yet succeeded in developing completely automatic segmentation techniques that can take full advantage of such prior knowledge to achieve segmentation accuracy and repeatability. Although it may be generally acknowledged that such techniques will require the incorporation of context-based information within a robust decision-making framework (Duncan and Ayache, 2000), we contend that prior frameworks of this sort are inflexible and do not operate at an appropriate level of abstraction, which limits their potential to deal with the most difficult data sets.

Deformable models demonstrated early promise in image segmentation, and they have become one of the most intensively researched segmentation techniques (McInerney and Terzopoulos, 1996). They are curve, surface, or solid models that change shape or deform in order to fit to and thereby delineate (segment) target structures in medical images. The measure of

goodness of fit for deformable models is usually represented by an energy functional: the better the segmentation, the lower the energy. The dynamic, energy-minimizing deformations of these shape models are typically simulated numerically through the time integration of associated Euler-Lagrange partial differential equations derived using variational calculus. In addition, the classical deformable model methodology, which is epitomized by "snakes" (Kass et al., 1988), is also based on an interactive strategy that relies on human expert initialization and guidance.

The difficult challenge in automating this approach is to develop intelligent initialization mechanisms, including the setting of free parameters, along with control mechanisms that can guide the optimization-driven segmentation process. Researchers have tried in vain to obtain the right global behavior (i.e., on the scale of the entire image) by embedding nuggets of contextual knowledge into the low-level optimization engine. As a result of such efforts, it has become painfully obvious that current deformable models have little to no explicit "awareness" of where they are in the image, how their parts are arranged, or to what structures they or any neighboring deformable models are converging during the optimization process. To make progress toward full automation, we must augment the deformable model framework by combining the powerful low-level feature detection and integration component with intelligent high-level control components.

To this end, we have proposed a new approach to automated MIA that augments deformable model methodologies with concepts from the field of artificial life (ALife) (see Terzopoulos, 1999). Our ALife approach to MIA was originally introduced in Hamarneh, McInerney, and Terzopoulos (2001), Hamarneh (2001), and McInerney and others (2002). In particular, we developed deformable organisms, which are autonomous agents capable of automatically segmenting, labeling, and quantitatively analyzing anatomical structures of interest in medical images. By architecting deformable organisms in a bottom-up, layered manner, we can separate the global, model-fitting control functionality from the local feature detection and integration functionality; thus, the deformable organism can make decisions about the segmentation process at the proper conceptual level.

The result is autonomous and, to a degree, intelligent segmentation algorithms that are aware of their progress and apply prior knowledge in a deliberative manner during different phases of the segmentation process. Analogous to natural organisms capable of voluntary movement, the artificial organisms possess deformable bodies with distributed sensors, as well as (rudimentary) brains with motor, perception, behavior, and cognition centers. Deformable organisms are perceptually aware of the image analysis process (Figure 15.1). Their behaviors, which manifest themselves in proactive movement and alteration of body shape, are based on sensed image features, prestored anatomical knowledge, and a deliberate cognitive plan.

The layered architecture as well as the reusable behavior routines and other components facilitate the rapid implementation of powerful, custom-tailored

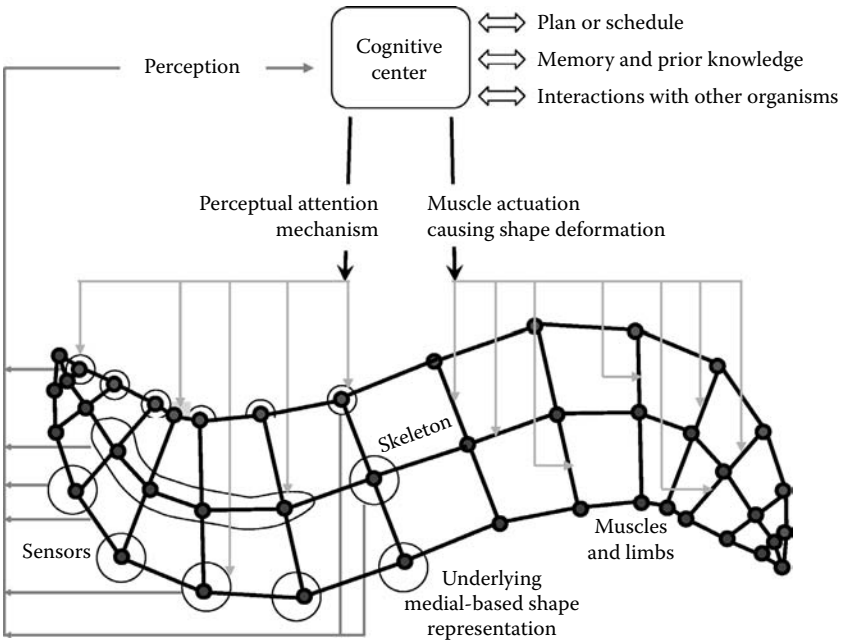


FIGURE 15.1: A deformable organism: the brain issues muscle actuation and perceptual attention commands. The organism deforms and senses image features whose characteristics are conveyed to the brain. The brain makes decisions based on sensory input, information in memory and prior knowledge, and a prestored plan, which may involve interaction with other organisms. (From Hamarneh, G., McInerney, T., Terzopoulos, D. 2001. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*, Utrecht, the Netherlands, 66–75. With permission.)

deformable organisms that can serve as new tools for automated segmentation, object-based registration, and the quantification of shape variation.

The remainder of this chapter is organized as follows. We begin with an overview of deformable organisms (Section 15.2) and present two illustrative examples. Following a detailed motivation and background, we review the ALife modeling framework, outline the notion of controlled deformations, and explain how intelligent decision and control mechanisms are incorporated into deformable organisms. The different layers of deformable organisms architecture are explained in Section 15.3. Recent research advances that utilize this framework, including extensions to 3-D physics-based deformations, vessel and spinal crawlers, evolutionary models, and an Insight Users Toolkit (ITK)-based deformable organisms software framework, are presented in Section 15.4. Examples and results are presented throughout the chapter.

15.2 Background and Related Work

Deformable organisms are deformable models equipped with higher-level control algorithms motivated by artificial life. The latter are responsible for controlling the shape deformations in a systematic manner based on a set of relevant criteria, a predefined set of situationally dependent actions, and the current system state, including the image and even the user.

15.2.1 Illustrative examples of deformable organisms

Several prototype deformable organisms based on an axis-symmetric body morphology were presented by McInerney and colleagues (2002). This geometric representation in conjunction with a set of multiscale deformation operators affords the motor center in the brain of the organism precise local growth and shape control processes at a variety of scales. Various reusable behavior routines, which support single organism behaviors as well as multiple, interacting organism behaviors, were also examined. Interaction among organisms may be as simple as (a) collision detection and the imposition of nonpenetration constraints between two or more organisms in contact; (b) one or more parent organisms spawning a new child organism and supplying it with appropriate initial conditions; or (c) the sharing of statistical shape constraints and/or image appearance information between organisms. More complex, rule-based interactions are also possible.

Figure 15.2 illustrates deformable organisms with a nontrivial example involving the detection and detailed segmentation of the lateral ventricle, caudate nucleus, and putamen in the left and right halves of a transverse 2-D MR image of the brain. Since the ventricles are the most discernible and stable structures in this imaging modality, the segmentation process begins with the release of two "ventricle organisms" in the black background region outside the cranium, at the upper left and right edges of the image in subfigure (1). Performing a coordinated scanning behavior, the organisms proceed first to locate the tops of the ventricles, as shown in the zoomed-in view of subfigure (2), and their inner and outer (with respect to the brain) boundaries (3)–(5). Next, both ends of each ventricle organism actively stretch to locate the upper and lower lobes of the ventricle (6), and then the organism fattens to finish segmenting the ventricle (7). Each organism employs the information that it has gleaned about the shape and location of the segmented ventricles to spawn and initialize a caudate nucleus organism in an appropriate location (8). Each caudate nucleus organism first stretches to locate the upper and lower limits of the caudate nucleus (9) then fattens until it has accurately segmented the caudate nucleus (10). From its bottommost point in the image, each caudate nucleus organism then spawns and initializes a putamen organism (11), which then moves laterally outward toward the low-contrast putamen (12). Each putamen organism then rotates and bends to latch onto the nearer

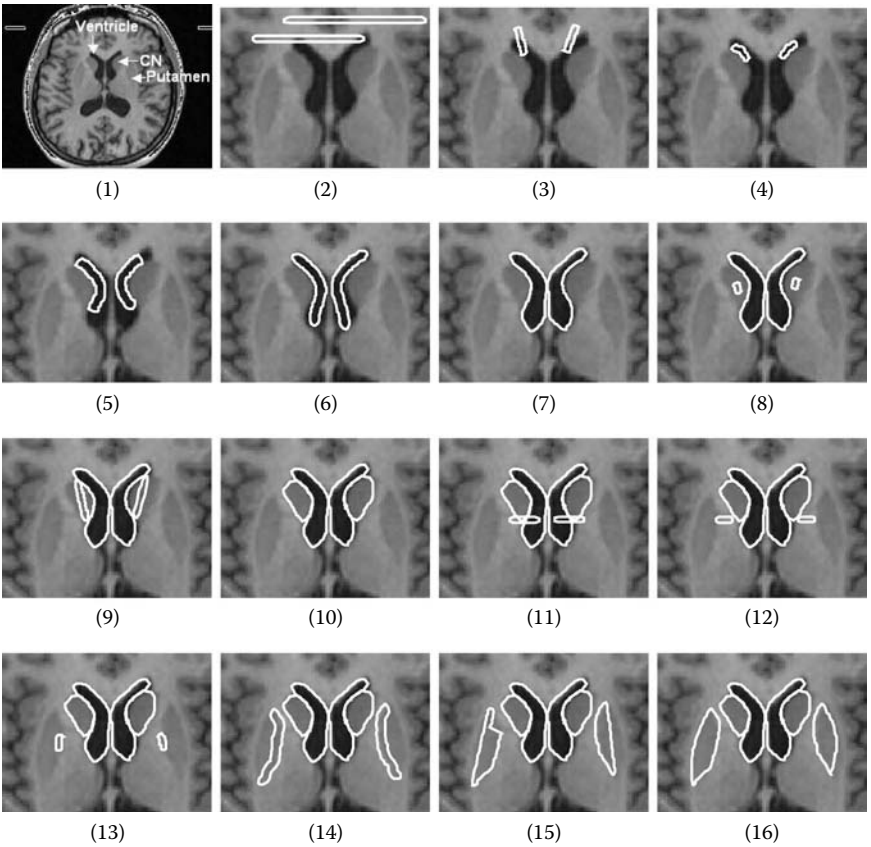


FIGURE 15.2: Automatic brain MR image segmentation by multiple deformable organisms. The sequence of images illustrates the temporal progression of the segmentation process. Deformable lateral ventricle (1–7), caudate nucleus (8–10), and putamen (11–16) organisms are spawned in succession and progress through a series of behaviors to detect, localize, and segment the corresponding structures in the MR image (see text). (From McInerney, T., Hamarneh, G., Shenton, M., Terzopoulos, D. 2002. *Medical Image Analysis*, 6(3): 251–266. With permission.)

putamen boundary (13). Next, it stretches and grows along the boundary until it reaches the upper and lower ends of the putamen (14), thus identifying the medial axis of the putamen (15). Since the edges of the putamen boundary near the gray matter are often weak, the organism activates an explicit search for an arc (parameterized by a single curvature parameter) that best fits the low-contrast intensity variation in that region, thus completing the segmentation (16).

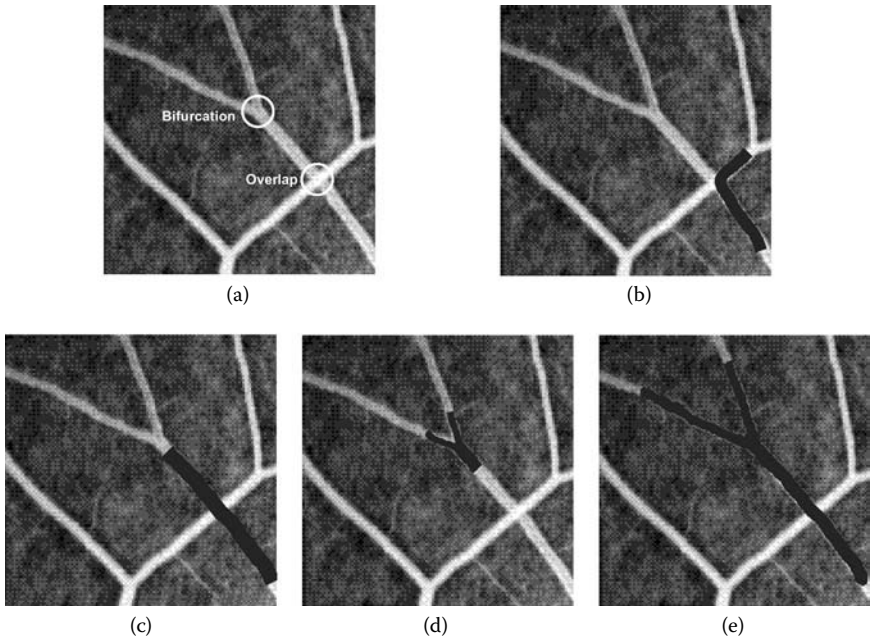


FIGURE 15.3: Multiple deformable organisms segmenting vascular structures in an angiogram. (a) Vessel overlap and bifurcation. (b) A simplistic vessel organism incorrectly bends into the more prominent overlapping vessel. (c) Appropriate high-level behaviors enable the vessel organism to identify the overlap and distinguish it from bifurcations. (d) Upon identifying a bifurcation, the organism spawns two new organisms, each of which proceeds along a branch. (e) The segmented vessel and branches. (From McInerney, T., Hamarneh, G., Shenton, M., Terzopoulos, D. 2002. *Medical Image Analysis*, 6(3): 251–266. With permission.)

As a second illustrative example, Figure 15.3 shows a different type of axis-symmetric deformable organism specialized to ribbonlike structures. This organism, called a vessel crawler, segments a vascular structure in a retinal angiogram. If given insufficient prior knowledge, the organism can latch onto the wrong overlapping vessel as shown in Figure 15.3b. However, given a suitable repertoire of perceptual capabilities and behavior routines (see Sections 15.3.3 and 15.3.4), the vessel organism can distinguish between overlapping vessels and deal with bifurcations (Figure 15.3c). When a bifurcation is encountered, the organism spawns two new child vessel organisms (Figure 15.3d), each of which extends along a branch (Figure 15.3e). These 2-D vessel crawlers can be extended to 3-D (Section 15.4.3).

15.2.2 Motivation and background

Current model-based MIA frameworks utilize geometric and often physical modeling layers. The models are fitted to images by minimizing energy functions, simulating dynamical systems, or applying probabilistic inference methods, but they do not control this optimization process other than in primitive ways, such as monitoring convergence or equilibrium. Some deformable models incorporate prior information to constrain shape and image appearance and the observed statistical variation of these quantities (Cootes et al., 1995; Cootes et al., 1999; Szekely et al., 1996). These models have no explicit awareness of where they or their parts are, and therefore the effectiveness of such constraints is dependent on appropriate model initialization. The lack of self-awareness may also prevent models from knowing when to trust the image feature information and ignore the prior constraint information or vice versa. The lack of optimization control can prevent these models from performing intelligent searches over their parameter spaces during the fitting process; that is, the constraint information is applied more or less indiscriminately and, once set in motion, the optimization process continues "mechanically" to completion.

Furthermore, because there typically is no active, deliberate search for stable image features, the models can latch onto nearby spurious features (Cootes et al., 1999). Their short-sighted decision-making abilities prevent these models from correcting missteps. Even if global optimization methods such as simulated annealing are employed to perform more extensive searches, the parameter space of the model is explored in a rather random fashion and there is no guarantee (other than an excruciatingly slow, asymptotic one) that the correct solution will be found. Moreover, it remains an open question whether suitable solution metrics can be defined for many MIA tasks using the "language" of objective functions and low-level optimization, or even of probabilistic inference.

There are several examples of earlier MIA systems that attempt to leverage top-down reasoning strategies against bottom-up feature detection, usually with limited success (e.g., the ALVEN cardiac left ventricular wall motion analyzer; Tsotsos et al., 1980). Other notable related work is the "visual routines" theory of Ullman (1984), which was proposed as a model of intermediate visual perception and includes a prominent top-down, sequential control mechanism. More recently, researchers have attempted to add some model-fitting control capability by constructing hierarchical deformable model-based systems (McInerney and Kikinis, 1998; Shen and Davatzikos, 2000). The idea here is that it is often possible to prioritize the strength or stability of different image features of the target structure(s). These systems attempt to control the fitting process such that the models systematically shift their focus from stable image features to less stable features. However, features may occur at different locations and scales and may vary from low-level landmark points to curves or surface patches to volumetric regions or to more complex features. Without

a proper nomenclature with which to define the high level features, and a system organization that provides more complete and easily programmable control over the model, it may be difficult to exploit this information effectively.

Alternatively, if a model can be made aware of itself and its environment, it can potentially be programmed to perform a more intelligent search for correct solutions by exploiting global contextual knowledge more effectively. For example, it may explore several alternative paths and choose the optimal result. To add this ability to deformable models, it seems prudent to investigate analogies with living systems.

15.2.3 Artificial life modeling

The modeling and simulation of living systems is central to an emerging scientific discipline known as artificial life (ALife).* In recent years, the ALife paradigm has had substantial impact in computer graphics, giving impetus to several important avenues of research and development, including artificial plants and animals, behavioral modeling and animation, and evolutionary modeling (Terzopoulos, 1999). These graphical models typically employ geometry- and physics-based techniques, as is characteristic of the deformable models used in MIA, but they also aspire to simulate many of the biological processes that characterize living systems—including birth and death, growth and development, natural selection, evolution, perception, locomotion, manipulation, adaptive behavior, learning, and cognition.

Most relevant to the MIA approach presented in this chapter is the ALife modeling of animals. The key components of artificial animals such as the prototypical "artificial fishes" (Terzopoulos et al., 1994) are synthetic bodies, including functional motor organs (contractile muscles), sensory organs (eyes, etc.), and most importantly, brains with motor, perception, behavior, and learning centers. In the motor center, motor controllers coordinate muscle actions to carry out specific motor functions, such as locomotion and sensor control. The perception center incorporates perceptual attention mechanisms, which support active perception that acquires information about the dynamic environment. The behavior center realizes an adaptive sensorimotor system through a repertoire of behavior routines that couple perception to action. The learning center in the brain enables the artificial animal to learn motor control and behavior through practice and sensory reinforcement.

To manage the complexity, artificial animals are best organized hierarchically, with each successive modeling layer adding to the more basic functionalities of underlying layers (Terzopoulos, 1999). At the base of the modeling

*See Levy (1992) for an entertaining perspective of the genesis of the artificial life field. Journals such as *Artificial Life* and *Adaptive Behavior* chronicle the state of the art.

hierarchy, a geometric modeling layer represents the morphology of the animal. Next, a physical modeling layer incorporates biomechanical principles to constrain geometry and emulate biological tissues. Further up the hierarchy, a motor control layer motivates internal muscle actuators to synthesize lifelike locomotion. Behavioral and perceptual modeling layers cooperate to support a reactive behavioral repertoire. The apex of the modeling pyramid, the domain of classical artificial intelligence, simulates the proactive or deliberative abilities of higher animals. Here, a cognitive modeling layer concerns how knowledge is represented and how automated reasoning and planning processes achieve high-level goals.

15.2.4 Controlling shape deformation

Top-down, knowledge-driven, model-fitting strategies require underlying shape representations that are responsive to high-level controlled shape deformation commands. As the organism's cognitive layer selects a behavior, the activated behavior routine will execute a series of shape deformations realized through high-level motor programs, which in turn are realized through simpler low-level motor skills. The goal here is to intelligently control the different types and extent of model deformations during the model-fitting process, focusing on the extraction of the most stable image features before proceeding to object regions with less well-defined features.

The choice of shape representation is undeniably crucial for segmentation, recognition, and interpretation of medical images. The study of shape has attracted a great deal of attention within the medical image analysis community (Bookstein, 1997; Costa and Cesar, 2000; Dryden and Mardia, 1998). A desirable characteristic in a shape representation is the ability to easily specify nonrigid object deformations at multiple locations and scales. Deformable shape models (McInerney and Terzopoulos, 1996), in particular, are mostly boundary based, and therefore multiscale deformation control is constructed from low-level boundary point sets and not upon object-relative geometry. Although they provide excellent local shape control, they lack the ability to undergo intuitive global deformation and to decompose shape variability into intuitive deformations, with some notable exceptions (Terzopoulos and Metaxas, 1991). It is therefore difficult to incorporate intelligent deformation control at the right level of abstraction into the conventional deformable model framework. Consequently, deformable models remain sensitive to initial conditions and spurious image features in image interpretation tasks. Hierarchical boundary-based shape models (Lachaud and Montanvert, 1999; Mandal et al., 1998; Miller et al., 1991; Montagnat and Delingette, 1997) and parameterized, free-form, and volumetric deformation mechanisms have been proposed (Barr, 1984; Coquillart, 1990; Sederberg and Parry, 1986; Singh and Fiume, 1998). They are also typically not defined in terms of the object; rather, the object is unnaturally defined in terms of the representation or deformation mechanism. Other 3-D shape representations

with similar drawbacks include spherical harmonics, finite element models, nonuniform rational B-splines, and wavelet-based representations (Burel and Henocq, 1995; Davatzikos et al., 2003; McInerney and Terzopoulos, 1995; Mortenson, 1997).

Shape models founded upon the use of the medial-axis transform (Blum, 1973) are a powerful alternative to boundary-based and volume-based techniques (Attali and Montanvert, 1997; Borgefors et al., 1999; Bouix et al., 2000; Dimitrov et al., 2003; Fritsch et al., 1997; Hamarneh and McInerney, 2003; Hamarneh et al., 2004; Leymarie and Levine, 1992; Pizer and Fritsch, 1999; Pizer et al., 2003; Sebastian et al., 2001; Siddiqi et al., 2002; Hamarneh et al., 2007). Medial representations provide both a local and global description of shape. Deformations defined in terms of a medial axis are natural and intuitive and they can be limited to a particular scale and location along the axis, while inherently handling smoothness and continuity constraints.

Statistical models of shape variability have been used for medical image interpretation (Cootes et al., 1995; Cootes et al., 2001; Duta et al., 1999; Leventon et al., 2000; Szekely et al., 1996). These typically rely on principal component analysis (PCA) and hence are only capable of capturing global shape variation modes (Cootes et al., 1995). This is because the eigen decomposition is typically performed on the covariance matrix of all the landmark locations representing the shape, without any spatial localization, as opposed to performing location- and deformation-specific dimensionality reduction (Hamarneh et al., 2004). The statistical analysis of medial-axis-based shape representation has been the focus of recent research (Fletcher et al., 2004; Grenander, 1963; Lu et al., 2003; Styner et al., 2003; Yushkevich et al., 2003).

15.2.5 Intelligent decision and control

The layered architecture enables a deformable organism to perform intelligent searches for anatomical features in images by allowing it to make appropriate high-level deformation decisions. The higher control layers afford the organism a nontrivial "awareness" (i.e., it knows where it is, where its parts are, and what it is seeking) and therefore it is more effectively able to utilize global contextual knowledge of the target anatomical structure, as well as local knowledge of the anatomical feature on which it is currently focused.

An organism typically begins by searching for the most stable anatomical features in the image, and then proceeds to the next best features, and so on. Alternatively, an organism may interact with other organisms to determine optimal initial conditions or resolve conflicting views of data. Once stable features are found and labeled, an organism can selectively use prior knowledge in regions known to offer little or no feature information. That is, the organism intelligently "fills in" the boundary in ways tailored to specific regions of interest in the target structure.

A deformable organism's awareness allows it to perform smart, flexible feature searches. It need not be satisfied with the nearest matching feature, but can look further within a region to find the best match, thereby avoiding globally suboptimal solutions. Furthermore, by carrying out explicit searches for features, correct correspondences between the organism and the data are more readily assured. If an expected feature cannot be found, an organism may "flag" the anomaly. If multiple plans exist, another plan can be selected and/or the search for the missing feature postponed until further information is available (e.g., from a neighboring organism). Alternatively, the organism can retrace its steps and return to a known state, and then inform the user of the failure. A human expert can intervene and put the organism back on course by manually identifying the feature. This strategy is possible because of the sequential and spatially localized nature of the deformable organism fitting process.

To support intelligent feature search, an organism needs powerful, flexible and intuitive model deformation control coupled with a flexible feature perception system. We currently achieve this with a set of "motor" (i.e., deformation) controllers and medial-axis-based deformation operators. Deformation controllers are parameterized procedures dedicated to carrying out a complex deformation function, such as successively bending a portion of the organism over some range of angle or stretching part of the organism forward some distance. They translate natural control parameters such as ⟨bend-angle; location; scale⟩ or ⟨stretch-length; location; scale⟩ into detailed deformations. The first deformable organisms used medial-profiles (Hamarneh et al., 2004) for their shape representation, which follow the geometry of the structure and describe general and intuitive shape variation (stretch, bend, thickness). Shape deformations result from the application on the medial profiles of deformation operators at certain locations and scales, or by varying the weights of the main variation modes obtained from a hierarchical (multiscale) and regional (multilocation) principal component analysis of the profiles (Hamarneh et al., 2004). Section 15.3 describes these organism components in detail.

Finally, an organism may begin in an "embryonic" state with a simple protoshape, and then undergo controlled growth as it develops into an "adult," proceeding from one stable object feature to the next. Alternatively, an organism may begin in a fully developed state and undergo controlled deformations as it carries out its model-fitting plan. The evolutionary corpus callosum organisms (Section 15.4.5) are an example of the latter.

Which type of organism to use, or whether to use some sort of hybrid organism, is dependent on the image and shape characteristics of the target anatomical structure. In summary, the ALife modeling paradigm provides a general and robust framework within which standard behavior subroutines may be employed to build powerful and flexible custom-tailored deformable organisms.

15.3 Deformable Organism Architecture

This section describes the layers of the deformable organism architecture, starting with the lowest layers (geometrical and physical) and working up to the higher layers (sensory, behavioral, and cognitive).

15.3.1 Geometrical layer (morphology and topology)

The geometrical layer of the deformable organism defines the shape representation as well as the morphological and topological constraints of the organism. The deformable organisms developed to date are based on medial-axis (Hamarneh et al., 2004) or medial-sheet (Hamarneh et al., 2007) deformable shape representations (Figure 15.1). The medial shape representation allows for a variety of intuitive and controlled shape changes that are described relative to the natural geometry of the object rather than as local displacements to its boundary. These changes in shape include stretching, thickening, or bending and are realized either as purely geometric deformations* or as deformations simulated at the physics layer of the organism (Hamarneh and McInerney, 2003; Hamarneh and McIntosh, 2005).

For example, medial profiles were used in the lateral ventricle, caudate nucleus, and putamen organisms, 2-D vessel crawler (Section 15.2.1), the corpus callosum (CC) organism (Section 15.3.4), and in the "evolutionary CC organism" (Section 15.4.5). Medial profiles are a set of functions that describe the geometry of the structure through general, intuitive, and independent shape measures (length, orientation, and thickness) taken with respect to the medial axis (Hamarneh et al., 2004). For instance, the thickness profile describes the distance from the medial axis to the boundary, while the orientation profile describes the angular offset between sequential medial nodes. In the physics-based CC organism (Section 15.4.1), the underlying geometry is a medial mesh of vertices and edges that is automatically constructed from the pruned skeletons of a training set of shapes (Hamarneh and McInerney, 2003). In 3-D vessel and spinal crawlers (Sections 15.4.3 and 15.4.4), the geometry is a tubular mesh with an underlying medial axis and a radial thickness measure connecting to a triangulated surface. The chosen representation imposes a strong tubular shape prior, while remaining general enough to allow for arbitrary stretching, bending, thickening, and branching, which is suitable for accurately representing tubular structures and complex vascular networks.

In the ITK Deformable Organisms software framework (Section 15.4.6), the developer of new deformable organism instantiations can substitute

*Deformations that manipulate the geometry of the model directly and without physical constraint.

various geometrical C++ classes, which conform to the predefined interfaces of abstract base classes. This restriction ensures that the user can employ different geometric layers (different shape parameterizations) while retaining all of the framework's most general abilities—those that are not specific to a particular parameterization. For example, the system can change between (Lagrangian) mesh models that explicitly store the locations of the surface and medial nodes or implicit (Eulerian) shape models, where the shape is the zero level-set of an embedding function.

15.3.2 Physical layer (motor system and deformations)

The physical layer of a deformable organism is analogous to the muscles in the human body. The actions carried out by the physical layer of the organism are divided into two classes: deformations and locomotions. The deformation capabilities of the organism constitute its low-level motor skills or basic actuations. These include global deformations such as affine transformations or localized (nonlinear) deformations such as a simple bulge or stretch. The more complex locomotion actions are the high-level motor skills of the organism. They result from several basic motor skills, which are parameterized procedures that carry out complex deformation functions, such as sweeping over a range of rigid transformation (translation, rotation) parameters; sweeping over a range of stretch, bend, or thickness parameters; bending at increasing scales; and moving a bulge along the boundary. Other high-level deformation capabilities include smoothing the medial axis or the boundary of the model, or reinitializing the medial axis to a position midway between the boundaries on both sides of the medial axis (sheet). See Figure 15.4 for examples.

Clearly the implementation of these deformations depends on the underlying geometry. Deformable organisms based on medial profiles can be deformed by directly modifying the profiles (Hamarneh et al., 2004). Mesh-based deformable organisms can be simulated as spring-mass systems, allowing deformations to take place through the modification of spring rest-lengths, or the application of forces to the masses or changing their velocities (Hamarneh and McIntosh, 2005). In principle, any shape representation can be used as long as it provides high-level, intuitive deformation control parameters.

15.3.3 Perceptual layer

The perception system of the deformable organism consists of a set of sensors that provide external information to the organisms. Using virtual sensors, an organism can collect data about its environment, such as image data or data from interaction with other organisms. For example, a single organism may have image intensity sensors, edge sensors, and texture sensors. Sensors can be focused or trained for a specific image feature in a task-specific way and, hence, the organism is able to disregard sensory information superfluous to its current behavioral needs.

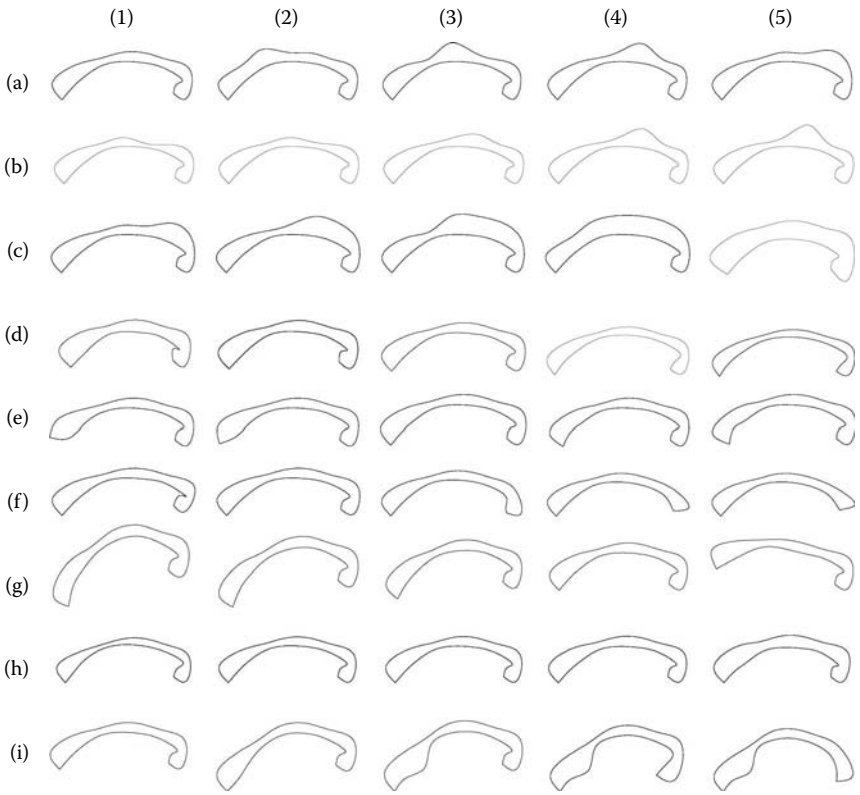


FIGURE 15.4: Examples of controlled deformations: (a)–(c) Operator-based bulge deformation at varying locations/amplitudes/scales. (d) Operator-based stretching with varying amplitudes over entire CC. (e)–(g) Statistics-based bending of left end, right end, and left half of CC. (h) Statistics-based bulge of the left and right thickness over entire CC. (i) From left to right: (1) mean shape, (2) statistics-based bending of left half, followed by (3) locally increasing lower thickness using operator, followed by (4) applying operator-based stretch and (5) adding operator-based bend to right side of CC. (From Hamarneh, G., McInerney, T., Terzopoulos, D. 2001. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*, Utrecht, the Netherlands, 66–75. With permission.)

Various parts of the organism are dynamically assigned sensing capabilities and thus act as sensory organs or receptors. A sensor can either be on-board or off-board. On-board sensors are confined to the organisms body, such as at its medial or boundary nodes, at curves or segments connecting different nodes, or at internal geometric subregions, while off-board sensors are free floating. Once the sensory data are gathered, they are fed to the cognitive center of the brain for processing.

A wide variety of image processing techniques can be used to enable the organism to perceive relevant features of its surroundings. In the ITK deformable organisms software framework (Section 15.4.6), the abundant image processing filters supported by the ITK can be applied. The results of potentially complex image processing pipelines can be used as perceptual input to the organisms.

As an example, consider one of the earliest deformable organisms created, the CC worm organism for segmenting the corpus callosum in midsagittal magnetic resonance (MR) images (Section 15.3.4). The geometry-based CC worm is equipped with sensors measuring image intensity, image gradient magnitude and direction, multiscale edge-preserving diffusion-filtered images (Perona and Malik, 1990; Weickert, 1998), a Canny-edge detected version of the image, and the result of a Hough transform applied to locate the top of the human skull in the brain MR image. The physics-based CC deformable organism (Section 15.4.1) was equipped with sensors for collecting statistical measures of image data such as mean and variance, in addition to edge strength and direction sensors. The vessel crawler organisms (Sections 15.2.1 and 15.4.3) used sensory input modules to make decisions about which direction to grow in, and where bifurcations are located. For example, the 2-D angiogram vessel crawler is equipped with an off-board arc sensor (Figure 15.5), which enables it to differentiate between overlapping vessels and authentic branch points. The 3-D crawler makes use of a similar spherical sensor, as well as the eigenvectors of an image intensity-based Hessian (Frangi et al., 1998) to estimate directions of curvature along the vasculature (Figure 15.6).

15.3.4 Behavioral layer

A behavior is defined as the reaction of a life form in response to external or internal stimuli. A deformable organism's behaviors are a set of actions performed in response to perceived input, the actions of other organisms, or input from a user, or they are self-initiated as part of a predetermined plan. Behavioral routines are designed based on the available organism motor skills, perception capabilities, and available anatomical landmarks.

For example, the routines implemented for the CC deformable organism introduced in Hamarneh et al. (2001) include find-top-of-head, find-upper-boundary-of-CC, find-genu, find-rostrum, find-splenium, latch-to-upper-boundary, latch-to-lower-boundary, find-fornix, thicken-right-side, thicken-left-side, and back-up. Each behavior routine subsequently activates the appropriate deformation or growth controllers to complete a stage in the plan and bring an organism closer to its target shape. The progression of these routines is shown in Figure 15.7. Starting from an initial default position shown in subfigure (1), the CC organism goes though different behaviors as it progresses toward its goal. As the upper boundary of the CC is very well defined and can be easily located with respect to the top of the head, the cognitive center of the CC organism activates behaviors to locate first

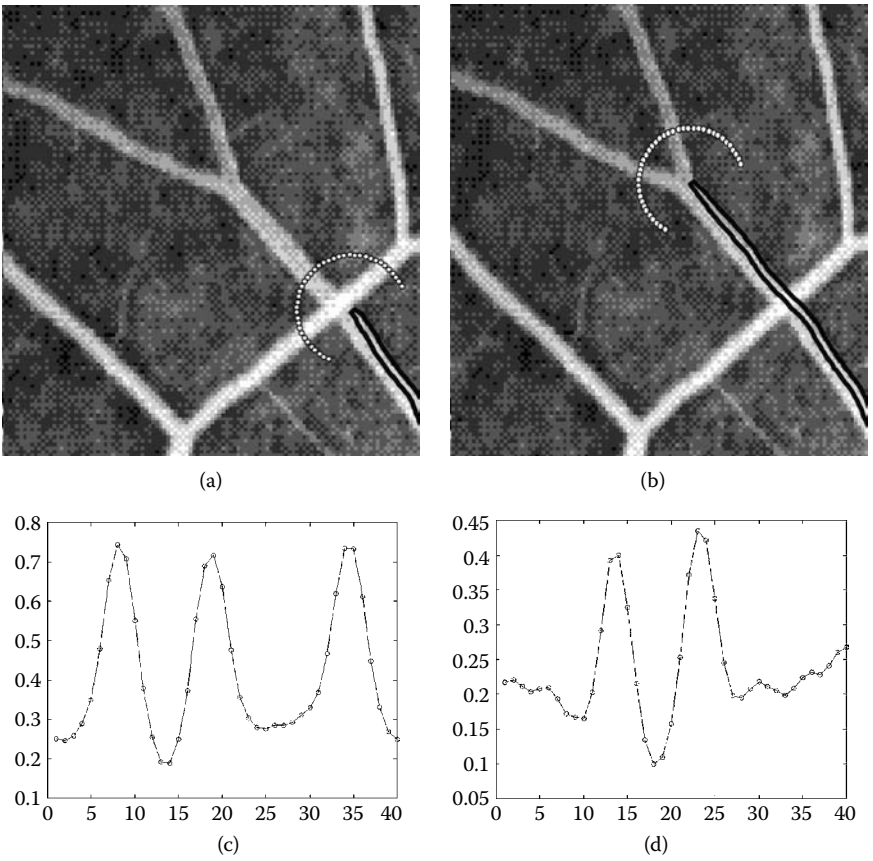


FIGURE 15.5: Off-board sensors (arc of white nodes in (a) and (b)) measure image intensity (along the arc). This results in an intensity profile exhibiting three distinct peaks when an overlapping vessel is ahead (c) and only two peaks in the case of a bifurcation (d).

the top of the head (subfigures 2–3) then move downwards through the gray and white matter in the image space to locate the upper boundary of the CC (4–7). The organism then bends to latch to the upper boundary (8) and activates a find-genu routine, causing the CC organism to stretch and grow along this boundary toward the genu (9–11). It then activates the find-rostrum routine causing the organism to back up, thicken (12), and track the lower boundary until reaching the distinctive rostrum (13–15). Once the rostrum is located, the find-splenium routine is activated and the organism stretches and grows in the other direction (15–16). The genu and splenium are easily detected by looking for a sudden change in direction of the upper boundary toward the middle of the head. At the splenium end of the CC, the organism backs up and finds the center of a circle that approximates

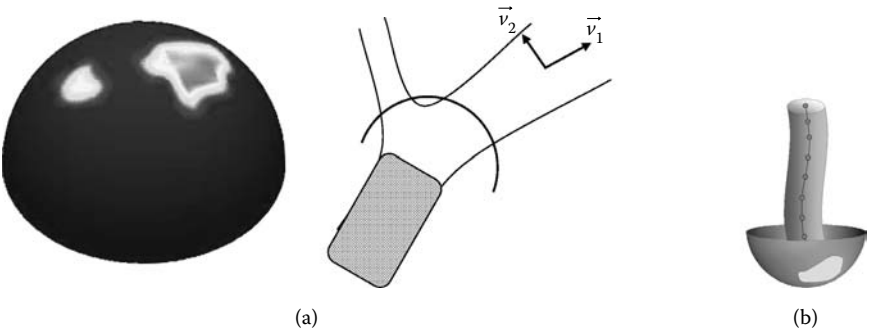


FIGURE 15.6: 3-D sensors for the crawlers. (a) A vessel crawler (gray) utilizing a hemispherical off-board sensor with example output (left), with the primary and secondary eigenvectors of the Hessian computed at optimal scale (top right). (b) A spinal crawler (gray) utilizing a hemispherical "off-board" sensor showing an idealized output (light gray). The crawlers use these sensors to determine the centerline of the vessel/spine tube by extracting the midpoint of the tube's intersection with the surface of the hemisphere. Once the centerline estimate is obtained, the crawlers can add a new layer to their geometry and, consequently, expand along the centerline of the structure (Figures 15.13 and 15.14). (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.) (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.)

the splenium end cap (17). The lower boundary is then progressively tracked from the rostrum to the splenium while maintaining parallelism with the organism's medial axis in order to avoid latching to the potentially connected fornix structure (18–21). Nevertheless, the lower boundary might still dip toward the fornix, so a successive step of locating where, if at all, the fornix connects to the CC is performed by activating the find-fornix routine (making use of edge strength along the lower boundary, its parallelism to the medial axis, and statistical thickness values). Thus, prior knowledge is applied only when and where required. If the fornix is indeed connected to the CC, any detected dip in the organism's boundary is repaired by interpolation using neighboring thickness values. The thickness of the upper boundary is then adjusted to latch to the corresponding boundary in the image (22–26). At this point the boundary of the CC is located (26) and the CC organism has almost reached its goal. However, at this stage the axis is no longer in the middle of the CC organism (27) so it is re-parameterized until the medial nodes are halfway between the boundary nodes (28–30). Finally, the upper and lower boundaries, which were reset in the previous step, are relocated (31–36) to obtain the final segmentation result shown in subfigure (36).

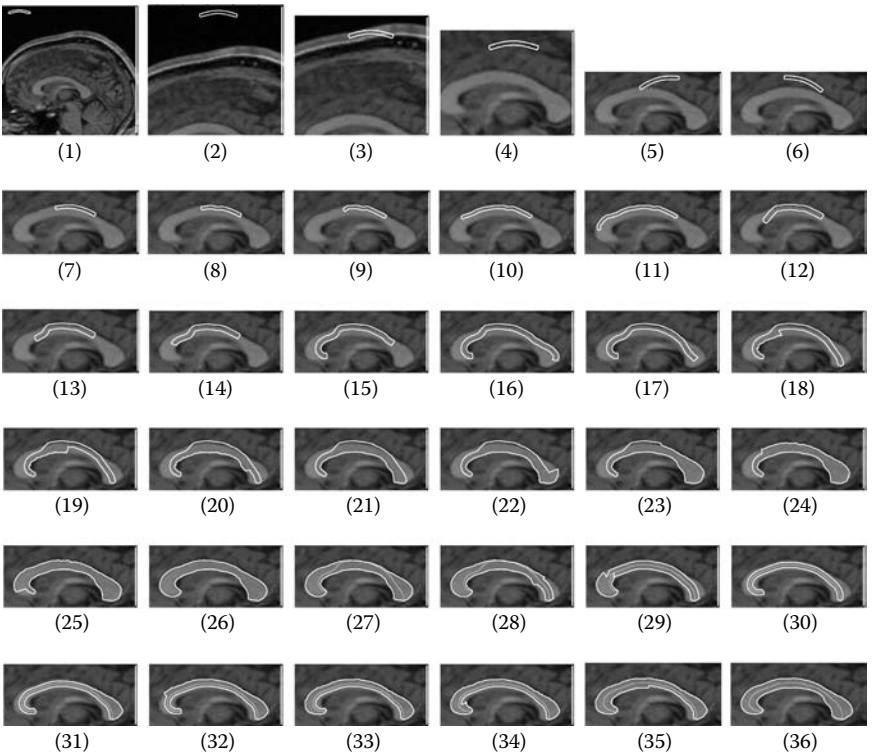


FIGURE 15.7: Deformable corpus callosum organism progressing through a sequence of behaviors to segment the CC (see text). (From McInerney, T., Hamarneh, G., Shenton, M., Terzopoulos, D. 2002. *Medical Image Analysis*, 6(3): 251–266. With permission.)

As a second example, the behaviors exhibited by the physics-based CC deformable organisms (Section 15.4.1) included initializing itself in the image, aligning its body to the CC (first globally, i.e., its whole body, then locally, i.e., its anatomical subregions), and detecting and repairing segmentation inaccuracies commonly caused by the similar image intensity of the fornix.

As a final example, the 3-D vessel crawler (Section 15.4.3) deformable organism included routines for growing or propagating itself along a vessel, fitting or conforming to the vascular boundaries, and spawning new vessel crawlers upon the detection of a branch in the vascular tree.

15.3.5 Cognitive layer

The cognitive layer of the ALife architecture combines memorized information, prior anatomical knowledge, a segmentation plan, and the organism's sensory data (Section 15.3.3) in order to initiate behaviors, carry out shape

deformations, change sensory input, and make decisions toward segmenting and analyzing the target structure (Figure 15.8). A single fixed path in a segmentation plan can be followed, or multiple paths with a plan selection scheme can be implemented. In the first case the organism exhibits a sequential flow of control, proceeding directly from one behavior to the next. In the second case, it decides between different options within the plan, potentially taking different paths given different images or initialization conditions.

Most often, the segmentation plan is subdivided into behavioral stages with subgoals that are easy to define and attain (e.g., locating the upper boundary of an anatomical structure). Consequently, it provides a means for human experts to intuitively incorporate global contextual knowledge. The segmentation plan contains instructions on how best to achieve a correct segmentation by optimally prioritizing behaviors. For example, if it is known that the superior boundary of the CC is consistently and clearly defined in an MR image, then the find-upper-boundary behavior should be given a higher priority as segmenting stable features will provide a good initialization for segmenting nearby less-stable features. Adhering to the segmentation plan and defining it at a behavioral level makes the organism aware of the segmentation process.

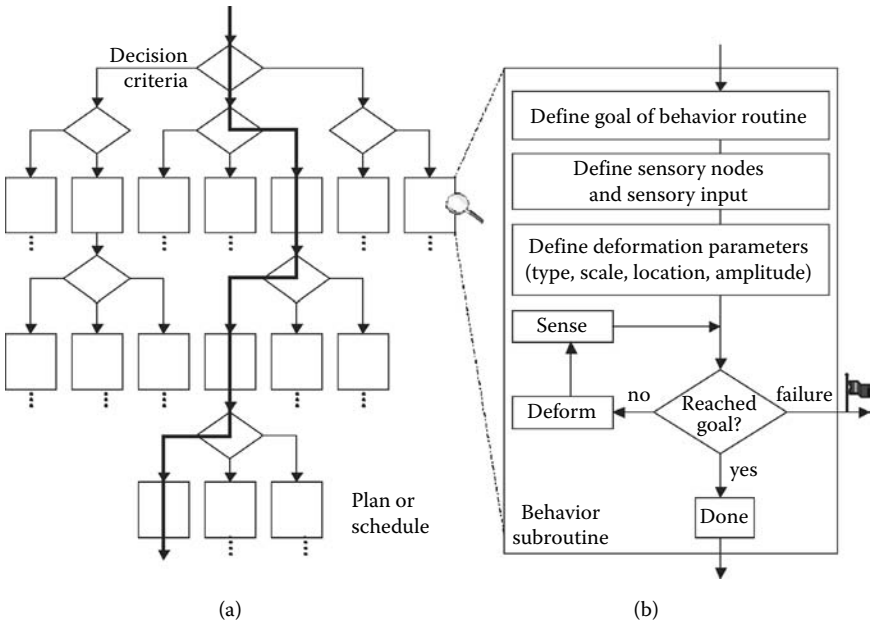


FIGURE 15.8: (a) A procedural representation of a fragment of a deformable organism's plan or schedule. The organism goes through several behavior routines (bold path in a). (b) A simple example of a standard behavior routine. (From McInerney, T., Hamarneh, G., Shenton, M., Terzopoulos, D. 2002. *Medical Image Analysis*, 6(3): 251–266. With permission.)

This enables the organism to make effective use of prior shape knowledge, applying it only in anatomical regions of the target object where a high level of noise or gaps in the objects boundary edge are known to exist.

15.4 Recent Advances

Since its inception in 2001, there have been a number of advances to the original deformable organisms framework. These advances, which are outlined in the following sections, include physical deformation layers, an extension to 3-D, an evolutionary component, and an open-source framework.

15.4.1 Physics-based deformations

A true physical layer was absent in the original work (Hamarneh et al., 2001; McInerney et al., 2002). To complete medical image segmentation tasks, deformable organisms relied on purely geometric shape deformations guided by sensory data, prior structural knowledge, and expert-generated schedules of behaviors. The geometrical deformations made intuitive, real-time user interaction difficult to implement and also necessitated the incorporation of extraneous geometric constraints for maintaining the integrity of the deformable shape model. An early extension to the original deformable organisms framework was to develop and incorporate physics-based shape deformations, which yield additional robustness by enabling intuitive real-time user guidance and interaction when necessary. The physics-based deformations also inherently imposed structural integrity; that is, they provided shape regularization via internal elastic forces.

In Hamarneh and McIntosh (2005), physics-based deformable organisms were presented and applied to the segmentation, labeling, and quantitative analysis of medical image data. External forces are provided by the image gradient and a drag force, while internal forces are supplied through Hookes law and spring damping forces. The results validated that the physics-based formalism allows the expert to provide real-time guidance through small, intuitive interactions, if needed, which further improves the segmentation accuracy and keeps the medical expert "in the loop." Figure 15.9 shows the progression of an example physics-based deformable organism for CC segmentation, Figure 15.10a shows the automatic anatomical labeling, and Figure 15.10b and Figure 15.10c show results before and after some minor user interaction.

15.4.2 Extension to 3-D

A further advancement is the extension of deformable organisms to 3-D for the segmentation and analysis of anatomy in volumetric medical images. To operate in a 3-D environment, deformable organisms need a 3-D geometry that

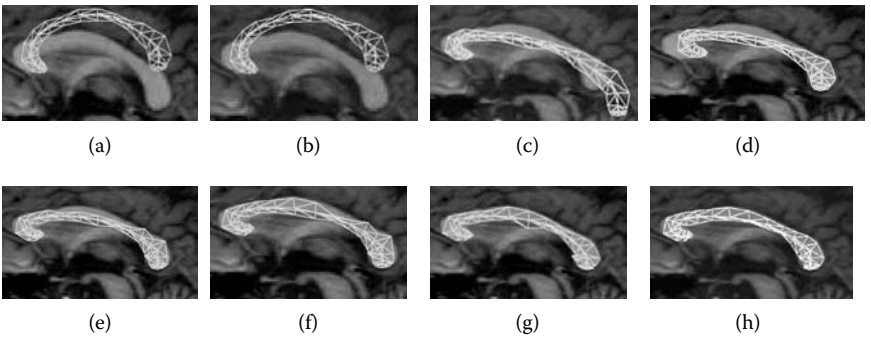


FIGURE 15.9: Progress of segmentation through its primary phases. (a) Global model alignment, (b) model part alignment through (c) expansion and (d) contraction, (e) medial axis alignment, (f) fitting to boundary, (g) detecting and (h) repairing fornix dip. (From Hamarneh, G., McIntosh, C. 2005. In *Proceedings of SPIE Medical Imaging: Image Processing*, 5747, 326–335. With permission.)

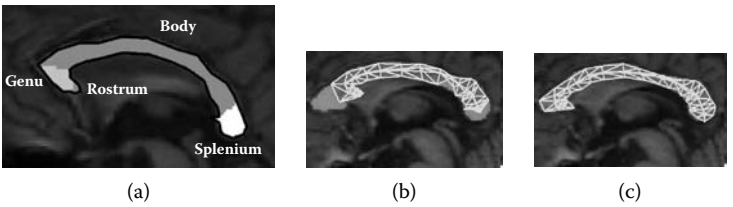


FIGURE 15.10: Physics-based CC organism. (a) Automatic labeling of important anatomical regions of the CC. (b) Before and (c) after intuitive manual intervention to improve the segmentation. (From Hamarneh, G., McIntosh, C. 2005. In *Proceedings of SPIE Medical Imaging: Image Processing*, 5747, 326–335. With permission.)

supports controlled 3-D deformations analogous to the controlled 2-D deformations previously outlined, a 3-D physics layer to carry out those deformations with corresponding behaviors that can operate appropriately in the higher-dimensional environment, 3-D sensors suitable for the 3-D environment, and a cognitive layer that can control the segmentation process. We use medial based shape representation of the organisms 3-D geometry to enable intuitive and localized deformation control, such as medial sheet (Hamarneh et al., 2007) or medial axis based in the case of tubular structures (Sections 15.4.3 and 15.4.4). If physics-based deformations are used (Section 15.4.1) then increasing the dimensionality of the position, force, velocity, and acceleration vectors from 2-D to 3-D enables the calculation of 3-D deformations (Hamarneh and McIntosh, 2005).

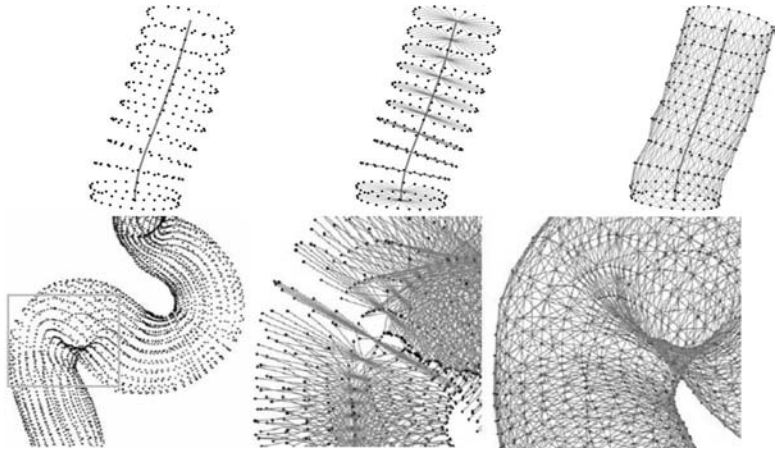


FIGURE 15.11: The geometric and physical layers of tubular shape models. Surface vertices or masses (left) are sampled in a layered fashion along the tubular surface. Radial springs (middle) connect the surface vertices with a medial axis. Stability springs (right) form a triangulated mesh, and act to stabilize the surface model against image noise. (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.). (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.)

The following examples present deformable organisms for the segmentation of tubular structures in volumetric images. The forces are simulated on a tubular geometry parameterized by the distance between neighboring medial masses and the number of circumferential boundary masses (Figure 15.11). This layered medial shape representation enables intuitive deformations (Hamarneh and McInerney, 2003; Hamarneh et al., 2004; O'Donnell, 1994) wherein the medial axis governs the bending and stretching and links to boundary nodes to control thickness. As deformable organisms are typically modeled after their target structures, this tubular branching topology has been employed to the task of vascular and spinal cord segmentation and analysis in volumetric medical images (Sections 15.4.3 and 15.4.4).

15.4.3 Vessel crawlers

In modern internal medicine, noninvasive imaging procedures are often crucial to the diagnosis of cardiovascular, pulmonary, renal, aortic, neurological, and abdominal diseases (Bullitt et al., 2003). Common to the diagnosis of these diseases is the use of volumetric imaging to highlight branching tubular structures such as CT for airways and angiography for vasculature. Without computer assistance, it is difficult to explore, segment, or analyze volumetric images of complex branchings (Figure 15.12).

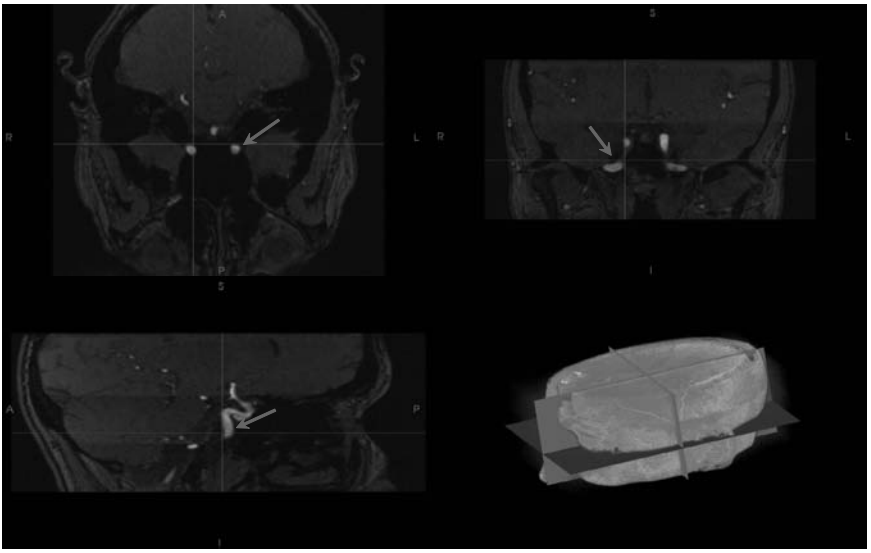


FIGURE 15.12: Slices from a volumetric magnetic resonance angiography image. Plane orientations are shown bottom right. Notice the difficulty in discerning the 3-D structure of the vessels (indicated by the white arrows). Compare this to the segmentations and high-level representations shown in Figures 15.14, 15.15, 15.16, and 15.17. Figure 15.15, for example, gives a quick and intuitive grasp of vessel radius and connectivity.

A new deformable organism for the segmentation and analysis of vasculature from volumetric medical image data was presented in (McIntosh and Hamarneh, 2006d). The bodies of these "vessel crawlers" were modeled as 3-D spring-mass systems (Section 15.4.1). The vessel crawlers use a new repertoire of sensory modules that enable them to detect the centerlines of adjacent vasculature and discern branching points. These sensors are an extension of the arclike sensors in 2-D (Figure 15.5) to a spherical sensor in 3-D (Figure 15.6). They take advantage of optimal vessel filtering, obtained by locally deriving vessel filtering parameters technique (Frangi et al., 1998). Its behavioral routines enable it to grow to the next center-point of a vessel, latch onto the walls of the vessel, and spawn new child vessel crawlers for the exploration of bifurcations in the vascular tree. Finally, a new set of decision-making strategies enables the crawler to decide where to grow to, when to spawn new child vessel crawlers, and when to terminate, all based on its sensory input and current state. Consequently, once initialized within a 3-D image, this new breed of deformable organisms crawls along vasculature, accurately segmenting vessel boundaries, detecting and exploring bifurcations, and providing sophisticated, clinically relevant visualization and structural analysis (Figures 15.13, 15.14, 15.15).

FIGURE 15.13: A 3-D vessel crawler (white arrows) shown as it progressively segments an MRA (shown volume-rendered in grayscale). During each iteration, an estimate of the vessel centerline is obtained from the sensor (Figure 15.6). A new layer is then added with a new medial node at the estimated centerline. Approximately 20 iterations have been performed between each image shown above.

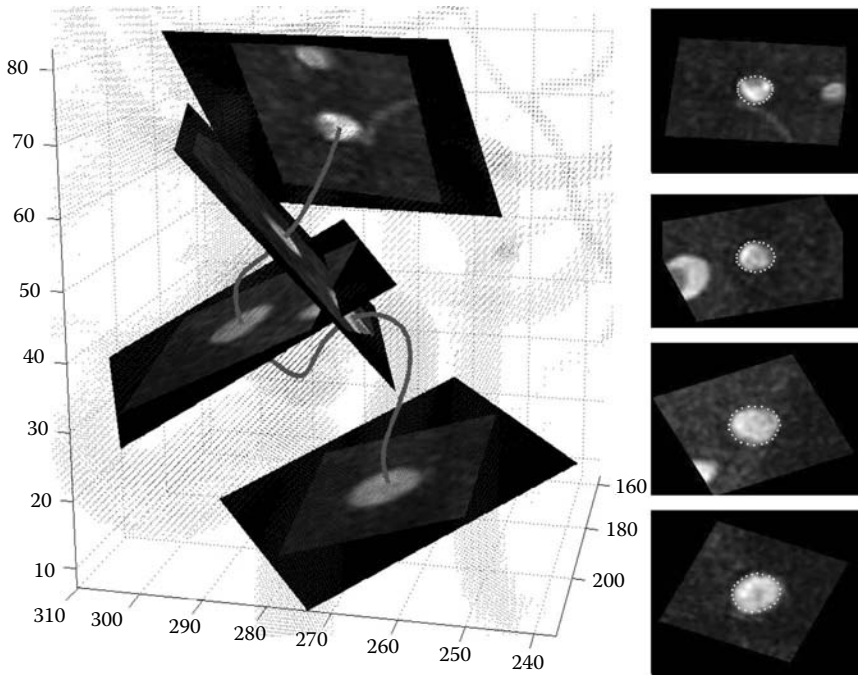


FIGURE 15.14: Multiple cross-sectional slices (left) taken throughout the segmentation process enable the measurement of vessel radii and intensity profiles. Boundary nodes of the vessel crawler after local vessel fitting are shown on extracted slices as white dots (right). (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.)

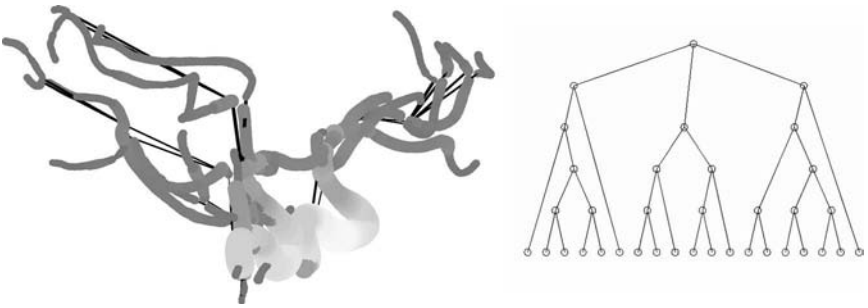


FIGURE 15.15: Intuitive vasculature representations automatically extracted from the MRA image shown in Figure 15.12 (left). Directed acyclic graph (in black) shown in 3-D context overlaid on a plot of the vessels with color corresponding to local radial thickness (right). Tree showing vascular hierarchy displayed out of context. In combination, the two graphs highlight thin vessels (darker gray in the left figure), and give an impression of the overall connectivity (right). Clicking on a node of the tree (right) can also highlight a section of the vasculature (left) along with the set of vessels that feed off of the initial selection. (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.)

The 3-D vessel crawler extracts clinically relevant features as it crawls though the vasculature, including the distance metric, sum of angles metric, and inflection count metric (Bullitt et al., 2003). In addition to these features, the vessel crawlers are able to label branch points and vessel segments, determine branching angles, cross-sectional radius, vessel segment length, and vessel volume, as well as highlight potential problem areas along with the vascular regions they affect, and the shortest path to those locations from any target point (Figure 15.15) (McIntosh and Hamarneh, 2006d).

For example, the vessel crawler was applied to segment and analyze a computed tomography angiogram (CTA) phantom (Figure 15.16) and a magnetic resonance angiogram (MRA) (Figure 15.17). The vessel crawler was initialized using a single seed point at each of the root vessels. As shown, it was able to detect and track all of the connected vessel segments in the phantom, and the vast majority of connected vessel segments in the MRA.

15.4.4 3-D spinal crawler

The spinal cord is a crucial part of the nervous system that resides within the vertebral canal of the spinal column and acts as a relay to convey information between the brain and the rest of the body. There are numerous clinical problems related to the spinal cord, including multiple sclerosis (MS), meningitis, neural tube defects, syringomyelia, transverse myelitis, and spinal

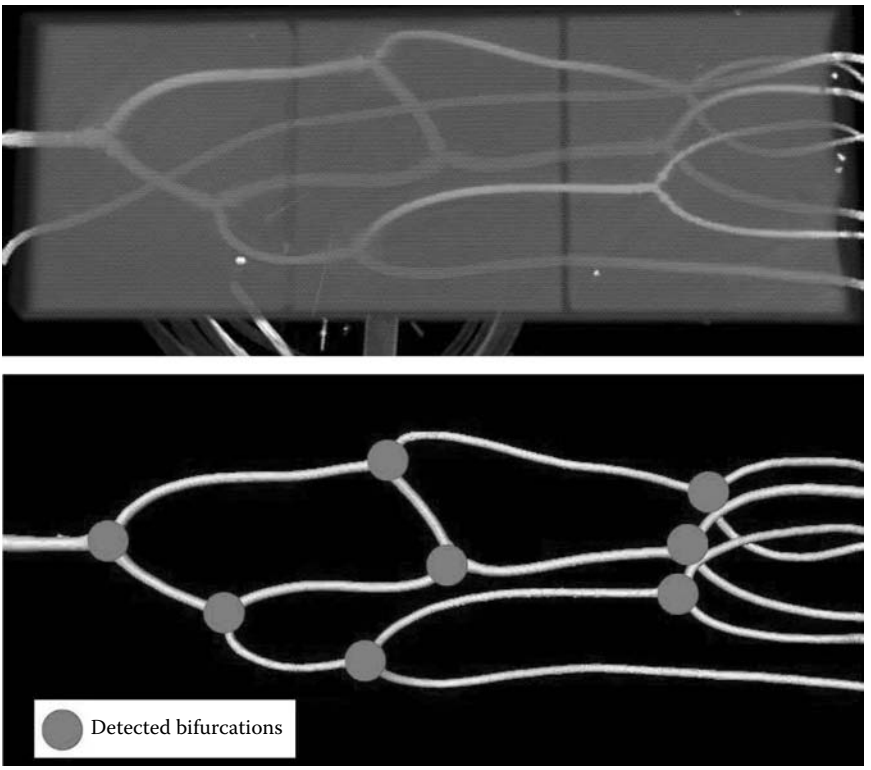


FIGURE 15.16: The figure shows a maximum intensity projection rendering of the phantom CTA (top). We acknowledge Luboz et al. (2005), for providing us with the phantom CTA. Also, it shows the segmentation obtained using vessel crawler (bottom). Notice that the vessel crawler is able to track all the vessels and detect all the bifurcations. (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.)

cord injury (SCI). MS, for example, which affects more than a quarter million people in the United States, is suspected of shrinking the spinal cord.

To statistically analyze the shape of the spinal cord, a segmentation of the cord must be obtained, which is not a trivial task. Though the cord is approximately cylindrical in shape, its diameter varies at different vertebral levels (Figure 15.18). Furthermore, when progressing from superior to inferior, the spinal cord's information transfer requirement decreases, resulting in decreased white matter and, therefore, decreased contrast. This change in both shape and contrast proves difficult for nonadaptive segmentation and filtering methods. To address this problem, we extended and adapted the vessel crawlers to spinal cord segmentation (McIntosh and Hamarneh, 2006c).

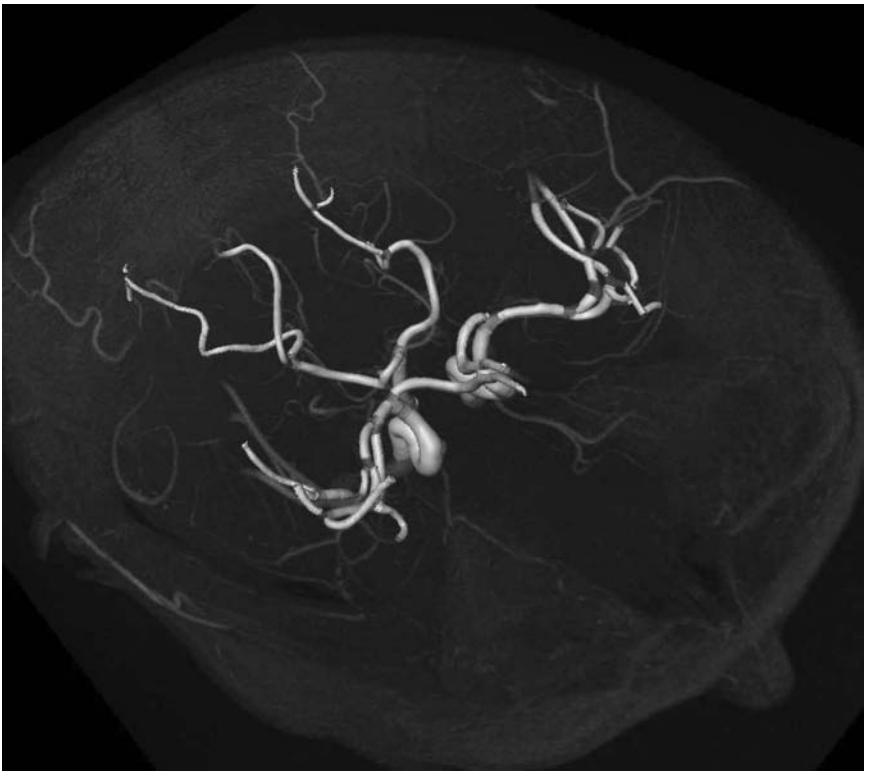


FIGURE 15.17: A volume rendering of a head MRA, viewed superior to inferior (nose appears in the bottom right corner), showing the vessel crawler in shaded white. In this example, the vessel crawler tracks vessel segments with high accuracy, but it fails to detect some of the smaller branch points, resulting in failure to capture the entire tree. (From McIntosh, C., Hamarneh, G. 2006d. *IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1084–1091. With permission.)

The new "spinal crawlers" are tubular spring-mass-based deformable organisms with an adaptive and predominantly elliptical cross section and appropriate anisotropic image data sensors.

By explicitly modeling the elliptical nature of the spinal cord, the spinal crawlers are able to locally detect its shape and intensity characteristics. This information is used to adaptively filter the image, highlighting the spinal cord, and enabling the crawler to perceive the direction in which to grow. The filter extends the single optimal scale in the vessel crawlers with circular cross section (McIntosh and Hamarneh, 2006d) to include two optimal scales, one related to the major axis and the other related to the minor axis of the spinal cord cross section. The result is an anisotropic (elliptical) filter that locally

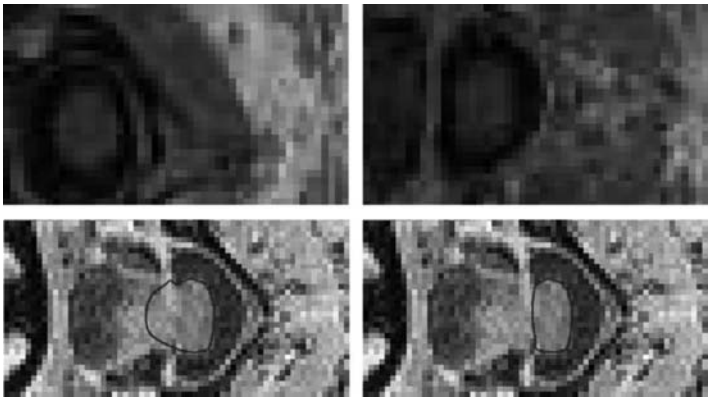


FIGURE 15.18: (See color insert following page 370.) Axial slices of a spinal cord illustrating the difficulty of the segmentation task. Top pair: Circular and elliptical cross-sections of the spinal cord; note the change in shape and lack of contrast. Bottom pair: Extraspinal structures with gray-level intensity higher than that of the spinal cord, which are often encountered in the minor-axis derivatives (horizontal direction). The bottom pair also shows how simple user interaction (Section 15.4.1) on a cross-sectional slice can help repair a false boundary (in red). (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.)

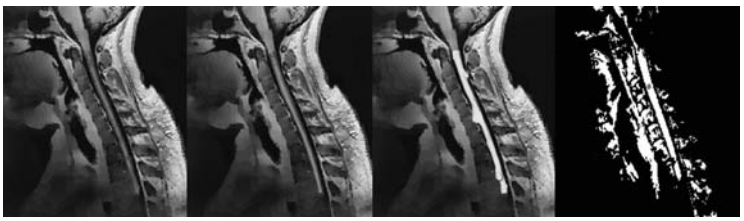


FIGURE 15.19: (See color insert following page 370.) Sample 3-D spinal cord segmentation results (left to right) manual, spinal crawler, SNAP level set, and a region grower with seed point shown in red that has leaked into nearby extraspinal structures. Note that without an adaptive shape prior, SNAP level set (third from the left) frequently leaks outside the spinal cord. (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.)

adapts to the spinal cord's shape (as its cross section progresses from elliptical to circular). The results demonstrate the ability of the method to accurately segment the spinal cord, even in areas where popular techniques such as ITK-SNAP (Yushkevich et al., 2006) and region-growing fail (Figure 15.19).

15.4.5 Evolving deformable organisms

A natural extension to the deformable organisms framework is to incorporate evolution. By using algorithms that model natural selection, the organisms can evolve toward improved segmentations of the image data. Specifically, genetic algorithms (GA), a simple survival-of-the-fittest model of evolution, can be employed to evolve a population of shape models where the fitness of a model is measured by how well it segments the image (Ballerini, 1999; McIntosh and Hamarneh, 2006a; MacEachern and Manku, 1998). Consequently, the organisms that produce the best segmentations will be more likely to survive and reproduce, enabling them to pass their parameters on to the next generation. Therefore, over the course of the algorithm, the shape models evolve into an improved segmentation.

For example, the evolution of a population of medial-axis-based CC segmentation organisms was modeled in McIntosh and Hamarneh (2006a). Shape deformations are caused by mutating individual members of the population (Figure 15.20), as well as through crossover (reproduction) operations. Figure 15.21 shows the beginning, middle, and end of the shape model's evolution for two images. The evolution was constrained, and thus the size of the search space reduced, by using statistically based deformable models represented by medial profiles. The deformations are intuitive (stretch, bulge, bend) and driven in terms of the principal modes of variation of a learned mean shape (Figure 15.22) (Hamarneh et al., 2004). Through the simultaneous evolution of a large number of models, the GAs alleviate the typical deformable model weaknesses pertaining to model initialization and energy functional local minima. The results validate that through simulated evolution, the deformable organisms can evolve into states that better segment the target structure than they could through the use of hand-crafted deformation schedules like those in Hamarneh and McIntosh (2005).

15.4.6 Software framework for deformable organisms

Previously, deformable organisms were restricted to a closed-source MATLAB® framework. Though straightforward and intuitive in their design,

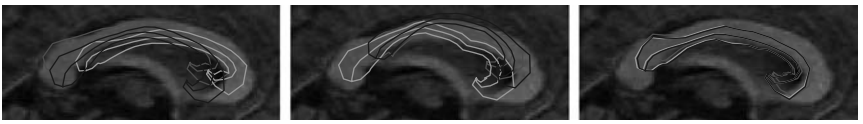


FIGURE 15.20: (See color insert following page 370.) Each image shows evolving deformable organisms whose shape changes result from mutation operators applied sequentially to the medial profiles in order of red, green, blue. Sequential mutations of an increasingly small scale are shown: (a) global scale, (b) medium scale, and (c) small scale.

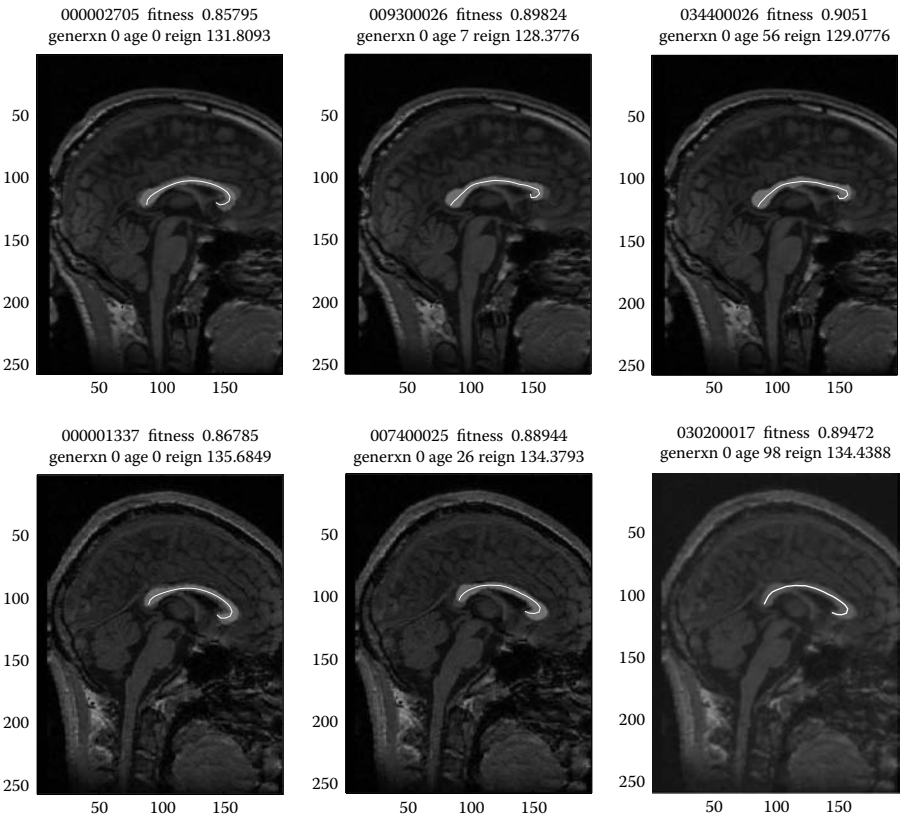


FIGURE 15.21: (See color insert following page 370.) Two example segmentation results progressing left to right, showing the fittest individual after automatic initialization (left), global deformations (middle), and local deformations (right). (From McIntosh, C., Hamarneh, G. 2006a. ACM Workshop on Medical Applications of Genetic and Evolutionary Computation Workshop (MedGEC) in conjunction with the Genetic and Evolutionary Computation Conference (GECCO), 1–8. With permission.)

they were not readily extensible in this form. In McIntosh and Hamarneh (2006b), we reimplemented the deformable organism framework in an open-source manner using the Insight Toolkit (ITK) (www.itk.org), which has a large user base of medical imaging researchers and developers. Furthermore, the incorporation of ITK affords deformable organisms access to faster processing, multithreading, additional image processing functions and libraries, and straightforward compatibility with the powerful visualization capabilities of the Visualization Toolkit (VTK) (www.vtk.org).

Another reason for choosing ITK is because deformable organisms are constructed through the realization of many abstract and independent

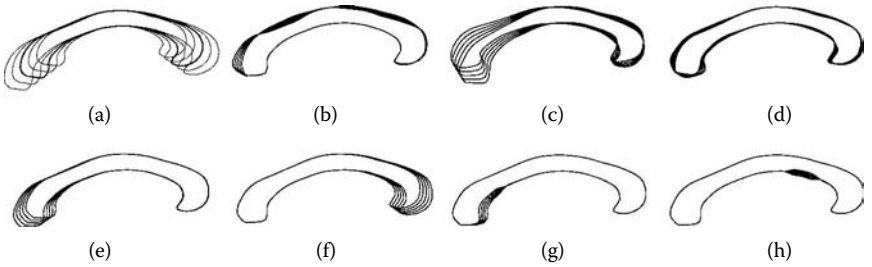


FIGURE 15.22: Statistical shape deformations: (a) global stretch, (b) global upper boundary thickness, (c–d) global bend along two main modes of variation with (c) representing a higher degree of variance, (e–f) localized bend, and (g–h) localized lower boundary thickness for two different scales and locations.

concepts/layers (cognitive, behavioral, physical, geometrical, sensing). Consequently, the deformable organism framework must reflect this modular design by allowing users to replace one implementation (layer) for another. For example, new shape representations should be easily introducible without redesigning existing cognitive layers. To this end, the interface between layers must be clearly defined and consistent across implementations (plug and play).

The software implementation must also be extendable, allowing it to grow and advance as the concept of deformable organisms does. That is, it should support current research into new types of deformable organisms with increasingly advanced decision-making and deformation abilities designed for different applications. These concepts are at the very core of the ITK design methodology.

The ITK Deformable Organisms (I-DO) framework (McIntosh and Hamarneh, 2006b) bridges the ITK framework and coding style with deformable organism design methodologies. It provides a basis for the medical image analysis community to develop new instances of deformable organisms and exchange components (spatial objects, dynamic simulation engines, image sensors, etc.), enabling fast future development of new custom deformable organisms for various clinical applications.

For example, a powerful deformable organism, the itkOrganism, is both a deformable organism and an ITK image filter. Consequently, these organisms can be easily incorporated as autonomous tools into existing ITK filtering pipelines (taking an image as input and producing a segmented image as output). Figure 15.23 demonstrates at a low level the steps required to implement deformable organisms. Although the code excerpts in the figure are specific to the I-DO framework, they represent the general steps we believe need to be taken regardless of the implementation framework or programming language used. The first step is to choose an appropriate shape representation. The

(a)
```
typedef itk::ItkOrganism <ImageType, ImageType, GradientImageType, float, 3> organismType;
organismType::Pointer testOrg = organismType::New();
testOrg->SetInput(reader->GetOutput());
```

(b)
```
typedef Phys_Euler<float,GradientImageType,3> PhysLayerType;
typedef Geom_MeshSpatialObject<float,3> GeometricType;

PhysLayerType::Pointer physLayer = PhysLayerType::New();
GeometricType::Pointer geomLayer = GeometricType::New();
physLayer->setExternalForces((void *) &(output->imageOut));
physLayer->setGeometry(geomLayer);
geomLayer->readTopologyFromFile(topologyInputFileName);
testOrg->setPhysicsLayer(physLayer);
testOrg->setGeometricLayer(geomLayer);
```

(c)
```
Ctrl_ScheduleDriven<float, 3>::Pointer cgL = Ctrl_ScheduleDriven<float, 3>::New();
cgL->setSchedule(scheduleFileName);
testOrg->setCognitiveLayer(cgL);
```

(d)
```
Beh_TranslateAll<float, 3>::Pointer beh1 = Beh_TranslateAll<float,3>::New();
Beh_UniformScale<float, 3>::Pointer beh2 = Beh_UniformScale<float,3>::New();
Def_Translation<float, 3>::Pointer def1 = Def_Translation<float,3>::New();
Def_UniformScale<float, 3>::Pointer def2 = Def_UniformScale<float,3>::New();
testOrg->addBehaviour(beh1);
testOrg->addBehaviour(beh2);
testOrg->addDeformation(def1);
testOrg->addDeformation(def2);
```

(e)
```
testOrg->setRunTime(120);
writer->SetInput(testOrg->GetOutput());
writer->Update();
```

FIGURE 15.23: Example ITK-Deformable Organisms code. Lines are used to (a) initialize the deformable organism, (b) set up and attach the physical and geometrical layers, in this case a spring-mass system, (c) instantiate and attach a cognitive layer, (d) create and attach a set of behaviors and corresponding deformations, and (e) run the deformable organism within an ITK pipeline. (From McIntosh, C., Hamarneh, G., Mori, G. 2007. In *IEEE Workshop on Motion and Video Computing*, 31–38, 2007. With permission.)

issues regarding what constitutes a suitable shape representation for the geometric layer were discussed in Section 15.3.1, most importantly, a shape that provides intuitive deformation handles (code in Figure 15.23b). The second step is to incorporate a method for simulating the deformation dynamics; for example, utilizing mass-spring, physics-based deformations (as discussed in Section 15.3.2; code in Figure 15.23b). The third step is to consult with the domain experts (e.g., clinicians) and identify what constitutes a good strategy for segmentation, thus dictating the sequence of behaviors. As discussed in Section 15.2.5, we recommend progressing from stable to less stable features to improve the robustness of the algorithm (code in Figure 15.23c). The sensors must then be implemented and attached to the deformable organism. The choice of sensors is driven by the goals of the behaviors and the type of features that need to be detected in each behavioral routine (Section 15.3.4). Finally, the different decision-making strategies need to be specified, including

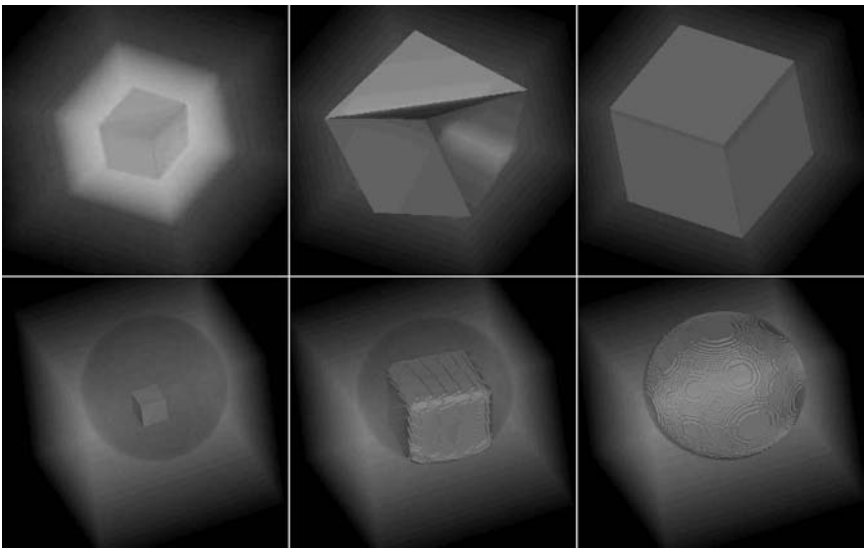


FIGURE 15.24: Two toy example deformable organisms (rendered in gray) progressing from left to right in volume rendered volumetric images. (top) A basic example initialized with a cube, performing a Beh_UniformScale behavior and coming to rest. (bottom) An example initialized with a cube, increasing its size using a Beh_UniformScale behavior simulated on a spring-mass system with the Phys Euler physical layer, then smoothing and coming to rest under image forces using level set–based deformations implemented in the Phys_LevelSet physical layer.

which sensory modules to activate and which behavioral routines to initiate (Section 15.3.5; code in Figure 15.23d).

Other example deformable organisms in the framework can automatically convert between physics-based deformations on spring-mass systems and implicit deformations via the level-sets technique. Two example deformable organisms are shown in Figure 15.24. The bottom row shows a cube is expanded via spring actuation, then automatically converted to a level-set representation and deformed using image gradients.

15.5 Discussion

The original ideas of deformable organisms were presented in Hamarneh, McInerney, and Terzopoulos (2001), Hamarneh (2001), and McInerney et al. (2002). In this original research, several examples of deformable organisms for segmenting and labeling anatomy from medical images were demonstrated.

The classical example is the CC deformable organism, also affectionately referred to as the *CC worm*, which was developed to segment the corpus callosum (CC) in midsagittal brain MR images. The underlying geometry and shape deformations (medial profiles), several of its behavioral routines (e.g., bending, stretching and thickening to crawl and latch to boundaries), and perception capabilities (e.g., edge-detectors) were reused to create interacting organisms (ventricles, putamen, and caudate nucleus) and then 2-D vessel crawlers that detect branches and spawn new child organisms.

As we discussed in the previous section, several extensions have been developed in recent years, from physics-based deformations to 3-D vessel and spinal crawlers to evolutionary organisms to the I-DO ITK-software framework for deformable organisms.

The deformable organisms framework provides several advantages not seen together in any other medical image analysis technique or class of techniques. These advantages include a design that bridges the gap between contextual knowledge and low-level image processing, a layered, modular framework that allows for component reuse and rapid building of new custom medical image analysis applications, a built-in consciousness exemplified by the organisms' awareness of the part of anatomy they are targeting to delineate or label at the current stage of the segmentation plan, the ability to incorporate human input either through user interaction or prestored expert knowledge, and the ability to flag a difficulty or failure should it arise.

15.6 Future Research Directions

Research into the ALife deformable organisms framework is by no means complete. Many improvements can be made in each of the layers, including improved shape representation and deformation capabilities, improved simulation of deformation dynamics, improved design of plans and strategies for medical image analysis, improved incorporation of domain knowledge, and the adoption of improved image processing algorithms used for the organisms' sensory input. Most importantly, all these extensions will be most meaningful if they are applied to practical, clinically motivated medical image analysis applications.

Our existing deformable organisms have been restricted to static 2-D or 3-D images. An important area for future research is dynamic deformable organisms that can analyze motion in time-varying medical imagery. Like conventional deformable models, deformable organisms can be extended to dynamic deformable organisms capable of tracking moving objects in spatiotemporal image data. In McIntosh, Hamarneh, and Mori (2007), we presented a dynamic deformable organism for tracking walking humans from color video (sequences of 2-D color images). The geometry of this "human tracker" deformable organism is a multifigural medial representation, where the torso

and each limb is based on a medial-axis description and the inter-limb connectivity follows the articulated human structure (wrist to elbow, elbow to shoulder, etc.). Using appropriate sensors and behaviors (deformation control sequences), the organisms can search for the person appearing in each frame and automatically delineate his or her boundaries and joint positions. Results are presented in McIntosh and others (2007). Dynamic deformable organisms along these lines remain to be developed for time-varying 2-D and 3-D medical image analysis.

15.7 Conclusions

In this chapter, we reviewed deformable organisms, a novel paradigm for highly automated medical image analysis. The deformable organisms framework is based on deformable models augmented by ALife modeling concepts. This is realized through the knowledge-based, top-down modeling of cognitive, perceptual, and behavioral processes that control and guide the common low-level, bottom-up, data-driven deformable models technique. The result is novel image analysis algorithms that incorporate high-level models of how medical experts approach a medical image analysis task.

Deformable organisms are autonomous agents that can automatically segment, label, and quantitatively analyze anatomical structures of interest in medical images. Rather than relying on expert human interaction to initialize deformable models, set their free parameters, and guide their evolution in order to attain accurate segmentations, deformable organisms incorporate an intelligent, artificial brain that automatically performs these tasks with excellent results. This chapter motivated and overviewed deformable organisms, explained the principal layers of their ALife modeling architecture, reviewed some of the original deformable organism instances, outlined recent extensions and advances, and suggested avenues for future research relevant to this promising new medical image analysis technique.

References

Attali, D., Montanvert, A. 1997. Computing and simplifying 2D and 3D continuous skeletons. *Computer Vision and Image Understanding* 67(3), 261–273.

Ballerini, L. 1999. Genetic snakes for medical images segmentation. In R. Poli, Ed., *Evolutionary Image Analysis, Signal Processing and Telecommunications.* Berlin, Springer, pp. 59–73.

Barr, A. 1984. Global and local deformations of solid primitives. In *Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH84)* 18, New York, 21–30.

Blum, H. 1973. Biological shape and visual science. *Theory of Biology* 38, 205–287.

Bookstein, F. 1997. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press.

Borgefors, G., Nystrom, I., Baja, G. S. D. 1999. Computing skeletons in three dimensions. *Pattern Recognition* 32, 1225–1236.

Bouix, S., Dimitrov, P., Phillips, C., Siddiqi, K. 2000. Physics-based skeletons. In *Proceedings of the Vision Interface Conference*, 23–30.

Bullitt, E., Gerig, G., Aylward, S. R., Joshi, S. C., Smith, K., Ewend, M., Lin, W. 2003. Vascular attributes and malignant brain tumors. In *Proceedings of the Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI03). Lecture Notes in Computer Science* 2878, Springer, Berlin, 671–679.

Burel, G., Henocq, H. 1995. Three-dimensional invariants and their application to object recognition. *Signal Process* 45(1), 1–22.

Cootes, T. F., Beeston, C., Edwards, G., Taylor, C. 1999. A unified framework for atlas matching using active appearance models. In *Proceedings of the Image Processing in Medical Imaging Conference*, Visegrad, Hungary, 322–333.

Cootes, T. F., Cooper, D., Taylor, C. J., Graham, J. 1995. Active shape models: their training and application. *Computer Vision and Image Understanding* 61(1), 38–59.

Cootes, T. F., Edwards, G. J., Taylor, C. J. 2001. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(1), 681–685.

Coquillart, S. 1990. Extended free form deformations: a sculpting tool for 3D geometric modeling. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH90)* 24, New York, 187–196.

Costa, L., Cesar, Jr R. 2000. *Shape Analysis and Classification: Theory and Practice*. Boca Raton, FL: CRC Press.

Davatzikos, C., Tao, X., Shen, D. 2003. Hierarchical active shape models: using the wavelet transform. *IEEE Transactions on Medical Imaging* 22(3), 414–423.

Dimitrov, P., Damon, J., Siddiqi, K. 2003. Flux invariants for shape. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR03)*, Washington, DC, 835–841.

Dryden, I., Mardia, K. 1998. *Statistical Shape Analysis*. New York: Wiley.

Duncan, J., Ayache, N. 2000. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1), 85–106.

Duta, N., Sonka, M., Jain, A. 1999. Learning shape models from examples using automatic shape clustering and Procrustean analysis. In *Proceedings of the 17th International Conference on Information Processing in Medical Imaging (IPMI99). Lecture Notes in Computer Science* 1613, Springer, Berlin, 370–375.

Fletcher, P., Conglin, L., Pizer, S., Joshi, S. 2004. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging* 23(8), 995–1005.

Frangi, A. F., Niessen, W. J., Vincken, K. L., Viergever, M. A. 1998. Multiscale vessel enhancement filtering. *Lecture Notes in Computer Science* 1496, 130–137.

Fritsch, D., Pizer, S., Yu, L., Johnson, V., Chaney, E. 1997. Segmentation of medical image objects using deformable shape loci. In *Proceedings of the 15th International Conference on Information Processing in Medical Imaging (IPMI97). Lecture Notes in Computer Science* 1230, Springer, Berlin, 127–140.

Grenander, U. 1963. *Probabilities on Algebraic Structures*. New York: Wiley.

Hamarneh, G. 2001. Toward Intelligent Deformable Models for Medical Image Analysis. Ph.D. Thesis, Chalmers University of Technology, Göteborg, Sweden.

Hamarneh, G., McInerney, T. 2003. Physics-based shape deformations for medical image analysis. In *Proceedings of the 2nd SPIE Conference on Image Processing: Algorithms and Systems (SPIEIST 03)* 5014, 354–362.

Hamarneh, G., McIntosh, C. 2005. Physics-based deformable organisms for medical image analysis. In *Proceedings of SPIE Medical Imaging: Image Processing* 5747, 326–335.

Hamarneh, G., McInerney, T., Terzopoulos, D. 2001. Deformable organisms for automatic medical image analysis. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*, Utrecht, the Netherlands, 66–75.

Hamarneh, G., Abu-Gharbieh, R., McInerney, T. 2004. Medial profiles for modeling deformation and statistical analysis of shape and their use in medical image segmentation. *International Journal of Shape Modeling* 10(2), 187–209.

Hamarneh, G., Ward, A., and Frank, R. 2007. Quantification and visualization of localized and intuitive shape variability using a novel medial-based shape representation. In *IEEE International Symposium on Biomedical Imaging (ISBI).*

Kass, M., Witkin, A., Terzopoulos, D. 1988. Snakes: active contour models. *International Journal of Computer Vision* 1(4), 321–331.

Lachaud, J., Montanvert, A. 1999. Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction. *Medical Image Analysis* 3(1), 1–21.

Leventon, M., Grimson, W., Faugeras, O. 2000. Statistical shape influence in geodesic active contours. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, Washington, DC, 316–323.

Levy, S. *Artificial Life.* New York: Pantheon, 1992.

Leymarie, F., Levine, M. D. 1992. Simulating the grassfire transform using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(1), 56–75.

Lu, C., Pizer, S., Joshi, S. 2003. A Markov random field approach to multiscale shape analysis. In *Proceedings of the Conference on Scale Space Methods in Computer Vision. Lecture Notes in Computer Science* 2695, Springer, Berlin, 416–431.

Luboz, V., Wu. X., Krissian, K., Westin, C.-F., Kikinis. R., Cotin, S., Dawson. S. A segment and reconstraction technique for 3D vascular structures. In MICCS, 2005, 43–50. Palm Springs, CA.

MacEachern L., Manku T. 1998. Genetic algorithms for active contour optimization. *IEEE Proceedings of the International Symposium on Circuits and Systems* 4, 229–232.

Mandal, C., Vemuri, B., Qin, H. 1998. A new dynamic FEM-based subdivision surface model for shape recovery and tracking in medical images. In *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI98). Lecture Notes in Computer Science* 1496, Springer, Berlin, 753–760.

McInerney, T., Kikinis, R. 1998. An object-based volumetric deformable atlas for the improved localization of neuroanatomy in MR images. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*, Cambridge, MA, 861–869.

McInerney, T., Terzopoulos, D. 1995. A dynamic finite-element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *Computerized Medical Imaging and Graphics* 19(1), 69–83.

McInerney, T., Terzopoulos, D. 1996. Deformable models in medical image analysis: a survey. *Medical Image Analysis* 1(2), 91–108.

McInerney, T., Hamarneh, G., Shenton, M., Terzopoulos, D. 2002. Deformable organisms for automatic medical image analysis. *Medical Image Analysis* 6(3), 251–266.

McIntosh, C., Hamarneh, G. 2006a. Genetic algorithm driven statistically deformed models for medical image segmentation. ACM Workshop on Medical Applications of Genetic and Evolutionary Computation Workshop (MedGEC) in conjunction with the Genetic and Evolutionary Computation Conference (GECCO), 1–8.

McIntosh, C., Hamarneh, G. 2006b. I-DO: A Deformable Organisms Framework for ITK. Medical Image Computing Analysis and Intervention Open Science Workshop 2006, 1–14.

McIntosh, C., Hamarneh, G. 2006c. Spinal crawlers: deformable organisms for spinal cord segmentation and analysis. *Medical Image Computing Analysis and Intervention* (1), 808–815.

McIntosh, C., Hamarneh, G. 2006d. Vessel crawlers: 3D physically-based deformable organisms for vasculature segmentation and analysis. *IEEE Conference on Computer Vision and Pattern Recognition* 1, 1084–1091.

McIntosh, C., Hamarneh, G., Mori, G. 2007. Human limb delineation and joint position recovery using localized boundary models. In *IEEE Workshop on Motion and Video Computing*, 31–38, 2007.

Miller, J., Breen, D., Lorensen, W., O'Bara, R., Wozny, M. J. 1991. Geometrically deformed models: a method for extracting closed geometric models from volume data. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH91)* 25, New York, 217–226.

Montagnat, J., Delingette, H. 1997. Volumetric medical image segmentation using shape constrained deformable models. In *Proceedings of the Second International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed-MRCAS97). Lecture Notes in Computer Science* 1205, Springer, Berlin, 13–22.

Mortenson, M. 1997. *Geometric Modeling*. New York: Wiley.

O'Donnell, T., Boult, T., Fang, X., Gupta, A. 1994. The extruded generalized cylinder: a deformable model for object recovery. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR94)*, Washington, DC, 174–181.

Perona, P., Malik, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(7), 629–639.

Pizer, S., Fritsch, D. 1999. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging* 18(10), 851–865.

Pizer, S., Gerig, G., Joshi, S., Aylward, S. R. 2003. Multiscale medial shape-based analysis of image objects. In *Proceedings of IEEE* 91(10), 1670–1679.

Sebastian, T., Klein, P., Kimia, B. 2001. Recognition of shapes by editing shock graphs. In *Proceedings of the International Conference on Computer Vision (ICCV01)*, Washington, DC, 755–762.

Sederberg, T. W., Parry, S. R. 1986. Free-form deformation of solid geometric models. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH86)* 4, New York, 151–160.

Shen, D., Davatzikos, C. 2000. An adaptive-focus deformable model using statistical and geometric information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 906–913.

Siddiqi, K., Bouix, S., Tannenbaum, A., Zucker, S. 2002. Hamilton-Jacobi skeletons. *International Journal of Computer Vision* 48(3), 215–231.

Singh, K., Fiume, E. 1998. Wires: A geometric deformation technique. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH98)* 99(1), New York, 405–414.

Styner, M., Gerig, G., Lieberman, J., Jones, D., Weinberger, D. 2003. Statistical shape analysis of neuroanatomical structures based on medial models. *Medical Image Analysis* 7(3), 207–220.

Szekely, G., Kelemen, A., Brechbuehler, C. H., Gerig, G. 1996. Segmentation of 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models. *Medical Image Analysis* 1(1), 19–34.

Terzopoulos, D., Metaxas, D. 1991. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(7), 703–714.

Terzopoulos, D., Tu, X., Grzeszczuk, R. 1994. Artificial fishes: autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life* 1(4), 327–351.

Terzopoulos, D. 1999. Artificial life for computer graphics. *Communications of the ACM* 42(8), 32–42.

Tsotsos, J., Mylopoulos, J., Covvey, H. D., Zucker, S. W. 1980. A framework for visual motion understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2(6), 563–573.

Ullman, S. 1984. Visual routines. *Cognition* 18, 97–159.

Weickert, J. 1998. *Anisotropic Diffusion in Image Processing.* ECMI Series. Stuttgart: Teubner.

Yushkevich, P., Joshi, S., Pizer, S. M., Csernansky, J., Wang, L. 2003. Feature selection for shape-based classification of biological objects. In *Proceedings of the 17th International Conference on Information Processing in Medical Imaging (IPMI03). Lecture Notes in Computer Science* 2732, Springer, Berlin, 114–125.

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., Gerig, G. 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage.*

Index

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

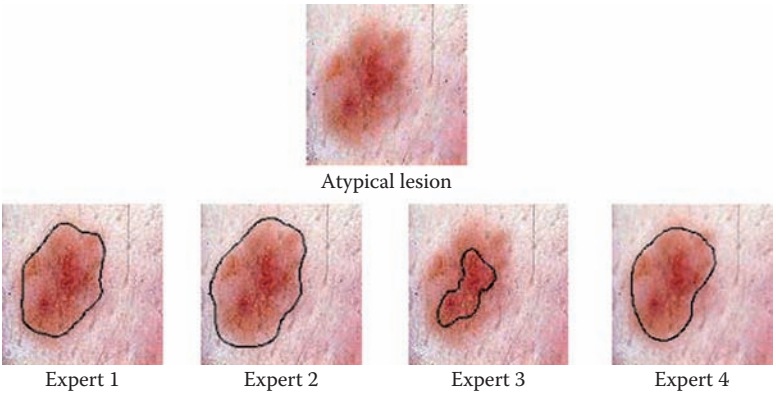


FIGURE 5.1: Segmentation of atypical lesion by four clinical experts. (From Xu et al., 1999. *Image Visions Computing*, 17: 65–74. With permission.)

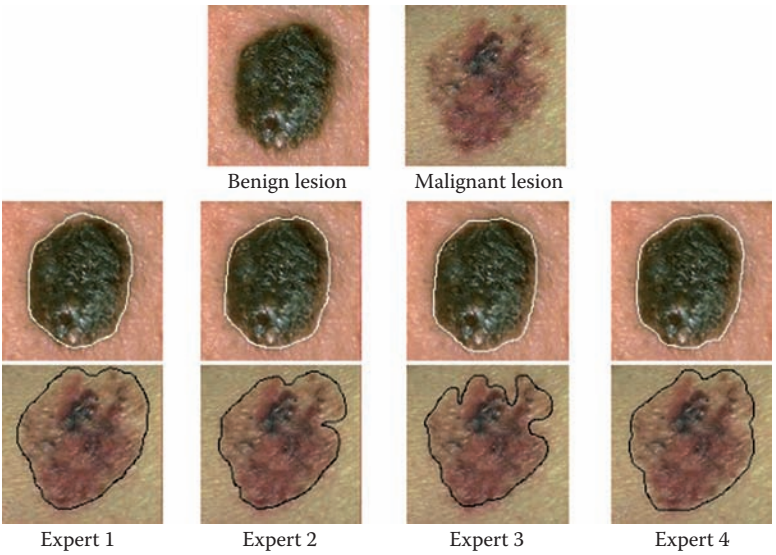


FIGURE 5.2: Segmentation of a benign lesion (white border) and a malignant lesion (black border) by four clinical experts. (From Xu et al., 1999. *Image Visions Computing*, 17: 65–74. With permission.)

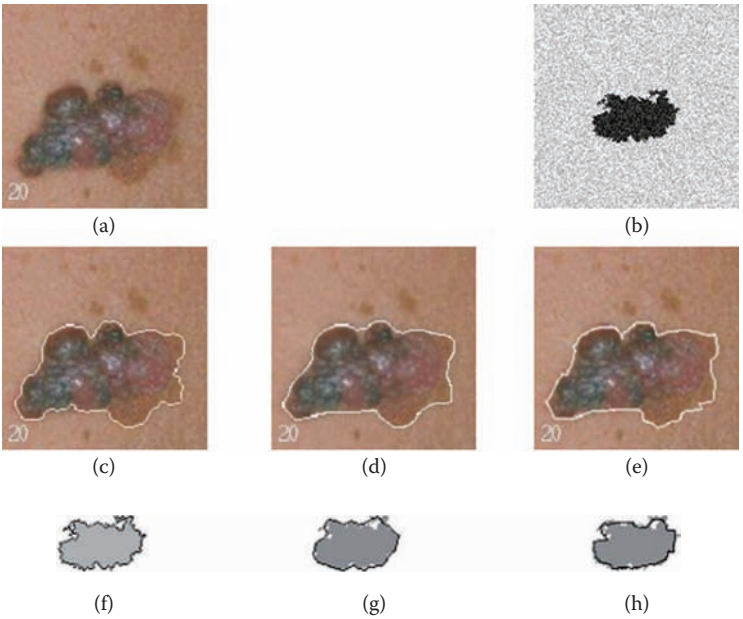


FIGURE 5.10: (a) A real lesion. (b) Irregular noisy synthetic lesion (S/N = 3.0). (c–e) The lesion boundaries produced by IS and ASS, and NNED, respectively when applied to the lesion in (a). Similarly, (f–h) are the output results for IS, ASS, and NNED, respectively, when applied to a synthetic lesion in (b).

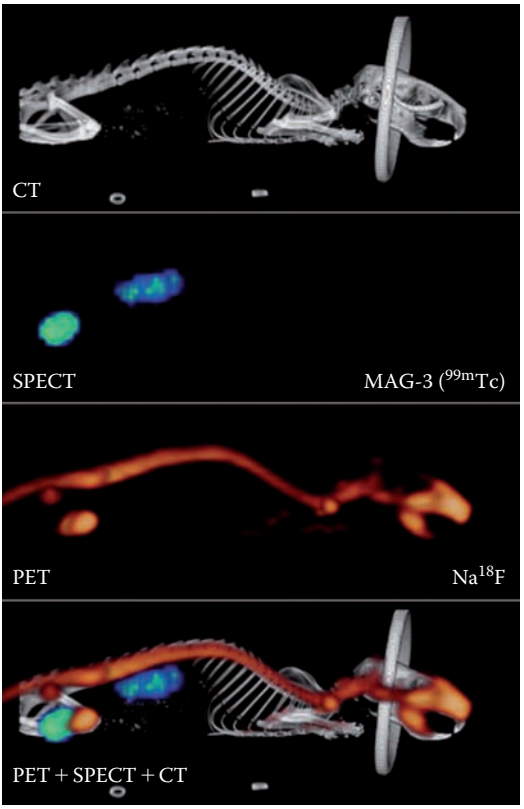


FIGURE 9.3: Multimodality imaging using microCT, SPECT, and PET in an individual live mouse. Initially, renal images were acquired by SPECT using 250 μCi of MAG-3 (^{99m}Tc), followed by CT imaging. The mouse then received 100 μCi of Na^{18}F to visualize bone by PET imaging. Data were collected using the microCAT II + SPECT and microPET P4 scanners. The ring visible in the CT is the nose cone used to deliver 2% isoflurane anesthesia. Each modality was coregistered manually using Amira^{TM} software.

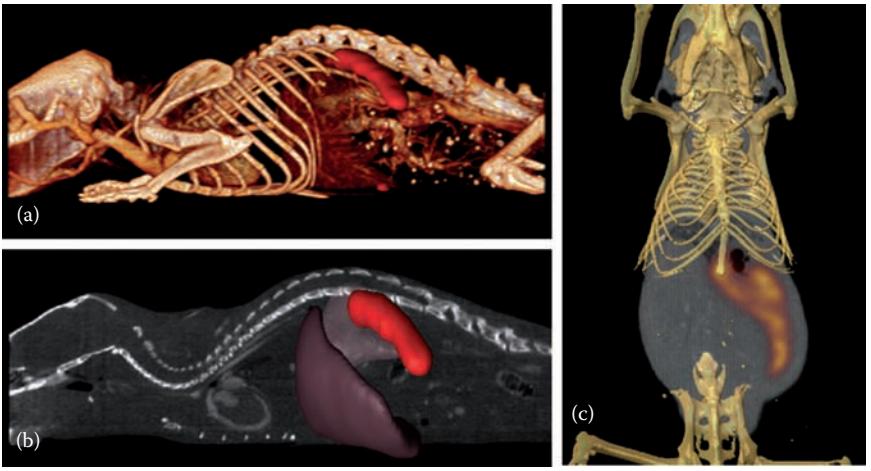


FIGURE 9.5: Volume-rendered, contrast-enhanced SPECT/CT visualization of hepatosplenic amyloid imaged using ^{125}I-serum amyloid P-component. (a) Isosurface representation of the SPECT activity in the spleen with volumetric CT data. (b) Sagittal view of the same mouse with isosurface representation of the segmented spleen and liver volume overlaid. (c) A single coronal plane through fused SPECT/CT image of another mouse with splenic uptake of ^{125}I-serum amyloid P-component in which the skeletal structure (visualization) is overlaid for visual interpretation.

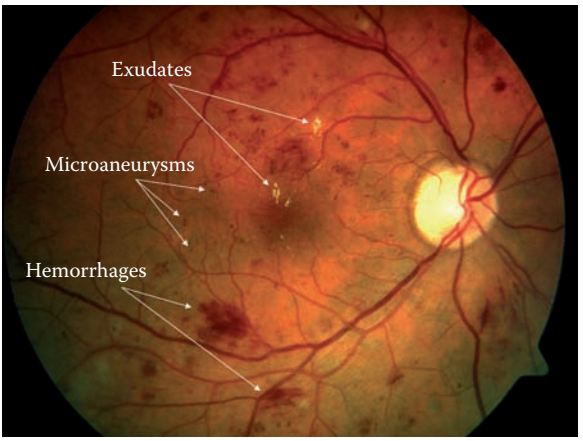


FIGURE 9.6: Color fundus image showing manifestations of nonproliferative diabetic retinopathy including exudates, hemorrhages, and microaneurysms as indicated.

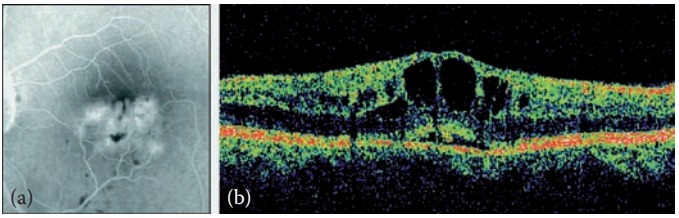


FIGURE 9.8: (a) Fluorescein angiogram showing fluid with cystic changes in the retina. (b) Optical coherence tomography image showing fluid with cystic changes in the same retina. Retinal thickness can be quantified accurately using the OCT cross-sectional method.

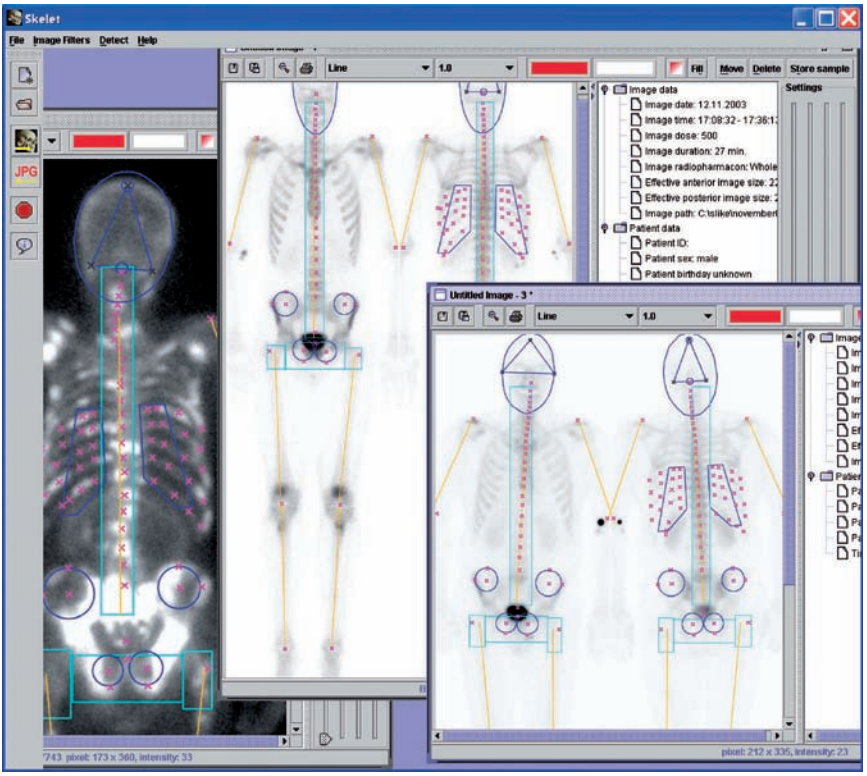


FIGURE 12.8: View of the program skeleton.

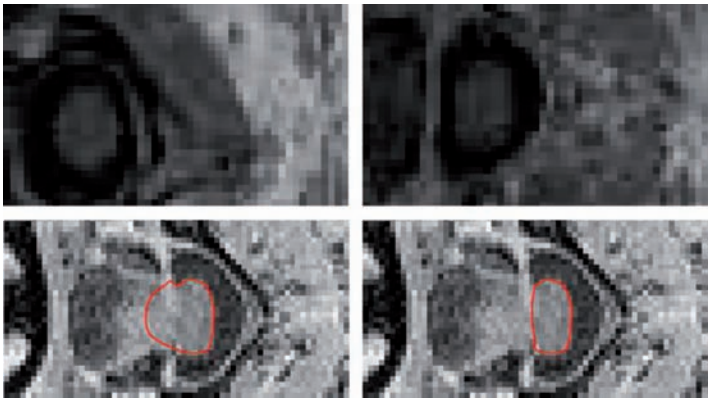


FIGURE 15.18: Axial slices of a spinal cord illustrating the difficulty of the segmentation task. Top pair: Circular and elliptical cross-sections of the spinal cord; note the change in shape and lack of contrast. Bottom pair: Extraspinal structures with gray-level intensity higher than that of the spinal cord, which are often encountered in the minor-axis derivatives (horizontal direction). The bottom pair also shows how simple user interaction (Section 15.4.1) on a cross-sectional slice can help repair a false boundary (in red). (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.)

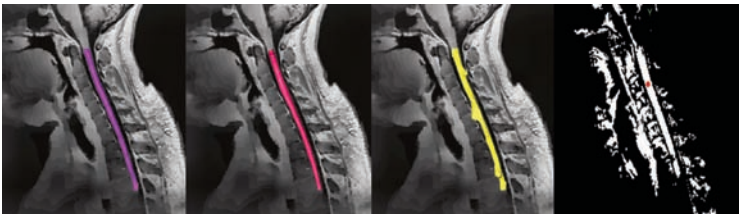


FIGURE 15.19: Sample 3-D spinal cord segmentation results (left to right) manual, spinal crawler, SNAP level set, and a region grower with seed point shown in red that has leaked into nearby extraspinal structures. Note that without an adaptive shape prior, SNAP level set (third from the left) frequently leaks outside the spinal cord. (From McIntosh, C., Hamarneh, G. 2006c. *Medical Image Computing Analysis and Intervention*, 1, 808–815. With permission.)

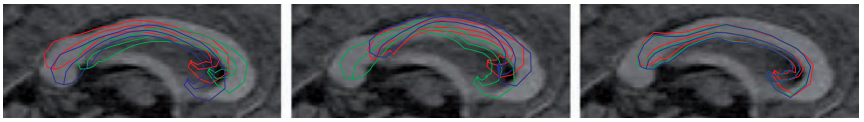


FIGURE 15.20: Each image shows evolving deformable organisms whose shape changes result from mutation operators applied sequentially to the medial profiles in order of red, green, blue. Sequential mutations of an increasingly small scale are shown: (a) global scale, (b) medium scale, and (c) small scale.

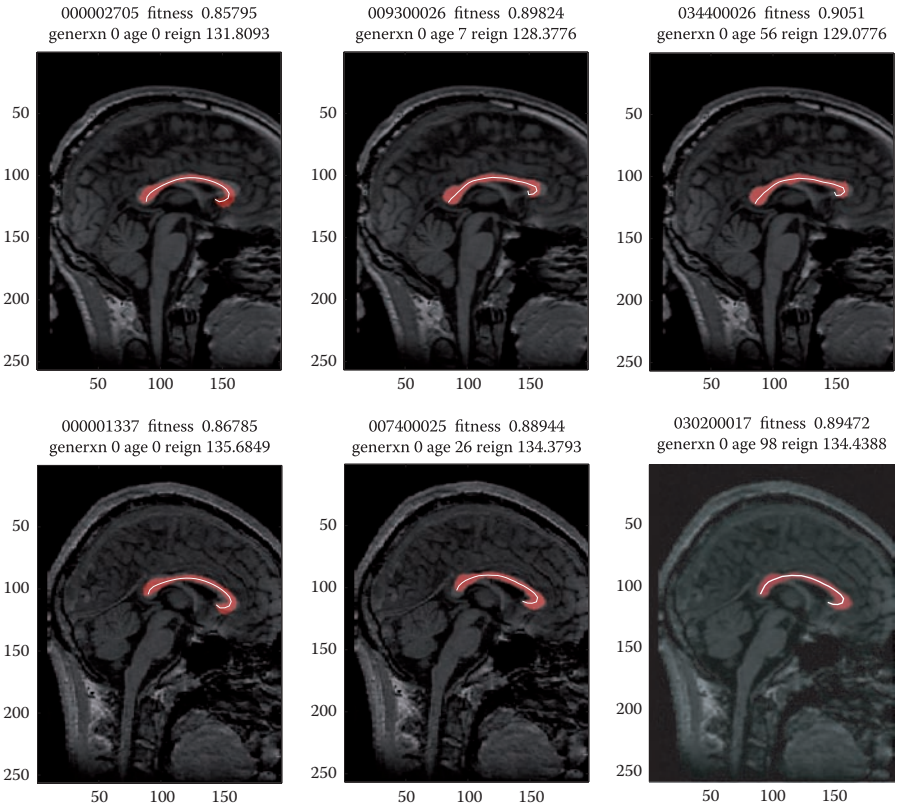


FIGURE 15.21: Two example segmentation results progressing left to right, showing the fittest individual after automatic initialization (left), global deformations (middle), and local deformations (right). (From McIntosh, C., Hamarneh, G. 2006a. ACM Workshop on Medical Applications of Genetic and Evolutionary Computation Workshop (MedGEC) in conjunction with the Genetic and Evolutionary Computation Conference (GECCO), 1–8. With permission.)

Computer Science/Computer Engineering/Computing

A compilation of the latest trends in the field, **Computational Intelligence in Medical Imaging: Techniques and Applications** explores how intelligent computing can bring enormous benefit to existing technology in medical image processing as well as improve medical imaging research. The contributors also cover state-of-the-art research toward integrating medical image processing with artificial intelligence and machine learning approaches.

The book presents numerous techniques, algorithms, and models. It describes neural networks, evolutionary optimization techniques, rough sets, support vector machines, tabu search, fuzzy logic, a Bayesian probabilistic framework, a statistical parts-based appearance model, a reinforcement learning-based multistage image segmentation algorithm, a machine learning approach, Monte Carlo simulations, and intelligent, deformable models. The contributors discuss how these techniques are used to classify wound images, extract the boundaries of skin lesions, analyze prostate cancer, handle the inherent uncertainties in mammographic images, and encapsulate the natural intersubject anatomical variance in medical images. They also examine prostate segmentation in transrectal ultrasound images, automatic segmentation and diagnosis of bone scintigraphy, 3-D medical image segmentation, and the reconstruction of SPECT and PET tomographic images.

Features

- Presents the current state of the art in various areas of computational intelligence
- Reflects forefront research in intelligent medical image processing
- Explores a range of computational algorithms and techniques, such as neural networks, fuzzy sets, and evolutionary optimization
- Encompasses many important applications, including the identification of melanoma and prostate cancer analysis
- Describes a comprehensive system for handling and using biomedical image databases

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

C6059

ISBN: 978-1-4200-6059-1
90000
9 781420 060591