Punam K. Saha · Ujjwal Maulik
Subhadip Basu   *Editors*

Advanced Computational Approaches to Biomedical Engineering

Springer

Advanced Computational Approaches
to Biomedical Engineering

Punam K. Saha • Ujjwal Maulik • Subhadip Basu
Editors

Advanced Computational Approaches to Biomedical Engineering

Springer

Editors

Punam K. Saha
Dept. of Electrical and Computer
Engineering
University of Iowa
Iowa City, IA, USA

Ujjwal Maulik
Subhadip Basu
Dept. of Computer Science and Engineering
Jadavpur University
Kolkata, India

To my late grandmother Manorama Saha
  Punam Kumar Saha

To Sanghamitra and Utsav
  Ujjwal Maulik

To my parents Shyamal Kumar Basu and Rita
Basu, my wife Koel, and my daughter Aishiki
  Subhadip Basu

Preface

Over the last few decades, advancements in biomedical imaging, modeling, and sensing systems, coupled with rapid growth in computational and networking technologies, analytic approaches, visualization and virtual-reality, man–machine interaction, and automation, have led to growing interest in biomedical engineering. While research interest in this field is rapidly emerging among international communities, most leading universities around the globe have already adopted academic curricula on biomedical engineering and related subjects.

In general, biomedical engineering relates to the application of engineering principles to advance the medical and biological sciences. It comprises several topics, including biomedicine, medical imaging, physiological modeling and sensing, instrumentation, real-time systems, microtechniques, automation and control, signal processing, image reconstruction, visualization, processing and analysis, pattern recognition and data mining, and biomechanics.

Biomedical engineering holds great promise for advancing clinical research in medicine, pharmaceuticals, and public health services focusing on prevention, diagnosis, and treatment of complex medical conditions and understanding the etiologies of different biological processes. In recent years, biomedical engineering has taken on a significant role in clinical research, facilitating the development of new drugs, innovative implants and prosthetics, novel imaging technologies, and improved tools for diagnostic and therapeutic mechanisms.

This book brings together contributions from leading research groups presenting recent advances in various active fields of biomedical engineering, including different imaging technologies, image processing and visualization, pattern recognition and data mining, and their clinical applications. Major sources of biomedical imaging data include X-ray computed tomography, magnetic resonance imaging, ultrasound imaging, and positron emission tomography. Often, biomedical data are associated with several significant challenges including limited signal-to-noise ratios and spatial and temporal resolution, misregistration of deformable data in cross-sectional and longitudinal imaging studies, the inherently complex and fuzzy nature of the target objects or information, and large image and data volumes. The overarching goal of most medical imaging research programs is to overcome

these challenges with a focus on accurate and reproducible information gathering, enhanced automation at various levels of imaging and data analysis, and the integration of multidisciplinary expertise toward facilitating diagnostic and therapeutic practices.

The chapters in this book are divided into two parts: Part I includes chapters primarily dedicated to computational approaches to medical imaging, while Part II consists of chapters devoted to task-specific biomedical imaging applications. The methods and algorithms presented in this book include medical image segmentation, information-theoretic clustering, multiobjective differential-evolution-based fuzzy clustering, spectral and nonlinear analysis, and metalearning approaches. The second part of this book covers a number of biomedical applications across various imaging modalities, including 3D ultrasound, fluorescent confocal microscopy, magnetic resonance imaging, and X-ray computed tomography.

Five chapters on advanced computational methods of biomedical engineering are included in the first part of this book. In the chapter "Graph Algorithmic Techniques for Biomedical Image Segmentation," Garvin and Wu present an in-depth overview of two state-of-the-art graph-based methods, viz., graph cuts as well as the layered optimal graph image segmentation of multiple objects and surfaces (LOGISMOS) approach, for segmenting three-dimensional structures in medical images. In each case, an overview of the underlying optimization problem is presented first, followed by the graph-based representation of the problem which enables the globally optimal solution to be found in polynomial time.

In the chapter "Information Theoretic Clustering for Medical Image Segmentation," Hill et al. review a recent approach to clustering under an information-theoretic framework that efficiently finds the suitable number of clusters in a medical image representing different tissue characteristics. The proposed clustering approach optimizes an objective function quantifying the quality of a given cluster configuration. In the chapter "Multiobjective Differential Evolution Based Fuzzy Clustering for MR Brain Image Segmentation," Saha and Maulik present a multi-objective differential-evolution-based fuzzy clustering technique for simultaneous optimization of multiple clustering measures. Here differential evolution is used as the underlying optimization technique and cluster centers are encoded in vectors in differential evolution. The technique is evaluated on several simulated T1-weighted, T2-weighted, and proton density-weighted magnetic resonance (MR) brain images for classifying normal and multiple sclerosis lesions.

In the chapter "Spectral and Non-linear Analysis of Thalamocortical Neural Mass Model Oscillatory Dynamics," Sen-Bhattacharya et al. first present the multimodal analytical techniques for investigating abnormal brain oscillations in Alzheimer's disease. They focus on a combined power spectral and nonlinear behavioral analysis of a neural mass model of the thalamocortical circuitry. They also present the power spectral analysis on a model that implements feedforward and feedback connectivities in a thalamocorticothalamic circuitry.

In the chapter "A Meta-learning Approach for Protein Function Prediction," Plewczynski and Basu present a review of multiscale protein biological function prediction algorithms. They also highlight the advantages of using a metalearning

approach for protein function prediction, particularly related to protein sequence analysis, 3D structure comparison, biological function annotation, and molecular interaction predictions. They include diverse computational methods for predicting the biological function for a given biomolecule using multiscale features and a metalearning prediction system to analyze the impact of microdynamics on the global behavior of selected biological systems.

The second part of the book contains four chapters on important biomedical applications. In the chapter "Segmentation of the Carotid Arteries from 3D Ultrasound Images," Ukwatta and Fenster present an accurate and robust method for segmentation of the carotid arteries from 3D ultrasound images allowing efficient quantification of carotid atherosclerosis. They describe the carotid vessel segmentation algorithm and demonstrate that 3D ultrasound is a viable technique for quantifying the progression and regression of carotid atherosclerosis. In the chapter "Contemporary Problems in Quantitative Image Analysis in Structural Neuronal Plasticity," Ruszczycki et al. study the brain structure at the microscopic level to elucidate the processes of structural neuronal plasticity to understand various neurodegenerative diseases. They used fluorescent confocal microscopy images for quantitative analysis of brain tissues.

In the chapter "Advanced MRI of Cartilage and Subchondral Bone in Osteoarthritis," Chang and Regatte provide an overview of MR imaging techniques to evaluate cartilage and subchondral bone microstructure, and cartilage biochemical composition, in order to provide a noninvasive alternative to detect pathological alterations in these two tissues for early diagnosis of osteoarthritis in at-risk subjects. The derived knowledge and capabilities should ultimately accelerate the discovery and testing of novel therapies to treat osteoarthritis, a disease that represents an enormous socioeconomic and health burden on society.

In the final chapter of the book, Chowdhury et al. address the problem of detection of hairline mandibular fractures from computed tomography (CT) images. It has been observed that such fractures are difficult to detect manually due to the absence of sharp surface and contour discontinuities and the presence of intensity inhomogeneity in the CT images. In the presented work, a two-stage method is adopted to first identify the 2D CT image slices of a mandible with hairline fractures from a fractured craniofacial skeleton. In the second stage, the hairline fracture is detected using a maximum-flow minimum-cut algorithm from the previously identified subset of images.

Overall, the book chapters cover a wide spectrum of biomedical engineering research related to advanced computational methods and their applications. Currently there are not many books available that address these important research discipline in one volume. This edited book, which is unique in its character, will be useful to graduate students and researchers in computer science, biomedical engineering, computational and molecular biology, electrical engineering, system science, and information technology, both as a text and a reference book for some parts of the curriculum. Researchers and practitioners in industrial and research laboratories will also benefit.

Finally, we take this opportunity to thank all the authors for contributing chapters related to their current research work that explain the state of the art in advanced computational approaches to biomedical engineering. Thanks are due to Indrajit Saha and Ayatullah Faruk Mollah who provided technical support in preparing this volume, as well as to our students who incessantly provided us with the necessary academic stimulus to go on. We are also grateful to Ronan Nugent of Springer for his constant support toward this project.

Iowa City, IA, USA   Punam Kumar Saha
September 2013   Ujjwal Maulik
                                                                              Subhadip Basu

Contents

Part I
Advanced Computational Methods

Graph Algorithmic Techniques for Biomedical Image Segmentation

Mona K. Garvin and Xiaodong Wu

Abstract After presenting an introduction to a more traditional graph-based 2D segmentation technique, this chapter presents an in-depth overview of two state-of-the-artgraph-based methods for segmenting three-dimensional structures in medical images: graph cuts and the Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces (LOGISMOS) approach. In each case, an overview of the underlying optimization problem is presented first (i.e., the formulation of an energy/cost function and the specified constraints), followed by the graph-based representation of the optimization problem which enables the globally optimal solution to be found in polynomial time. In particular, in the 2D case, a 2D boundary segmentation optimization problem is transformed into that of finding a minimum-cost path in a graph. In the graph-cuts approach, a 3D object/background labeling problem is transformed into that of finding a minimum s–t cut in a graph, and in the LOGISMOS approach, a single or multiple 3D surface segmentation problem is first transformed into that of finding a minimum-cost closure in a graph (which is further transformed into finding a minimum s–t cut in a graph). For each approach, example applications and extensions are also presented.

M.K. Garvin (✉)
Iowa City VA Health Care System, Iowa City, IA 52246, USA

The University of Iowa, Iowa City, IA 52242, USA
e-mail: mona-garvin@uiowa.edu

X. Wu
The University of Iowa, Iowa City, IA 52242, USA
e-mail: xiaodong-wu@uiowa.edu

1   Introduction

Segmentation problems in medical imaging often require the precise identification of multiple boundaries/regions in volumetric images. Graph-based segmentation approaches are becoming increasingly popular for the automated delineation of medical structures/objects of interest within three-dimensional images, in part, due to the ability of many of the more recent approaches (such as graph cuts [1–6] and the Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces (LOGISMOS) approach [7–11]) to efficiently produce globally optimal three-dimensional segmentations in a single pass (and correspondingly not get stuck in local optima). In addition, graph-based approaches such as LOGISMOS enable the simultaneous optimal detection of multiple surfaces in volumetric images, which is important in many medical image segmentation applications.

In this chapter, after presenting an introduction to a more traditional 2D segmentation technique in Sect. 2 (which is still very useful in its own right for 2D boundary delineation problems), we present an in-depth overview of two state-of-the-art-graph-based methods for segmenting three-dimensional structures in medical images: graph cuts (Sect. 3) and the LOGISMOS approach (Sect. 4). In each case, an overview of the underlying optimization problem is presented first (i.e., the formulation of an energy/cost function and the specified constraints), followed by the graph-based representation of the optimization problem which enables the globally optimal solution to be found in polynomial time. In particular, in the 2D case, a 2D boundary segmentation optimization problem is transformed into that of finding a minimum-cost path in a graph. In the graph-cuts approach, a 3D object/background labeling problem is transformed into that of finding a minimum s–t cut in a graph, and in the LOGISMOS approach, a single or multiple 3D surface segmentation problem is first transformed into that of finding a minimum-cost closure in a graph (which is further transformed into finding a minimum s–t cut in a graph). For each approach, we also present example applications and extensions.

2   2D Boundary Delineation as a Minimum-Cost Path Problem

2.1   The 2D Single Boundary Optimization Problem

For our first example graph-based segmentation approach, we will describe how one can formulate a single 2D boundary segmentation problem as a minimum-cost path problem. While many of the high-level concepts discussed in this section (such as the formulation of an optimization problem with constraints) and the transformation to a well-known graph problem are similar to that discussed in later sections, the graph-based path problem representation in the 2D case is often much more intuitive
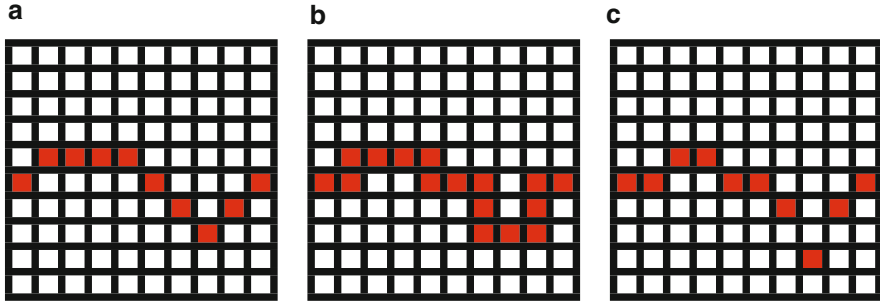
Fig. 1 Three candidate boundaries (in *red*) for the minimum-cost-path-based approach shown on a small image grid. While all three examples seem to contain reasonable boundaries to separate the top and bottom of the image, only (**a**) would be a feasible boundary using a smoothness constraint, Δ, of 1. (**b**) Is not feasible because some of the columns have more than one boundary point defined. (**c**) Is not feasible because the change in boundary position from column 6 to column 7 and from column 7 to column 8 has a magnitude of 2 and the smoothness constraint requires a maximum change of 1 unit. However, (**c**) would be feasible if Δ were defined to be 2

to understand. For simplicity, we will only consider the case in which we desire to find an optimal boundary that is defined from the leftmost column of an image to the rightmost column of an image. In particular, we assume that only one boundary point should exist from each column of the image so that the desired boundary can be written by the function $f(x)$ mapping each column (x-position) to a desired y-value within the image. We also define a smoothness constraint that requires that the boundary's position between neighboring columns does not change more than a given constant: $|f(x+1) - f(x)| \leq \Delta$ for $0 \leq x \leq N_x - 2$, where N_x is the number of columns of the image. Figure 1 provides an example of a feasible boundary and two non-feasible boundaries in a small image when $\Delta = 1$. While such feasibility constraints may initially seem quite limiting, it is important to remember that many common transforms, such as a Cartesian-to-polar transform, may be used to enable the desired boundaries to satisfy these constraints. For example, use of polar coordinates can make the boundaries of roughly circular objects, such as the boundaries of the heart, the boundaries of vessel walls, and the boundaries of the optic disc/cup, satisfy such constraints (Figs. 2b, c and 3).

From the original image $I(x, y)$, we also assume that we can define a cost image, $c(x, y)$, reflecting each pixel's unlikeliness of belonging to the boundary. The cost $E(f)$ of a feasible boundary $f(x)$ can correspondingly be defined as follows:

$$E(f) = \sum_{x=0}^{N_x-1} c(x, f(x)). \tag{1}$$

Thus, the optimization problem we desire to solve is to find the minimum-cost (according to Eq. (1)) boundary $f(x)$ subject to the feasibility constraints.
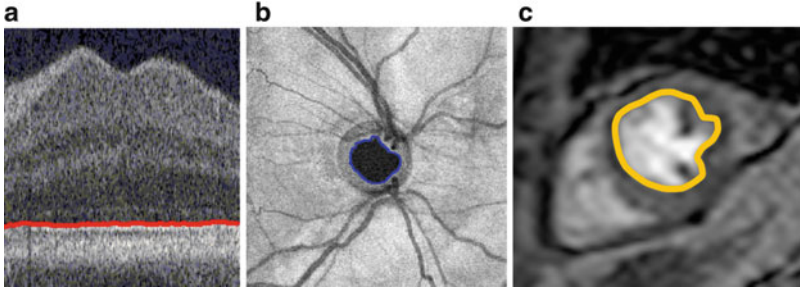
Fig. 2 Example applications for the 2D single boundary optimization problem. (**a**) Single retinal boundary (in *red*) on spectral-domain optical coherence tomography slice (no transformation necessary). (**b**) Optic cup boundary (in *blue*) on spectral-domain optical coherence tomography projection image (Cartesian-to-polar transform necessary). (**c**) Left ventricular boundary (in *yellow*) on magnetic resonance imaging slice (Cartesian-to-polar transform necessary)
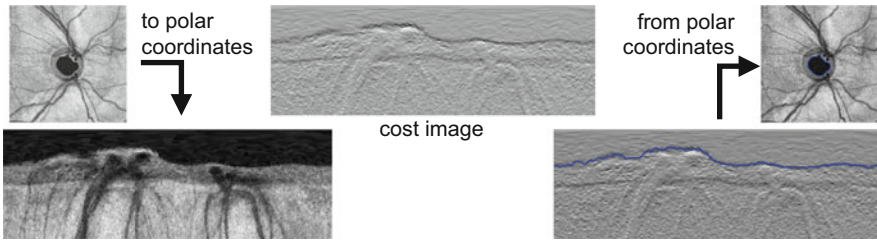
Fig. 3 Example use of Cartesian-to-polar transform to enable the desired boundary to satisfy the feasible constraints required for the 2D single boundary optimization problem (e.g., that one boundary point exists in each column of the image and that the boundary is sufficiently smooth). The boundary is found from a cost image defined in polar coordinates using a minimum-cost path approach, with the final boundary being obtained with a polar-to-Cartesian transform

2.2 Graph Representation of the 2D Single Boundary Optimization Problem

The 2D single boundary optimization problem as defined in Sect. 2.1 can be readily transformed into a minimum-cost path problem within a node-weighted graph (Fig. 4). Each pixel in the cost image, $c(x, y)$, becomes a node with weight given by the intensity value. Because of this direct correspondence between pixels and nodes, we can label each node with its corresponding (x, y) position. Directed edges are added for each "feasible" transition from a boundary point from one column to the next column. In particular, given a smoothness constraint Δ, for each node (x, y), a directed edge is added to nodes $(x+1, y-\Delta), (x+1, y-\Delta+1), \ldots, (x+1, y+\Delta)$ (as long as these destination nodes exist). If a destination node does not exist (as on the boundaries), no directed edge is added. Finding the minimum-cost path in such
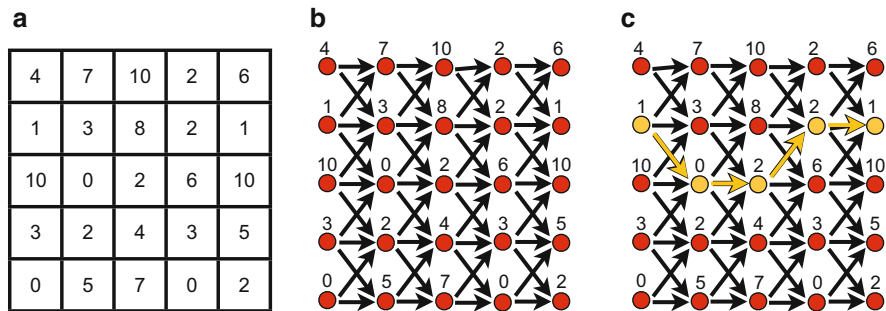
Fig. 4 Example construction of graph for finding minimum-cost boundary according to Eq. (1). (**a**) Small example cost image $c(x, y)$. (**b**) Node-weighted graph representation when the smoothness constraint, Δ, is 1. Edges reflect allowed transitions between boundary points. The minimum-cost path from any node in the *leftmost column* to any node in the *rightmost column* will correspond to the minimum-cost 2D boundary in the image. (**c**) Highlighted (in *yellow*) minimum-cost path in graph

a node-weighted graph directly corresponds to finding the minimum-cost boundary in the original image.

While a number of approaches exist for finding minimum-cost paths in general graphs, the simple structure of this graph (including its acyclic nature) makes a dynamic programming approach a good choice for finding the minimum-cost path. In fact, in implementing this approach, one does not even need to explicitly define the nodes/edges and can simply use an implicit definition of the underlying graph (e.g., given a node position, you can use a simple formula to determine the neighboring nodes). Figure 5 illustrates the basic concepts of a dynamic programming approach. When $\Delta = 1$, you can define the following recursive definition for the cost of the minimum-cost path $p(x, y)$ from any node in the first column ($x = 0$) to node at position (x, y):

$$
\begin{cases}
c(x, y) & x=0 \text{(basecase)} \\
\min[p(x-1, y), p(x-1, y+1)] + c(x, y) & x > 0; y = 0 \\
\min[p(x-1, y-1), p(x-1, y), p(x-1, y+1)] + c(x, y) & x > 0; 0 < y < N_y-1 \\
\min[p(x-1, y-1), p(x-1, y)] + c(x, y) & x > 0; y=N_y - 1
\end{cases}
$$

$$(2)$$

Here, N_y is the number of rows in the image. A similar recursive definition can be defined for other values of Δ. Note that the nodes/edges are defined "implicitly" by directly writing the recursive formula for the path costs.

Because of this recursive definition, the path costs can be efficiently found by computing the path-cost values $p(x, y)$ one column at a time from $x = 0$ to $x = N_x - 1$. Conceptually, in computing each path-cost value, one also keeps track of the predecessor "node" that provided the minimum (although this is not
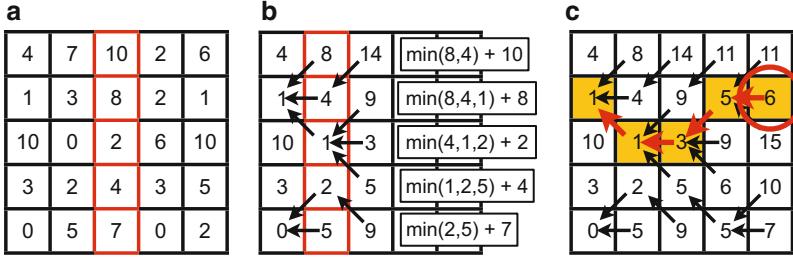
Fig. 5 Use of dynamic programming to find minimum-cost path corresponding to the optimal boundary when $\Delta = 1$. (**a**) Cost image, $c(x, y)$ with highlighted column 2 (note that the leftmost column is column 0). (**b**) Computation of minimum-cost path costs for all of the nodes in column 3. For each node, the minimum-cost path value is computed by adding the cost image value to the path-cost value of the predecessor [for a non-boundary case and $\Delta = 1$, the predecessors of node (x, y) are nodes $(x - 1, y - 1)$, $(x - 1, y)$ and $(x - 1, y + 1)$] with the minimum path-cost value. *Arrows* are shown to indicate each selected predecessor node. (**c**) Highlighted minimum-cost path. The minimum-cost path is found by selecting the node in the last column of $p(x, y)$ with the smallest value and then backtracking (i.e., following the *predecessor arrows*) until reaching a node in the first column

absolutely necessary as determining this node can be done in constant time by simply re-examining all of the values). Finding the actually minimum-cost path involves simply finding the smallest value in the last column and then backtracking through the predecessor nodes until reaching a node in the first column.

If the underlying boundary is roughly circular such that a Cartesian-to-polar transform was used in creating the cost image (meaning that boundaries define one radial value per angle), it is also useful to require that the found boundary be "closed" (i.e., a smoothness constraint exists from the last column to the first column). With a dynamic programming approach, one option for enforcing this is to separately find the minimum-cost path from each of the starting nodes in the first column to the set of nodes in the last column that would obey the constraint by setting the node weights of all other nodes in the first column and the weight of all non-feasible nodes in the last column to ∞. The overall minimum-cost path can then be found as the smallest out of these N_y minimum-cost paths.

3 3D Object Segmentation Using Graph Cuts

3.1 The Graph-Cut Optimization Problem (A Labeling Problem)

In this section, we describe the graph-cut approach of Boykov et al. [1–6] for segmenting a 3D object within a 3D image. While the 2D single boundary segmentation problem discussed in Sect. 2 and the single and multiple surface

segmentation problems to be discussed in Sect. 4 are inherently formulated to focus on the detection of an object boundary or set of object boundaries (with the number of boundaries specified a priori), the optimization problem associated with the graph cut approach is inherently formulated as a labeling problem (i.e., determining the set of pixels/voxels that should be labeled as object). One example corresponding difference is that it is possible that the "object" to be segmented will be disconnected using the standard formulation of the graph-cut approach, whereas the feasibility constraints prevent this in the surface-based methods. (In some applications, preventing such disconnections is desirable, and in other applications, it is not desirable.)

   More specifically, in the graph-cut approach, if $\mathscr{P}$ is the set of voxels in the image, we assume that each voxel $p \in \mathscr{P}$ has a cost $R_p(\text{"obj"})$ and $R_p(\text{"bkg"})$ associated with being labeled as an object voxel or background voxel, respectively. We furthermore assume that each neighboring pair of voxels $\{p, q\} \in \mathscr{N}$, where $\mathscr{N}$ is the set of pairs of neighboring voxels, has a cost $B_{p,q}$ associated with the pair of voxels having different labels. Our goal is to find the optimal labeling $A = (A_1, A_2, \ldots, A_{|\mathscr{P}|})$ to assign each voxel p to a label A_p where A_p is "bkg" or "obj" to minimize the following "energy" or "cost" function:

$$E(A) = \lambda \cdot R(A) + B(A) , \tag{3}$$

where $R(A)$ is the regional term given by

$$R(A) = \sum_{p \in \mathscr{P}} R_p(A_p) , \tag{4}$$

and $B(A)$ is the boundary term given by

$$B(A) = \sum_{\{p,q\} \in \mathscr{N}} B_{p,q} \cdot \delta_{A_p \neq A_q} , \tag{5}$$

where $\delta_{A_p \neq A_q}$ is 0 if the label of A_p is equal to that of A_q and 1 otherwise. λ provides the relative weighting between the regional cost terms and the boundary cost terms.

   Figure 6 provides example regional (i.e., object/background) and boundary costs associated with a small toy image and Fig. 7 provides two example labelings and their associated costs based on Eq. (3).

3.2   Direct Representation of the Graph-Cut Optimization Problem as a Minimum s–t Cut Problem

The graph-cut optimization problem discussed in Sect. 3.1 can be readily formulated as a minimum s–t cut problem in an edge-weighted graph [2, 6]. In an s–t cut
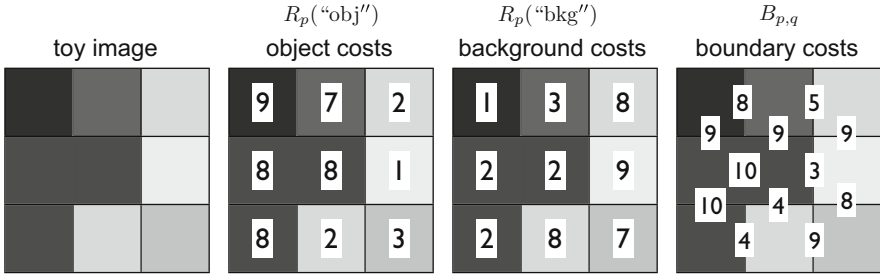
Fig. 6 Example toy image and object (R_p("obj")), background (R_p("bkg")), and boundary costs ($B_{p,q}$) for use in a graph-cuts approach
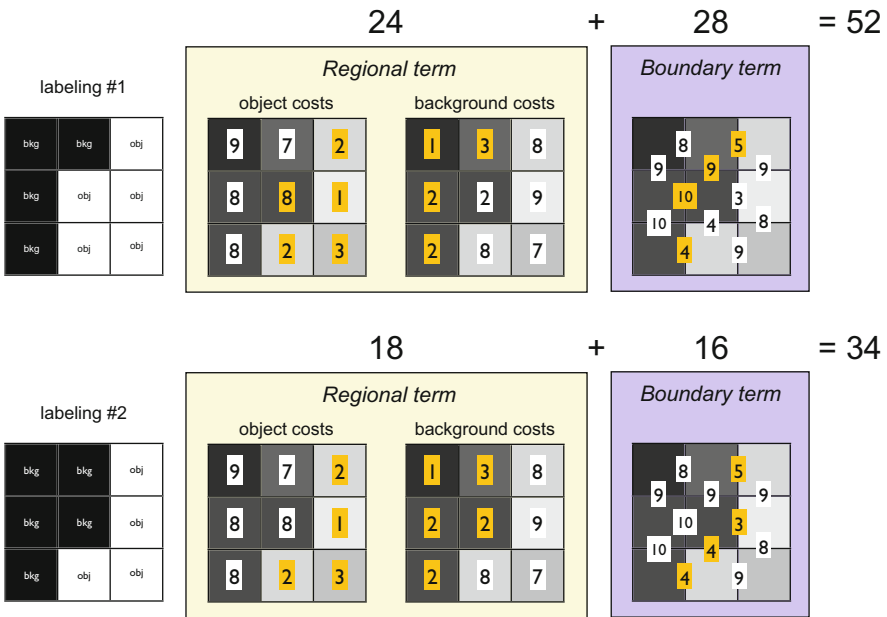
Fig. 7 Two example labelings and associated total costs from toy example from Fig. 6. For simplicity, the weighting λ in Eq. (3) between region and boundary costs is set to 1

formulation, you have an edge-weighted graph with two terminal nodes: s and t. A cut of (S, T) of the graph is a partition of the nodes V into two disjoint sets: S and T ($T = V - S$) with the requirement that $s \in S$ and $t \in T$. The cost of a cut is defined as the summation of the weights of the edges from nodes in S to nodes in T (intuitively the edges that are "cut" in the separation) [12]. The goal in a minimum s–t cut problem is to find the cut with the minimum cost. Once the problem is formulated as a minimum s–t cut problem, many algorithms exist for finding the globally optimal solution in low-order polynomial time [4].
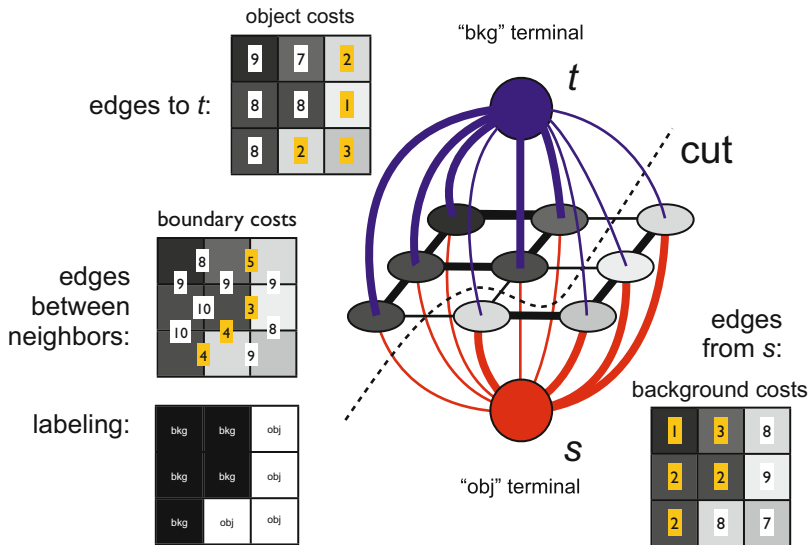
Fig. 8 Example graph representation and minimum cut from regional and boundary costs shown in Figs. 6 and 7. Graph example adapted from examples presented by Boykov et al. [2, 6]

In transforming the graph-cut optimization problem into a minimum s–t cut problem, first, a node p is added for every voxel in the image and edges are added between each pair of nodes corresponding to neighborhood voxels with weight $B_{p,q}$ (these edges between neighbors are often referred to as "n-links" as their weights correspond to the neighbor discontinuity costs). Next, two extra "terminal" nodes, s (the "obj" terminal) and t (the "bkg" terminal), are added to the graph and for each voxel in the image (corresponding to node p in the graph), two edges are added (often called "t-links" for "terminal"): $\{s, p\}$ and $\{p, t\}$. The weight of each edge between "obj" terminal s and node p, $\{s, p\}$, is given by the background cost value for the associated voxel (times a constant): $\lambda \cdot R_p(\text{"bkg"})$. The weight of each edge between node p and "bkg" terminal t, $\{p, t\}$, is given by the object cost value for the associated voxel (times a constant): $\lambda \cdot R_p(\text{"obj"})$. When determining the minimum s–t cut of this constructed graph, the nodes remaining connected to the "obj" terminal s correspond to the voxels belonging to the segmented object (i.e., the voxels that should be given a label "obj"). Figure 8 illustrates the graph and associated minimum cut associated with the toy example shown in Figs. 6 and 7. The associated labeling (segmentation) associated with the cut is also shown. The edges cut include the edges from the nodes in the final object set (with a label "obj") to the background terminal t (with each p in this set contributing $\lambda \cdot R_p(\text{"obj"})$ to the cut cost), the edges from object terminal s to the nodes in the final background set (with label "bkg", with each p in this set contributing $\lambda \cdot R_p(\text{"bkg"})$ to the cut cost) and each of the edges between a node in the object set and a node in the background set (with each such pair of edges $\{p, q\}$ contributing $B_{p,q}$ to the cut cost). Thus,

with the above graph construction, the value of the minimum-cut corresponds to the value of the minimal labeling according to Eq. (3).

Note that the above formulation assumes an undirected edge-weighted graph, but many *s–t* cut algorithms are inherently designed for directed edge-weighted graphs. However, the conversion to the corresponding directed graph is straightforward. In particular, each edge $\{s, p\}$ between terminal s and node p becomes directed with s as the tail and p as the head. Similarly, each edge $\{p, t\}$ between node p and terminal t becomes directed with p as the tail and t as the head. Each edge $\{p, q\}$ between neighbors p and q becomes two edges (each with the same weight): one from node p to node q and one from node q to node p.

When using graph cuts approaches, it is also common to specify a set of "seed" object and background voxels (i.e., voxels you wish to constrain as being labeled as object and background, respectively). Let $\mathscr{O}$ be the set of object seeds and $\mathscr{B}$ be the set of background seeds. You can enforce this constraint in the minimum *s–t* cut graph representation by modifying the edge weight from s to every node in $\mathscr{O}$ to have a weight of infinity (intuitively, ensuring the edges from the object terminal to those nodes that correspond to seed object voxels will not be cut) and modifying the edge weight of every node in $\mathscr{B}$ to t to have a weight of infinity (intuitively, ensuring the edges from nodes that correspond to seed background voxels and the background terminal will not be cut). It is also common to modify the weight of edges $\{s, p\}$ where $p \in \mathscr{B}$ and $\{p, t\}$ where $p \in \mathscr{O}$ to have a weight of zero, but this is not theoretically necessary (assuming the original weights were also not infinity).

3.3 Example Applications and Extensions

3.3.1 Multi-Label (Multi-Region) Extensions

A multi-label extension of the original graph cuts energy function in Eq. (3) follows:

$$E(A) = \lambda \cdot \sum_{p \in \mathscr{P}} R_p(A_p) + \sum_{\{p,q\} \in \mathscr{N}} B_{p,q}(A_p, A_q) \cdot \delta_{A_p \neq A_q} \,, \tag{6}$$

where $A \in \{\text{obj}_1, \text{obj}_2, \ldots, \text{obj}_N\}^{|\mathscr{P}|}$ is a multi-object image labeling. However, unfortunately, solving this particular multi-label problem is NP-hard in the general case [3, 6]. Thus, approximation algorithms, such as the $\alpha\beta$-swap algorithm and the α-expansion algorithm, have been proposed for efficiently obtaining a close-to-optimal solution [3]. At a high level, both the $\alpha\beta$-swap algorithm and the α-expansion algorithm are examples of a local search approach (Fig. 9), where, after assigning an arbitrary initial labeling as the current candidate solution, a best "neighboring" solution is generated and saved as the current candidate solution if it has a lower cost based on Eq. (6). The generation of "neighboring" candidate
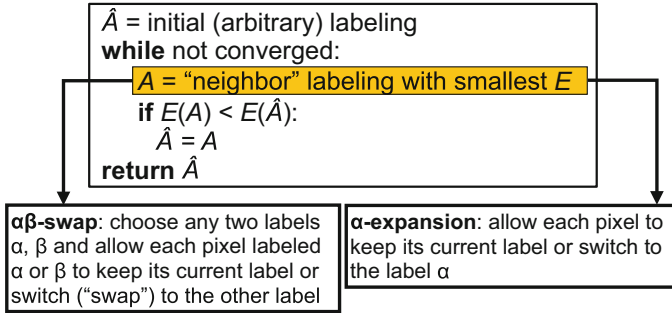
Fig. 9 The $\alpha\beta$-swap algorithm and the α-expansion algorithm are two local search approaches for efficiently determining a close-to-optimal solution to the multi-label graph cuts problem with the energy function of Eq. (6) [3]. In both cases, determination of the best "neighboring" solution can be obtained by solving a single graph cuts problem

solutions (and saving them as the current solution if they have a lower cost) is repeated until convergence. The key clever idea behind both approaches is in the efficient generation of the best "neighboring" solution. In the $\alpha\beta$-swap algorithm, a neighboring candidate solution is one in which, after an arbitrary selection of any two labels α and β, every pixel with label α or β maintains its current label or switches ("swaps") to the other label (Fig. 10). In the α-expansion algorithm, a neighboring candidate solution is one in which after arbitrary selection of label α every pixel maintains its current label or switches to the label α (Fig. 11). With either definition of a neighboring solution, determination of the best neighboring solution can be obtained by solving a single graph cut problem as only a binary labeling choice needs to be made at each voxel location (i.e., α versus β for the $\alpha\beta$-subset of voxels in the $\alpha\beta$-swap algorithm and α versus "original label" for all voxels in the α-expansion algorithm). In practice, the α-expansion algorithm is often more effective than the $\alpha\beta$-swap algorithm, but care must be taken to ensure that submodularity requirements of the cost function are satisfied [3, 13].

While approximation algorithms discussed above may work well for solving the multi-label optimization problem for certain application domains, by recognizing special cases (e.g., through additional geometric constraints) of the multi-label/multi-region labeling problem, certain multi-label/multi-region problems can be actually solved optimally. (Of course, the binary graph cuts formulation using the energy term of Eq. (3) is an obvious example of a special case.) For example, through the addition of specific geometric interaction constraints (in part, motivated from the multi-surface graph-based approach discussed in Sect. 4) such as the fact that "region C contains region A," "region D contains region B," and "region C excludes region D," Delong and Boykov [13] have proposed a formulation of the multi-region graph cuts approach for which an optimal solution can be found with a single graph cut.
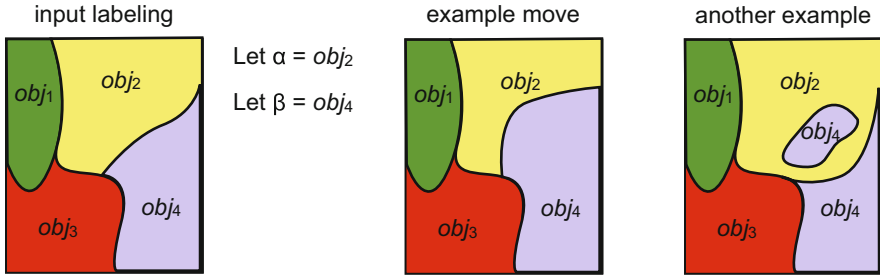
Fig. 10 Two example $\alpha\beta$-swap moves from initial input labeling on the left. In this example, α and β were arbitrarily selected as obj_2 and obj_4, respectively. Valid neighboring solutions are thus labelings where any pixels originally labeled as obj_2 keep their original label or change to label obj_4 and any pixels originally labeled as obj_4 keep their original label or change to label obj_2. (Pixels with labels other than obj_2 and obj_4 must keep their original label.) The binary nature of these moves means that the best move can be determined with a single graph cut
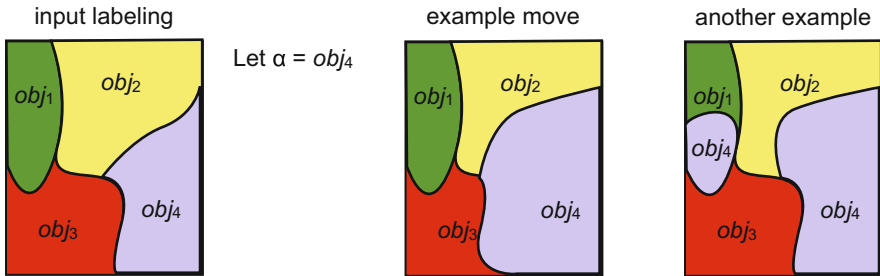
Fig. 11 Two example α-expansion moves from initial input labeling on the left. In this example, α was arbitrarily selected as obj_4. Valid neighboring solutions are thus labelings where any pixels keeps its original label or changes to label obj_4 (thus "expanding" the number of pixels labeled as obj_4). The binary nature of these moves means that the best move can be determined with a single graph cut

3.3.2 Co-segmentation of Tumors in PET–CT Images

Positron emission tomography–computed tomography (PET–CT) has revolutionized modern cancer imaging. The integrated PET–CT, by adding precise anatomic localization to functional imaging, currently provides the most accurate information available on tumor extent and distribution for various common cancers. It has been increasingly playing an indispensable role in cancer treatment planning, therapy response assessment, and tumor staging. This section presents an optimal co-segmentation method for tumor delineation from PET–CT scans with the extension of the graph-cut method [14].

Modeling PET–CT Co-segmentation. The basic idea is to formulate the simultaneous segmentation of tumor from PET–CT scans as minimizing the Markov Random Field (MRF) energy function, which consists of a data fidelity term and

a boundary smoothness term. The data fidelity term measures how well voxels fit into the object or background model. The boundary term penalizes the discontinuity between the object and background. More precisely, the MRF energy functions commonly used for segmentation on the individual PET and CT images (denoted by $\mathscr{I}_{\text{PET}}$ and $\mathscr{I}_{\text{CT}}$, respectively) are as follows:

$$\mathscr{E}_{\text{PET}}(f^P) = \sum_{p \in \mathscr{I}_{\text{PET}}} \mathscr{D}_p^P(f_p^P) + \lambda_1 \sum_{(p,q) \in \mathscr{N}_P} \mathscr{V}_{pq}^P(f_p^P, f_q^P) \tag{7}$$

$$\mathscr{E}_{\text{CT}}(f^C) = \sum_{p \in \mathscr{I}_{\text{CT}}} \mathscr{D}_p^C(f_p^C) + \lambda_2 \sum_{(p,q) \in \mathscr{N}_C} \mathscr{V}_{pq}^C(f_p^C, f_q^C) \tag{8}$$

$\mathscr{D}_p^P(f_p^P)$ is the individual penalty for assigning a voxel p to "object" or "background" in the PET $\mathscr{I}_{\text{PET}}$; $\mathscr{V}_{pq}^P(f_p^P, f_q^P)$ is the penalty for assigning labels f_p^P and f_q^P to two neighboring voxels p and q according to the neighborhood setting $\mathscr{N}_P$. $\mathscr{D}_p^C(f_p^C)$ and $\mathscr{V}_{pq}^C(f_p^C, f_q^C)$ have the same meaning for $\mathscr{I}_{\text{CT}}$. Most close voxels are expected to have the same label ("object" or "background"). Thus, one may expect no penalty if neighboring voxels have the same label and a penalty w_{pq} otherwise. We can then apply the graph cut method in Sect. 3.2 to segment the target tumor from PET and CT, respectively, by minimizing the energy functions $\mathscr{E}_{\text{PET}}(f^P)$ and $\mathscr{E}_{\text{CT}}(f^C)$, which, however, does not utilize the information from the other modality.

To co-segment the tumor from both PET and CT scans, an additional PET–CT context term $\mathscr{E}_{\text{PET–CT}}$ is introduced to the energy function, which penalizes the segmentation difference between two image datasets. Without loss of generality (WLOG), assume that the PET and CT images are well registered. Let (p, p') denote a pair of corresponding voxels in $\mathscr{I}_{\text{PET}}$ and $\mathscr{I}_{\text{CT}}$. The label difference is penalized with $\delta_{pp'}(f_p^P, f_{p'}^C)$ for p and p', with

$$\delta_{pp'}(f_p^P, f_{p'}^C) = \begin{cases} g(p, p') & \text{if } f_p^P \neq f_{p'}^C \\ 0 & \text{if } f_p^P = f_{p'}^C \end{cases}, \tag{9}$$

where $g(p, p') > 0$ is employed to penalize the disagreement between labels of corresponding voxels p and p'. The PET–CT context term then takes the form:

$$\mathscr{E}_{\text{PET--CT}}(f^P, f^C) = \sum_{(p,p') \text{ with } p \in \mathscr{I}_{\text{PET}}, p' \in \mathscr{I}_{\text{CT}}} \delta_{pp'}(f_p^P, f_{p'}^C) \tag{10}$$

Note that the PET–CT context constraint is soft with $\delta_{pp'}(f_p^C, f_{p'}^P) < +\infty$. It is not assumed that the tumor contour in CT is identical to that in PET. Hence, the uncertainties of imaging and registration are accommodated in this model. The energy function of the co-segmentation is then defined as follows:

$$\mathscr{E}_{\text{cs}}(f^P, f^C) = \mathscr{E}_{\text{PET}}(f^P) + \mathscr{E}_{\text{CT}}(f^C) + \mathscr{E}_{\text{PET–CT}}(f^P, f^C) \tag{11}$$
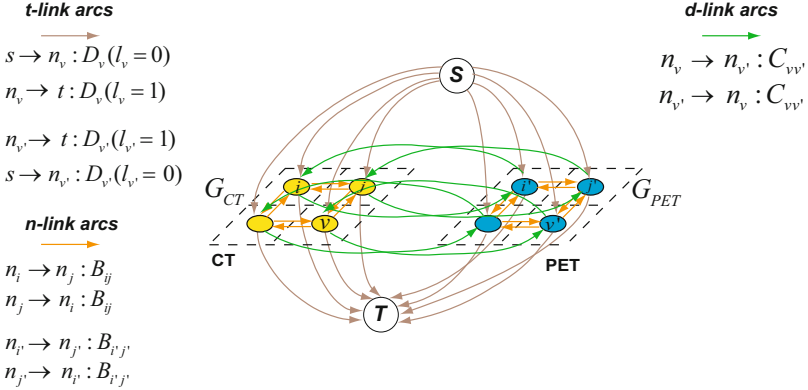
Fig. 12 Graph construction of G with two subgraphs G_{CT} and G_{PET} for the co-segmentation of PET–CT images [14]. Three types of arcs are introduced. The *orange arcs* (*n*-links) encode the boundary terms; the *brown arcs* (*t*-links) encode data terms; and the *green arcs* (*d*-links) enforce the PET–CT context term in the co-segmentation energy function

Most previous co-segmentation algorithms in computer vision [15–18] work on a pair of images with similar (or nearly identical) foregrounds and unrelated backgrounds, and explicitly make use of histogram matching, which makes the models computationally intractable. Batra et al. recently applied the graph cut method for co-segmentation [19], in which a common appearance model is assumed across all the segmented images. However, the PET and CT images may not have such a common appearance model. The co-segmentation scheme here provides a more flexible PET–CT context term $\mathscr{E}_{\text{PET–CT}}$ to make use of the dual modalities information.

Optimization. The co-segmentation problem is solved by extending the graph cut method in Sect. 3.2 with a computation of a minimum-cost s–t cut in a transformed graph G, which admits a globally optimal solution in low-order polynomial time.

The graph G consists of two node-disjoint subgraphs G_{PET} and G_{CT}, each being used for the search of the tumor in the PET $\mathscr{I}_{\text{PET}}$ and the CT $\mathscr{I}_{\text{CT}}$. The construction of each G_{PET} and G_{CT} follows the graph cut method in Sect. 3.2 to encode the energy terms $\mathscr{E}_{\text{PET}}$ and $\mathscr{E}_{\text{CT}}$, respectively. To enforce the PET–CT context term $\mathscr{E}_{\text{PET–CT}}$, additional inter-subgraph arcs are introduced between G_{PET} and G_{CT}. For every pair of corresponding voxels (p, p') with $p \in \mathscr{I}_{\text{CT}}$ and $p' \in \mathscr{I}_{\text{PET}}$, two directed arcs are added between the corresponding nodes of the two subgraphs in opposite directions. The weight of each arc is assigned to penalize the labeling difference of the two voxels p and p' (i.e., $\delta_{pp'}(f_p^P, f_{p'}^C)$ with $f_p^P \neq f_{p'}^C$). Figure 12 illustrates the construction of the graph G, with green arcs encoding the PET–CT context term $\mathscr{E}_{\text{PET–CT}}$.
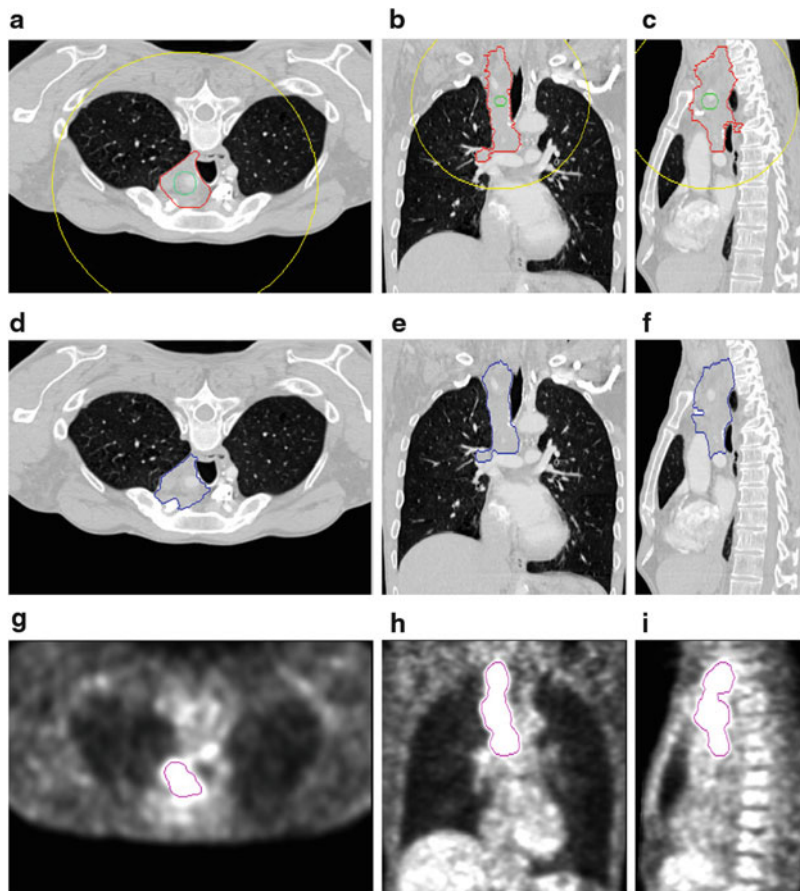
Fig. 13 Typical tumor segmentation in the transverse (*left*), coronal (*middle*), and sagittal (*right*) views [14]. (**a–c**) 2D slices of a 3D CT image with the reference standard (*red*) and outlines of spherical initialization (*green* and *yellow*). (**d–f**) Proposed co-segmentation results in the CT image. (**g–i**) Co-segmentation results in the PET image

As shown in [14], the minimum-cost s–t cut $\mathscr{C}^* = (A^*, \overline{A}^*)$ in G defines an optimal delineation of tumor in both PET and CT images. The target tumor volume on the CT image includes those voxels whose corresponding nodes in G_{CT} belong to the source set A^*. Similarly, the segmented tumor volume on the PET image is given by those voxels whose associated nodes in G_{PET} belong to the source set A^*.

Typical lung tumor segmentation in three (transverse, coronal and sagittal) views are shown in Fig. 13. The co-segmentation method demonstrated accurate PET and CT segmentation results [14].

Fig. 14 Optimal surface
detection. (**a**) An example 3D
retinal optical coherence
tomography (OCT) image
with surfaces of interest
specified. (**b**) The schematic
surfaces and the orientation
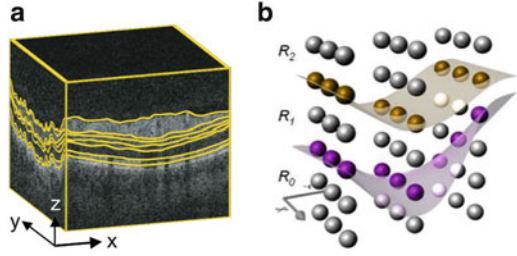
4   Optimal Layered Graph Search for Image Segmentation

This section presents the theory of the optimal layered graph search method[7–
11], also sometimes called the LOGISMOS approach, which is applicable to the
multi-surface segmentation of terrain-like surfaces (represented by orthogonal n-D
graphs), tubular objects (cyclic regular and irregular graphs), objects with complex
shapes, and multiple partially interacting objects. WLOG, we focus on using terrain-
like surfaces and 4-neighborhoods to discuss the optimal layered graph search
method. However, the simpler principles used for this illustration are directly
applicable to arbitrarily irregular surfaces.

4.1   Optimal Surface Detection Problems

Let $\mathscr{I}$ be a given 3D volumetric image with size of $n = X \times Y \times Z$. For each (x, y)
pair, $0 \leq x < X$ and $0 \leq y < Y$, the voxels with different z-coordinates, that is, the
voxel subset $\{\mathscr{I}(x, y, z) \mid 0 \leq z < Z\}$, form a *column* parallel to the **z**-axis, denoted
by $\mathrm{Col}(x, y)$. Two columns are *neighboring* if their (x, y)-coordinates satisfy some
neighborhood conditions. For example, under the 4-neighboring setting, the column
$\mathrm{Col}(x, y)$ is neighboring to $\mathrm{Col}(x', y')$ if $|x - x'| + |y - y'| = 1$. Henceforth,
a model of the 4-neighboring setting is used; this simple model can be easily
extended to other neighboring settings. Each of the target *terrain-like surfaces*
contains one and only one voxel in each column of $\mathscr{I}$ (Fig. 14). The feasibility of
the target surfaces is governed by the surface smoothness and separation constraints.
The surface *smoothness constraint* is specified by two *smoothness parameters*,
$\Delta_{\mathbf{x}}$ and $\Delta_{\mathbf{y}}$, which define the maximum allowed change in the **z**-coordinate of a
surface along each unit distance change in the **x** and **y** dimensions, respectively.
If $\mathscr{I}(x, y, z')$ and $\mathscr{I}(x + 1, y, z'')$ (resp., $\mathscr{I}(x, y + 1, z'')$) are two (neighboring)
voxels on a feasible surface, then $|z' - z''| \leq \Delta_{\mathbf{x}}$ (resp., $|z' - z''| \leq \Delta_{\mathbf{y}}$) (Fig. 15a). In
multiple surface detection, for each pair of the target surfaces S and S', we use two
parameters, $\delta^l \geq 0$ and $\delta^u \geq 0$, to represent the surface *separation constraint*, which
defines the relative positioning and the distance range of the two surfaces. That is,
if $\mathscr{I}(x, y, z) \in S$ and $\mathscr{I}(x, y, z') \in S'$, we have $\delta^l \leq z' - z \leq \delta^u$ for every (x, y)
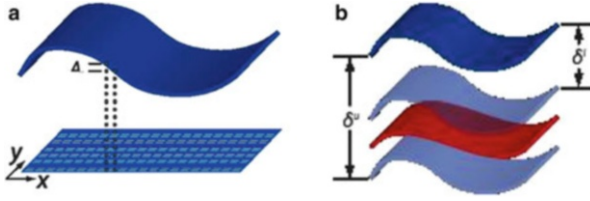
Fig. 15 The surface feasibility. (**a**) The surface smoothness constraint. (**b**) The surface separation constraint
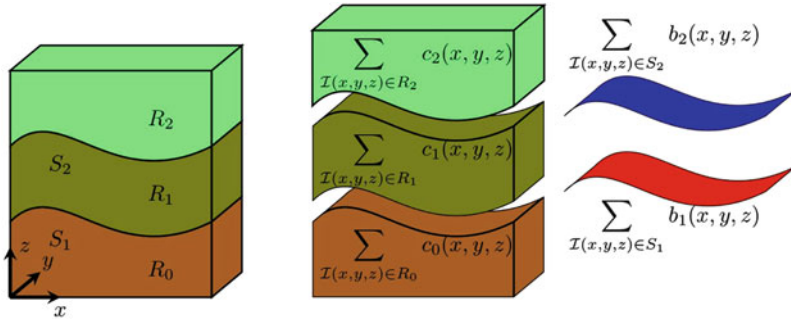
Fig. 16 Example schematic cost of two surfaces for the optimal surface detection problem. The two surfaces divide the volume into three regions

pair (Fig. 15b). A set of λ surfaces $\mathscr{S} = \{S_1, S_2, \ldots, S_\lambda\}$ are considered *feasible* if each individual surface in the set satisfies the given surface-specific smoothness constraints and if each pair of surfaces satisfies the surface separation constraints.

Each voxel $\mathscr{I}(x, y, z)$ may be considered as having an *edge-based* real valued cost $b_i(x, y, z)$ (called an *on-surface cost*) associated with it for each target surface S_i. As shown in Fig. 14b, the λ surfaces form $\lambda + 1$ regions $\{R_0, R_1, \ldots, R_\lambda\}$. For each region R_i $(i = 0, 1, \ldots, \lambda)$, every voxel $\mathscr{I}(x, y, z)$ is assigned a real-valued *region-based cost* $c_i(x, y, z)$ (called an *in-region cost*). The edge-based cost of each voxel in $\mathscr{I}$ is inversely related to the likelihood that it may appear on a desired surface, while the region-based costs $c_i(\cdot)$ $(i = 0, 1, \ldots, \lambda)$ measure the inverse likelihood of a given voxel preserving the expected regional properties (e.g., homogeneity) of the partition $\{R_0, R_1, \ldots, R_\lambda\}$. Then, the total energy $\mathscr{E}(\mathscr{S})$ induced by the λ surfaces in $\mathscr{S}$ is defined as (Fig. 16)

$$\mathscr{E}(\mathscr{S}) = \sum_{i=1}^{\lambda} b_i(S_i) + \sum_{i=0}^{\lambda} c_i(R_i) = \sum_{i=1}^{\lambda} \sum_{\mathscr{I}(x,y,z) \in S_i} b_i(x, y, z) + \sum_{i=0}^{\lambda} \sum_{\mathscr{I}(x,y,z) \in R_i} c_i(x, y, z).$$

$$(12)$$

The *optimal surface detection* (*OSD*) problem on medical images is: Given a 3D image $\mathscr{I}$ and an integer $\lambda > 0$, find a feasible set of λ surfaces $\mathscr{S} = \{S_1, S_2, \ldots, S_\lambda\}$, such that the total cost $\mathscr{E}(\mathscr{S})$ is minimized.

The OSD problem (for which $\lambda = 1$) was first proposed for the cardiac border detection by Thedens et al. [20, 21]. In the past decades, variations of the OSD problem have been studied extensively in theoretical computer science, computer vision, and operations research. The most related problems include the metric labeling problem in theoretical computer science or so-called MRF optimization problem in computer vision, which introduces an additional pairwise penalty energy in the objective function. The metric labeling problem captures a broad range of classification problems where the quality of a labeling depends on the pairwise relations between the underlying set of objects, such as image restoration [3,22], image segmentation [23–26], visual correspondence [27,28], and deformable registration [29]. After introduced by Kleinberg and Tardos [30], it has been studied extensively in theoretical computer science [30–34]. The best known approximation algorithm for the problem is an $O(\log L)$ (L is the number of labels) [30, 33] and has no $\Omega(\sqrt{\log L})$ approximation unless NP has quasi-polynomial time algorithms [32]. Due to the application nature of the problem, researchers in image processing and computer vision have also developed a variety of good heuristics that use classical combinatorial optimization techniques, such as network flow and local search (e.g., [22–24, 27, 35]), for solving some special cases of the metric labeling problem. This section focuses on polynomial-time solutions to the OSD problem.

4.2 Overview of the OSD Algorithms

The basic idea for solving the OSD problem is to formulate it as computing a minimum-cost closed set in a directed graph with arbitrary node weights. A *closed set $\mathscr{C}$* in a directed graph with real-valued node weights is a subset of graph nodes such that there is no arc from a node in $\mathscr{C}$ to a node in the complement of $\mathscr{C}$. The cost of a closed set $\mathscr{C}$ is simply the total sum of the weights of all nodes in $\mathscr{C}$. It is well known that a minimum-cost closed set can be computed in low-order polynomial time by solving a minimum s–t cut problem [25, 36]. The solution to the OSD problem is built upon novel observations that capture the *self-closure structures* of the OSD problem, which relate the target problem to the minimum-cost closed set problem. The OSD algorithm uses the following three major steps:

- *Graph construction.* Construct a node-weighted directed graph $G = (V, E)$, which consists of λ node-disjoint subgraphs $G_i = (V_i, E_i)$. Each subgraph G_i is used for identifying the ith surface in $\mathscr{I}$. The construction hinges on the self-closure structures of the OSD problem. The surface smoothness and separation constraints are enforced with graph arcs; while the optimality of the solution is encoded with node weights.

- *Computing a minimum-cost closed set.* Compute a minimum-cost nonempty closed set $\mathscr{C}^*$ in G, which is solvable by computing a minimum s–t cut in an arc-weighted directed graph transformed from G.
- *Surface recovery.* The set of λ optimal surfaces is recovered from the minimum-cost closed set $\mathscr{C}^*$ with each surface being specified by the envelop of $\mathscr{C}^* \cap V_i$.

4.3 Optimal Single Surface Detection

This section presents the algorithm for solving the optimal single surface detection (OSSD) problem (in which $\lambda = 1$).

To facilitate the discussion of transforming the OSD to a minimum-cost closed set problem, we introduce the concept of bottom-most neighbors and the intra-layer self-closure structure of the OSSD problem. For a voxel $\mathscr{I}(x, y, z)$ and each neighboring column $\mathrm{Col}(x', y')$ of $\mathrm{Col}(x, y)$, the *bottom-most neighbor* of $\mathscr{I}(x, y, z)$ on $\mathrm{Col}(x', y')$ is the voxel on $\mathrm{Col}(x', y')$ with the smallest **z**-coordinate that can appear together with $\mathscr{I}(x, y, z)$ on the same feasible surface in $\mathscr{I}$ (Fig. 17a). Note that the bottom-most neighbor of $\mathscr{I}(x, y, z)$ on $\mathrm{Col}(x + 1, y)$ (resp., $\mathrm{Col}(x, y + 1)$) is the voxel $\mathscr{I}(x+1, y, \max\{z-\Delta_{\mathbf{x}}, 0\})$ (resp., $\mathscr{I}(x, y+1, \max\{z-\Delta_{\mathbf{y}}, 0\})$). For a feasible surface S, denote by $S(x, y)$ the z-coordinate of the voxel $\mathscr{I}(x, y, z)$ on the surface S. We say that a voxel $\mathscr{I}(x, y, z)$ is *below* a surface S if $S(x, y) > z$, and denote by $BL(S)$ all the voxels of $\mathscr{I}$ that are on or below S. A key observation is that *for any feasible surface S in $\mathscr{I}$, the bottom-most neighbors of every voxel in $BL(S)$ are also contained in $BL(S)$* (Fig. 17b). This intra-layer self-closure property relates our target problem to the minimum closed set problem. In our approach, instead of finding an optimal surface S^* directly, we seek in $\mathscr{I}$ a voxel set $BL(S^*)$, which uniquely defines the surface S^*.

4.3.1 Graph Construction

The construction of the directed node-weighted graph $G = (V, E)$ hinges on the intra-layer self-closure structure of the OSSD problem. Every node $v(x, y, z) \in V$ represents one and only one voxel $\mathscr{I}(x, y, z) \in \mathscr{I}$. G can thus be viewed as a geometric graph defined on a 3D grid. Arcs are added to G to make sure that each closed set includes all the nodes associated with the corresponding surface voxels plus all those below the surface. This is done by adding two types of arcs: *intracolumn arcs* and *intercolumn arcs*. The intracolumn arcs ensure that all nodes below a given node (within one column) are also included in the closed set. The intercolumn arcs ensure that the smoothness constraints are satisfied. As an example in Fig. 17c, we consider the added arcs for one node v not involved in boundary conditions to avoid cluttering the exposition of our key ideas. It will be associated with two intracolumn arcs: one directed towards the node below it in the column and one from the node above it (red arcs in Fig. 17c). Two intercolumn arcs will
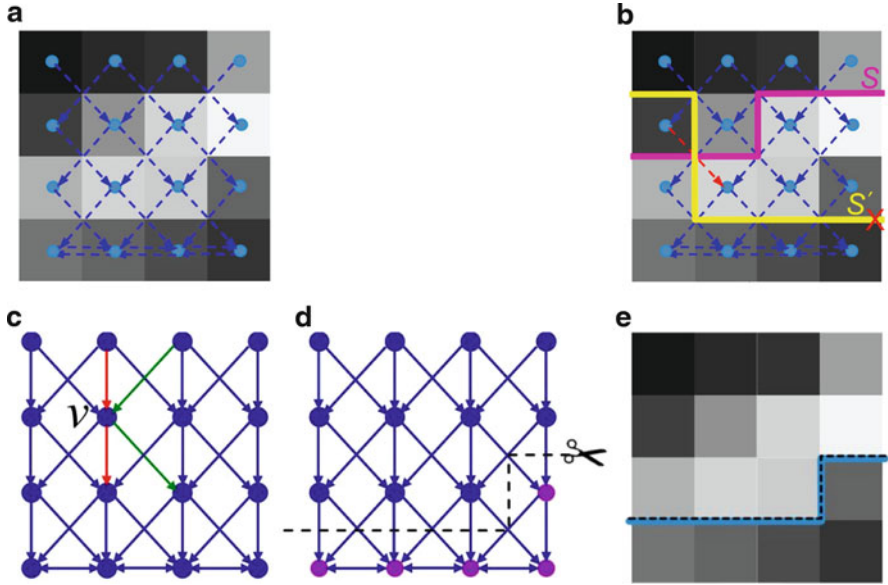
Fig. 17 Modeling single surface detection. An example 2D slice from a 3D image is used. The surface smoothness parameter $\Delta_{\mathbf{x}} = 1$. (**a**) *Dashed line arrows* point to lowest-neighbors of a voxel. (**b**) Illustrating the intra-layer self-closure structure. The surface S is a feasible one while S' is not. The *red line arrow* indicates the violation of the smoothness constraint for S'. (**c**) The constructed graph enforces the surface geometry. The minimum-cost closed set in (**d**) consisting of all *purple nodes* defines the optimal surface in (**e**)

also exist for each neighboring column in the **x**-direction (**y**-direction): one directed to the bottom-most neighbor of v on the neighboring column and one from the node on the neighboring column whose bottom-most neighbor on the column of v is node v (green arcs in Fig. 17c).

To finish the construction of the graph G, we need to encode the energy function Eq. (12) by assigning appropriate weights to the nodes in G. For each node $v(x, y, z) \in V$ corresponding the voxel $\mathscr{I}(x, y, z)$, its weight $w(x, y, z)$ is defined as follows:

$$
w(x, y, z) = \begin{cases} b(x, y, z) + [c_0(x, y, z) - c_1(x, y, z)] & \text{if } z = 0, \\ [b(x, y, z) - b(x, y, z - 1)] + [c_0(x, y, z) - c_1(x, y, z)] & \text{if } z > 0. \end{cases}
$$
(13)

The following theorem is the key for solving the OSSD problem as computing a minimum-cost nonempty closed set in G.

Theorem 1 ([7,8]). *(1) Any finite nonempty closed set $\mathscr{C}$ in G specifies a feasible surface S in $\mathscr{I}$ whose total cost $\mathscr{E}(S)$ differs from the cost of $\mathscr{C}$ by a constant. (2) Any feasible surface S in $\mathscr{I}$ corresponds to a nonempty closed set $\mathscr{C}$ in G whose cost differs from $\mathscr{E}(S)$ by a constant.*

4.3.2 Computing a Minimum-Cost Nonempty Closed Set

It is well known that a minimum-cost closed set in a directed node-weighted graph can be computed by solving a minimum s–t cut problem [25, 36]. However, if there is no "negative" cost closed set, the minimum-cost closed set is empty, which offers little useful information for recovering the target surface. In fact, we need to find a minimum-cost "nonempty" closed set.

Based on the construction of G, the bottom-most nodes of all columns form a closed set $\mathscr{C}_0$, that is, $\mathscr{C}_0 = \{v(x, y, 0) \mid 0 \leq x < X, 0 \leq y < Y\}$. Note that $\mathscr{C}_0$ is a subset of any nonempty closed set in G. If the minimum-cost closed set in G is empty, it indicates that the cost of every nonempty closed set in G is non-negative. We do the following transformation to reduce the cost of each closed set by a constant and to make sure that at least one closed set whose cost is negative after the transformation. Let C be the total cost of nodes in $\mathscr{C}_0$; If $C \geq 0$, choose arbitrarily a node $v(x', y', 0) \in \mathscr{C}_0$ and assign a new weight $w(x', y', 0) - C - 1$ to $v(x', y', 0)$. After this transformation, the total cost of the closed set $\mathscr{C}_0$ is negative. Since $\mathscr{C}_0$ is a subset of any nonempty closed set in G, the cost of any nonempty closed set is reduced by $(C + 1)$. Hence, the minimum-cost nonempty closed set in G is not changed after the transformation and it is not empty. Thus, we can simply apply the algorithm in [25, 36] to find the minimum-cost closed set $\mathscr{C}^*$ in G after performing the transformation. $\mathscr{C}^*$ is then a minimum-cost nonempty closed set in G before the transformation.

4.3.3 Surface Recovery

From a minimum-cost nonempty closed set $\mathscr{C}^*$, an optimal surface S^* in $\mathscr{I}$ can be defined as follows. For each (x, y)-pair, let $\mathscr{C}(x, y) \subset \mathscr{C}$ be the set of nodes on the column $\mathrm{Col}(x, y)$ of G. Due to the construction of G, we have $\mathscr{C}_0 \subset \mathscr{C}^*$. Thus, $\mathscr{C}(x, y)$ is not empty. Let $z_{x,y}$ denote the largest z-coordinate of the nodes in $\mathscr{C}(x, y)$. Then, the optimal surface S^* is defined as $S(x, y) = z_{x,y}$ for every (x, y)-pair. Figure 17e shows an example of an optimal surface defined by a minimum-cost closed set in Fig. 17d.

4.4 Optimal Multiple Surface Detection

This section presents the algorithm for solving the Optimal Multiple Surface Detection (OMSD) problem, i.e., simultaneous detection of $\lambda > 1$ interrelated surfaces in a 3D image $\mathscr{I}$ such that the total cost of the λ surfaces is minimized.

In simultaneously detecting multiple distinct but interrelated surfaces, the optimality is not only determined by the inherent costs and smoothness properties of the individual surfaces, but also confined by their interrelations (i.e., the surface separate constraints). Obviously, computing each of the λ surfaces individually using our

algorithm in Sect. 4.3 does not work well. The solution thus obtained might even be infeasible. In addition, the integration of the region-based costs makes the problem more complicated.

To efficiently solve the OMSD problem, the intrinsic common characteristics of the smoothness and the separation constraints are explored to extend the OSSD technique to multiple surface detection. Assume that $\mathscr{S} = \{S_1, S_2, \ldots, S_\lambda\}$ is a feasible set of λ surfaces in $\mathscr{I}$ with S_{i+1} being "on top" of S_i. For each pair of the sought surfaces, S_i and S_{i+1}, two parameters $\delta_i^l \geq 0$ and $\delta_i^u \geq 0$ are used to specify the surface separation constraint (note that the separation constraint can be defined between any pair of the surfaces in $\mathscr{S}$). First, consider each individual surface $S_i \in \mathscr{S}$. Recall that $BL(S_i)$ denotes the subset of all voxels of $\mathscr{I}$ that are on or below S_i. As in Sect. 4.3, we observe that each $BL(S_i)$ also has the intra-layer self-closure structure. However, the OMSD problem is more involved since the λ surfaces in $\mathscr{S}$ are inter-related. The following observation reveals the common essential structure between the smoothness and the separation constraints, leading to further study the closure structure *between* the $BL(S_i)$ and $BL(S_{i+1})$. One may view the 3D image $\mathscr{I}$ as a set of X 2D slices embedded in the **yz**-plane. Thus, a feasible surface S of $\mathscr{I}$ is participated into X **z**-monotone curves with each in a 2D slice. Observe that each feasible **z**-monotone curve is subjective to the smoothness constraint in the corresponding slice, and any pair of adjacent **z**-monotone curves expresses to meet the analogical separation constraints. This observation suggests that the surface separation constraint in a d-D image may be viewed as the surface smoothness constraint in the $(d + 1)$-D image consisting of the stack of a sequence of λ d-D images. Hence, we intend to map the detection of λ optimal surfaces in d-D to the problem of finding a single optimal surface in $(d + 1)$-D. To distinguish the self-closure structures, define below the *upstream* and *downstream* voxels of any voxel $\mathscr{I}(x, y, z)$ in $\mathscr{I}$ for the given surface separation parameters δ_i^l and δ_i^u: if $\mathscr{I}(x, y, z) \in S_i$, then the ith upstream (resp., downstream) voxel of $\mathscr{I}(x, y, z)$ is the voxel on column $\mathrm{Col}(x, y)$ with the smallest **z**-coordinate that can be on S_{i+1} (resp., S_{i-1}). Together with the intra-layer self-closure structures, the following *inter-layer self-closure* structure is the key for solving the OMSD problem: *Given any set $\mathscr{S}$ of λ feasible surfaces in $\mathscr{I}$, the ith upstream (resp., downstream) voxel of each voxel in $BL(S_i)$ is in $BL(S_{i+1})$ (resp., $BL(S_{i-1})$), for every $1 \leq i < \lambda$ (resp., $1 < i \leq \lambda$).* Both the intra-layer and inter-layer self-closure structures together bridge the OMSD problem and the minimum closed set problem.

4.4.1 Graph Construction

The graph $G = (V, E)$ constructed for solving the OMSD problem consists of λ node-disjoint subgraphs $\{G_i = (V_i, E_i) : i = 1, 2, \ldots, \lambda\}$. Each G_i is constructed as in Sect. 4.3 to reflect the intra-layer self-closure structure of each surface and is used for searching the ith surface S_i. Every node $v_i(x, y, z) \in V_i$ represents one and only one voxel $\mathscr{I}(x, y, z)$. The intracolumn and intercolumn arcs are added in E_i to reflect the intra-layer self-closure structure of surface S_i, which enforce the
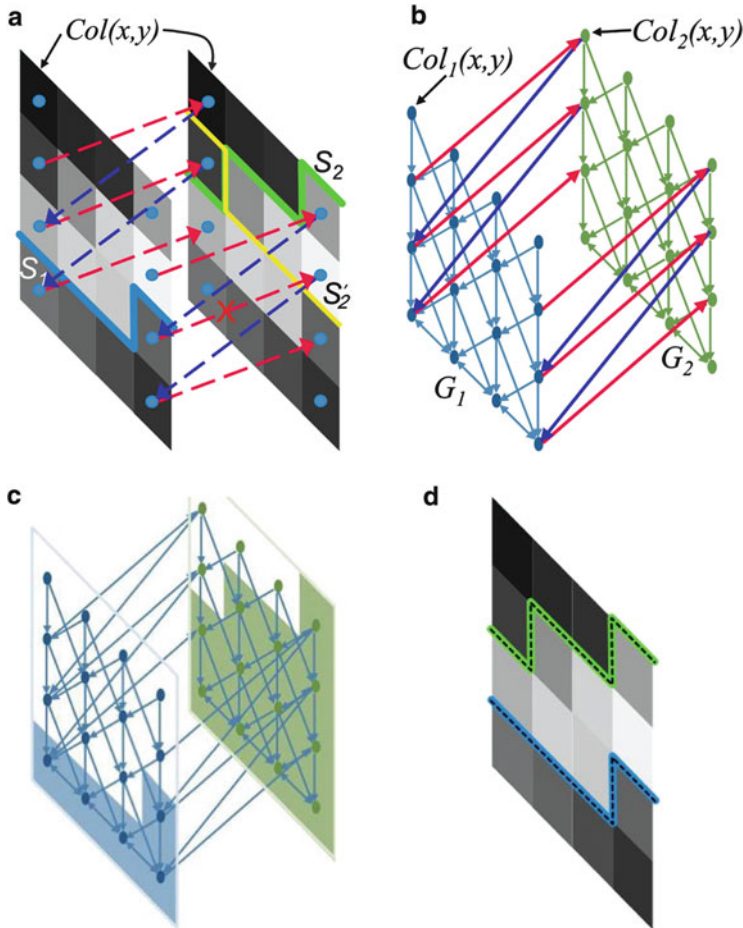
Fig. 18 Modeling multiple surface detection. An example 2D slice from a 3D image is used. Two surfaces are sought and the distance in between ranges from $\delta^l = 1$ to $\delta^u = 2$. The surface smoothness parameter $\Delta_x = 1$. (**a**) Illustrating the inter-layer self-closure structure. For the visualization purpose, the slice is duplicated and the sought two surfaces are visualized in the two separated slices. The *red dashed line arrows* point to the upstream voxels and the *blue dashed line arrows* point to the downstream voxels. S_1 and S_2 are two feasible surfaces, but S_1 and S_2' are not. The *red line arrow* with a mark "X" indicates the violation of the surface separation constraint. (**b**) The constructed graph to enforce the surface separation constraints (arcs are only shown for the first and the last columns). The envelopes of the minimum-cost closed set in (**c**) consisting of all shaded nodes defines the optimal surfaces in (**d**)

monotonicity and the smoothness constraint of each sought surface. The separation constraints between any two surfaces S_i and S_j are enforced in G by a set of arcs E_s, connecting the corresponding subgraphs G_i and G_j, in a way to reflect the inter-layer self-closure structure. Suppose for the two sought surfaces S_i and S_j, the prior knowledge puts S_i below S_j (Fig. 18a in which $i = 1$ and $j = 2$).

Fig. 19 Schematic showing how the assignment of node weights encodes the total in-region cost of the objective function $\mathcal{E}(\mathcal{S})$

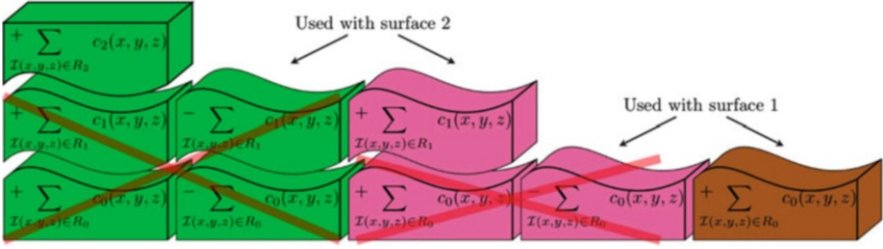Assume that the distance between S_i and S_j ranges from $\delta_{ij}^l = 1$ to $\delta_{ij}^u = 2$. For convenience, denote by $\text{Col}_i(x, y)$ ($\text{Col}_j(x, y)$) the column of nodes in G_i (G_j) corresponding to the column $\text{Col}(x, y)$ of voxels in $\mathcal{I}$. The separation constraints between S_i and S_j are incorporated into G in the following way. For each node $v_i(x, y, z)$ on $\text{Col}_i(x, y)$ with $z < Z - \delta_{ij}^l$, as illustrated in Fig. 18b, an arc is put in E_s from $v_i(x, y, z)$ to $v_i(x, y, z + \delta_{ij}^l)$ in G_i (red arcs in Fig. 18b). Note that $\mathcal{I}(x, y, z + \delta_{ij}^l)$ is the upstream voxel of $\mathcal{I}(x, y, z)$ (red dashed line arrows in Fig. 18a). On the other hand, each node $v_j(x, y, z)$ on $\text{Col}_j(x, y)$ with $z \geq \delta_{ij}^l$ has an arc in E_s to $v_i(x, y, z^o) \in \text{Col}_i(x, y)$ with $z^o = \max\{0, z - \delta_{ij}^u\}$ (blue arcs in Fig. 18b). $\mathcal{I}(x, y, z^o)$ is the downstream voxel of $\mathcal{I}(x, y, z)$ (blue dashed line arrows in Fig. 18a). The node set of G is the union of the node sets of the λ subgraphs (i.e., $V = \cup_{i=1}^{\lambda} V_i$), and the arc set of G is the union of the arc sets of all the subgraphs plus E_s (i.e., $E = \cup_{i=1}^{\lambda} E_i \cup E_s$).

To encode the objective function $\mathcal{E}(\mathcal{S})$ in the graph model, we assign the node weight for each node $v_i(x, y, z)$ as follows:

$$w_i(x, y, z) = \begin{cases} b_i(x, y, z) + [c_{i-1}(x, y, z) - c_i(x, y, z)] & \text{if } z = 0, \\ [b_i(x, y, z) - b_i(x, y, z-1)] + [c_{i-1}(x, y, z) - c_i(x, y, z)] & \text{for } z > 0. \end{cases}$$
(14)

Figure 19 illustrates the intuition behind the node weight assignment to encode the objective function, in which two surfaces are simultaneously detected. Readers are referred to [8] for the formal proof.

4.4.2 Computing a Minimum-Cost Nonempty Closed Set

Note that our goal is to compute a minimum-cost nonempty closed set in G, which can be used to define an optimal set of λ surfaces in $\mathcal{I}$. However, the constructed graph G so far does not yet work for this purpose. In the graph construction above for G_i and G_j, the node $v_i(x, y, z)$ with $z \geq Z - \delta_{ij}^l$ has no arc to any node on $\text{Col}_j(x, y)$, and there is no arc from the node $v_j(x, y, z)$ with $z < \delta_{ij}^l$ to any node on $\text{Col}_i(x, y)$. Those nodes are called *deficient nodes*. The voxels corresponding to

the deficient nodes cannot be on any corresponding feasible surfaces. We need to remove those deficient nodes from G. Otherwise, if a closed set $\mathscr{C}$ in G includes a deficient node as the topmost node on a column that is in $\mathscr{C}$, which is possible, then it does not define a set of λ feasible surfaces in $\mathscr{I}$. However, the removal of deficient nodes may in turn cause other nodes in G to become deficient (i.e., the voxels associated with those nodes cannot be on any corresponding feasible surfaces). We can apply the deficient node pruning scheme in [8] to remove all the deficient nodes as well as those recursively caused by their removal, resulting in a graph G' without any deficient nodes. Note that in the resulting graph G' if any column $\mathrm{Col}_i(x, y) = \emptyset$, then there is no feasible solution to the OMSD problem. We thus assume in the rest of this section that the OMSD problem admits feasible solutions. For each (x, y)-pair and every $i = 1, 2, \ldots, \lambda$, let $z_i^{\mathrm{bot}}(x, y)$ and $z_i^{\mathrm{top}}(x, y)$ be the smallest and largest z-coordinates of the nodes on the column $\mathrm{Col}_i(x, y)$ of G', respectively. The deficient node pruning process may remove some useful arcs whose ending nodes need to be changed. For any node $v_i(x, y, z)$ with $z < z_i^{\mathrm{bot}}(x, y)$ in G, if it has an incoming arc, the ending node of the arc needs to be changed to the node $v_i(x, y, z_i^{\mathrm{bot}}(x, y))$ in G'.

Denote by Z_0 the set of the bottom-most nodes (i.e., whose z-coordinate is $z < z_i^{\mathrm{bot}}(x, y)$) of every column $\mathrm{Col}_i(x, y)$ in G'. Z_0 in fact is a closed set in G', and for any nonempty closed set $\mathscr{C}'$ in G', Z_0 is a subset of $\mathscr{C}'$ [8]. Thus, the same transformation as that for the OSSD problem can be applied to compute a minimum-cost nonempty closed set $\mathscr{C}'^*$ in G'.

4.4.3 Surface Recovery

From the minimum-cost nonempty closed set $\mathscr{C}'^*$ in G', we define an optimal set $\mathscr{S}^*$ of λ surfaces, $\mathscr{S}^* = \{S_1^*, S_2^*, \ldots, S_\lambda^*\}$, as follows. Let $\mathscr{C}^*$ denote the corresponding node set of $\mathscr{C}'^*$ in G. Recall that we search for each surface S_i^* in the subgraph $G_i = (V_i, E_i)$. For each $i = 1, 2, \ldots, \lambda$, let $\mathscr{C}_i = \mathscr{C} \cap V_i$. Considering any (x, y)-pair, denote by $\mathscr{C}_i(x, y)$ the set of nodes of $\mathscr{C}_i$ that are on $\mathrm{Col}_i(x, y)$ of G_i. Then, the voxel $\mathscr{I}(x, y, z)$ corresponding to the "top" node in $\mathscr{C}_i(x, y)$ (i.e., the node whose z-coordinate is the largest among those in $\mathscr{C}_i(x, y)$) is on the ith optimal surface S_i^*. We thus define an optimal set $\mathscr{S}^*$ of λ surfaces from the minimum-cost nonempty closed set $\mathscr{C}'^*$.

Theorem 2 ([8]). *The OMSD problem is solvable by computing a minimum-cost closed set in a derived graph.*

4.5 *Optimal Surface Detection with Shape and Context Priors*

Up to this point, in the OSD framework, only node weights are employed in the graph to represent the desired segmentation properties, and the desired surface

smoothness is hard-wired as connectedness of neighboring columns. This representation limits the ability to incorporate a broader variety of a prior knowledge. The connectedness of one voxel to the voxels of its neighboring columns is basically of equal importance in the OSD framework discussed above, which prevents us from fully utilizing image edge information as well as from taking full advantage of shape priors. To make full use of prior information, an arc-weighted graph representation has been proposed, which utilizes the weights of both graph nodes and arcs to incorporate a wider spectrum of constraints [37]. Two additional pairwise terms are added into the energy function, which encode the shape and the context prior information using a set of convex functions. For optimization, the new pairwise terms are enforced by adding specific weighted arcs in the graph. A globally optimal solution is then computed by solving a single maximum flow problem in the graph, which corresponds to optimal surfaces.

4.5.1 Incorporation of Shape and Context Priors

To simplify the notation, we use $p(x, y)$ or simply p to denote a column $\mathrm{Col}(x, y)$ in $\mathscr{I}$. For surface S_i, let $S_i(p)$ denote the z-coordinate of the voxel of the column p on S_i.

The shape prior of a surface is incorporated by enforcing the surface height changes between pairs of neighboring columns. More specifically, for any pair of neighboring columns p and q, the shape change of surface S_i between p and q can be measured by $h_{pq}^i = S_i(p) - S_i(q)$. Assume that $\overline{h}_{pq}^i$ represents the learned shape change model. The deviation of the shape changes from the learned model (that is, $|h_{pq}^i - \overline{h}_{pq}^i|$) is penalized by a convex function $f_s(h_{pq}^i - \overline{h}_{pq}^i)$. In addition, we can enforce the possible ranges of the shape change deviations with $|h_{pq}^i - \overline{h}_{pq}^i| \leq \Delta_{pq}^i$ (namely, a *hard shape constraint*), where Δ_{pq}^i is a given shape change deviation parameter between columns p and q for surface S_i.

For a set of surfaces, the context prior is enforced by penalizing the surface distance changes between two adjacent surfaces. Let S_i and S_j be two adjacent surfaces in the set of λ target surfaces, and without loss of generality, assume that S_i is on the top of S_j. The surface distance between S_i and S_j on column p is defined as $d_p^{ij} = S_i(p) - S_j(p)$. Denote by $\overline{d}_p^{ij}$ the learned prior surface distance model. As for the shape prior constraint, the deviation of the surface distances from the learned model (i.e., $d_p^{ij} - \overline{d}_p^{ij}$) is penalized by a convex function $f_c(d_p^{ij} - \overline{d}_p^{ij})$. We may also enforce the possible range of the surface distance deviation on each column p, for instance, $|d_p^{ij} - \overline{d}_p^{ij}| \leq \delta_p^{ij}$, which is called a *hard context constraint*.

Two additional terms are added into the energy function $\mathscr{E}(\mathscr{S})$ (see Eq. (16)) to incorporate the shape and context priors, with

$$\mathcal{E}(\mathscr{S}) = \sum_{i=1}^{\lambda} \sum_{\mathscr{I}(x,y,z)\in S_i} b_i(x,y,z) + \sum_{i=0}^{\lambda} \sum_{\mathscr{I}(x,y,z)\in R_i} c_i(x,y,z) +$$

$$\underbrace{\sum_{i=1}^{\lambda} \sum_{(p,q)\in \mathscr{N}_s} f_s(h_{pq}^i - \overline{h}_{pq}^i)}_{\text{shape prior penalty term}} + \underbrace{\sum_{p} \sum_{(i,j)\in \mathscr{N}_c} f_c(d_p^{ij} - \overline{d}_p^{ij})}_{\text{context prior penalty term}}, \qquad (15)$$

where $\mathscr{N}_s$ denotes the pairwise neighboring relation of all columns in the image $\mathscr{I}$ and $\mathscr{N}_c$ specifies a set of interacting surface pairs for which the surface context constraint needs to be enforced. The problem is to find an optimal set $\mathscr{S}$ of λ surfaces, such that (1) each surfaces satisfies the hard shape constraint; (2) each pair of surfaces satisfies the hard context constraint; and (3) the energy $\mathscr{E}(\mathscr{S})$ in Eq. (15) is minimized. We call this problem an OMSD with both shape and context priors (OMSD-P).

The basic idea for solving the OMSD-P problem is to transform it to computing a so-called *minimum-cost s-excess set* in a directed graph. Instead of forcing no arc leaving the sought node set as in the minimum-cost closed set problem, the minimum s-excess problem [7, 25, 38] charges a penalty onto each arc leaving the set (i.e., the tail of the arc is in the set while the head is not). More precisely, given a directed graph $G' = (V', E')$, each node $v' \in V'$ having an arbitrary weight $w'(v')$ and each edge $e' = (u', v') \in E'$ having a non-negative cost $c'(u', v')$, the minimum-cost s-excess problem seeks a node subset $\mathscr{X}' \subseteq V'$ such that the cost $\gamma(\mathscr{X}')$ of $\mathscr{X}'$, with $\gamma(\mathscr{X}') = \sum_{v'\in\mathscr{X}'} w'(v') + \sum_{\substack{(u',v')\in E' \\ u'\in\mathscr{X}', v'\in V'-\mathscr{X}'}} c'(u', v')$, is minimized. Our goal is to construct a both arc- and node-weighted graph so that: (1) the feasibility of the sought surfaces is enforced with the graph structure; (2) the total on-surface and in-region cost of the surfaces is encoded as the total node-weight of the s-excess set; and (3) the total shape-prior penalty and the context-prior penalty is encoded as the total cost of the cut induced by the s-excess set. The minimum-cost s-excess problem can be solved by using a maximum flow algorithm [25].

4.5.2 Arc-Weighted Graph Construction

We construct a graph $G = (V, E)$ consisting of λ node-disjoint subgraphs $\{G_i = (V_i, E_i) \mid i = 1, 2, \ldots, \lambda\}$ to transform the OMSD-P problem into computing a minimum-cost s-excess set in G. Note that the hard shape and context constraints in the OMSD-P problem are essentially the same as the surface smoothness and separation constraints in the OMSD problem. Each subgraph G_i is constructed as in Sect. 4.4 for the OMSD problem. The hard shape constraint is enforced between any two neighboring columns in G_i; and the hard context constraint is encoded between the corresponding columns in G_i and G_j for any two interacting surfaces S_i and S_j. To translate to the minimum-cost s-excess problem, we assign a $+\infty$ weight to

each arc added so far to the graph G. We next put additional arcs in G to incorporate both the shape prior penalty term and the context prior penalty term.

Let $p(x, y)$ and $q(x', y')$ denote any two neighboring columns in $\mathcal{I}$. For the shape prior penalty term, we want to "distribute" the shape prior penalty $f_s(h_{pq}^i - \overline{h}_{pq}^i)$ to arcs between the two adjacent columns $\mathrm{Col}_i(p)$ and $\mathrm{Col}_i(q)$ in G_i. Two intertwined questions need to be answered: (1) how to put arcs between the two columns; and (2) how to assign a non-negative cost to each arc (negative arc weights make the minimum-cost s-excess problem computationally intractable)? Fortunately, the convexity of $f_s(\cdot)$ can be used to resolve the problems. Define the (discrete equivalent of) second derivative of $f_s(\cdot)$ as

$$[f_s(\kappa)]'' = [f_s(\kappa + 1) - f_s(\kappa)] - [f_s(\kappa) - f_s(\kappa - 1)].$$

The first derivative of $f_s(\cdot)$, $[f_s(\kappa)]' = f_s(\kappa + 1) - f_s(\kappa)$, will be used to guide the introduction of new arcs between $\mathrm{Col}_i(p)$ and $\mathrm{Col}_i(q)$. Consider each $\kappa = h_{pq}^i - \overline{h}_{pq}^i$ with $-\Delta_{pq}^i < \kappa < \Delta_{pq}^i$. We distinguish the following three cases for all possible z ($0 \leq z < Z$). Note that we do not show the boundary conditions to avoid cluttering the exposition of the key ideas.

- If $[f_s(\kappa)]' \geq 0$, an arc with a weight of $[f_s(\kappa)]''$ is added from $v_i(x, y, z)$ to $v_i(x', y', z - \overline{h}_{pq}^i - \kappa)$.
- If $[f_s(\kappa)]' \leq 0$, an arc with a weight of $[f_s(\kappa)]''$ is added from $v_i(x', y', z)$ to $v_i(x, y, z + \overline{h}_{pq}^i + \kappa)$.
- Assume that $f_s(\kappa)$ has its minimum at κ_0. We put in an arc from $v_i(x, y, z)$ to $v_i(x', y', z - \overline{h}_{pq}^i - \kappa_0)$ whose weight is $[f_s(\kappa_0^+)]'' := f_s(\kappa_0 + 1) - f_s(\kappa_0)$. In addition, an arc with a weight of $[f_s(\kappa_0^-)]'' := f_s(\kappa_0 - 1) - f_s(\kappa_0)$ is introduced from $v_i(x', y', z)$ to $v_i(x, y, z + \overline{h}_{pq}^i + \kappa_0)$.

Figure 20 shows an example graph construction for incorporating the shape prior penalty term.

The context prior penalty term is enforced in a similar way by introducing weighted arcs between corresponding subgraphs. Suppose S_i and S_j are two interacting surfaces. For each column $p(x, y)$ in $\mathcal{I}$, by making use of the (discrete equivalent of) second derivative of $f_c(\cdot)$, the context prior penalty $f_c(d_p^{ij} - \overline{d}_p^{ij})$ is distributed to the arcs between the corresponding columns $\mathrm{Col}_i(p)$ and $\mathrm{Col}_j(p)$ in the subgraphs G_i and G_j, respectively. An example construction is shown in Fig. 21.

4.5.3 Computing a Minimum-Cost s-Excess Set

From the construction of the graph G and the argument for the OMSD problem in Sect. 4.4, to show that a minimum-cost s-excess set can be used to define an optimal set of λ surfaces for the OMSD-P problem, we need to demonstrate that: (1) the total weight of the arcs cut by a feasible surface S_i between two neighboring columns
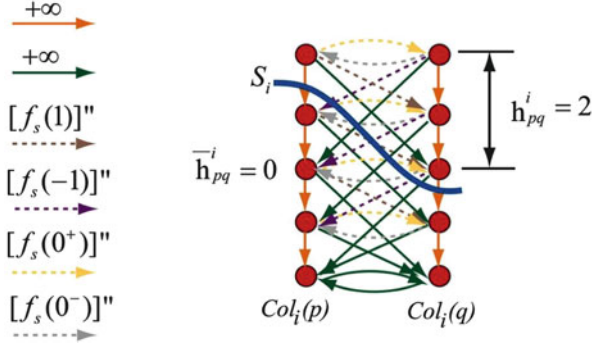
Fig. 20 Arc-weighted graph construction for the incorporation of the shape prior penalty on surface S_i between neighboring columns p and q [37]. The intra-column arcs are shown in *orange* with $+\infty$ weight. The hard shape constraint $|h^i_{pq} - \overline{h}^i_{pq}| \le \Delta^i_{pq}$ is enforced by *green arcs*. Here we suppose $\overline{h}^i_{pq} = 0$ and $\Delta^i_{pq} = 2$. The shape prior penalty is incorporated by arcs with *dashed lines* (*brown, purple, yellow,* and *gray*). Here we assume that $[f_s(0)]' = 0$ and $f_s(0) = 0$. The target surface S_i cuts arcs with weight $[f_s(1)]''$ (*brown*) and $[f_s(0^+)]'' = f_s(1) - f_s(0)$ (*yellow*). The total weight is equal to $f_s(2)$
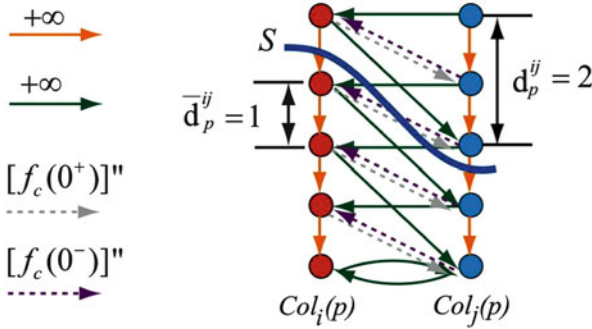

Fig. 21 Arc-weighted graph construction for the incorporation of the context prior constraints between subgraphs G_i (*red*) and G_j (*blue*) on column p [37]. The hard context constraint $|d^{ij}_p - \overline{d}^{ij}_p| \le \delta^{ij}_p$ is incorporated by *green arcs*. Here $\overline{d}^{ij}_p = 1$, $\delta^{ij}_p = 1$. The context-prior penalty is enforced by *gray* and *purple arcs*. Assume that $[f_c(0)]' = 0$ and $f_c(0) = 0$. The pseudo-surface S (connecting S_i and S_j) cuts arcs with weight $[f_c(0^+)]''$ (*gray*). The total weight is equal to $f_c(1)$

$\mathrm{Col}_i(p)$ and $\mathrm{Col}_i(q)$ equals the shape prior penalty $f_s(h^i_{pq} - \overline{h}^i_{pq})$; and (2) the total weight of the arcs cut by the pseudo-surface of two interacting surfaces S_i and S_j (connecting the node $v_i(p) \in \mathrm{Col}_i(p)$ on the surface S_i and the corresponding node $v_j(p) \in \mathrm{Col}_j(p)$ on the surface S_j for all the columns p) equals the context prior penalty $f_c(d^i_{pq} - \overline{d}^i_{pq})$. The reader is referred to [7] for the detailed proof.

A minimum s-excess set $\mathcal{X}^*$ in G can then be computed by using a minimum s–t cut algorithm [7, 25]. If the minimum s-excess set in G thus obtained is empty, we

can first apply on G the same transformation as in Sect. 4.4 to compute a minimum nonempty s-excess set in G.

4.5.4   Surface Recovery

The minimum s-excess set $\mathcal{X}^*$ can be used to specify an optimal set of λ surfaces as in Sect. 4.4.

## 4.6   Layered Graph Search for Segmentation of Objects with Complex Shapes

This section generalizes the OSD framework for solving the problem of detecting boundary surfaces for objects with complex topologies. Several key issues need to be resolved for the success of this generalization: (i) how to obtain relevant information about the target object boundaries; (ii) how to capture such information by a graph; and (iii) how to search the graph for the optimal surfaces of the target objects. The following five steps constitute the general strategy to address these three key issues.

Step 1: Pre-segmentation.   Image segmentation frequently starts with the localization of the object of interest (to distinguish the object from other objects in the image). Given an input 3D/4D image, a pre-segmentation obtains an approximation to the (unknown) surfaces for the target object boundaries. Such approximate surface detection methods are generally available, e.g., using parametric deformable models (cf. [39–42]), geometric deformable models (cf. [43–45]), and other similar approaches. The pre-segmentation gives useful information about the topological structures of the target objects.

Step 2: Mesh Generation.   From the resulting approximate surfaces, a mesh is constructed. The mesh is used to specify the structure of a graph G_B, called *base graph*. G_B defines the neighboring relations among voxels on the sought (optimal) surfaces. Trivial surface geometries (e.g., terrain-like, tubular, or spherical surfaces) may not need a pre-segmentation and allow a direct definition of a mesh.

Step 3: Image Resampling.   For each voxel v on the sought surfaces, a vector of voxels is created that is expected to contain v. This is done by resampling the input image along a ray intersecting every vertex u of the mesh (one ray per mesh vertex). These voxel vectors produced by the resampling form a new image. A big challenge for image resampling is how to avoid column interference (see Fig. 22). A few methods based on medial axes [46, 47] and the electric lines of force (ELF) [11] have been developed for resolving this problem. Steps 1–3 were developed for solving the issue (i) from the (i)–(iii) list presented at the beginning of this section.

Fig. 22 Illustrating image resampling with column interferences [46]. Interferences caused by inappropriate column lengths introduce disordered structures in the constructed graph
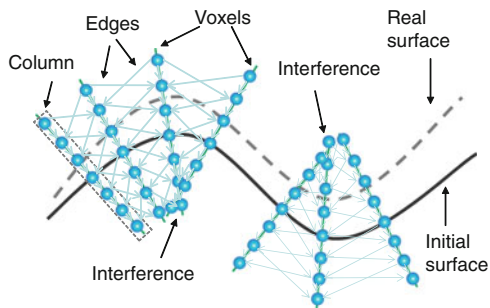
Step 4: Graph Construction. A weighted directed graph G is built using the vectors of voxels (columns) in the resampled image. Each column corresponds to a list of nodes in G. G is a *geometric* graph since it is naturally embedded in a d-D space ($d \geq 3$). The neighboring relations among voxels on the sought surfaces are represented by the adjacency relations among the columns of G, as specified by the edges in the base graph G_B. Each column contains exactly one voxel on the sought surface. The edges of G enforce constraints on the sought surfaces, such as smoothness and inter-surface separation. The node costs of G can encode edge-based and region-based cost functions. This step solves the issue (ii).

Step 5: Graph Search. The OSD scheme ensures that the sought optimal surfaces correspond to a structure of interest in the weighted directed graph G (as proven in [7–9]). The sought optimal surfaces are obtained by searching for a minimum-cost closed set in G. This step solves the issue (iii).

4.7 Example Applications and Extensions

4.7.1 Segmentation of Retinal Layers in Optical Coherence Tomography Volumes

Ophthalmology has recently witnessed a transformation with the relatively recent (since 2007) commercial availability of volumetric spectral-domain optical coherence tomography (SD-OCT) images of the eye. The retinal layers are one important structure within SD-OCT images of the back of the eye as these layers often change in the presence of blinding diseases such as glaucoma, diabetic retinopathy, and age-related macular degeneration. For example, the retinal nerve fiber layer and ganglion cell layer are known to thin in glaucoma. The need for multiple layered surfaces within SD-OCT volumes makes the graph search approach discussed in Sect. 4.4 an excellent choice for their segmentation [10]. For example, Fig. 23 illustrates the use of the graph search approach (with both on-surface and in-region cost terms) for the simultaneous segmentation of seven major surfaces within an SD-OCT slice of the macular region. While the segmentation result of Fig. 23 was obtained through the
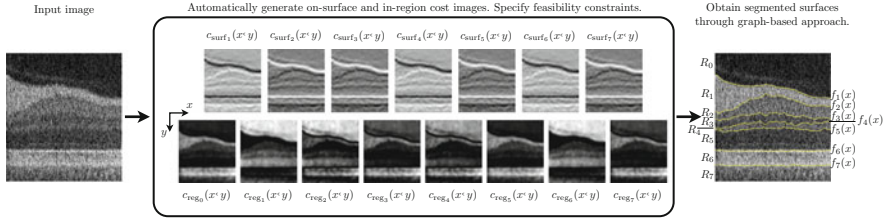
Fig. 23 Example simultaneous segmentation of seven layered surfaces from SD-OCT slice using the multisurface graph-theoretic approach [7, 9, 10]. Given the input image, seven on-surface cost images and eight in-region cost images are first automatically generated. Given these cost images, along with feasibility constraints, the graph-based approach finds the optimal set of surfaces. While 2D images are shown for illustrative purposes, in practice the cost functions and generated and layer segmentations are obtained in 3D

simultaneous segmentation of all seven surfaces using the graph search approach, in practice, the surface segmentation often occurs in groups with the "easier" bounding surfaces being simultaneously segmented first (e.g., surfaces f_1, f_6, and f_7 from Fig. 23) followed by the simultaneous segmentation of the more difficult interior surfaces [10, 48]. In addition, a multi-resolution approach is often used for efficiency purposes [49, 50] with up to 11 surfaces commonly being segmented, depending on the ophthalmic application. The cost functions for use in the segmentation of the retinal layers have been both designed primarily "by hand" [10, 51] (e.g., specifying an edge-based filter for bright-to-dark and dark-to-bright on-surface edge-based terms) in possible combination with learning the parameters for specifying the relative importance between in-region and on-surface cost terms as in [10] and/or using a machine-learning approach for a more complete automated design on the cost function terms [52].

In specifying the feasibility constraints for use for the simultaneous segmentation of the retinal layers, instead of specifying global smoothness and surface-interaction constraints (as described in the early versions of the graph search approach), one can specify varying constraints such that each location (i.e., each pair of neighboring columns for smoothness constants and each column for surface-interaction constraints) has its own (potentially different) values for the constraint [10]. The particular values of these location-specific constraints can be learned from a training set [10] to enable, for example, greater changes in surface position to be allowed near the expected location of the fovea of the macula and the cup of the optic nerve head. Soft smoothness constraints can also be imposed through the addition of arc-weighted edges [37] as discussed in Sect. 4.5.

4.7.2 Segmentation of Prostate and Bladder in Computed Tomography Scans

This section shows how to apply the OSD framework for the segmentation of prostate and bladder in CT images [53–55]. The segmentation of pelvic structure is
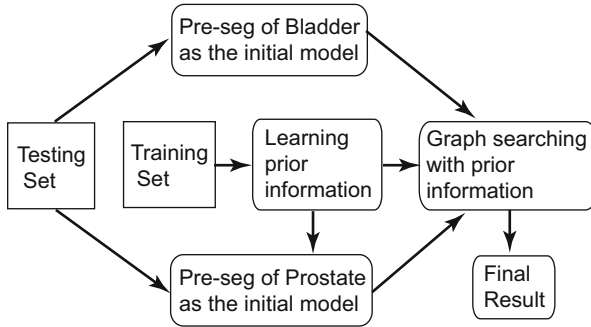
Fig. 24 Workflow for simultaneous segmentation of the bladder and the prostate

particularly difficulty. It involves soft tissues that present a large variability in shape and size. Those soft tissues also have similar intensity and have mutual influence in position and shape. To overcome these difficulties, multiple constraints (i.e., the shape prior and the context prior) are incorporated into the segmentation process using the method in Sect. 4.6. The workflow is shown in Fig. 24.

Initial Model. First, a pre-segmentation step is performed to construct an initial model for each target object, which contains the basic topological information. A 3D geodesic active contour method [56] is conducted for pre-segmentation of the bladder. Three user-defined points are required as an initial input. The prostate shows a much better coherency in shape than the bladder. Hence the mean shape of the prostate is computed from the training set of manual contours. Then an approximate bounding box of interest for the prostate is interactively defined and the obtained mean shape is roughly fitted into the never-before seen CT images using rigid transformations as the initial model of the prostate. Note that the model only serves to provide basic topological structural information; thus accurate segmentation is not required at this stage. Overlapping between the models of the bladder and the prostate is also allowed, which can be resolved in the graph optimization step. From the pre-segmentation results, two triangulated meshes $M_1(V_1, E_1)$ and $M_2(V_2, E_2)$ are constructed respectively using an isosurfacing algorithm (Fig. 25), where V_i ($i \in 1, 2$) denotes the vertex set of M_i and E_i denotes the edge set of M_i.

Graph Construction. The weighted graph $G_i(N, A)$ is built from the triangulated mesh M_i as follows. For each vertex $v \in V_i$, a column of K nodes $n(v, k)$ is created in G_i, denoted by $p(v)$ (Fig. 25b). The positions of nodes reflect the positions of corresponding voxels in the image domain. The length of the column is set according to the required search range. The number of nodes K on each column is determined by the required resolution. The direction of the column is set as the triangle normal. The nodes on the same column are connected by the intra-column arc from $n(v, k)$ to $n(v, k - 1)$ with an infinity weight, as described in Sect. 4.3.
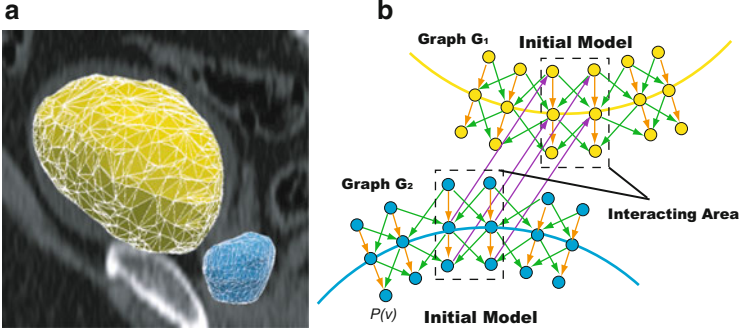
Fig. 25 (**a**) Triangulated meshes for the bladder (*yellow*) and the prostate (*blue*) based on initial models. (**b**) Corresponding graph construction. An example 2D slice is presented. $p(v)$ represents the column with respect to the vertex v on the mesh. *Dots* represent nodes $n_i \in G_i$. Two subgraphs G_1 and G_2 are constructed for the segmentation of the bladder and the prostate, respectively. Note that in the interacting region (*dashed box*), for each column $p(v_1) \in G_1$, there exists a corresponding column $p(v_2) \in G_2$ with the same position. The inter-surface arcs (*purple*) between corresponding columns enforce the surface context constraints in the interacting region

The intra-column arcs make sure that the surface containing exactly one node in each column.

Each column also has a set of neighbors, i.e., if $(v, u) \in E_i$, then $p(v)$ and $p(u)$ are neighboring columns. The shape prior penalty serves to keep the topology of the original shape model. Specifically, for any pair of neighboring columns p and q, the shape change is defined by $h_{pq}^i = S_i(p) - S_i(q)$. Note that the graph is constructed based on the mesh of the initial shape model. Thus the original shape prior $\overline{h}_{pq}^i$ is equal to 0. The shape prior penalty of surface S_i is set as $f_s(h_{pq}^i)$, where $f_s(\cdot)$ is a convex function penalizing the shape changes of S_i on p and q. For bladder and prostate segmentation, the penalty function is set with the form: $f_s(h_{pq}^i) = \beta \cdot (h_{pq}^i)^2$, where β is a parameter learned from the training data. The shape prior is enforced by introducing additional arcs as in Sect. 4.5.

To avoid the overlapping of two target surfaces, a "partially interacting area" is defined according to the distance between two meshes, which indicates that the two target surfaces may mutually interact with each other at that area. To model the context relation, for each column $p(v_1) \in G_1$ in the partially interacting area, there exists a corresponding column $p(v_2) \in G_2$ with the same position in $\mathscr{I}$, and the target surfaces S_1 and S_2 both cut those columns, as shown in Fig. 25b. For implementation, a one-to-one correspondence between two surface meshes needs to be computed on the partially interacting region. We project the pre-segmented prostate surface mesh on the interacting region to the mesh of the pre-segmented bladder boundary surface. Then we use the projected mesh patch to replace the original bladder surface mesh on the interacting region. Thus, a one-to-one mesh correspondence on the interacting region is established since the two new meshes on that area have exactly the same topological structure. In this way, two graphs share

the same nodes' positions in the partially interacting area. The non-overlapping constraint is enforced in the area as the hard context prior constraint by adding arcs between corresponding columns using the approach proposed in Sect. 4.5. No soft context prior penalties are introduced in our energy.

The optimal set $\mathscr{S}$ of two surfaces corresponding to the bladder and the prostate can then be found by minimizing the following energy through the constructed graph:

$$\mathscr{E}(\mathscr{S}) = \sum_{i=1}^{2} \mathscr{E}_{\text{boundary}}(S_i) + \sum_{i=0}^{2} \mathscr{E}_{\text{region}}(R_i) + \sum_{i=1}^{2} \mathscr{E}_{\text{shape}}(S_i) \qquad (16)$$

The boundary energy term serves as an external force, which drives the mesh towards the best fit to the image data. The shape energy term functions as an internal force, which keeps the shape of the original model and restricts the flexibility of the mesh. To incorporate the learned regional information, an additional region energy term is added in our energy function. Specifically, two surfaces for the bladder and the prostate naturally divide the volume into three regions denoted by R_0, R_1, and R_2, which corresponds to the region enclosed by the bladder surface S_1, one between S_1 and the prostate surface S_2 at the partially interacting area, and the region enclosed by S_2, respectively. Our region energy term $E_{\text{region}}(R_i)$ reflects the region property of all voxels inside R_i. An example cost function design for this application can be referred to [55].

The illustrative results in three views are displayed in Fig. 26a–d. The 3D representation is shown in Fig. 26e. From all views, the proposed algorithm produces a very good delineation of both the bladder and the prostate in the 3D space. The shape prior constraints succeed to keep the original topological structure of the target organs. No overlapping of the bladder and the prostate is found due to the enforcement of the context constraints.

4.7.3 Robust Delineation of Pulmonary Tumors Using Surface-Region Context

This section presents a segmentation method that integrates the graph cut method with the OSD method for segmenting the target objects of an arbitrary shape mutually interacting with terrain-like surfaces [57]. This scenario widely exists in the medical imaging field. For instance, lung tumors may be in touch with lung parenchyma or close to the diaphragm (Fig. 27a), and fluid and fluid-associated abnormalities may appear in between the retinal layers (Fig. 27b). The delineation of such a target object could be very challenging due to the similar intensity profiles of the surrounding tissues. However, in that setting, the boundary surfaces of the adjacent structures of the target object can serve as valuable prior information to help separate the target object from those structures. It is expected to be promising to simultaneously segment those boundary surfaces together with the target object.
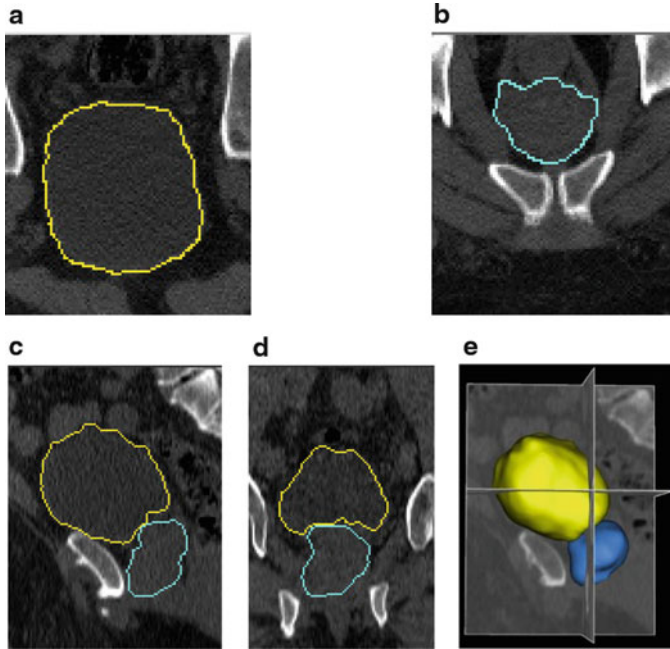
Fig. 26 Typical slices of the simultaneous segmentation result of the bladder (*yellow*) and the prostate (*blue*) in 3D CT images. (**a** and **b**) Transverse view. (**c**) Sagittal view. (**d**) Coronal view. (**e**) 3D representation of the segmentation result
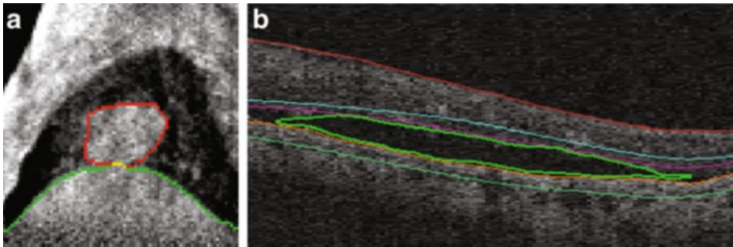

Fig. 27 Examples of surface and region interactions [57]. (**a**) Example slices of a lung tumor (*red*) in megavoltage cone-beam CT. The tumor is attached/adjacent to the diaphragm (*green*). (**b**) Example slices of a fluid-filled region (*green*) in retinal optical coherence tomography (OCT). The fluid region is surrounded by intraretinal layers

The computational feasibility is achieved by integrating the graph cut method in Sect. 3 and the OSD method in Sect. 4. The integration of those two methods into one single optimization process, yet still admitting globally optimal solutions, will certainly become a more powerful tool for medical image analysis.

We illustrate the method using lung tumor segmentation on Mega-Voltage Cone Beam CT (MVCBCT) images. WLOG, assume that the CBCT image $\mathscr{I}$ is oriented

Fig. 28 A sketched example of a 2D slice from a 3D image [57]. The terrain-like surface S is shown in *blue* and the tumor region R of arbitrary shape is shown in *brown*. Col(p) and Col(q) indicate two neighboring columns in the image, and $S(p) = 2$ and $S(q) = 3$
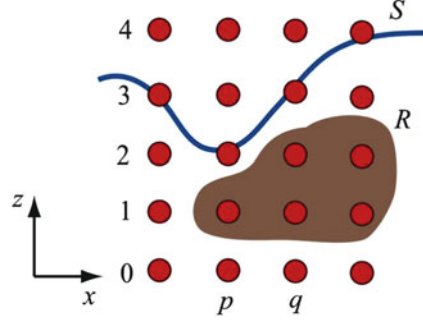
such that the target surface S intersects each column Col(p) of $p(x, y)$ in $\mathscr{I}$, and the target tumor is below the surface (Fig. 28). The same principles used for this illustration are directly applicable to multiple pairs of surfaces and regions with interactions between those two types of targets. The method can also handle the case when the tumor region is above the target surface [57].

The neighborhood setting of those columns is specified by $\mathscr{N}_s$. Let $S(p)$ denote the height (z-coordinate) of the surface S on the column Col(p). The OSD method is adopted to compute the surface S. Each voxel $v(x, y, z) \in \mathscr{I}$ is associated with an on-surface cost $b(x, y, z)$, which is inversely related to the likelihood that the desired surface S indeed contains the voxel. We can also incorporate the shape prior of the target surface S as in Sect. 4.5. The optimal surface S can be computed by minimizing the following energy function:

$$\mathscr{E}_s(S) = \sum_{v(x,y,z)\in S} b(x, y, z) + \sum_{(p,q)\in \mathscr{N}_s} h_{pq}(S(p) - S(q)), \qquad (17)$$

where $h_{pq}(\cdot)$ is a convex function.

The graph cut method in Sect. 3 is employed to segment the target tumor region R. Let f_v denote the label of each voxel $v \in \mathscr{I}$: $f_v = 1$ means that v belongs to the target tumor region R and $f_v = 0$ means that v is in the background. The MRF energy $\mathscr{E}_c$ for segmenting the tumor region R can be expressed as follows:

$$\mathscr{E}_c(f) = \sum_{v\in \mathscr{I}} \mathscr{D}_v(f_v) + \sum_{(v_i,v_j)\in \mathscr{N}_c} \mathscr{V}_{i,j}(f_{v_i}, f_{v_j}), \qquad (18)$$

where $\mathscr{N}_c$ defines the neighboring setting between voxels, and $\mathscr{D}_v(\cdot)$ and $\mathscr{V}_{i,j}(\cdot, \cdots)$ are the data term and the boundary term in the graph cut method (Sect. 3), respectively.

The geometric interaction relation between the boundary surface S and the target tumor region R is to enforce that R is at least $d > 0$ voxels "below" S. Denote by γ_v the penalty of a tumor voxel $v \in R$ that violates the interaction constraint. We introduce a surface-region interaction penalty term $\mathscr{E}_{\text{surf-tumor}}$, which is the total
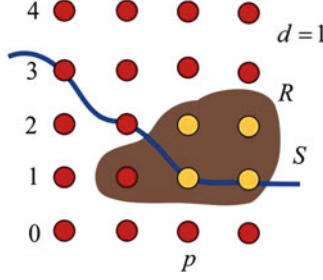
Fig. 29 The surface-region interaction constraint in which the region R tends to be positioned "lower" than the terrain-like surface S [57]. For voxels $v \in R$ and $S(p) - z(v) < d$ with $v \in \text{Col}(p)$ (*yellow voxels*), a penalty γ_v is enforced; d is set as 1

penalty of those tumor voxels that violates the interaction constraint (Fig. 29). More specifically, let $z(v)$ denote the height (z-coordinate) of voxel $v \in \mathcal{I}$, and v is on the column $\text{Col}(p)$ of $p(x, y)$ (that is, the x- and y-coordinates of v is x and y, respectively). Then, if $v \in R$ and $S(p) - z(v) < d$, a penalty γ_v is enforced. Hence,

$$\mathcal{E}_{\text{surf-tumor}}(S, f) = \sum_{v \in \mathcal{I}} \sum_{\substack{v \in \text{Col}(p) \\ S(p) - z(v) < d}} \gamma_v \cdot f_v. \qquad (19)$$

To enforce the boundary surface prior, we co-segment the tumor region as well as the boundary surface by minimizing the energy function $\mathcal{E}(S, f)$, with

$$\mathcal{E}(S, f) = \mathcal{E}_s(S) + \mathcal{E}_c(f) + \mathcal{E}_{\text{surf-tumor}}(S, f). \qquad (20)$$

To optimize the energy function $\mathcal{E}(S, f)$, two subgraphs, $G_c = (N_c, A_c)$ and $G_s = (N_s, A_s)$, are constructed to encode the terms $\mathcal{E}_c(f)$ and $\mathcal{E}_s(S)$, respectively, using the approaches in Sects. 3 and 4.5. For each voxel $v(x, y, z) \in \mathcal{I}$, denote by $n_c(x, y, z)$ ($n_s(x, y, z)$) the node in G_c (G_s) corresponding to v. The surface-region penalty term $\mathcal{E}_{\text{surf-tumor}}$ is incorporated by adding directed arcs between the corresponding nodes of the two subgraphs with a weight of the penalty for violating the interaction constraint. More specifically, for each voxel $v(x, y, z)$, a directed arc from $n_c(x, y, z)$ to $n_s(x, y, z+d)$, whose weight is γ_v, is added between the two subgraphs G_c and G_s (Fig. 30). Note that we do not consider the boundary conditions here to avoid cluttering the exposition of the key ideas.

Once the graph is constructed, a globally optimal solution can be found by solving a single maximum flow problem, which minimizes the total energy $\mathcal{E}(S, f)$ in a low-order polynomial time [57].
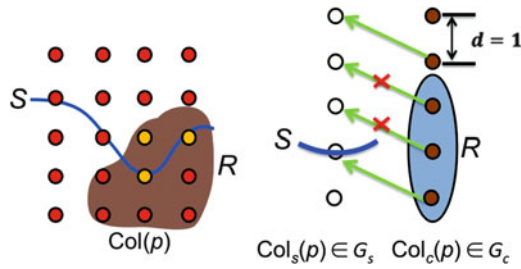
Fig. 30 The introduced arcs in the constructed graph to enforce the surface-region interaction constraint. The node column $\text{Col}_s(p)$ ($\text{Col}_c(p)$) in the subgraph G_s (G_c) corresponds to the column $\text{Col}(p)$ of voxels in the input image. The *green arcs* introduced to enforce the interaction constraint. The two arcs with a *red cross* indicate the penalties for the present surface violating the interaction constraint
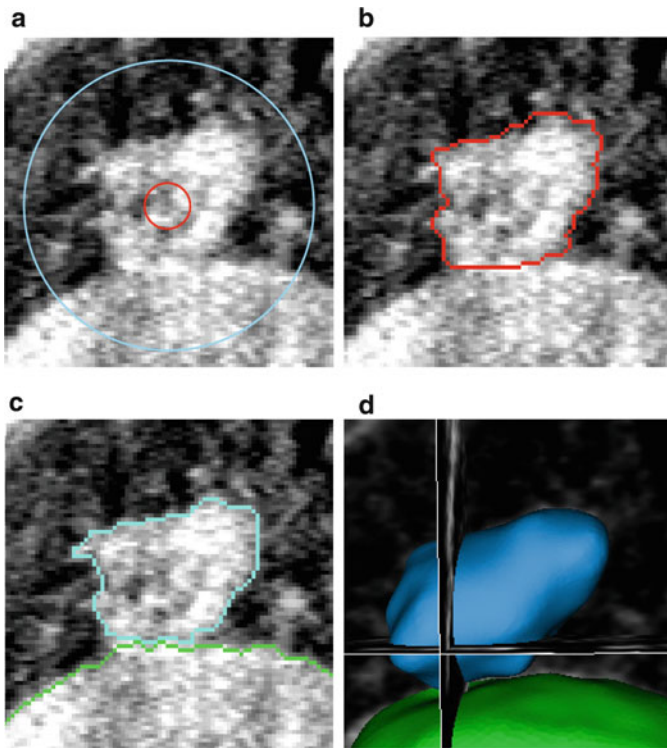
Fig. 31 A typical tumor segmentation result [57]. (**a**) One 2D slice of 3D MVCBCT image with outlines of spherical initialization. (**b**) Manual segmentation of the lung tumor—independent standard. (**c**) Simultaneous region-and-surface segmentation of the diaphragm (*green*) and the lung tumor (*blue*) using the reported approach showing excellent segmentation performance—the Dice similarity coefficient (DSC) is 0.878. (**d**) The 3D representation of the diaphragm (*green*) and the tumor (*blue*)

Fig. 32 Performance comparison on a difficult case. (**a**) Independent standard obtained by manual segmentation and shown in one 2D slice of the 3D volume. (**b**) Tumor segmentation failure resulting from the conventional graph cut method—DSC = 0.70. (**c**) Tumor segmentation obtained using the method that simultaneously segments the tumor (*blue*) and the lung boundary surface (shown in *yellow* and *green* in two orthogonal directions)—DSC = 0.84

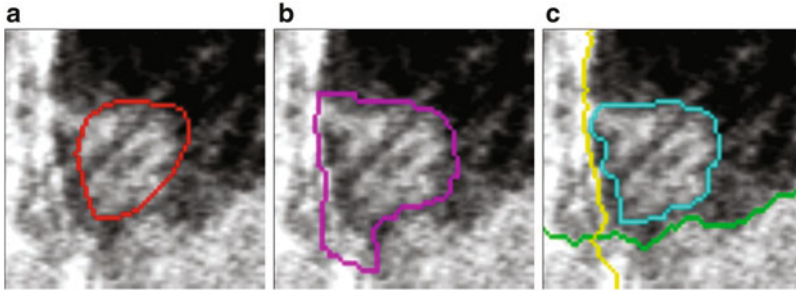Figure 31 shows a typical tumor segmentation result on a pulmonary MVCBCT dataset. Figure 32 demonstrates a segmentation result on a difficult case, in which the tumor is closely adjacent to the lung boundary surface from two directions.

References

1. Boykov, Y., Jolly, M.-P.: Interactive organ segmentation using graph cuts. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2000. Lecture Notes in Computer Science, vol. 1935, pp. 276–286. Springer, Berlin (2000)
2. Boykov, Y., Jolly, M.-P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: IEEE International Conference on Computer Vision (ICCV) 2001, pp. 105–112. IEEE, New York (2001)
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. **23**(11), 1222–1239 (2001)
4. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Pattern Anal. Mach. Intell. **26**(9), 1124–1137 (2004)
5. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 147–159 (2004)
6. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient N-D image segmentation. Int. J. Comput. Vis. **70**(2), 109–131 (2006)
7. Wu, X., Chen, D.Z.: Optimal net surface problems with applications. In: Widmayer, P., Eidenbenz, S., Triguero, F., Morales, R., Conejo, R., Hennessy, M. (eds.) Automata, Languages and Programming. Lecture Notes in Computer Science, vol. 2380, pp. 775–775. Springer, Berlin (2002)
8. Wu, X., Chen, D.Z., Li, K., Sonk, M.: The layered net surface problems in discrete geometry and medical image segmentation. Int. J. Comput. Geom. Appl. **17**(3), 261–296 (2007)
9. Li, K., Wu, X., Chen, D.Z., Sonka, M.: Optimal surface segmentation in volumetric images—a graph-theoretic approach. IEEE Trans. Pattern Anal. Mach. Intell. **28**(1), 119–134 (2006)

10. Garvin, M.K., Abràmoff, M.D., Wu, X., Russell, S.R., Burns, T.L., Sonka, M.: Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE Trans. Med. Imaging **28**(9), 1436–1447 (2009)
11. Yin, Y., Zhang, X., Williams, R., Wu, X., Anderson, D.D., Sonka, M.: LOGISMOS—layered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint. IEEE Trans. Med. Imaging **29**(12), 2023–2037 (2010)
12. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn. MIT, Cambridge (2001)
13. Delong, A., Boykov, Y.: Globally optimal segmentation of multi-region objects. In: International Conference on Computer Vision (ICCV), Tokyo, pp. 285–292 (October 2009)
14. Song, Q., Bai, J., Han, D., Bhatia, S., Sun, W., Rockey, W., Bayouth, J., Buatti, J., Wu, X.: Optimal co-segmentation of tumor in PET-CT images with context information. IEEE Trans. Med. Imaging **32**(9), 1685–1697 (2013)
15. Mu, Y., Zhou, B.: Co-segmentation of image pairs with quadratic global constraint in MRFs. In: Proceedings of the 8th Asian Conference on Computer Vision, Part II, Tokyo, pp. 837–846 (2007)
16. Mukherjee, L., Singh, V., Dyer, C.R.: Half-integrality based algorithms for cosegmentation of images. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, pp. 2028–2035 (2009)
17. Hochbaum, D.S., Singh, V.: An efficient algorithm for co-segmentation. In: Proceedings of the International Conference on Computer Vision, Kyoto, pp. 269–276 (2009)
18. Rother, C., Minka, T., Bake, A., Kolmogorov, V.: Cosegmentation of image pairs by histogram matching—incorporating a global constraint into MRFs. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, pp. 993–1000 (2006)
19. Batra, D., Kowdle, A., Parikh, D., Luo, J., Chen, T.: Interactively co-segmentating topically related images with intelligent scribble guidance. Int. J. Comput. Vis. **93**(3), 273–292 (2011)
20. Thedens, D.R., Skorton, D.J., Fleagle, S.R.: A three-dimensional graph searching technique for cardiac border detection in sequential images and its application to magnetic resonance image data. In: Proceedings of Computers in Cardiology, Chicago, pp. 57–60 (1990)
21. Thedens, D.R., Skorton, D.J., Fleagle, S.R.: Methods of graph searching for border detection in image sequences with applications to cardiac magnetic resonance imaging. IEEE Trans. Med. Imaging **14**(1), 42–55 (1995)
22. Boykov, Y., Veksler, O., Zabih, R.: Markov Random Fields with efficient approximations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, pp. 648–655 (1998)
23. Ishikawa, H., Geiger, D.: Segmentation by grouping junctions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, pp. 125–131 (1998)
24. Greig, D., Porteous, B., Seheult, A.: Exact maximum a posteriori estimation for binary image. J. R. Stat. Soc. Ser. B **51**, 271–279 (1989)
25. Hochbaum, D.S.: An efficient algorithm for image segmentation, Markov random fields and related problems. J. ACM **48**(4), 686–701 (2001)
26. Glocker, B., Komodakis, N., Paragios, N., Glaser, C., Tziritas, G., Navab, N.: Primal/dual linear programming and statistical atlases for cartilage segmentation. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007). Lecture Notes in Computer Science, vol. 4792, pp. 536–543. Springer, Berlin (2007)
27. Roy, S., Cox, I.: A maximum-flow formulation of the n-camera stereo correspondence problem. In: Proceedings of International Conference on Computer Vision (ICCV), Bombay, pp. 492–499 (1998)
28. Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions using graph cuts. In: Proceedings of International Conference on Computer Vision (ICCV), Vancouver, pp. 508–515, July 2001

29. Glocker, B., Komodakis, N., Paragios, N., Tziritas, G., Navab, N.: Inter and intra-modal deformable registration: continuous deformations meet efficient optimal linear programming. In: Proceedings of the 20th International Conference on Information Processing in Medical Imaging (IPMI). Lecture Notes in Computer Science, vol. 4584, pp. 408–420. Springer, Berlin (2006)
30. Kleinberg, J., Tardos, E.: Approximation algorithms for classification problems with pairwise relationships: metric labeling and Markov random fields. In: Proceedings of the 40th IEEE Symposium on Foundations of Computer Science, New York City, pp. 14–23 (1999)
31. Archer, A., Fakcharoenphol, J., Harrelson, C., Krauthgamer, R., Talvar, K., Tardos, E.: Approximate classification via earthmover metrics. In: Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, pp. 1079–1089 (2004)
32. Chuzhoy, J., Naor, S.: The hardness of metric labeling. SIAM J. Comput. **36**(5), 1376–1386 (2007)
33. Chekuri, C., Khanna, A., Naor, J., Zosin, L.: A linear programming formulation and approximation algorithms for the metric labeling problem. SIAM J. Discrete Math. **18**(3), 608–625 (2005)
34. Gupta, A., Tardos, E.: Constant factor approximation algorithms for a class of classification problem. In: Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC), Portland, pp. 652–658 (2000)
35. Dubes, R., Jain, A.: Random field models in image analysis. J. Appl. Stat. **16**, 131–164 (1989)
36. Picard, J.C.: Maximal closure of a graph and applications to combinatorial problems. Manag. Sci. **22**, 1268–1272 (1976)
37. Song, Q., Bai, J., Garvin, M.K., Sonka, M., Buatti, J., Wu, X.: Optimal multiple surface segmentation with shape and context priors. IEEE Trans. Med. Imaging **32**(2), 376–386 (2013)
38. Ahuja, R.K., Hochbaum, D., Orlin, J.B.: A cut based algorithm for the convex dual of the minimum cost network flow problem. Algorithmica **39**, 189–208 (2004)
39. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **1**(4), 321–331 (1987)
40. Amini, A.A., Weymouth, T.E., Jain, R.C.: Using dynamic programming for solving variational problems in vision. IEEE Trans. Pattern Anal. Mach. Intell. **12**(9), 855–867 (1990)
41. Amini, A.A., Weymouth, T.E., Jain, R.C.: On active contour models and balloons. CVGIP: Image Underst. **53**(2), 211–218 (1991)
42. Amini, A.A., Weymouth, T.E., Jain, R.C.: A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. Comput. Med. Imaging Graph. **19**(1), 69–83 (1995)
43. Caselles, V., Catte, F., Coll, T., Dibos, F.: A geometric model for active contours. Numerische Mathematik **66**, 1–31 (1993)
44. Malladi, R., Sethian, J.A., Vemuri, C.: Shape modeling with front propagation: a level set approach. IEEE Trans. Pattern Anal. Mach. Intell. **17**(2), 158–175 (1995)
45. Osher, S., Paragios, N.: Geometric Level Set Methods in Imaging, Vision, and Graphics. Springer, New York (2003)
46. Liu, X., Chen, D., Tawhai, M., Wu, X., Hoffman, E., Sonka, M.: Optimal graph search based segmentation of airway tree double surfaces across bifurcations. IEEE Trans. Med. Imaging **32**(3), 493–510 (2013)
47. Liu, X., Chen, D.Z., Wu, X., Sonka, M.: Optimal graph-based segmentation of 3D pulmonary airway and vascular trees across bifurcations. In: First International Workshop on Pulmonary Image Analysis, at the 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), New York City, pp. 103–111 (September 2008)
48. Garvin, M.K., Abràmoff, M.D., Kardon, R., Russell, S.R., Wu, X., Sonka, M.: Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. IEEE Trans. Med. Imaging **27**(10), 1495–1505 (2008)
49. Lee, K., Abràmoff, M.D., Niemeijer, M., Garvin, M.K., Sonka, M.: 3-D segmentation of retinal blood vessels in spectral-domain OCT volumes of the optic nerve head. In: Proceedings of SPIE Medical Imaging 2010: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 7626, San Diego, p. 76260V-8. SPIE (2010)

50. Quellec, G., Lee, K., Dolejsi, M., Garvin, M.K., Abràmoff, M.D., Sonka, M.: Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula. IEEE Trans. Med. Imaging **29**(6), 1321–1330 (2010)
51. Lee, K., Niemeijer, M., Garvin, M.K., Kwon, Y.H., Sonka, M., Abràmoff, M.D.: Segmentation of the optic disc in 3-D OCT scans of the optic nerve head. IEEE Trans. Med. Imaging **29**(1), 159–168 (2010)
52. Antony, B.J., Abràmoff, M.D., Lee, K., Sonkova, P., Gupta, P., Kwon, Y., Niemeijer, M., Hu, Z., Garvin, M.K.: Automated 3D segmentation of intraretinal layers from optic nerve head optical coherence tomography images. In: Proceedings of SPIE Medical Imaging 2010: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 7626, San Diego, p. 76260U-12. SPIE (2010)
53. Song, Q., Wu, X., Liu, Y., Smith, M., Buatti, J., Sonka, M.: Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2009. Lecture Notes in Computer Science, vol. 5762, pp. 827–835. Springer, Berlin (2009)
54. Song, Q., Wu, X., Liu, Y., Sonka, M., Garvin, M.K.: Simultaneous searching of globally optimal interacting surfaces with shape priors. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, pp. 2879–2886 (June 2010)
55. Song, Q., Liu, Y., Liu, Y., Saha, P., Sonka, M., Wu, X.: Graph search with appearance and shape information for 3-D prostate and bladder segmentation. In: Jiang, T., Navab, N., Pluim, J., Viergever, M. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010. Lecture Notes in Computer Science, vol. 6363, pp. 172–180. Springer, Berlin (2010)
56. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vis. **22**, 61–97 (1997)
57. Song, Q., Chen, M., Bai, J., Sonka, M., Wu, X.: Surface-Region context in optimal multi-object Graph-Based segmentation: robust delineation of pulmonary tumors. In: Szekely, G., Hahn, H. (eds.) Information Processing in Medical Imaging. Lecture Notes in Computer Science, vol. 6801, pp. 61–72. Springer, Berlin (2011)

Information Theoretic Clustering for Medical Image Segmentation

Jason Hill, Enrique Corona, Jingqi Ao, Sunanda Mitra, and Brian Nutter

Abstract Analysis of medical images involves robust computational approaches to various optimization problems prior to interpretation of the embedded pathological changes. Major computational efforts are essential in unsupervised learning of structures in various medical images. Some unsupervised learning techniques that take advantage of information theory concepts provide a different perspective on the solution of automated learning methods. This chapter will review a recent approach to clustering examined under an information theoretic framework that efficiently finds a suitable number of clusters representing different tissue characteristics in a medical image. The proposed clustering approach optimizes an objective function which quantifies the quality of particular cluster configurations. Recent developments involving interesting relationships between spectral clustering (SC) and kernel principal component analysis (kPCA) are used in this technique to include the nonlinear domain. In this novel SC approach, the data is mapped to a new space where the points belonging to the same cluster are collinear if the parameters of a radial basis function (RBF) kernel are adequately selected. The effectiveness of this nonlinear approach is demonstrated in the segmentation of uterine cervix color images for early identification of cervical neoplasia, as an aid to cervical cancer diagnosis. The limitations of this method in the segmentation of specific medical images such as brain images with multiple sclerosis lesions and a strategy to overcome them are discussed.

J. Hill • J. Ao • S. Mitra (✉) • B. Nutter
Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409, USA
e-mail: sunanda.mitra@ttu.edu

E. Corona
Whirlpool Corporation, 750 Monte Road, MD 5310, Benton Harbor, MI 49022, USA

1   Introduction

The discovery of underlying structures in medical images in a robust and efficient manner is of crucial significance in many applications, such as the diagnostic classification of precancerous uterine cervix lesions and the identification of anomalies in magnetic resonance image (MRI) brain structures. Often, with medical images, an adequate ground truth is not available, and unsupervised learning techniques must be used. One of the fundamental questions in the identification process is the number of distinguishable clusters or segments that comprise the image data. The machine learning methods that address these problems form a highly interdisciplinary field, where a non-interacting machine can be developed to learn in either a *supervised* or *unsupervised* manner.

With supervised learning, the system is "taught" by providing a labeled training set, which is a sequence of desired output $\mathbf{y} = \{y_1, y_2, \ldots\}$. The goal of the machine is to learn to produce the correct output (or class label in a classification context) given a new input. Although such a system may be quite fast after training, it can be limited by a subjective bias that depends on the training set, i.e. it may perform poorly when presented with a data pattern that is considerably different from the patterns encountered during training. Furthermore, some applications cannot provide a universally accepted "ground truth" to label training data in the first place. In unsupervised learning, the machine only receives the input data $\mathbf{x} = \{x_1, x_2, \ldots\}$ but obtains neither supervised target outputs nor rewards from its environment. It may seem mysterious that a system can learn to recognize patterns without being informed of the features it should analyze or even the correct number of clusters. With unsupervised learning, the machine is not explicitly taught; instead, its goal is to build representations of the input that can be used for decision making. Instead of relying upon a correctly labeled training set, the machine learns typically via a general objective function or a set of update rules. In one sense, unsupervised learning can be thought of as discovering patterns that are more significant than noise. Hence, a standard example of unsupervised learning is clustering, which becomes segmentation when applied to image data [1].

The two competing formalisms of machine learning methods are *discriminative* and *generative* approaches. Generative or informative approaches produce a probability density model over all variables to compute classification in the manner of parametric clustering. In parametric clustering, a predefined distribution for the dataset is assumed, and then sufficient statistics or parameters are calculated. For example, when assuming a multidimensional normal distribution, the sufficient statistics are the mean and the covariance matrix [2]. Furthermore, some information about the statistical dependencies of the data can be gained through unsupervised learning [3, 4]. Another approach is discriminative learning, which lends itself to nonparametric clustering. Discriminative approaches provide a direct attempt to compute the input–output mappings for classification, while eschewing any direct modeling of the underlying distributions. Nonparametric clustering divides the data into groups of points that possess strong internal similarities. A criterion function

Table 1 A summary of method for various scales of discriminative and generative approaches in machine learning (Jebara 2004)

Scale/approach	Generative	Discriminative
Local	Maximum likelihood	Empirical risk minimization
	Maximum mutual information	
Local + prior	Maximum a posteriori	Support vector machines (SVM)
		Regularization theory
Model averaging	Bayesian inference	Maximum entropy discrimination
		Bayes point machines

is used to measure these similarities and to seek the grouping that optimizes the criterion [2] (Jebara 2004).

Another approach to unsupervised learning uses an information theoretic approach, which in a sense transcends both the generative and discriminative approaches. For example, from Table 1 both the maximum mutual information and maximum entropy discrimination approaches to learning are connected to information theory. Information theory addresses the trade-off between memory storage cost and accuracy. Almost all work in unsupervised learning can be viewed in terms of learning a probabilistic model of the data. The machine can build a model of the data which assumes that the data points are independently drawn from some distribution $P(x)$. In general, the true distribution of the data is unknown, but we can learn a model of this distribution, say $Q(x)$. Information theory is used to estimate the accuracy of such a model at a given memory storage cost. Such a model can be used for classification. Furthermore, an information theoretic measure can be used as a similarity criterion for clustering. This approach is appealing, because it can capture descriptions beyond second order (variance) statistics. The major problem of information theoretic based clustering has been the difficulty to evaluate the metric without imposing unrealistic assumptions on the data distributions. Successful implementations carefully estimate the information theoretic measure [2].

In this information theoretic approach to medical image analysis, where the number of significant segments in an image is unknown a priori, we apply a method of information theoretic unsupervised learning. A simple information theoretic approach to clustering, proposed by Sugar and James [5], is called the "jump" method as explained below. The "jump" method performs well for high dimensional Gaussian data, but it behaves unreliably for clusters corresponding to natural image segments. To make this approach more practical, Corona developed an improved "jump" method (IJM) [6]. IJM performs well with synthetic data and has been validated with standard clustering datasets as well as with medical images with known ground truths. In this chapter, we will briefly introduce the original jump method and improvements to address some of its drawbacks. The applications of

the IJM will be described in segmentation of medical images. The potential of the IJM in compression and 3-D visualization of medical images will also be addressed.

2 The "Jump" Method

Suppose one has p-dimensional data with n points to be clustered, but the exact number of clusters K is unknown a priori. First, a method is needed to quantify the quality of the clustering given the number of clusters, that is, a measure of the inter-cluster dispersion. One such measure is the minimal achievable distortion, d_K. This distortion is defined as the mean squared Mahalanobis distance of every point in the cluster to the *nearest* cluster's centroid (or cluster representative), i.e.

$$d_K = \frac{1}{p} \min_{c_1, c_2, \ldots, c_K} \left\{ \frac{1}{n} \sum_{k=1}^{K} \sum_{i \in k} (x_i - c_k)^T \Sigma^{-1} (x_i - c_k) \right\}, \quad (1)$$

where x_i is the ith data point that belongs to the kth cluster, c_k is the centroid or representative point of the kth cluster, Σ^{-1} is the inverse covariance matrix for each cluster (assumed to be approximately equal for all clusters), and K is the number of clusters considered. One can create a monotonically decreasing curve by gradually increasing the value of K, and re-estimating Eq. (1) at each step; that is, consider the sequence $d_K = \{d_{K=1}, d_{K=2}, \ldots, d_{K=K\max}\}$. This curve contains important information that can be used to estimate a suitable number of representatives for the data.

Sugar and James developed a simple information theoretic approach where a classic clustering algorithm (e.g., k-means with $\Sigma = I$) is used to generate a curve of the number of clusters, K, versus an estimate of the associated distortion, d_K [5]. What they observed is that there tends to be a discontinuous "jump" in this curve at the appropriate number of clusters K (see Figs. 1 and 3). The jump operator can be defined as

$$J_K^Y = \frac{1}{d_K^Y} - \frac{1}{d_{K-1}^Y}, \quad (2)$$

where $Y > 0$ is a suitable power transformation acting on d_K [5]. It can be theoretically shown that for Gaussian clusters sharing the same covariance matrix, the distortion curve becomes linear for $Y = p/2$, if the dataset dimensionality p is large enough. Furthermore, if one defines $d_{K=0}^{-Y} = 0$, the detection of the absence of clustering structure (i.e., $K = 1$) is feasible. Finally, the number of clusters where the largest jump occurs provides the estimate of the number of clusters in the dataset, i.e.
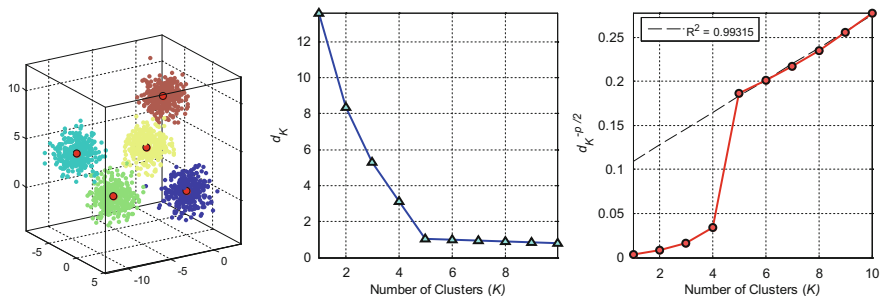
$$K^* = \max_K \{J_K\}. \quad (3)$$

Fig. 1 The "jump" method, with linear behavior of the transformed d_K curve for a mixture of five spherical Gaussians. The *box on the left* shows the mixture of Gaussian components. The *middle plot* exhibits the estimated d_K curve. The *rightmost plot* illustrated the transformed distortion curve of d_K^{-Y} with $Y = p/2$ with a fitted line for $K > K^*$. The jump operator calculates the difference of successive values on the *rightmost plot*, thus converting the problem into finding a maximum [7]

Finally, it should be noted that this "jump" method approach is hierarchical. That is, the relative height arrangement of the jump operator reflects more likely cluster configurations. This means that a proper sub-clustering may show up as a distinct second highest jump operator peak (see Fig. 2).

3 The Improved "Jump" Method

3.1 The Major Drawbacks of the "Jump" Method

The major drawbacks of the "jump" method are rooted in the fixing of the value $Y = p/2$ by applying the restrictive assumptions:

- Because the *k*-means algorithm performance is extremely dependent on the choice of the initial cluster centers, the accuracy of the distortion curve may be compromised by a poor initialization; i.e., the standard *k*-means implementation can get "trapped" in local optima, which yields a poor estimation of d_K using *k*-means and results in very noisy or inaccurate "jump" selections.
- The method assumes high dimensionality ($p \to \infty$), accurate d_K values, well balanced clusters, identity covariance matrix, sufficiently large distance between centroids, etc., and the power transformation can only be determined under these assumptions for underlying Gaussian distributions. Real-world data easily violate these assumptions.
- If strong correlations are present among variables, or if the clusters have intriguing or complicated shapes, linear algorithms such as *k*-means will not capture the data essence [6].
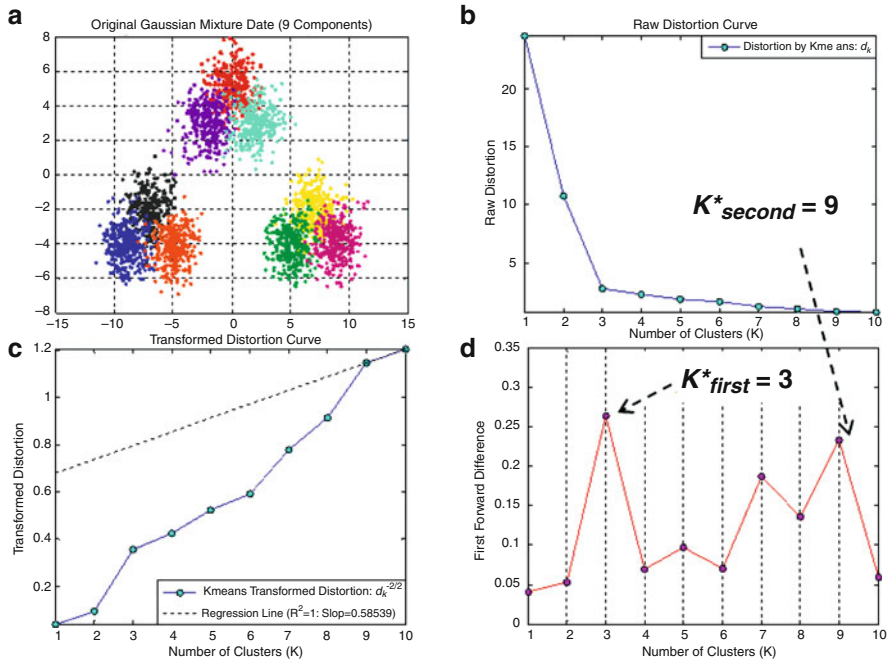
Fig. 2 The hierarchical aspect of the "jump" method. (**a**) The raw data generated from nine three-dimensional Gaussians that form three super-clusters. (**b**) The raw distortion curve. (**c**) The transformed distortion curve with $Y = p/2 = 3/2$. (**d**) The jump statistic, defined as the first forward difference of the transformed distortion. Note that the most obvious configuration is the three super-clusters, with the nine individual clusters at a close second in terms of the jump operator [6]

3.2   The Improvements to the "Jump" Method

The IJM works by first fixing the cluster number K and then searching for its best associated transformation, finding $Y = p_{\text{eff}}/2$ for each K, where p_{eff} is called the *effective dimension*:

- Spectral clustering (SC) and probabilistic initialization techniques can be used to alleviate the suboptimal convergence of the k-means algorithm.
- Handling realistic data necessitates that the power transform be adaptively modified according to the data at hand. One can make use of the hierarchical property of the "jump" method to create a criterion, the margin, which measures the quality of each of the K partitions. This approach allows the measurement of the operator performance in terms of linearity and the relative "jump" peak height. By maximizing this new objective function, one obtains a suitable power transformation, Y, which ensures a more robust solution to the number of clusters in a dataset.

Fig. 3 The jump and highest margin operator for data shown in Fig. 1 [6]
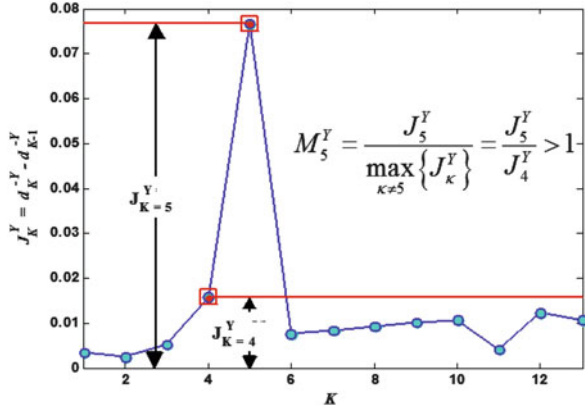
- Find a suitable kernel method to map the original dataset onto a "better" space where the mapped data is well behaved, such that the "jump" method can be used there effectively.

3.3 The Margin Operator of the IJM

The original "jump" method possesses a hierarchical nature, in the sense that the relative heights of the peaks of the jump operator J_K^Y indicate the degree of validity of the solution (or a valid sub-clustering). With this in mind, one can further define another operator called the margin operator,

$$M_K^Y = \frac{J_K^Y}{\max_{\beta \neq K} \left\{ J_\beta^Y \right\}}. \qquad (4)$$

For a fixed Y, the margin operator M_K^Y attempts to quantify the quality of a configuration with K clusters with respect to its closest competitor. Therefore, the larger the value of the margin operator becomes (for any positive Y), the better the cluster configuration is. Thus, by maximizing the margin value with respect to Y, as in

$$M_K^* = \max_{Y>0} \left\{ M_K^Y \right\}, \qquad (5)$$

it becomes possible to quantify the quality of a clustering configuration with K partitions in terms of cluster compactness and the degree of decorrelation of its variables. The higher the margin, the better the configuration is with respect to the others. Hence, the margin measures the quality of the configuration. The maximum jump peaks must have margin values greater than unity (see Fig. 3).

After solving the eigenproblem posed by the spectral clustering (SC) algorithm for a number of training points, the cluster indicator vectors (i.e., eigenvectors) and the extended kernel matrix are used to expand the solution to the rest of the points [8]. To find the labels for K clusters, one performs kernel principal component analysis (kPCA), solving for at least the $K - 1$ eigenvectors associated with the highest eigenvalues. The ith entries of these eigenvectors represent the components of the ith point in a new $K - 1$ dimensional space. In principle, when using a properly parameterized radial basis function (RBF) kernel, the points in the same cluster become collinear in this new space. Thus, by projecting all points onto the unit hyper sphere, one can use spherical k-means [8] and compute the required values to build the d_K curve (see Fig. 4).

Using the d_K information, one can calculate the corresponding M_K^* for each configuration tested (i.e., $M_{K=0}^*$, $M_{K=1}^*$, ..., $M_{K=K\,max}^*$). A simple comparison among the values of M_K^* is sufficient to determine a suitable number of clusters and appropriate parameters for the construction of the kernel matrix (i.e., model selection). The clustering is readily obtained from the identified parameters.

3.4   The Image Segmentation Technique: Application of IJM

In our IJM experiments, the RGB uterine cervix image is first transformed to the $L^*a^*b^*$ space, which is a more perceptual color space (this step is not applied to non-visual medical data) [9]. Then, the image color levels are quantized to only N_q levels using the criteria of minimum variance. Next, the color-quantized image is subdivided into windows centered at each pixel, and a normalized histogram is computed for each window. A χ^2 score can then be used to compare the pixels as computed from these histograms [10]. The χ^2 score is defined as

$$\chi_{ij}^2 = \frac{1}{2} \sum_{k=1}^{K} \frac{\left[h_i(k) - h_j(k) \right]^2}{h_i(k) + h_j(k)}, \tag{6}$$

where h_i and h_j are histograms for the windows located on the ith and jth pixels, and $h_i(k)$ indicates the repetition frequency of the kth bin (out of N_q bins) in the corresponding histogram for the ith pixel. Finally, a χ^2 (chi-squared) RBF kernel and its kernel matrix can be computed from $k(h_i, \ h_j) = \exp(-\chi_{ij}^2/\sigma_\chi)$, where $k(h_i, h_j)$ is the element in the ith row and jth column of the kernel matrix, and σ_χ is the kernel standard deviation, as determined previously from the model selection.

The information theoretic IJM is applied to image segmentation by examining the local statistics of an indexed version of the image. Corona details many tests of IJM on synthetic data and evaluates this method's performance on standard clustering benchmark datasets in his dissertation [6]. The performance of the IJM in segmenting uterine cervix images is also presented. This application is important
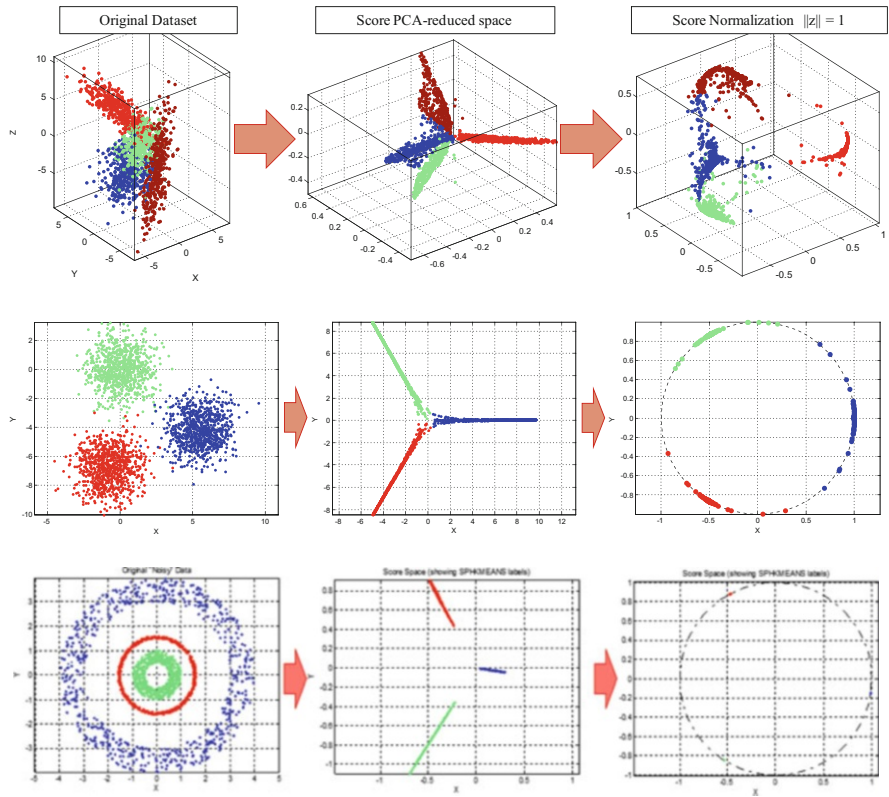
Fig. 4 The collinearity property in the new space (*middle boxes*) approximately transforms each of the clusters into lines radiating from the origin. After mapping these collinear points onto the surface of a hyper sphere (*rightmost boxes*), the clustering problem in most cases reduces in complexity, and an algorithm such as spherical k-means can be applied to find the d_K curve and, subsequently, the number of clusters. The quality of the line formation depends on the RBF parameter σ and the shape of the original clusters. One can sweep σ and select the value that generates the best clusters according to the IJM criteria. The $p = 3$ raw data in the *top row* is highly correlated with $N = 400$ in each cluster, and it would be handled poorly by standard k-means. The $p = 2$ raw data in the *bottom row* could not be correctly segmented by any linear algorithm such as k-means [6, 7]

because these kinds of medical images have no established ground truth, and even segmentations derived by experts differ due to subjective bias. It remained for IJM to be validated for an application involving medical images. This validation is carried out with MRI datasets of the human brain that have a labeled ground truth available.

4 IJM Validation Results

The IJM has been shown to perform well with benchmark clustering datasets. This is remarkable because IJM proves to be one of the few methods available that can cluster real-world data without knowing the number of clusters a priori. We desire eventually to apply the IJM to the segmentation of medical images to aid in the creation of automated diagnostic tools. One such application proposed here is the segmentation of uterine cervix images. However, the remaining issue is that the pathological uterine cervix images lack an established ground truth. Therefore, there is a need to validate the application of IJM to medical images, using medical image data with an available ground truth. This validation is done with a standard benchmark MRI dataset.

4.1 MRI Benchmark Test

Experiments are conducted with the standard benchmark MRI data of the human brain (BrainWeb: Simulated Brain Database; http://www.bic.mni.mcgill.ca/brainweb/), where the gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) are segmented from the background using the proposed technique. This dataset contains MRI cross-sectional images of a normal brain with no noise or field inhomogeneity [11]. The MRI raw data is given in terms of T_1, T_2, and the proton density, which were all available in both 8-bit and 12-bit formats, where the data has been stretched from the minimum to maximum value. These three modalities can be used to generate a pseudo-RGB space [12]. Analysis of the means and variances of the data for the various tissues shows that the space is not fully spanned, so an appropriate pseudo-color transformation can be performed on the raw data to improve the perception between the various tissue types (see Fig. 5; [7]).

The same procedure is followed for the MRI experiment as for the uterine cervix images. However, in this case, the number of clusters is selected using the second highest margin, instead of the maximum margin [6, 10]. The reason for this is the hierarchical nature of the results. Put simply, the maximum margin M_K^* reflects the most obvious cluster configuration in terms of cluster separation. The second highest margin is associated with less obvious and, sometimes, more interesting configurations. For example, the background and foreground often tend to strongly separate into two clusters, while we are actually interested in different types of foreground. Prior to the application of the clustering algorithm, the MRI data is preprocessed [11, 13]. The anatomical labels are used to create a mask that automatically removes most of the skull. The three modalities can be treated as a quasi-RGB image [12]. The original dataset does not span the entire range of possible values, so a transformation (to a pseudo-color space) is made to stretch out the ranges and to increase the perceptibility between the different tissue types. The transformed image is segmented using the procedure previously described (without
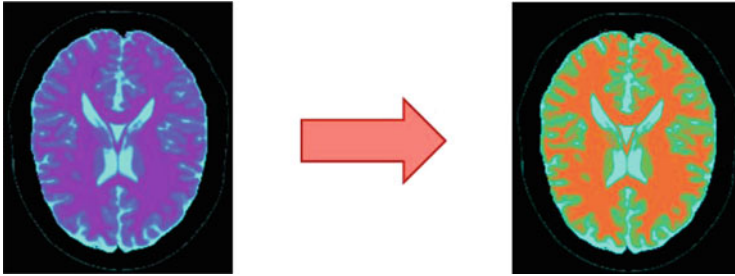
Fig. 5 The original eight bit MRI data in pseudo-RGB (*left*) transformed into a perceptually enhanced image (*right*)
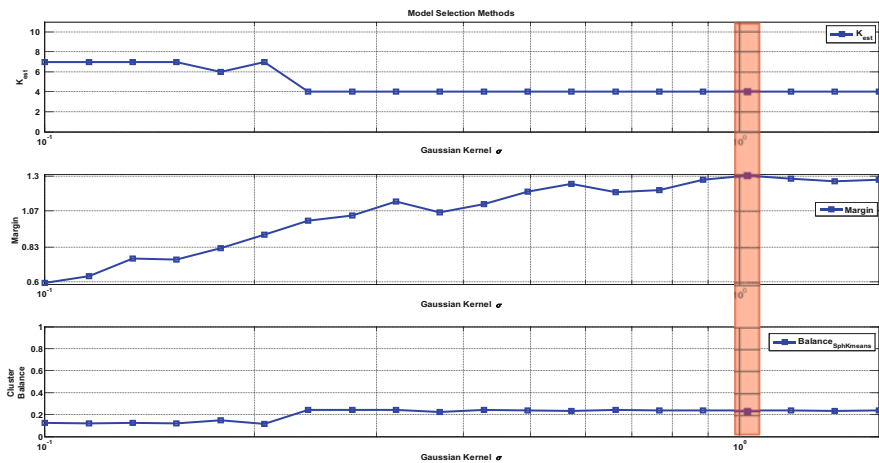

Fig. 6 Model selection methods as applied to an MRI of the human brain. The *top plot* displays the number of clusters K_{est} using the second best margin scores, where 15 out of 20 are for $K = 4$. (the best margin scores are always for $K = 2$. due to the striking difference between the tissue data and the background). The middle plot is of the corresponding margin scores, where the maximum margin at $\sigma = 1.025$ is highlighted in red (also for $K = 4$.). The *bottom plot* is the cluster balance. Note the 20 trial standard deviation values are on a logarithmic scale [7]

the $L^*a^*b^*$ transformation). The anatomically labeled data can provide a ground truth to compare with the segmentation results of the IJM.

Figure 6 shows an example output of the model selection plots for such an MRI; the number of clusters with the second highest margin scores is $K = 4$, with a peak value located at the kernel standard deviation of $\sigma = 1.025$ in this run. Due to internal randomness of the training and testing sets, the algorithm's results do somewhat vary from run to run. Recall that the jump method is hierarchical, so the second peak in J_K^Y can correspond to a valid sub-clustering of the data. For the model selection plots, the horizontal axes represent the standard deviation of the RBF kernel used. The top plots represent the selection of the number of clusters,
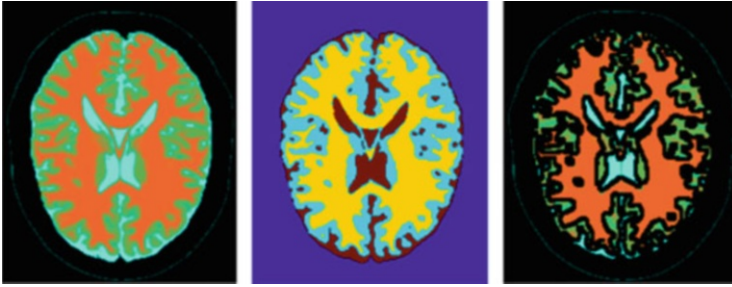
Fig. 7 The original MRI after preprocessing (*left*), the false-colored cluster labels (*middle*) and the resulting segmented image (*right*) [7]

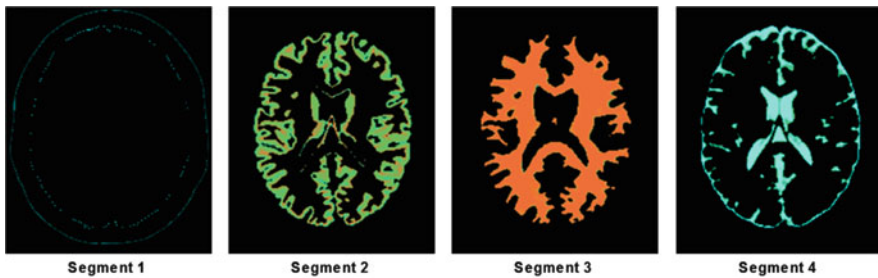Segment 1   Segment 2   Segment 3   Segment 4

Fig. 8 The successful segmentation results from applying the IJM algorithm, with the background (segment 1), the gray matter (segment 2), the white matter (segment 3), and the cerebrospinal fluid (segment 4) [7]

called K_{est}. The middle plots depict the values for the margin, while the bottom plots show the cluster balance values. The cluster balance is a measure of uniformity of the size of the clusters. Equal-size cluster configurations have a balance equal to one. Very different cluster sizes are associated with a balance close to zero. Figures 7 and 8 illustrate an example of MRI segmentation using the parameters selected from the model selection plot in Fig. 6. The confusion matrix for this experiment is shown in Table 2. The segmentation results using only eight color levels and a window size of three pixels for the four clusters are 100 % correct for the background, 94.58 % correct for the white matter, and only 92.75 and 90.82 % correct for the CSF and gray/glial matter associated segments, respectively. The overall accuracy is 95.48 %.

The IJM has been shown to perform at 95 % agreement with a labeled MRI dataset, taking into account hierarchical issues. An MRI of the human brain has a distinct background, so it is most obviously clustered into two segments; the sub-clustering of the brain tissue is the issue of importance. The variation in illumination in the uterine cervix images is handled by factoring in cluster balance. The segmentation of the human brain MRI neglects the consideration of cluster balance, yet it still achieves acceptable results.

Table 2 The confusion matrix for MRI segmentation using a window size of three pixels, $N_q = 8$, and χ^2 RBF kernel with standard deviation of $\sigma = 1.025$

True/predicted	Segment 1	Segment 2	Segment 3	Segment 4	Accuracy (%)
Background	11,234	0	0	0	100.00
Gray matter	4	6,790	120	562	90.82
White matter	0	508	8,861	0	94.58
CSF	127	73	0	2,560	92.75

The overall accuracy is 95.48 % [7]

4.2 Segmentation of Color Uterine Cervix Images

In the uterine cervical images, changes in three distinct and usually concentric regions are of importance to specialists for early detection of cervical cancer. Because cervical cancer slowly develops, it causes a detectable and treatable precursor condition known as cervical intraepithelial neoplasia (CIN). CIN causes abnormal regions and changes in the cervix that are of interest to specialists for several reasons. The columnar epithelium (CE) region immediately surrounds the central cervix opening, and it is near the center of the image. The CE region is bright red and is usually the darkest in the luminance plane. The acetowhite (AW) region is the most common abnormality immediately surrounding the (CE) region. It appears white or very pale pink and is usually the brightest region in the luminance plane. The outermost region is the squamous epithelium (SE), which has a reddish hue usually between that of AW and CE [6]. The acetowhite region is produced by applying a 5 % acetic acid solution to the cervix, which causes the abnormal cells to become turgid (swollen) and increases their reflectivity relative to the normal CE. They become turgid partly because their cellular machinery is not working normally, causing them to produce different kinds of cytokeratin (cellular structural proteins), hence they are unable to hold their normal shape [14].

The IJM was also used to segment color uterine cervix images into these three regions for an abnormal case. The model selection results are shown in the form of plots (see Fig. 9) as described in the previous subsection. Figure 10 shows an example of uterine cervix image segmentation. The model selection plots indicate that $K = 3$ is a good configuration, which is located at $\sigma = 0.866$. The segmentation using these parameters is shown in Fig. 11. It should also be noted that in the generation of this segmentation cluster balance was also factored in. Cluster balance is a measure of size similarity between clusters. For MRI data, one does not expect the clusters to similar in size to each other, while for uterine cervix images one might. Thus, the cluster balance was calculated for each configuration and multiplied by the margin to determine the overall quantity to be maximized. This balancing tends to stabilize the segmentation of the uterine cervix images against smaller sub-clusters of the same region, such as the brightest region in Segment 2 of Fig. 11.
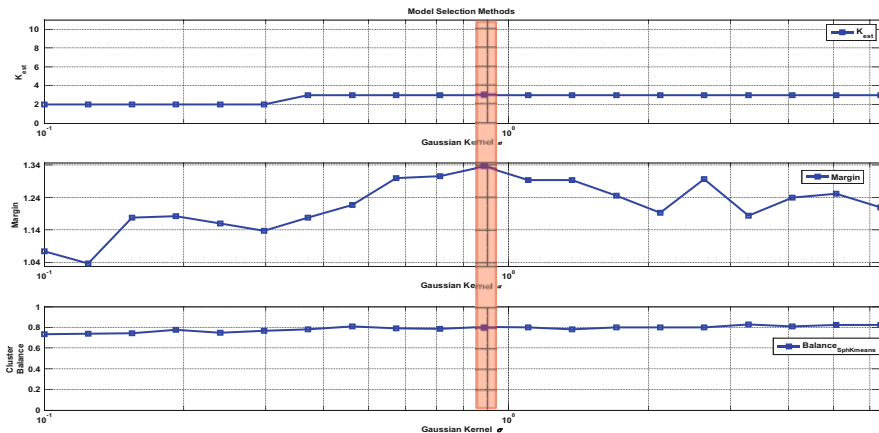
Fig. 9 Model selection methods as applied to a uterine cervix image. The *top plot* displays the number of clusters K_{est} with the overall best margin scores, where 15 out of 20 are for $K = 3$. The *middle plot* is of the corresponding margin scores, where the maximum margin at $\sigma = 0.886$ is highlighted in *red* (also for $K = 3$). The *bottom plot* is of the cluster balance. Note that the 20 trial standard deviation values are on a logarithmic scale [7]
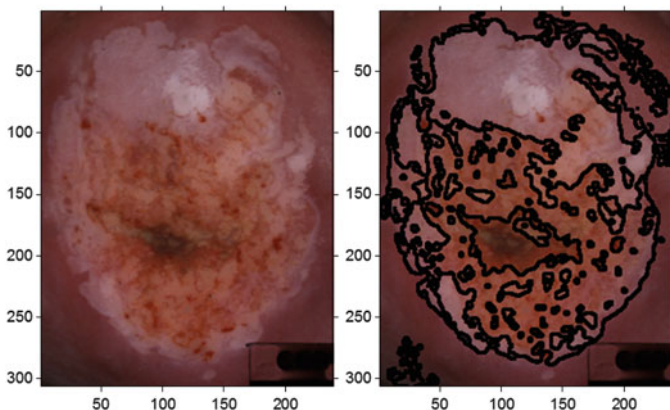

Fig. 10 The original uterine cervix image after preprocessing (*left*) and the segmented image (*right*) [7]

The direction of ongoing and future work is two-pronged. One is to apply the IJM to the segmentation of MRIs of abnormal human brains, and the other is to apply the IJM to the segmentation of cell images. Abnormal brain features tend to be small and distinct. This poses quite a challenging problem for the current implementation of the IJM. Evaluations have identified limitations of the IJM in such cases. To overcome these limitations, an iterative hierarchical approach is suggested along with a dynamic window size. These limitations and this proposed solution are detailed below.
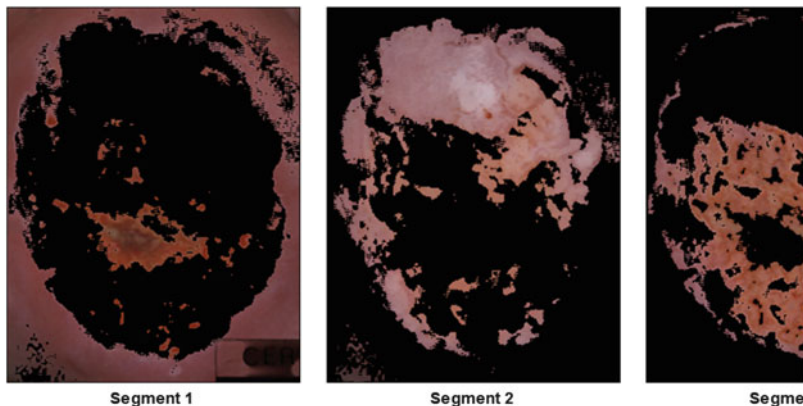
Fig. 11 Demonstration of uterine cervix image segmentation using the model selection in IJM [7]

In addition, much work has been done in automating the entire process for diagnosing cervical cancer. In addition to segmenting the uterine cervix images, there is also a need to preprocess and segment micrograph slides of cell smears. Many methods have been applied to this problem, and there is an aim to bring the IJM to bear on the preprocessing and segmentation of such cell images. This technique could aid in the development of a diagnostic tool for cervical cancer.

5 MRIs with MS Lesions

5.1 The IJM Failure with MRIs with MS Lesions

After successfully segmenting MRIs of a normal brain with the IJM, the next step of the validation process is to attempt to segment the MRI of a brain with multiple sclerosis (MS) lesions. This is a more challenging segmentation problem not just because of the existence of another cluster ($K = 5$ is expected), but also because the MS lesions are relatively small compared to the other tissue areas. Upon performing experiments, indeed five clusters are weakly indicated with the third highest margin ($K = 8$ is the found for second highest margin), however, the segmentation results are incorrect (see Figs. 12 and 13). The gray matter segment splits into the two additional transition zones between gray matter/white matter, and the gray matter/CSF, respectively, to form five clusters; the MS lesions remain undistinguished. The bulk of the MS lesion is included with the CSF cluster, while the transition zone between the MS lesion and the white matter is included with the transition zone between the gray and white matter. This failure prompted a study of the limitations of the implementation of the IJM.
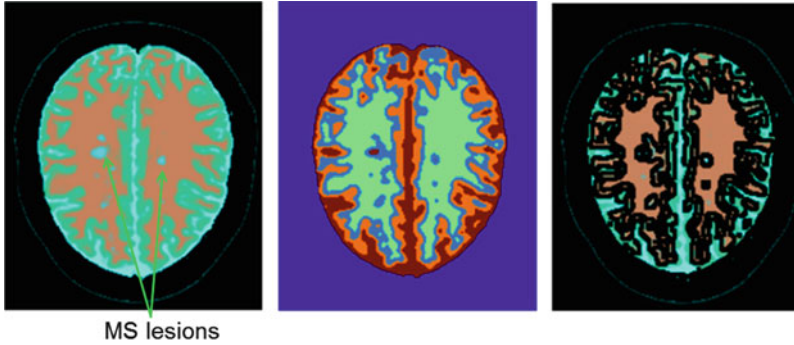
MS lesions

Fig. 12 The original MRI with MS lesions after preprocessing (*left*), the false-colored cluster labels (*middle*), and the resulting segmented image (*right*) [15]
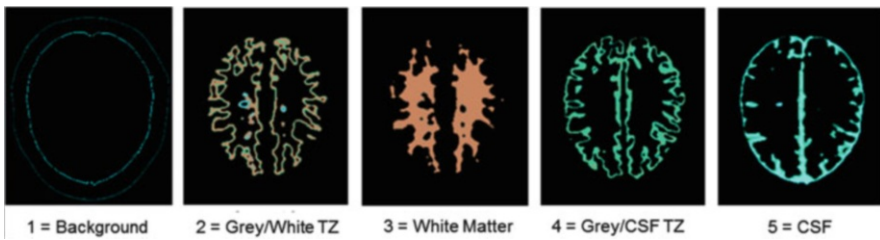

1 = Background   2 = Grey/White TZ   3 = White Matter   4 = Grey/CSF TZ   5 = CSF

Fig. 13 The failed segmentation results from applying the IJM algorithm, with the background (segment 1), the gray matter/white matter transition zone (segment 2), the white matter (segment 3), the gray matter/CSF transition zone (segment 4), and the cerebrospinal fluid (CSF) (segment 5) [15]

5.2 Limitations of the Proposed Image Segmentation Technique

In the above segmentation algorithm, the image color levels are quantized to only N_q levels; this color-quantized image is subdivided into windows centered at each pixel, and a normalized histogram is computed for each window. The estimation of the distortion and hence the margin operator is then calculated from these histograms. One limitation of this implementation is that the number of quantization levels must be less than or equal to the number of pixels in a window. Two issues are at the root of the failure of this IJM implementation. One issue is due to the size of the window and the other is due to the number of quantization levels.

Relying on the window to determine local statistics is good for a relatively smoothly varying image, such as the uterine cervix images. However, for images with a sharp background, discontinuities, or rapidly changing edges, the local statistics can become peculiar in the transition zones between regions. If such

boundaries are long and crenelated like those between the internal tissues of the brain, then local statistics from the windows can "trick" the algorithm into detecting another statistically significant segment. This new segment is the aggregate of the transition zone, which in the case of the MRIs can have an area larger than the area of the MS lesion tissue. This pseudo-clustering tendency also helps to explain an unexpected observation. Sometimes improving the perception contrast between regions can actually lead to poorer segmentation results. Again this is due to having more sharp transition zones. Decreasing the window size will minimize this effect for all pixels, except for those directly on the edge boundary.

Reducing the window size can mitigate the transition zone effect, but this can have an unintended consequence; it reduces the number of quantization levels. This reduction is significant because the ability to discern between regions in an image is limited by pixel indexing. Limiting quantization goes back to the core issue of this application of information theory, storage efficiency versus the ability to discriminate. As it turns out, the normal brain MRIs are handled well with a 3×3 window with eight quantization levels (labeled as zero through seven). Examining the raw index image shows that the different tissues are distinctly labeled. However, examining the MS lesion image reveals that the MS lesions are not distinctly labeled (with index 19) until there are at least 20 quantization levels. This number of levels is required to accommodate the transition zones between regions and the MS lesion tissue. All the various color regions that are comparable in size to the MS lesion area would bring the effective cluster number to at least seven or eight. Thus the MS lesions are not even distinguishable unless at least a 5×5 window is used. However, using a window size and number of quantization levels so large also turns the transition regions into statistically significant pseudo-clusters. The current IJM implementation here fails in the proper segmentation of the MS lesions. This issue was not previously encountered when applying the IJM to synthetic test data and medical images.

5.3 The Proposed Solution: An Iterative Hierarchical Scheme

One means to deal with the problematic transition zones is to divide them before they become a problem. Hence, this approach is to "divide and conquer." This scheme would involve utilizing the hierarchical information of the IJM in a different manner. Recall, that for the MRIs, the largest margin indicated $K = 2$ clusters. This result suggests that the correct approach is to first segment the MRI into two clusters, "cleaving" the background from the brain tissue cluster. Next, one reapplies the IJM to further segment the brain tissue super-cluster using a masked image, etc. In such an iterative scheme, the various segments can be progressively cleaved from each other, dividing the transition zones along the way and making them less significant. This process can be continued until some parameter threshold or statistical feature of the subsequent segmented image signals the iterative program to stop. This process could be automated to iteratively segment the image according to the number of clusters with the largest margin as obtained from the IJM, using the full power of its
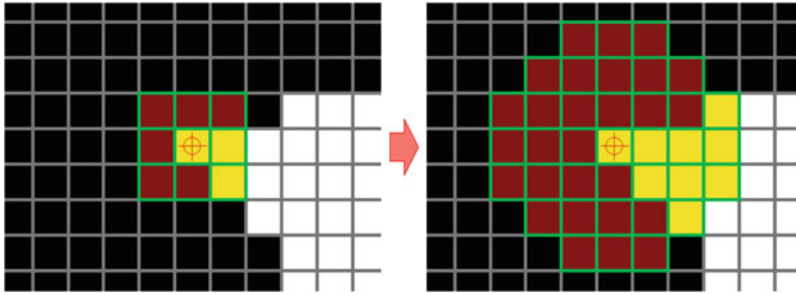
Fig. 14 *Left*: The original 3×3 window (*yellow*) superimposed on a sample point selected from the mask (*white*). *Right*: The window adaptively enlarged to include nine local points to form a comparable local histogram

hierarchical character. In this manner, the quantization restriction can be partially decoupled from the transition zone problem.

5.4 An Adaptive Window Size

The problem with iteratively segmenting an image is that the shapes of the future iterations of image masks are unknown. Consider the cases of a small "peninsula" or "island" of pixels with fewer neighbors than the window size (see Fig. 14). Another case would be an edge with a reduced number of neighbors. These samples near the mask boundary would degrade the subsequent statistical analysis of the histograms. Furthermore, as the sub-clusters get smaller and demonstrate more complicated edges, the performance of the IJM will deteriorate. In a sense, the transition zone problem is replaced with a mask edge effect problem.

The root of this problem is that for pixels near an already cleaved boundary, the effective window size has changed. In order to counter this effect, the window size must be adaptive for the segmentation of the masked images. The purpose of the window is to provide a consistent pool of neighboring pixels in order to characterize the local statistics. When segmenting a single image into multiple clusters, every sample pixel has the same window size; those on the edge of the image can be used with simple symmetric padding. However, this uniformity is broken in the subsequent steps of the iterative hierarchical scheme, and the effects of quasi-symmetric extrapolations could actually amplify the transitions near boundaries. In order to carry out the original purpose of the window, its size must be adaptive for the masked images. The suggested implementation enlarges the original square window into a successively more circular shape until a sufficient number of neighboring pixels are sampled.

The following description outlines the procedure to adaptively enlarge the window size. A pixel contained within the mask is selected. The initial window

is applied, and the number of existing neighbors is counted. If the number falls below the default window size, then the window size is enlarged. Allow at most three enlargements or until the default window size is passed. Randomly remove any excess pixels or randomly duplicate the pixels found until the window has the same number of pixels as the default window size. Now the histogram statistics for almost every sample pixel derived from the segment mask should be comparable. In this manner, the effects of transition zones are again reduced, because more of the pixels in the "bulk" are included in the histograms of sample pixels near boundaries. This procedure should also allow relatively small clusters like the MS lesion tissue to be robustly detectable, provided sufficient quantization levels. Hence, using an iterative hierarchical scheme and an adaptive window size should allow the proper segmentation of MRIs containing MS lesions using the IJM with at least 20 quantization levels.

6   Application to Pap Test Images

Another medical segmentation imaging problem of interest is the processing of cytological images in separating the nucleus from the cytoplasm. An example of this is seen in the micrographs of stained cells derived from the Papanicolaou (Pap) smear test. In the diagnosis of cervical cancer, cervical uterine photographic images are examined, and if a suspicious region is identified, then it is swabbed. The collected cells are then smeared onto a slide, stained, magnified through a microscope, and photographed. Precancerous or cancerous cells often exhibit a condition called polyploidy, where they have multiple copies of the same chromosome in their nuclei. Polyploidy usually causes their nuclei to be larger than normal. Hence, an automated system that can measure the ratio of the area of the nuclei to cytoplasm in Pap test images could be useful as a diagnostic aid. There are many approaches to such a system; one is reviewed below.

Current routine clinical tests for cervical cancer include the cytology-based Pap smear test and HPV testing to identify DNA or RNA in cervical cells. Because the recently developed HPV vaccination does not protect against all HPV types, cervical screening tests are still required [16]. Because cervical cancer develops slowly and has a precursor condition known as dysplasia or CIN, it can be detected through screening of at-risk women and treated to prevent the development of invasive cancer [17]. However, the true sensitivity of the conventional Pap smear has been reported to be as low as 51 %, requiring annual screenings to detect CIN, making the Pap smear impractical to implement in resource-poor regions [18, 19]. The development of automated diagnostic tools could increase the universal effectiveness and availability of the Pap test.

Several combinations of imaging modalities and classification algorithms have been used by researchers to detect some of the abnormalities due to CIN [20]. However, the problem of isolating and segmenting specific abnormalities, such as mosaicism and punctations, is not addressed by most methods. In previous work,

an automated computer vision system has been developed for the detection and characterization of CIN in an attempt to overcome the inadequacies of existing approaches used for automated diagnosis of CIN [9]. This system is capable of automatically segmenting the cervix region of interest (ROI) from raw cervico-graphic images; removing specular reflection (SR); segmenting the cervix ROI into acetowhite (AW), columnar epithelium (CE), and squamous epithelium (SE); classifying AW regions into AW, mosaic, or punctation tiles; segmenting mosaic and punctation from AW tiles; and assessing of disease severity [9, 18].

We intend to utilize the IJM [6] for precise segmentation of the macro-features as well as a hybrid watershed model for segmentation of cytological cell images acquired by Pap test to assess abnormalities in the cell nuclei [21]. The significance of including the cell structure analysis is to provide a cross-validation of cervical abnormalities assessed by two different approaches, namely from the analysis of the color cervix images and of the cervical cell images. Such an automated system for analyzing both cervix images and cervical cell images would provide an improved diagnostic tool for cross-validating the results from both imaging modalities. Based upon the validation of the IJM unsupervised learning algorithm, the high cost of generating manual ground truth from specialists, who are not in general readily available in resource-poor regions, can be reduced [18]. With the proposed modifications to its implementation discussed in Sects. 5.3 and 5.4, it should be possible to segment these images using the IJM. The clusters in the Pap test images have sharp edges and are better balanced than the MS lesion case in the MRIs.

6.1 A Hybrid Watershed Method for Generating Cervical Cell Profiles

A combination of global histogram thresholding and local watershed segmentation was used to generate the cervical cell profiles for final separation of the cell nuclei from the cell cytoplasm to assess the abnormalities in the cells. The preprocessing steps are described in detail in reference [21]. The selection of useful cervical cell images to be analyzed can be automatically achieved by the IJM clustering, and subsequently the cell nuclei can be precisely segmented. Figure 15 shows some examples of cervical cell segmentation results with this technique.

6.2 Experimental Results

All uterine cervix images and cervical cell cytology images were provided by the National Library of Medicine from the data collected by the National Cancer Institute at the National Institutes of Health. Figure 15 shows typical segmentation
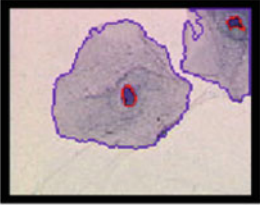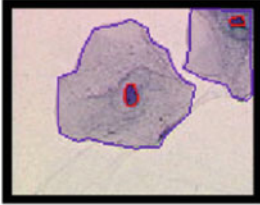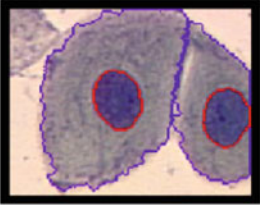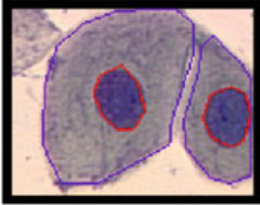
Fig. 15 Examples of segmentation of cervical cell images and the corresponding Dice coefficients (D_C refers to cell Dice coefficient and D_{CY} refers to cytoplasm Dice coefficient) as measures of segmentation quality [18, 21]

results of cervical cell images with semiautomated and manual segmentation, where corresponding Dice coefficients [22] are used as the segmentation quality measures. Dice's coefficient is defined as:

$$D = \frac{2 \times |S_{\text{Auto}} \cap S_{\text{Manu}}|}{|S_{\text{Auto}} + S_{\text{Manu}}|}, \tag{7}$$

in which S_{Auto} is the pixel set from automatic segmentation, and S_{Manu} is the pixel set from manual segmentation. The proposed IJM algorithm is expected to improve these segmentation results by automatically identifying and discarding the background and objects of no interest [18].

7 Conclusion

7.1 Further Development

The application of information theoretic techniques such as the IJM to medical image analysis is promising. The current outlook is to successfully apply the IJM to the segmentation of MRIs of abnormal human brains, such as brains with MS lesions. One scheme could involve scanning through the MRI slices for MS

lesions using IJM. Once found, a simple Gaussian model of the pixel values can be constructed and applied to the full volume, allowing an estimation of the total MS lesion volume that can be tracked over time. Alternatively, a three-dimensional voxel segmentation of the MS lesion could be developed by applying the IJM.

Finally, the IJM also is promising in its application to segmenting cell images, such as those derived from the Pap smear test. In the preprocessing step, the IJM could aid in discarding "junk" slide images from those containing useful images. When useful images are found, the IJM offers an additional method to segment such images into nucleus and cytoplasm clusters. Cell overlaps could even be identified to correct the estimation of the cytoplasm area. Then the ratio of nuclear to cytoplasm area can be calculated and used as a diagnostic aid. These steps combined with the uterine cervix image segmentation described earlier can be used as components of an automated, integrated diagnostic system for cervical cancer.

7.2 Outlook for Application

The IJM effectively performs an optimized transformation of the distortion-rate curve D(R) to reveal a distinct signature of the cluster configurations in an unlabeled image. This improved clustering technique, based on an information theoretic approach, demonstrates high accuracy in segmentation of the standard benchmark MRI data of a normal human brain.

Limitations of the current implementation of the IJM are found that cause it to fail to correctly segment the MRI data of a human brain with MS lesions. The root source of the failure was identified as the joint effect of the requirement for a relatively high number of quantized levels (20) and the problematic transition zones between tissues that can be sources of pseudo-clusters. An iterative hierarchical scheme with adaptive window sizes is proposed to make the IJM more robust in its segmentation.

This chapter discusses an element of a computer vision system with embedded novel segmentation models for analyzing both cervix and cervical cell images for detection of precancerous lesions. Application of the proposed IJM in segmentation of uterine cervix images, with ground truth unavailable in most cases, has potential to identify abnormal tissue characteristics at an early stage, thus providing a diagnostic aid to assessment of uterine cervical cancer. Such a system allows cross-validation of the CIN abnormalities detected by two different testing methods and may reduce frequent false-positive results obtained by the routine Pap test. This system may also provide a low-cost yet valuable analysis tool to aid in preventive assessment of cervical cancer in resource-poor regions [18].

References

1. Ghabhramani, Z.: "Unsupervised Learning." Advanced Lectures on Machine Learning, pp. 72–112. Springer, New York (2004)
2. Gokcay, E., Principe, J.C.: Information theoretic clustering. IEEE Trans. Pattern Anal. Mach. Intell. **24**(2), 158–171 (2002)
3. Bartlett, M.S.: Face Image Analysis by Unsupervised Learning. Kluwer Academic, Boston (2001)
4. Becker, S., Plumbley, M.: Unsupervised neural network learning procedures for feature extraction and classification. J. Appl. Intell. **6**, 1–21 (1996)
5. Sugar, C.A., James, G.M.: Finding the number of clusters in a dataset: an information-theoretic approach. J. Am. Stat. Assoc. **98**(463), 750–763 (2003)
6. Corona, E.: Unsupervised learning methods: an efficient clustering framework with integrated model selection. Ph.D. Dissertation, Texas Tech University, Lubbock (2012)
7. Corona, E., Hill, J., Ao, J., Nutter, B., Mitra, S.: An information theoretic approach to automated medical image segmentation. In: Proceeding of SPIE, vol. 8669. Medical Imaging 2013: Image Processing (2013)
8. Alzate, C., Suykens, J.A.K.: Multiway spectral clustering with out-of-sample extensions through weighted kernel PCA. IEEE Trans. Pattern Anal. Mach. Intell. **32**(2), 335–347 (2010)
9. Srinivasan, Y., Corona, E., Nutter, B., Mitra, S., Bhattacharya, S.: A unified model-based image analysis framework for automated detection of precancerous lesions in digitized uterine cervix images. IEEE J. Sel. Top. Signal Process. **3**(1), 101–111 (2009) (Special Issue on Digital Image Processing Techniques for Oncology)
10. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the Nystrom method. IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 214–225 (2004)
11. Kwan, R., et al.: An extensible MRI simulator for post-processing evaluation. In: BrainWeb: Simulated Brain Database. Proceedings of VBC'96, pp. 135–140. Springer (1996). Available online: http://www.bic.mni.mcgoll.ca/brainweb/
12. Lecoeur, J., Wang, F., Chen, L.M., Li, R., Avison, M., Dawant, B.: Sub millimeter coregistration of functional maps across imaging sessions. In: Proceedings of ISMRM Annual Conference, vol. 1625, Montreal, May 2011
13. Ge, Z., Mitra, S.: Optimized statistical modeling of MS lesions on MRI voxel outliers for monitoring the effects of drug therapy. Proc. SPIE **4684**, 873–882 (2002)
14. Maddox, P., Szarewski, A., Dyson, J., Cuzick, J.: Cytokeratin expression and acetowhite change in cervical epithelium. J. Clin. Pathol. **47**(1), 15–17 (1994)
15. Hill, J.E.: Application of information theoretic unsupervised learning to medical image analysis. Master Thesis, Texas Tech University, Lubbock (2013)
16. Ferlay, J., et al.: GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. IARC CancerBase No. 5. Version 2.0. Online. Available: http://www.iarc.fr/ (2002)
17. Koss, L.G.: The Papanicolaou test for cervical cancer detection. A triumph and tragedy. JAMA **261**, 773–774 (1987)
18. Corona, E., Hill, J., Ao, J., Nutter, B., Mitra, S.: A novel unsupervised learning model for automated detection of precancerous abnormalities in uterine cervix with unified analysis of cervical cells and digital uterine cervix images. In: First IEEE Healthcare Technology Conference: Translational Engineering in Health & Medicine, 7–9 November 2012. IEEE HIC 2012 Paper Abstract, Paper ThCT2.13 (2012)
19. Ferris, D.G.: Cervicography-an adjunct to Papanicolaou screening. Am. Fam. Physician **50**, 363–370 (1994)
20. Gordon, S., Zimmerman, G., Long, R., Antani, S., Jeronimo, J., Greenspan, H.: Content analysis of uterine cervix images: initial steps towards content based indexing and retrieval of cervigrams. Proc. SPIE Conf. Med. Imaging **6144**, 1549–1556 (2006)

21. Ao, J., Mitra, S., Long, R., Nutter, B., Antani, S.: A hybrid watershed method for cell image segmentation. In: Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation, Santa Fe, 22–24 April 2012
22. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)

Multiobjective Differential Evolution-Based Fuzzy Clustering for MR Brain Image Segmentation

Indrajit Saha and Ujjwal Maulik

Abstract The problem of classifying magnetic resonance (MR) brain images into different tissue classes has gained significant attention during the last couple of decades. In this regard, unsupervised clustering methods are used to group the pixels of images in the intensity space. Most of these clustering methods optimize a single clustering measure. However, to improve the clustering further multiple conflicting clustering measures can be used. In this chapter, the *Multiobjective Differential Evolution based Fuzzy Clustering* technique is used to optimize multiple clustering measures simultaneously. Here differential evolution is used as the underlying optimization technique and the cluster centers are encoded in a vector of differential evolution. In the final generation, it produces a set of nondominated solutions, from which the most promising one is selected based on a *Minkowski Score* as a final clustering solution. The proposed technique is applied on several simulated T1-weighted, T2-weighted, and proton density for normal and multiple sclerosis lesion magnetic resonance brain images. The performance of the differential evolution-based fuzzy clustering technique is demonstrated quantitatively by comparing it with other state-of-the-art methods. Apart from this, the segmented MR brain images produced by the proposed technique are also compared with the available ground truth information. Finally, a statistical significance test is conducted to judge the superiority of the results produced by the proposed technique.

I. Saha (✉) · U. Maulik

Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, West Bengal, India

e-mail: indra@icm.edu.pl; umaulik@cse.jdvu.ac.in

1   Introduction

Magnetic resonance imaging (MRI) has the capability of soft-tissue characterization and three-dimensional (3D) visualization, which helps in effective diagnosis. It can produce a sequence of multiple spectral images using three MR parameters, namely spin-lattice (T1), spin–spin (T2), and dual echo–echo proton density (PD). Moreover in vivo, the accurate classification of various brain tissues like gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), obtained from MR images using supervised and unsupervised algorithms has become an important tool for clinicians and researchers. For this purpose, during the last 15 years a number of supervised and unsupervised techniques for tissue segmentation have been developed. In the supervised approaches [10, 17, 27, 60, 63], a priori information is provided, which forms the basis of the segmentation process. Expectation maximization (EM) [27, 60] and support vector machines [64] belong to this class of algorithms. The unsupervised approach [37, 43] proceeds without any initial information, and segmentation is achieved with the information extracted from the image itself. Examples of this class of approaches include k-mean clustering [50], fuzzy c-mean (FCM) clustering, [2, 54, 62], etc. Apart from this, many other computer-assisted methods for analyzing MR images, such as orthogonal subspace projection (OSP) [58, 59], filter-based methods [57], neural networks [3, 28, 43], and knowledge-based techniques [9], have also been reported.

In general, supervised techniques for segmentation suffer from the drawback of manual intervention for providing a priori information. These approaches assume a particular intensity distribution function of the pixels which may not always accurately fit the actual image intensity distribution. While the unsupervised clustering method has high reproducibility because its results are mainly based on the information of the image data itself. Therefore, it doesn't require any assumption of the model as well as the distribution of the image data. In a MR brain image, a pixel represents an area inside the brain, which may not necessarily belong to a single tissue class. Hence, as a result a large amount of imprecision and uncertainty can be associated with the pixels in such images. Hence, it is natural to apply the principles of fuzzy set theory in the domain of pixel classification.

The use of FCM for MR brain image segmentation depends on its initial choice of centers, which may be trapped into some local optimum solution. To avoid this problem, meta-heuristic algorithms like differential evolution (DE) [52, 53], genetic algorithms (GA) [18, 19], and simulated annealing (SA) [24, 26] are used along with the clustering method [4, 5, 7, 29, 35]. Generally, these clustering techniques use a single objective to optimize. Simultaneous optimization of multiple cluster validity measures helps to cope with different characteristics of the partitioning and leads to higher quality solutions. This fact motivated us to use the Multiobjective Differential Evolution-Based Fuzzy Clustering technique, which optimizes both J_m [8] and Xie-Beni (XB) [61] indices simultaneously. Unlike single objective optimization, MODEFC provides a set of Pareto-optimal solutions. The best solution is selected from this set of solutions using a Minkowski Score [23]. The clustering results of

the MODEFC are reported in comparison with differential evolution-based fuzzy clustering (DEFC) [35], genetic algorithm-based fuzzy clustering (GAFC) [4, 7], simulated annealing-based fuzzy clustering (SAFC) [5], the well-known FCM [8] algorithm, and a widely used hierarchical clustering algorithm (average linkage (AL)) [56]. The effectiveness of the proposed technique is established quantitatively, visually, and statistically.

2 Basics of Multiobjective Optimization

In this section, the mathematical formulation of multiobjective optimization (MOO) [11, 12, 14–16, 21, 25, 51, 65], nondominated sorting [13], and the crowding distance [13] measure are discussed.

2.1 *Mathematical Formulation of MOO*

The goal of multiobjective optimization is to find the vector $\overline{x}^* = [x_1^*, x_2^*, \ldots, x_n^*]^T$ of decision variables which will satisfy the l inequality constraints:

$$a_i(\overline{x}) \geq 0, \quad i = 1, 2, \ldots, l, \tag{1}$$

the u equality constraints

$$b_i(\overline{x}) \geq 0, \quad i = 1, 2, \ldots, u, \tag{2}$$

and optimizes the vector function

$$\overline{f}(\overline{x}) = [f_1(\overline{x}), f_2(\overline{x}), \ldots, f_h(\overline{x})]^T \tag{3}$$

The constraints in Eqs. (1) and (2) define the feasible region $\mathscr{F}$ which contains all the admissible solutions. Any solution outside this region is inadmissible since it violates one or more constraints. The vector $\overline{x}^*$ denotes an optimal solution in $\mathscr{F}$. In the context of multiobjective optimization, the difficulty lies in the definition of optimality, since it is only rare that one will find a situation where a single vector $\overline{x}^*$ represents the optimum solution to all the objective functions.

The concept of Pareto-optimality comes handy in the domain of multiobjective optimization. A formal definition of Pareto-optimality from the viewpoint of minimization problem may be given as follows: A decision vector $\overline{x}^*$ is called Pareto-optimal if and only if there is no $\overline{x}$ that dominates $\overline{x}^*$, i.e., there is no $\overline{x}$ such that

$$\forall i \in \{1, 2, \ldots, h\}, \ f_i(\overline{x}) \leq f_i(\overline{x}^*) \tag{4}$$

and

$$\exists i \in \{1, 2, \ldots, h\}, \ f_i(\overline{x}) < f_i(\overline{x}^*) \tag{5}$$

In other words, $\overline{x}^*$ is Pareto-optimal if there exists no feasible vector $\overline{x}$ which causes a reduction on some criterion without a simultaneous increase in at least another. In this context, two other notions viz. *weakly nondominated* and *strongly nondominated* solutions are defined [11]. A point $\overline{x}^*$ is a weakly nondominated solution if there exists no $\overline{x}$ such that $f_i(\overline{x}) < f_i(\overline{x}^*)$, for $i = 1, 2, \ldots, h$. A point $\overline{x}^*$ is a strongly nondominated solution if there exists no $\overline{x}$ such that $f_i(\overline{x}) \leq f_i(\overline{x}^*)$, for $i = 1, 2, \ldots, h$, and for at least one i, $f_i(\overline{x}) < f_i(\overline{x}^*)$. In general, Pareto-optimum usually provides a set of solutions called nondominated solutions.

2.2 Nondominated Sorting

In this approach [13], for each solution i of a set S, two entities are computed: (1) domination count n_i, the number of solutions which dominate the solution i, and (2) S_i, a set of solutions that the solution i dominates. At the end of this procedure, all solutions in the first nondominated front F_1 have their domination count as zero. Now, for each solution i with $n_i = 0$, it visits each member (j) of its set S_i and reduces its domination count by one. While doing so, if for any member j the domination count becomes zero then it is put in a separate list P. These members belong to the second nondominated front F_2. The above procedure is continued with each member of P and the third front F_3 is identified. This process continues until all fronts are identified.

2.3 Crowding Distance

Crowding distance [13] is used to get an estimate of the density of solutions surrounding a particular solution i in the population. It computes the average distance of two solutions on either side of solution i (i.e., $i + 1$ and $i - 1$) along each of the objectives. The crowding-distance computation requires sorting the population according to each objective function value in ascending order of magnitude. Thereafter, for each objective function, the boundary solutions (solutions with smallest and largest function values) are assigned an infinite distance value. All other intermediate solutions are assigned a distance value equal to the absolute normalized difference in the function values of two adjacent solutions. This calculation is continued with other objective functions. The overall crowding-distance value is calculated as the sum of individual distance values corresponding to each objective.

3 Multiobjective Differential Evolution-Based Fuzzy Clustering

In this section, the proposed Multiobjective Differential Evolution-Based Fuzzy Clustering (MODEFC) is discussed in detail.

3.1 Vector Representation and Initial Population

In MODEFC, the length of each vector is a sequence of real numbers representing the K cluster centers. For an d-dimensional space, the length of a vector is $l = d \times K$, where the first d positions represent the first cluster center, the next d positions represent those of the second cluster center, and so on. The K cluster centers encoded in each vector are initialized to K randomly chosen points from the data set. This process is repeated for each of the P vectors in the population, where P is the size of the population.

3.1.1 Evaluation of Objectives

Here the two objectives, J_m and XB, are associated with each vector. The J_m is computed as follows:

$$J_m = \sum_{k=1}^{K} \sum_{i=1}^{n} u_{k,i}^m D^2(v_k, x_i) \qquad (6)$$

where n is the number of data objects, K represents number of clusters, $u_{k,i}$ is cluster membership of ith point in the kth cluster, and m denotes the fuzzy exponent. $D(v_k, x_i)$ denotes the distance of point x_i from the kth cluster center v_k. The fuzzy membership of each data point is computed using the following equation:

$$u_{k,i} = \frac{\left(\frac{1}{D(v_k, x_i)}\right)^{\frac{2}{m-1}}}{\sum_{j=1}^{K} \left(\frac{1}{D(v_j, x_i)}\right)^{\frac{2}{m-1}}}, \quad \text{for } 1 \leq k \leq K, \ 1 \leq i \leq n \qquad (7)$$

for $1 \leq k \leq K$; $1 \leq i \leq n$, where $D(v_k, x_i)$ and $D(v_j, x_i)$ are the distances between x_i and v_k, and x_i and v_j respectively. The value of m, the fuzzy exponent, is taken as 2. Note that while computing $u_{k,i}$ using Eq. (7), if $D(v_j, x_i)$ is equal to zero for some j, then $u_{k,i}$ is set to zero for all $k = 1, 2, \ldots, K, k \neq j$, while $u_{j,i}$ is set equal to one.

For computing XB, a function of the ratio of the total variation Υ to the minimum separation Φ of the clusters is computed. Here Υ and Φ can be written as follows:

$$\Upsilon(U, V; X) = \sum_{k=1}^{K} \sum_{j=1}^{n} u_{kj}^2 \ D(v_k, x_j), \qquad (8)$$

and

$$\Phi(V) = \min_{j \neq i} D(v_k, v_i), \tag{9}$$

The XB index is then defined as

$$XB(U, V; X) = \frac{\Upsilon(U, V; X)}{n \times \Phi(V)} \tag{10}$$

Subsequently, the centers encoded in a vector are recomputed using the following equation:

$$v_k = \frac{\sum_{i=1}^{n} u_{k,i}^m \, x_i}{\sum_{i=1}^{n} u_{k,i}^m} \quad \text{for } 1 \leq k \leq K \tag{11}$$

From Eq. (6), it can be noted that J_m computes the global cluster variance, i.e., it considers the within cluster variance summed up over all the clusters. Thus lower value of J_m implies better clustering solution. On the other hand, the XB index in Eq. (10) is a combination of global (numerator) and local (denominator) situations. Therefore, the XB index is minimized by minimizing Υ and by maximizing Φ. In this chapter, the J_m and XB have been chosen as two objectives to be optimized.

3.2 Mutation

The ith individual vector of the population at time-stamp (generation) t has l components, i.e., $\chi_{i,l}^{(t)}$. For each target vector $\chi_{i,l}^{(t)}$, randomly three other distinct individuals, i.e., $\chi_{m,l}^{(t)}$, $\chi_{o,l}^{(t)}$, and $\chi_{p,l}^{(t)}$ from the same generation are opted. Thereafter, component wise difference is computed and scaled it by a scalar factor F (usually $\in [0, 1]$) to create a trial offspring $\varpi_{i,l}^{(t+1)}$. Therefore, the lth component of each vector is stated as follows:

$$\varpi_{i,l}^{(t+1)} = \chi_{m,l}^{(t)} + F\,(\chi_{o,l}^{(t)} - \chi_{p,l}^{(t)}) \tag{12}$$

3.3 Crossover

In order to increase the diversity in mutant offspring vectors, crossover is introduced. For this purpose, the trial offspring vector $\tau_{i,K}^{(t+1)}$ is formed as follows:

$$\tau_{i,l}^{(t+1)} = \begin{cases} \varpi_{i,l}^{(t+1)} & \text{if } \text{rand}_i\,(0, 1) \leq \text{CR or } i = \text{rand}(l) \\ \chi_{i,l}^{(t)} & \text{if } \text{rand}_i\,(0, 1) > \text{CR and } i \neq \text{rand}(l) \end{cases} \tag{13}$$

```
1: Vector Representation and Population Initialization.
  (Here each vector encodes K centres)
2: Evaluate objective values (Jm and XB) of each target vector.
Repeat
 3: Perform Mutation using Eqn. 12.
 5: Perform Crossover using Eqn. 13.
 6: Evaluate objective values (Jm and XB) of each trial offspring vector.
 7: Combine trial offspring pool with its target pool to create new population.
 8: Perform nondominated sorting for assigning rank.
 9: Compute the crowding distance between members of each front.
10: Select the vectors, same as population size, from the combined pool of
    population based on lowest rank and least crowding distance for the next
    generation.
Until (termination criteria are met)
```

Fig. 1 Algorithm of MODEFC

In Eq. (13), $\text{rand}_i(0, 1)$ is the ith evaluation of a uniform random number generator with outcome $\in [0, 1]$. CR is the crossover constant $\in [0, 1]$ which has to be determined by the user. $\text{rand}(i)$ is a randomly chosen index $\in 1, 2, \ldots, l$ which ensures that $\tau_{i,l}^{(t+1)}$ gets at least one parameter from $\varpi_{i,l}^{(t+1)}$.

3.4 Selection

To decide the member of generation, the trial offspring pool $\tau_{il,P}^{(t+1)}$ is combined with its target pool $\chi_{il,P}^{(t)}$, where P is the population size, for performing the nondominated sort and crowding distance. The selection of P vectors has been performed based on the lowest rank assigned as well as least crowding distance. These processes are executed for a fixed number of generations. The final nondominated solutions or Pareto-optimal solutions are the result of the MODEFC algorithm. The best solution is selected from this set of solutions using best Minkowski Score. The different steps of MODEFC are shown in Fig. 1.

4 Results and Discussion

In this section, datasets, performance metrics, and input parameter for the different algorithms are described. Subsequently the visual and quantitative results of comparison of different algorithms are provided.

4.1 MR Brain Images

The MR images of the brain chosen for the experiment is downloaded from the Brainweb database [1]. Two types of simulated brain images are considered: normal

and multiple sclerosis (MS) lesions images. The images are available in three bands: T1-weighted, T2-weighted, and proton density (PD)-weighted. The images correspond to the 1 mm slice thickness, 3 % noise (relative to the brightest tissue), and with 20 % intensity nonuniformity. The images of size 217×181 is available in 181 different Z planes. For the normal image data, the images of the Z planes Z10, Z60, and Z130 are considered. For the MS lesions image data, the images of the Z planes Z40, Z90, and Z140 are used for experiments. The ground truth information for these images is also available from the Brainweb web site [1].

4.2   Performance Metrics

The performance of the clustering methods is evaluated by measuring the metrics like Minkowski Score (MS) [23], Adjusted Rand Index (ARI) [22], and $\mathscr{I}$ index [30]. Note that for MS, the optimum score is 0, with lower score signifies better clustering results. On the other hand, the value of ARI lies between zero and one, where higher value indicates that the resulting clustering is more close to the actual clustering. Finally, for $\mathscr{I}$ index, higher value signifies better clustering.

4.3   Input Parameters

The algorithms are applied to a particular Z plane at a time to get the clusters and the cluster centers. The population size and number of generation used for MODEFC, DEFC, and GAFC algorithms are 20 and 100, respectively. The crossover probability (CR) used for MODEFC, DEFC, and GAFC is 0.8, while the scalar factor (F) for MODEFC and DEFC is 0.6. The mutation probabilities for GAFC is taken to be 0.3. The parameters of the SAFC algorithm are as follows: $T_{\max} = 100$, $T_{\min} = 0.01$, $r = 0.9$, and $k = 100$. The FCM algorithm is executed till it converges to the final solution. Note that the input parameters used here are fixed either following the literature or experimentally. It is observed from ground truth information that each of the Z planes 10, 60, and 130 for MR images has 9 classes. Similarly, for MR brain images of the multiple sclerosis lesions, Z planes 40, 90, and 140 have classes 9, 11 and 9, respectively. Therefore, the algorithms are executed for three Z planes of each MR brain images.

4.4   Results

Table 1 shows the result in terms of average MS, ARI, and $\mathscr{I}$ values of 20 runs of different algorithms for MR normal brain images. Results show that the MS, ARI, and $\mathscr{I}$ values corresponding to the partitioning provided by the MODEFC

Table 1 Average MS, ARI, and $\mathscr{I}$ values over 20 runs of different algorithms for MR images of normal brain

Algorithms	Z10 (no. of clusters "9")			Z60 (no. of clusters "9")			Z130 (no. of clusters "9")		
	MS	ARI	$\mathscr{I}$	MS	ARI	$\mathscr{I}$	MS	ARI	$\mathscr{I}$
MODEFC	0.3252	0.8915	41,432.53	0.3822	0.8527	43,238.02	0.2608	0.9371	71,456.33
DEFC	0.3807	0.8501	41,356.02	0.4501	0.8051	43,155.55	0.3302	0.9022	71,342.46
GAFC	0.4022	0.8182	41,230.51	0.4941	0.7701	43,105.04	0.3653	0.8761	71,204.22
SAFC	0.4461	0.7908	41,156.66	0.5272	0.7383	42,954.22	0.4172	0.8354	71,113.05
FCM	0.5022	0.7263	41,078.42	0.6021	0.6871	42,878.62	0.4904	0.7607	70,992.83
AL	0.6305	0.5963	40,783.52	0.7383	0.5516	42,553.72	0.6155	0.6409	70,677.14
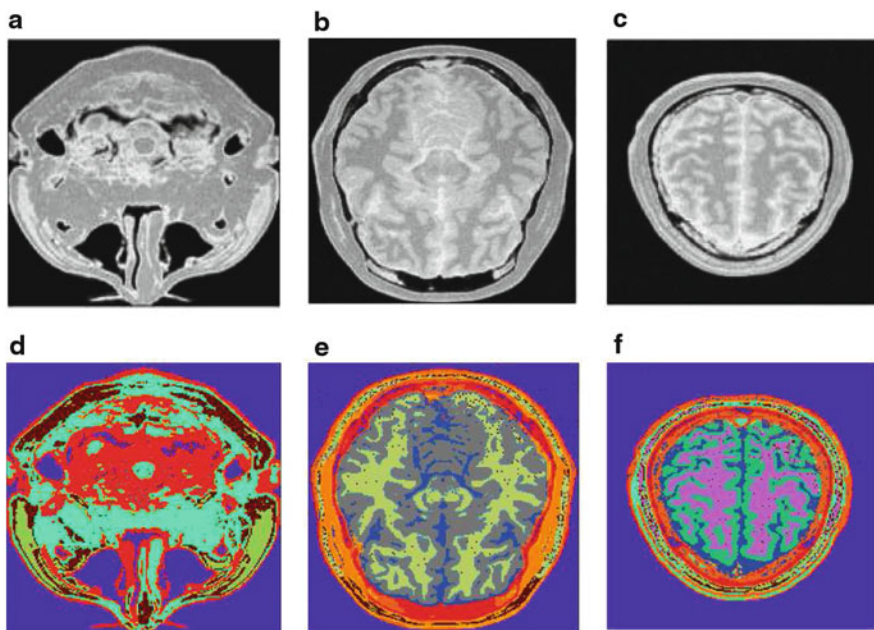
Fig. 2 (**a**), (**b**), and (**c**) are the original T1-weighted MR images of the normal brain in Z10, Z60, and Z130 planes, and (**d**), (**e**), and (**f**) are the corresponding segmented images produced by MODEFC

clustering are better among all the partitions. This implies the superior performance of MODEFC for detecting the proper partitioning from the MR normal brain images. Figure 2a–c shows the original MR normal brain images for T1 band in Z10, Z60, and Z130 planes, respectively. Figure 2d–f shows the corresponding segmented images obtained after application of MODEFC algorithm. Moreover, MODEFC algorithm is also executed on the images of three different Z-planes for MR multiple sclerosis lesions brain images. As the ground truth information is available, MS, ARI, and $\mathscr{I}$ values are calculated after application of MODEFC, DEFC,

Table 2 Average MS, ARI, and $\mathscr{I}$ values over 20 runs of different algorithms for MR images of multiple sclerosis lesions brain

Algorithms	Z40 (no. of clusters "10")			Z90 (no. of clusters "11")			Z140 (no. of clusters "9")		
	MS	ARI	$\mathscr{I}$	MS	ARI	$\mathscr{I}$	MS	ARI	$\mathscr{I}$
MODEFC	0.3826	0.9072	59,167.28	0.3315	0.8972	56,107.42	0.2902	0.9117	107,267.43
DEFC	0.4214	0.8602	59,042.52	0.4021	0.8388	56,024.25	0.3572	0.8806	107,198.75
GAFC	0.4655	0.8235	58,913.66	0.4301	0.8022	55,942.06	0.4052	0.8404	107,087.26
SAFC	0.4972	0.7861	58,845.51	0.4683	0.7602	55,796.44	0.4361	0.8115	106,987.51
FCM	0.5655	0.7011	58,711.47	0.5624	0.6864	55,668.53	0.5225	0.7384	106,864.72
AL	0.6882	0.5633	58,441.82	0.6974	0.5702	55,405.66	0.6417	0.5938	106,643.73
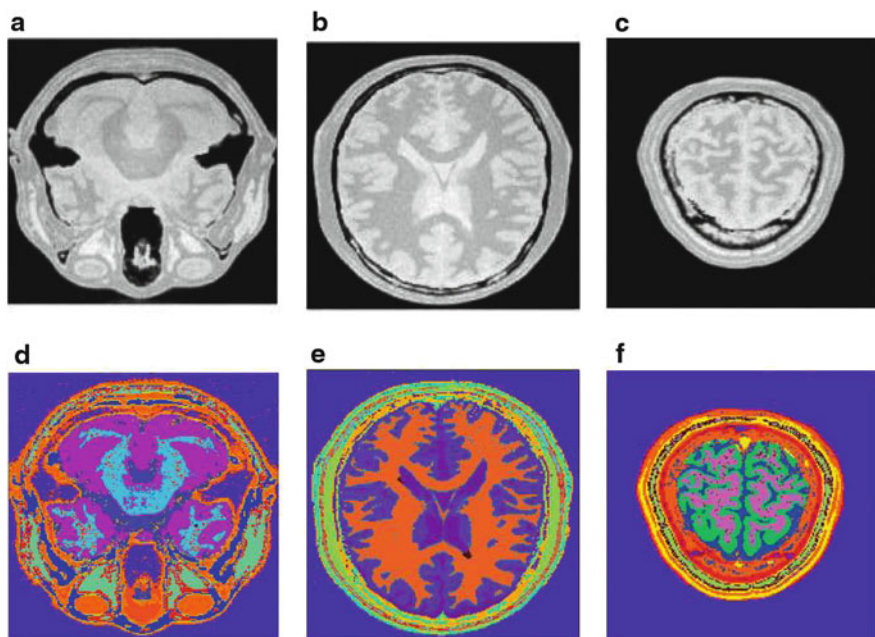
Fig. 3 (**a**), (**b**), and (**c**) are the original T1-weighted MR images of the multiple sclerosis lesions brain in Z40, Z90, and Z140 planes, and (**d**), (**e**), and (**f**) are the corresponding segmented images produced by MODEFC

GAFC, SAFC, FCM, and AL algorithms in order to measure the "goodness" of the solutions. Table 2 reports the result for three different MS Lesion Brain images, which show that the MS, ARI, and $\mathscr{I}$ values corresponding to the partitioning provided by the MODEFC is better. This is again revealing the effectiveness of the MODEFC algorithm for segmenting the MR brain image with multiple sclerosis lesions. Figure 3a–c shows the original MS Lesion Brain image in T1 band projected on Z40, Z90, and Z140 planes. Figure 3d–f shows the corresponding segmented images obtained after application of MODEFC algorithm respectively. Figures 4
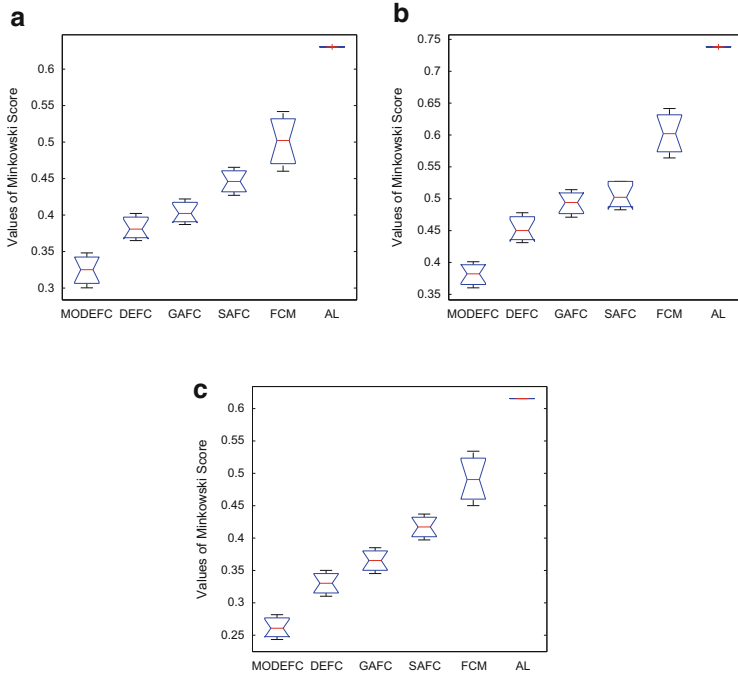
Fig. 4 Boxplot of MS for different clustering algorithms for MR images of the normal brain as (**a**) Z10, (**b**) Z60, and (**c**) Z130 planes

and 5 demonstrate the boxplot of different algorithms. As can be seen from the figures, the performance of the proposed MODEFC is the best for segmenting both normal and multiple sclerosis lesions brain images.

Furthermore, a non-parametric statistical significance test called Wilcoxon rank sum test [20] for independent samples is conducted at the 5 % significance level to judge the superiority of the results produced by MODEFC. For this purpose, six groups are created with 20 MS values each that correspond to the six algorithms such as (1) MODEFC, (2) DEFC, (3) GAFC, (4) SAFC, (5) FCM, (6) AL, for each of the Z planes of MR brain images. Table 3 reports the *p-values* produced by Wilcoxon rank sum test after comparing two groups (one group corresponding to MODEFC and another group corresponding to some other algorithm) at a time. For example, the rank sum test between the algorithms MODEFC and GAFC for Z130 of MR normal brain image provides a *p-value* of 1.3083e−004, which is very small. This is a strong evidence that values of performance metric, MS, produced by MODEFC is statistically significant and is not occurred by chance. Similar results are obtained for all other datasets and for all other algorithms compared to MODEFC which establish the significant superiority of the proposed technique.
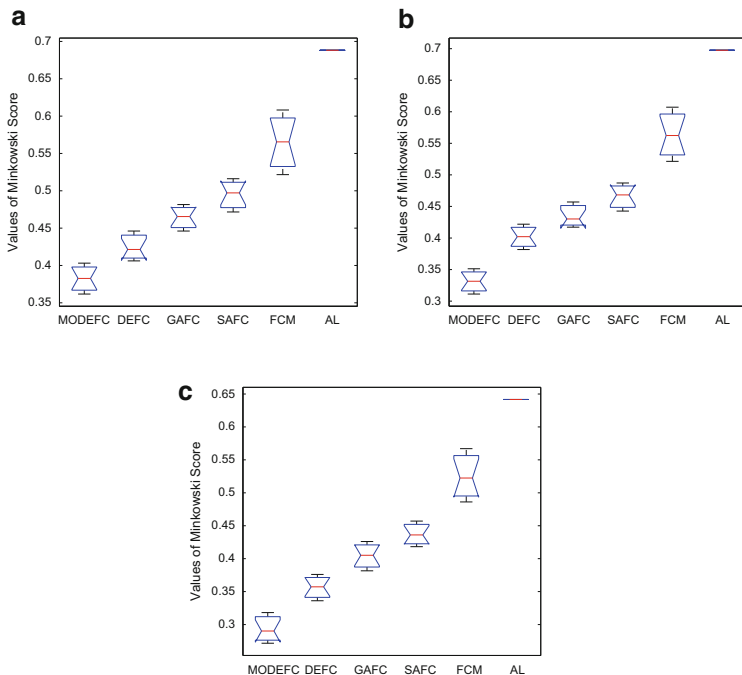
Fig. 5 Boxplot of MS for different clustering algorithms for MR images of multiple sclerosis lesions brain as (**a**) Z40, (**b**) Z90, and (**c**) Z140 planes

Table 3 *p-Values* produced by Wilcoxon Rank Sum test comparing MODEFC with other algorithms

MR brain images		*p*-Value				
		DEFC	GAFC	SAFC	FCM	AL
Normal brain	Z10	1.1025e−003	1.3152e−004	1.7164e−004	3.5427e−005	5.0362e−007
	Z60	1.1142e−003	1.5614e−004	1.8625e−004	3.4162e−005	4.8641e−007
	Z130	1.0281e−003	1.3083e−004	1.8015e−004	3.3482e−005	4.8753e−007
MS lesion brain	Z40	1.0726e−003	1.2581e−004	1.5872e−004	3.7122e−005	5.3562e−007
	Z90	1.1028e−003	1.5614e−004	1.8927e−004	3.2281e−005	4.6714e−007
	Z140	1.0742e−003	1.3041e−004	1.6104e−004	4.0162e−005	4.9514e−007

5   Conclusion

In this chapter, the problem of segmentation of MR brain images is posed as one of multiobjective clustering in the intensity space. The uncertainty in medical image segmentation comes from imprecision in computations and vagueness in class definitions. To overcome this, the searching capability of differential evolution is utilized to develop Multiobjective Differential Evolution-Based Fuzzy Clustering.

It simultaneously optimizes J_m and XB cluster validity indices. In MODEFC, assignment of points to different clusters are made based on Euclidean distance. The MODEFC algorithm requires the a priori specification of the number of clusters present in the dataset. In the final generation, it produces a set of solutions, from which the best solution is selected based on the least Minkowski Score. The effectiveness of the developed algorithm is shown in segmenting several MR brain images. The segmentation results are then compared with the available ground truth information. For the purpose of comparison, differential evolution-based fuzzy clustering, genetic algorithm-based fuzzy clustering, simulated annealing-based fuzzy clustering, well-known FCMs, and widely used average linkage algorithms are used. Experimental results show that MODEFC technique is not only able to segment the MR brain images into different tissue classes, but the segmentation result is also superior compared to other clustering algorithms quantitatively and statistically.

As a scope of further research, the work can be extended to solve clustering problems where the number of clusters is not known a priori [31, 36]. Moreover, MODEFC can also be applied for different problems including clustering of categorical data [33, 40, 49], pixel classification of satellite imagery [35, 38, 39, 45], finding co-expressed genes in microarray data [6, 32, 34, 44, 46], edge detection using thresholding [42], protein translational modification site prediction [41, 47], human leukocyte antigen class II binding peptide prediction [48], etc. The work can also be extended for other imaging techniques like computed tomography (CT) and positron emission tomography (PET).

References

1. Brainweb: Simulated Brain Database. http://www.bic.mni.mcgill.ca/brainweb
2. Ahmed, M.N., Yamany, S.M., Mohamed, N., Farag, A., Moriarty, T.A.: Modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans. Med. Imaging **21**(3), 193–199 (2002)
3. Alirezaie, J., Jernigan, M.E., Nahmias, C.: Automatic segmentation of cerebral MR images using artificial neural networks. IEEE Trans. Nucl. Sci. **45**(4), 2174–2182 (1998)
4. Bandyopadhyay, S., Maulik, U.: Non-parametric genetic clustering: comparison of validity indices. IEEE Trans. Syst. Man Cybern. C **31**(1), 120–125 (2001)
5. Bandyopadhyay, S., Maulik, U., Pakhira, M.K.: Clustering using simulated annealing with probabilistic redistribution. Int. J. Pattern Recognit. Artif. Intell. **15**(2), 269–285 (2001)
6. Bandyopadhyay, S., Mukhopadhyay, A., Maulik, U.: An improved algorithm for clustering gene expression data. Bioinformatics **23**(21), 2859–2865 (2007)
7. Bandyopadhyay, S., Pal, S., Maulik, U.: Incorporating chromosome differentiation in genetic algorithms. Inf. Sci. **104**(3–4), 293–319 (1998)
8. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York (1981)

9. Clark, M.C., Hall, L.O., Goldgof, D.B., Velthuizen, R., Murtagh, F.R., Silbiger, M.S.: Automatic tumor segmentation using knowledge-based techniques. IEEE Trans. Med. Imaging **17**(2), 187–201 (1998)
10. Clarke, L.P., Velthuizen, R.P., Phuphanich, S., Schellenberg, J.D., Arrington, J.A., Silbiger, M.: MRI: stability of three supervised segmentation techniques. Magn. Reson. Imaging **11**(1), 95–106 (1993)
11. Coello, C.A.C.: An empirical study of evolutionary techniques for multiobjective optimization in engineering design. Ph.D. Thesis, Department of Computer Science, Tulane University, New Orleans (1996)
12. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, Chichester (2001)
13. Deb, K., Agrawal, S., Pratab, A., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
14. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable test problems for evolutionary multi-objective optimization. Technical report, Computer Engineering and Networks Lab (TIK), ETH Zurich (2001)
15. Fonseca, C.M., Fleming, P.J.: Genetic algorithms for multi-objective optimization: formulation, discussion, and generalization. In: Proceedings of the Fifth International Conference on Genetic Algorithms, pp. 416–423 (1993)
16. Fonseca, C.M., Fleming, P.J.: An overview of evolutionary algorithms in multiobjective optimization. Evol. Comput. **3**(1), 1–16 (1995)
17. Glass, J.O., Ji, Q., Glas, L.S., Reddick, W.E.: Prediction of total cerebral tissue volumes in normal appearing brain from sub-sampled segmentation volumes. Magn. Reson. Imaging **21**(9), 977–982 (2003)
18. Holland, J.H.: Outline for a logical theory of adaptive systems. J. ACM **9**(3), 297–314 (1962)
19. Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975)
20. Hollander, M., Wolfe, D.A.: Nonparametric Statistical Methods. Wiley, New York (1999)
21. Horn, J.: Multi criteria decision making. Handbook of Evolutionary Computation **97/1**, F1.9 (1997)
22. Hubert, L., Arabie, P.: Comparing partitions. J. Classif. **2**(2), 193–218 (1985)
23. Jardine, N., Sibson, R.: Mathematical Taxonomy. Wiley, New York (1971)
24. Kirkpatrik, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science **220**, 671–680 (1983)
25. Knowles, J.D., Corne, D.W.: The Pareto archived evolution strategy: a new baseline algorithm for pareto multiobjective optimisation. In: Proceedings of the IEEE Congress on Evolutionary Computation, pp. 98–105 (1999)
26. Laarhoven, P.J.M., Aarts, E.H.: Simulated Annealing: Theory and Applications. Kluwer Academic, Dordrecht (1987)
27. Leemput, K.V., Maes, F., Vandermeulen, D., Suetens, P.: Automated model-based bias field correction of MR images of the brain. IEEE Trans. Med. Imaging **18**(10), 885–896 (1999)
28. Lin, J.S., Cheng, K.S., Mao, C.W.: Multispectral magnetic resonance images segmentation using fuzzy Hopfield neural network. Int. J. Biomed. Comput. **42**(3), 205–214 (1996)
29. Maulik, U.: Medical image segmentation using genetic algorithms. IEEE Trans. Inf. Technol. Biomed. **13**(2), 166–173 (2009)
30. Maulik, U., Bandyopadhyay, S.: Performance evaluation of some clustering algorithms and validity indices. IEEE Trans. Pattern Anal. Mach. Intell. **24**(12), 1650–1654 (2002)
31. Maulik, U., Bandyopadhyay, S.: Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification. IEEE Trans. Geosci. Remote Sens. **41**(5), 1075–1081 (2003)
32. Maulik, U., Bandyopadhyay, S., Mukhopadhyay, A.: Multiobjective Genetic Algorithms for Clustering: Applications in Data Mining and Bioinformatics. Springer, Heidelberg (2011)
33. Maulik, U., Bandyopadhyay, S., Saha, I.: Integrating clustering and supervised learning for categorical data analysis. IEEE Trans. Syst. Man Cybern. A **40**(4), 664–675 (2010)

34. Maulik, U., Mukhopadhyay, A., Bandyopadhyay, S.: Combining Pareto-optimal clusters using supervised learning for identifying co-expressed genes. BMC Bioinformatics **10:27** (2007)
35. Maulik, U., Saha, I.: Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery. Pattern Recognit. **42**(9), 2135–2149 (2009)
36. Maulik, U., Saha, I.: Automatic fuzzy clustering using modified differential evolution for image classification. IEEE Trans. Geosci. Remote Sens. **48**(9), 3503–3510 (2010)
37. Mehta, S.B., Chaudhury, S., Bhattacharyya, A., Jena, A.: Handcrafted fuzzy rules for tissue classification. Magn. Reson. Imaging **26**(6), 815–823 (2008)
38. Mukhopadhyay, A., Maulik, U.: Unsupervised pixel classification in satellite imagery using multiobjective fuzzy clustering combined with SVM classifier. IEEE Trans. Geosci. Remote Sens. **47**(4), 1132–1138 (2009)
39. Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S.: Multiobjective genetic clustering for pixel classification in remote sensing imagery. IEEE Trans. Geosci. Remote Sens. **45**(5), 1506–1511 (2007)
40. Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S.: Multiobjective genetic algorithm-based fuzzy clustering of categorical attributes. IEEE Trans. Evol. Comput. **13**(5), 991–1005 (2009)
41. Plewczynski, D., Basu, S., Saha, I.: AMS 4.0: Consensus prediction of post-translational modifications in protein sequences. Amino Acids **43**(2), 573–582 (2012)
42. Rakesh, R.R., Chaudhuri, P., Murthy, C.A.: Thresholding in edge detection: a statistical approach. IEEE Trans. Image Process. **13**(7), 927–936 (2004)
43. Reddick, W.E., Glass, J.O., Cook, E.N., Elkin, T.D., Deaton, R.J.: Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. IEEE Trans. Med. Imaging **16**(6), 911–918 (1997)
44. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Improvement of new automatic differential fuzzy clustering using SVM classifier for microarray analysis. Expert Syst. Appl. **38**(12), 15122–15133 (2011)
45. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: SVMeFC: SVM ensemble fuzzy clustering for satellite image segmentation. IEEE Geosci. Remote Sens. Lett. **9**(1), 52–55 (2011)
46. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Unsupervised and supervised learning approaches together for microarray analysis. Fundam. Inform. **106**(2), 45–73 (2011)
47. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Fuzzy clustering of physicochemical and biochemical properties of amino acids. Amino Acids **43**(2), 583–594 (2012)
48. Saha, I., Mazzocco, G., Plewczynski, D.: Consensus classification of human leukocyte antigens class II proteins. Immunogenetics **65**(2), 97–105 (2013)
49. Saha, I., Mukhopadhyay, A.: Improved crisp and fuzzy clustering techniques for categorical data. IAENG Int. J. Comput. Sci. **35**(4), 438–450 (2008)
50. Saha, I., Plewczynski, D., Maulik, U., Bandyopadhyay, S.: Real-coded differential crisp clustering for MRI brain image segmentation. In: Proceedings of the IEEE Congress on Evolutionary Computation, pp. 3912–3919 (2010)
51. Schaffer, J.D.: Multiple objective optimization with vector evaluated genetic algorithm. In: Proceedings of the First International Conference on Genetic Algorithms, pp. 93–100 (1985)
52. Storn, R., Price, K.: Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical Report TR-95-012. URL http://http.icsi.berkeley.edu/~storn/litera.html (1995)
53. Storn, R., Price, K.: Differential evolution: A simple and efficient heuristic strategy for global optimization over continuous spaces. J. Glob. Optim. **11**, 341–359 (1997)
54. Suckling, J., Sigmundsson, T., Greenwood, K., Bullmore, E.T.: A modified fuzzy next term clustering algorithm for operator independent brain tissue classification of dual echo MR images. Magn. Reson. Imaging **17**(7), 1065–1076 (1999)
55. Sur, A., Patra, N., Chakraborty, S., Saha, I.: A new wavelet based edge detection technique for iris imagery. In: Proceedings of the International Conference on IEEE IACC, Patiala, pp. 120–124, March 2009
56. Tou, J.T., Gonzalez, R.C.: Pattern Recognition Principles. Addison-Wesley, Reading (1974)

57. Wang, C.M., Chen, C.C.C., Chung, Y.N., Yang, S.C., Chung, P.C., Yang, C.W., Chang, C.I.: Detection of spectral signatures in multispectral MR images for classification. IEEE Trans. Med. Imaging **22**(1), 50–61 (2003)
58. Wang, C.M., Chen, C.C.C., Yang, S.C., Chung, P.C., Chung, Y.N., Yang, C.W., Chang, C.I.: Unsupervised orthogonal subspace projection approach to magnetic resonance image classification. Opt. Eng. **41**(7), 1546–1557 (2002)
59. Wang, C.M., Yang, S.C., Chung, P.C., Chang, C.I., Lo, C.S., Chen, C.C., Yang, C.W., Wen, C.H.: Orthogonal subspace projection-based approaches to classification of MR image sequences. Comput. Med. Imaging Graph. **25**(6), 465–476 (2001)
60. Wells, W.M., Grimson, W.L., Kikinis, R., Jolesz, F.A.: Adaptive segmentation of MRI data. IEEE Trans. Med. Imaging **15**(4), 429–442 (1996)
61. Xie, X.L., Beni, G.: A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. **13**(8), 841–847 (1991)
62. Yang, M.S., Hu, Y.J., Lin, K.C., Lin, C.C.: Segmentation techniques for tissue differentiation in MRI of ophthalmology using fuzzy clustering algorithms. Magn. Reson. Imaging **20**(2), 173–179 (2002)
63. Yang, M.S., Lin, K.C., Liu, H.C., Lirng, J.F.: Magnetic resonance imaging segmentation techniques using batch-type learning vector quantization algorithms. Magn. Reson. Imaging **25**(2), 265–277 (2007)
64. Zhou, J., Chan, K., Chong, V., Krishnan, S.: Extraction of brain tumor from MR images using one-class support vector machine. In: Proceedings of the 27th Annual Conference IEEE Engineering in Medicine and Biology, vol. 6, pp. 6411–6414 (2005)
65. Zitzler, E., Laumanns, M., Thiele, L.: An evolutionary algorithm for multiobjective optimization: the strength Pareto approach. Technical Report 43. Swiss Federal Institute of Technology, Zurich (1998)

Spectral and Non-linear Analysis of Thalamocortical Neural Mass Model Oscillatory Dynamics

Basabdatta Sen-Bhattacharya, Neslihan Serap-Sengor, Yuksel Cakir, Liam Maguire, and Damien Coyle

Abstract The chapter is organised in two parts: In the first part, the focus is on a combined power spectral and non-linear behavioural analysis of a neural mass model of the thalamocortical circuitry. The objective is to study the effectiveness of such "multi-modal" analytical techniques in model-based studies investigating the neural correlates of abnormal brain oscillations in Alzheimer's disease (AD). The power spectral analysis presented here is a study of the "slowing" (decreasing dominant frequency of oscillation) within the alpha frequency band (8–13 Hz), a hallmark of electroencephalogram (EEG) dynamics in AD. Analysis of the non-linear dynamical behaviour focuses on the bifurcating property of the model. The results show that the alpha rhythmic content is maximal at close proximity to the bifurcation point—an observation made possible by the "multi-modal" approach adopted herein. Furthermore, a slowing in alpha rhythm is observed for increasing inhibitory connectivity—a consistent feature of our research into neuropathological oscillations associated with AD. In the second part, we have presented power spectral analysis on a model that implements multiple feed-forward and feed-back connectivities in the thalamo-cortico-thalamic circuitry, and is thus more advanced in terms of biological plausibility. This study looks at the effects of synaptic connectivity variation on the power spectra within the delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz) and beta (14–30 Hz) bands. An overall slowing of EEG with decreasing synaptic connectivity is observed, indicated by a decrease of power

B. Sen-Bhattacharya (✉)
University of Lincoln, Lincoln, UK
e-mail: basab@ieee.org

N. Serap-Sengor · Y. Cakir
Istanbul Technical University, Istanbul, Turkey
e-mail: sengorn@itu.edu.tr; cakiryu@itu.edu.tr

L. Maguire · D. Coyle
University of Ulster, Magee Campus, Northern Ireland, UK
e-mail: lp.maguire@ulster.ac.uk; dh.coyle@ulster.ac.uk

within alpha and beta bands and increase in power within the theta and delta bands. Thus, the model behaviour conforms to longitudinal studies in AD indicating an overall slowing of EEG.

1 Introduction

Over the past few decades, power spectra analysis of EEG has been a popular tool in the quest for biomarkers in various pathological as well as neuropathological disorders including AD [2, 14, 16, 18, 34, 38, 45, 47]. It is now well known that the EEG of AD patients shows a definite "slowing" (decrease in peak power) within the alpha frequency band (8–12 Hz) [55, 62] and is considered to be the hallmark in an AD-affected brain. Furthermore, longitudinal studies have shown an overall slowing of EEG in AD [36] indicated by an increase in power within the delta (0.5–3 Hz) and theta (4–7.5 Hz) bands and a decrease in power within the alpha (8–10 Hz) and beta (14–30 Hz) bands. However, a definitive prediction or diagnosis of AD based solely on EEG power spectra analysis has not yet been possible, primarily due to large inter-individual differences in recorded EEG [18, 26]. In this context, non-linear dynamical analysis[1] of EEG is proposed as an important tool for better understanding and prediction of AD [30, 34, 35] as well as other neurological disorders [5, 57, 58, 66]. Indeed, dynamical system analysis has been termed a new "paradigm" towards investigating EEG in neurological disorders [48]. In this context, it has been speculated that a combined study of EEG power spectra analysis as well as non-linear dynamical analysis would be a promising approach for identifying viable early markers of AD in EEG [18].

The brain oscillatory activity (commonly referred to as "brain rhythms") recorded in EEG is believed to be generated by neuronal populations at the mesoscopic scale (10^4–10^7) [23] in the thalamic and cortical tissues [59, 61], interconnected with intricate feed-forward and feed-back pathways [31, 37, 39, 54]. Furthermore, cortical atrophy is a structural hallmark in the AD-affected brain, while recent studies have reported an atrophy of the thalamus in AD [19, 74]. Neural mass computational models mimicking mesoscopic-scale mass of thalamocortical neurons and pathways are often used to study brain oscillations [4, 17, 20, 33, 44, 56, 63]. The term "neural mass" was coined by Freeman [23, 24], while the "neural mass modelling" paradigm is based on the mathematical framework proposed by Wilson and Cowan [72, 73]. Each cell population in a neural mass model represents a neuronal "ensemble" of mesoscopic scale, which are densely packed in space and work at the same temporal-scale, so that for all practical purposes they can be mathematically treated as a single entity, whence "mass". We have studied alpha rhythm slowing in AD [7–11] using neural mass models, which are based on a

---

[1]Interested readers may refer to [32, 71] for a background on non-linear dynamical analysis used in neuroscience research. For a tutorial on non-linear dynamical analysis applied to EEG, please see [48].

classic computational model (Alpha Rhythm model, ARm) of the thalamo-cortical circuitry proposed by Lopes da Silva [17] to mimic thalamocortical alpha rhythms. Our study with the ARm indicate that significant slowing of alpha rhythm is a result of an increase in inhibitory connectivity [6, 7]. To bring the model up to date with recent research [69], we have modified the structure of a single neuronal population in the ARm [8]. Our study with this modified ARm, referred to herein as the modARm, also shows an increase in inhibitory connectivity parameter as a direct correlate of alpha rhythm slowing in the model. Subsequently, we have proposed a biologically plausible thalamo-cortico-thalamic circuitry [10] based on the synaptic layout described in [53] and experimental data on synaptic connectivity obtained from [29, 37, 54].

In all the studies mentioned above, our work involved power spectra analysis aimed at observing how depleted synaptic connectivity affects the model output oscillatory behaviour, mainly within the alpha band. The current trend is towards "multi-modal" analytical approaches in investigating AD conditions [18], in contrast to the largely "unimodal" approach followed thus far. In this context, recent studies have proposed non-linear dynamical analysis of thalamocortical neural mass models to study EEG characteristics in neurological disorders; for example a non-linear dynamical analysis on Suffczyński et al.'s neural mass thalamocortical model [64] was presented in [40]. These and other studies [25] have shown the usefulness of bifurcation analysis towards predictive technologies in epileptic absence seizures. Bifurcation analysis on Jansen's neural mass model [33] was performed by Grimbert and Faugeras [27] to obtain a better understanding of the model's parameter space and its usefulness in cortical columnar modelling. A combination of dynamical systems analysis and power spectral analysis on a model of hippocampal theta rhythm [76, 77] has shown remarkable agreement with experimental observations, and has implied possible ionic pathways responsible for aberrant theta rhythmic behaviour in AD. Along these lines, we have recently proposed a combined study of non-linear dynamical analysis and power-spectra analysis of a thalamocortical neural mass model mimicking alpha rhythmic slowing in AD [12]. Our results show that the modARm has two distinct modes of oscillation within the alpha band, viz. the point-attractor mode and the limit-cycle mode, and is similar to the observations made with the ARm [58]. Further, the combined study of the bifurcation and power spectral behaviour of the model indicate that the vicinity of the bifurcation point is the region of interest in studying alpha rhythmic slowing in AD. Inhibition plays a key role in modulating the bifurcation behaviour of the model. A decrease in inhibitory connectivity below a threshold effects a low frequency limit cycle oscillatory behaviour in the model similar to EEG recordings in absence seizures. This observation conforms to biological studies reporting a decrease in inhibition inducing epileptic seizure in rats [64]. Conversely, an increase in inhibition beyond a certain threshold diminishes power within the alpha band and the model output is in the point-attractor mode; this is similar to EEG slowing observed in AD. Such a crucial role of the bifurcation point in implicating the power spectral behaviour emphasises the importance of a combined non-linear dynamical analysis and power spectral analysis in model-based studies.

Here, we present a combined analyses of the power spectra and non-linear dynamical behaviour of the modARm as presented in [12]. For brevity, this study looks only at the alpha rhythmic behaviour of the model and ignores the other frequency bands which are also relevant in AD. Subsequently, we present a more biologically plausible thalamo-cortico-thalamic circuitry, referred to herein as the thalamo-cortico-thalamic (TCT) model, that we have proposed in [10]. A power spectra analysis of the model output within the delta, theta, alpha and beta bands demonstrates the overall EEG slowing such as observed in AD. In Sect. 2, we present a concise description of the modARm followed by results of power spectra and non-linear dynamical analyses on the model. In Sect. 3, we present the TCT-model followed by a power spectra analysis of the model within the delta, theta, alpha and beta bands. The results of both the modARm and the TCT-model are discussed in Sects. 2 and 3, respectively. We conclude the chapter in Sect. 4 and outline the significance of the findings.

2   The Modified Alpha Rhythm Model

The ARm proposed by Lopes da Silva has an excitatory cell population representing the thalamocortical relay cells (TCR) of the lateral geniculate nucleus (LGN—the thalamic nucleus in the visual pathway), and an inhibitory cell population representing the interneurons (IN). The synaptic connectivity layout and parameterisation in the model was based on Tombol's [68] experimental study on the cat thalamus. Newer experimental studies have revealed a lack of consensus on the feed-back pathway from the IN to the TCR cell population [54]. Moreover, the feed-forward and feed-back connectivities between the TCR cells and those of the thalamic reticular nucleus (TRN) are now known to influence brain oscillations significantly [43]. Thus, in more recent research based on ARm [63], the inhibitory cell population is assumed to be those of the TRN. We have introduced modifications to the ARm (modARm) [8] by introducing a slightly altered single-neuronal structure based on more recent research [69]. The dynamical behaviour of the retinal, TCR and the TRN cell populations can be mathematically defined as a set of first order differential equations of the state variables x_{ret}, x_{tcr} and x_{trn}, respectively, and are defined in (1)–(3):

$$\dot{x}_{\mathrm{ret1}} = x_{\mathrm{ret2}}$$

$$\dot{x}_{\mathrm{ret2}} = \frac{H_{\mathrm{e}}}{\tau_{\mathrm{e}}} P(t) - \frac{2}{\tau_{\mathrm{e}}} x_{\mathrm{ret2}} - \frac{1}{\tau_{\mathrm{e}}^2} x_{\mathrm{ret1}} \tag{1}$$

$$\dot{x}_{\mathrm{tcr1}} = x_{\mathrm{tcr2}}$$

$$\dot{x}_{\mathrm{tcr2}} = \frac{H_{\mathrm{e}}}{\tau_{\mathrm{e}}} S(C_3 x_{\mathrm{ret1}} - C_2 x_{\mathrm{trn1}}) - \frac{2}{\tau_{\mathrm{e}}} x_{\mathrm{tcr2}} - \frac{1}{\tau_{\mathrm{e}}^2} x_{\mathrm{tcr1}} \tag{2}$$

Table 1 Values of the parameters defined in (1)–(5)

H_e	H_i	τ_e	τ_i	ν	s_0	e_0	C_1	C_2	C_3
mV		ms		mV^{-1}	mV	s^{-1}	%age		
3.25	22	10	25	0.56	6	25	35	15	7

The mean ξ_M of the input $P(t)$ is varied from 0 to 100 to study the bifurcation behaviour of the model using the XPP*Aut* software [22]. For power spectra analysis, $P(t)$ is a white Gaussian noise with mean ξ_M varied from 74 to 101 (around the bifurcation point of ≈ 86) and standard deviation $\xi_D = 1$

$$\dot{x}_{trn1} = x_{trn2}$$
$$\dot{x}_{trn2} = \frac{H_i}{\tau_i} S(C_1 x_{tcr1}) - \frac{2}{\tau_i} x_{trn2} - \frac{1}{\tau_i^2} x_{trn1} \qquad (3)$$

Brain alpha rhythms are most prominent in EEG from the occipital lobe (seat of the visual cortex) under conditions of relaxed wakefulness and with eyes closed. This can be thought to be a resting state with no sensory input. Thus, the model input $P(t)$ in (1) represents the background firing rate (unit of measure is spikes per second (sps)) of the retino-geniculate neuronal populations and is simulated by a Gaussian white noise [17] with mean ξ_M and standard deviation ξ_D. The synaptic strength and time constant of the excitatory (e) and inhibitory (i) post-synaptic potential are defined as $H_{e/i}$ and $\tau_{e/i}$, respectively. Each connectivity parameter $C_i : i \in \{1, 2, 3\}$ represents the synaptic contact made by a pre-synaptic cell population on a single dendritic terminal of the post-synaptic cell population, and is expressed as a percentage of the total number of synaptic contacts made by all pre-synaptic cell populations on the terminal. The connectivity parameter values in the model are based on latest experimental data available from the literature [29, 37]. The sigmoid functions $S(\cdot)$ in (2) and (3) transform the membrane potential $V_{cell} : \text{cell} \in \{\text{tcr, trn}\}$ of a post-synaptic cell population into firing rate and is defined in (4):

$$S(V_{cell}) = \frac{2e_0}{1 + e^{\nu(s_0 - V_{cell})}} \qquad (4)$$

where $2e_0$ is the maximum firing rate of the neuronal population, s_0 is the firing threshold, ν is the sigmoid steepness parameter and the post-synaptic membrane potentials $V_{cell} : \text{cell} \in \{\text{tcr, trn}\}$ are defined as:

$$V_{trn} = C_1 x_{tcr1}$$
$$V_{tcr} = C_3 x_{ret1} - C_2 x_{trn1} \qquad (5)$$

The output of the model is the membrane potential of the TCR cell population (V_{tcr}) as defined in (5). The parameters defined in (1)–(5) are mentioned in Table 1.

2.1 Simulation Methods

The bifurcation and state-space analysis is performed using the XPP*Aut* software proposed by Bard Ermentrout [21,22]. All parameters are initialised to an arbitrarily small value of $2e^{-4}$ and the Dormand Prince solver is used to simulate the model in XPP*Aut*. The input $P(t)$ is represented by a constant mean value ξ_M, i.e. there is no added noise. The bifurcation diagrams are obtained by the *AUTO* software within XPP*Aut*.

For power spectra analysis, the model is solved in Matlab [41] using the 4th/5th order Runge–Kutta method. Equivalence of the solutions of the equations in Matlab and XPP*Aut* is verified. Total simulation time is 60 s with a resolution of 1 ms. A total of 10 simulations is carried out for each set of parameter values, thus mimicking sampling over ten subjects with similar disease conditions. An epoch of the model output from the start of the 3rd to the end of the 58th second is abstracted for each simulation. Further, the noisy input $P(t)$ is generated with random seeds for each simulation. The output vector is the average of the outputs of the ten simulations, which is then down-sampled at 250 Hz and bandpass filtered using a Butterworth filter of order 10, and lower and upper cut-off frequencies of 7.5 and 13.5 Hz (alpha band), respectively. The power spectral density of the filtered output is computed at a resolution of 0.25 Hz using a Welch periodogram with a Hamming window of segment length $1/2$ the size of the sampling frequency and overlap of 50 % [16].

2.2 Results

We study three model parameters by varying their values around their basal values as mentioned in Table 1—(a) the mean of the extrinsic input (ξ_M), (b) the inhibitory connectivity parameter from TRN to TCR (C_2) and (c) the excitatory connectivity parameter from the retinal cells to TCR (C_3).

2.2.1 Varying the Mean of the Extrinsic Input

It is well known that the firing rate of spiking neurons in the central nervous system is $\leqslant 100$ sps. Also, it is reported in [49] that the firing rate of the retinal neurons when the eyes are open (i.e. in the presence of sensory input) is ≈ 11 sps. A "wide" parameter range was studied for bifurcation analysis of the respective models in [27, 58, 63] as well as in our previous research [7–9]. Here, we study the modARm behaviour within the range $0 \leqslant \xi_M \leqslant 100$ sps so that the deviation from biologically plausible values is minimal.

The bifurcation analysis plot is shown in Fig. 1a along with the time series behaviour (inset) at the indicated values of ξ_M. For values of $\xi_M < 85.89$, V_{tcr} settles to a stable equilibrium state after an initial transient. At $\xi_M = 85.89$, V_{tcr}

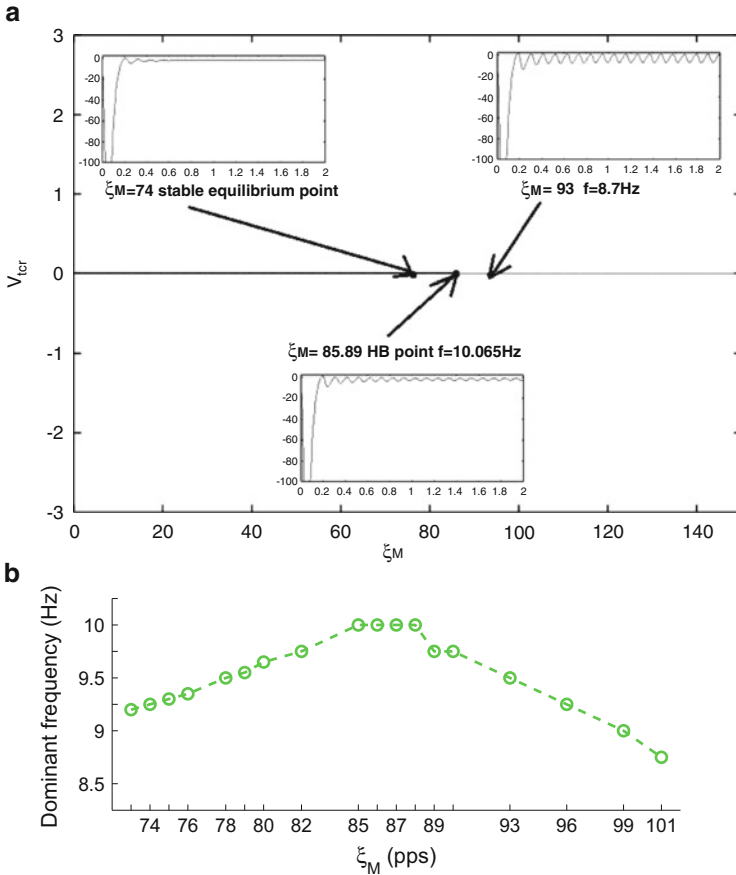**Fig. 1** (**a**) Bifurcation analysis for varying values of the mean ξ_M of the extrinsic model input $P(t)$. To avoid complexity, noise is not added for bifurcation analysis using the *AUTO* software within XPP*Aut* (as done in [27]). The points of bifurcation are shown with *arrows* within the figures; the dominant frequency of oscillation of the model output corresponding to these points as generated in XPP*Aut* are also mentioned. (**b**) Dominant frequency of oscillation within the alpha band for varying values (shown along the abscissa) of the mean ξ_M of the noisy model input $P(t)$ (simulated in Matlab)

undergoes a super-critical Hopf bifurcation and enters a limit cycle mode, oscillating with a dominant frequency of ≈ 10 Hz. The time series plot (embedded) of V_{tcr} at an arbitrarily selected value of $\xi_M = 74$, which is well below the point of bifurcation, show a fairly short transient phase. Three pairs of complex conjugate eigenvalues with real parts on the left complex plane imply the stability of the equilibrium point. Next, we select arbitrarily a value of $\xi_M = 93$ when the output (V_{tcr}) plot is well into the limit cycle mode with a constant amplitude of oscillation. Two of the three pairs of complex conjugate eigenvalues have their real parts on the left-hand side of

the complex plane; the real parts of the third pair of complex conjugate eigenvalues lie on the right-hand side of the complex plane, implying an unstable equilibrium point and bifurcation into a stable limit cycle mode. To observe the power spectra behaviour of the model at and around the bifurcation point as indicated in Fig. 1a ξ_M is varied through the range 70–100 sps. The dominant frequency of oscillation (DFO) of V_{tcr} within the alpha band and at each value of ξ_M is shown in Fig. 1b. We observe that the plots are characterised by three distinct regions: As ξ_M increases, the DFO within the alpha band increases steadily until for $85 \leqslant \xi_M \leqslant 88$ (i.e. values near the bifurcation point 85.89), the DFO ceases to increase any further. For $\xi_M > 88$, the DFO within the alpha band decreases progressively. Thus, with ξ_M as the control parameter and when it is below the bifurcation point ($\xi_M < 86$), the time series plots show that the model is in a point attractor mode, and we see an approximately linear decrease in the DFO, indicating slowing of the alpha rhythm with decreasing values of the sensori-geniculate input firing rate.

2.2.2 Varying the TRN to TCR Connectivity

The parameter C_2 representing the inhibitory synaptic connectivity from the TRN to TCR cells is varied through a range 0–30, which is in the range $\pm15\%$ of its basal value (Table 1). Figure 2a, b shows the model dynamical behaviour with varying C_2 for $\xi_M = 74$ and 93, respectively. In Fig. 2a, we observe that for $C_2 < 0.5836$ and $C_2 > 11.98$, the model output converges to a stable point attractor. However, the system undergoes a super-critical Hopf bifurcation for $0.5836 \leqslant C_2 \leqslant 11.98$ and V_{tcr} enters a limit cycle mode, oscillating with a dominant frequency within the alpha band. We note that the basal value of $C_2 = 15$ falls on a region of stable equilibrium. While the overall system behaviour in Fig. 2b is similar to that in Fig. 2a, we observe a right shift in the points of bifurcation so that the basal value of $C_2 = 15$ now falls within the region of stable limit cycles, i.e. $0.7388 \leqslant C_2 \leqslant 16.81$. The results provide insight into the model behaviour as a function of both inhibitory connectivity and the extrinsic input mean—a decrease in inhibitory connectivity throws the system out of a stable point attractor mode; this happens earlier for higher values of the excitatory extrinsic input. In other words, inhibition should increase in a directly proportional manner with the extrinsic input mean in order to maintain a stable point attractor mode in the system. The lower point of bifurcation observed in Fig. 2a, b are regions with very low values of C_2 and may be approximated as a condition of an open loop circuit with minimal negative feedback.

   The power spectra around the point of bifurcation is presented in Fig. 2c. We note that the DFO is maximum at around the point of bifurcation; for example in the case $\xi_M = 74$, the peak in the dominant frequency plot occurs at $C_2 = 12$, which is the point of bifurcation as indicated in Fig. 2a. Further, we note that the dominant frequency plots for all values of ξ_M have a similar shape, albeit with a right shift, which agrees with the right shift in bifurcation point with increasing values of ξ_M as shown in Fig. 2a, b. There is an approximately linear decrease in the DFO within the alpha band for $C_2 > 12$, which continues until $C_2 = 16$, after which
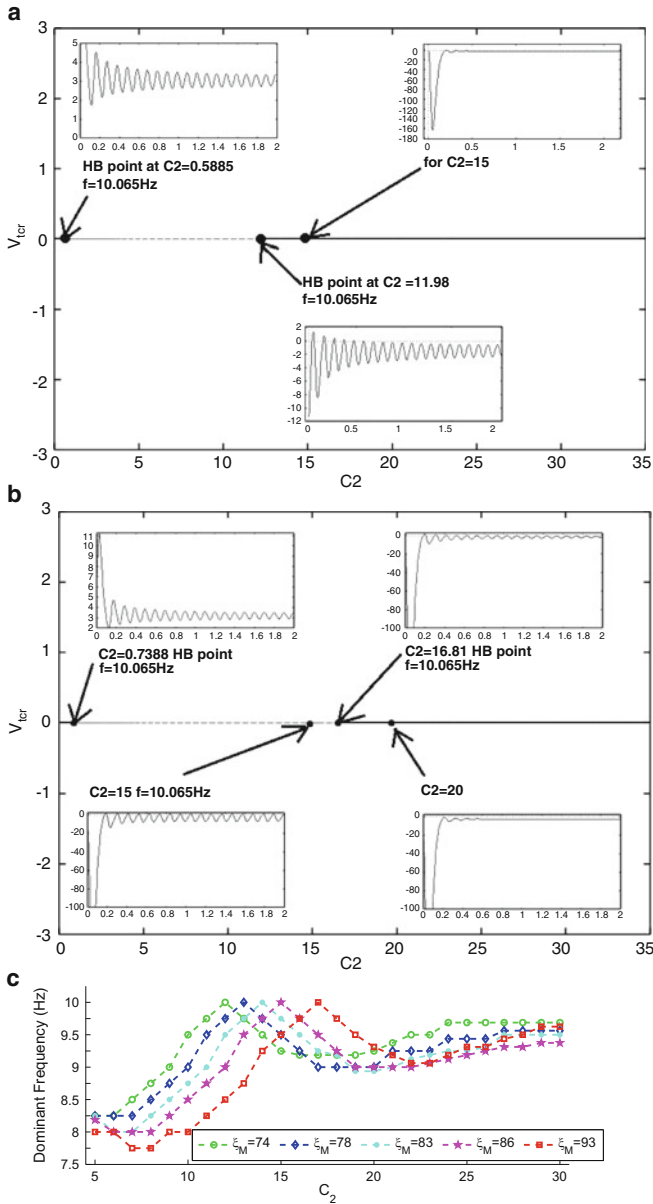
Fig. 2 Bifurcation analysis for varying values of the connectivity parameter C_2 when (**a**) $\xi_M = 74$ and (**b**) $\xi_M = 93$. The bifurcation analysis and figure generation are done in XPP*Aut*. The points of bifurcation are shown with *arrows* within the figures, while the dominant frequency of oscillation of the model output corresponding to these points as generated in XPP*Aut* are also mentioned. (**c**) Dominant frequency of oscillation with varying values of C_2 for $\xi_M = \{74, 76, 83, 86, 93\}$. The values of ξ_M are selected by visual inspection of the corresponding time series plots of V_{tcr} to demonstrate the bifurcation in model behaviour

the plots show no more slowing and is fairly flat. Thus, our results imply a slowing in alpha rhythm for increasing values of the inhibitory connectivity parameter (up to a certain range) and has been a consistent feature of our research while investigating alpha rhythmic slowing in AD using thalamocortical neural mass models. However, an important and novel direction implied by the results in this study is that the alpha rhythmic content is maximal at close proximity to the bifurcation point in the dynamical behaviour of the model.

2.2.3   Varying the Retinal to TCR Connectivity

Figure 3a, b shows a left shift of the bifurcation point with varying values of C_3 and with increasing values of ξ_M. The system undergoes a Hopf bifurcation at $C_3 \approx 8$ when $\xi_M = 74$ and at $C_3 \approx 6.5$ when $\xi_M = 93$; the DFO is at $\approx 10\,\mathrm{Hz}$ for both cases. The dominant frequency plot with varying C_3 in Fig. 3c shows a peak at the respective bifurcation points for both $\xi_M = 74$ and 93. As the model undergoes bifurcation and enters a limit cycle mode, there is a sharp fall in the dominant frequency within the alpha band. On the other hand, for values of C_3 prior to the bifurcation point, the DFO plots show a very slight fall of less than 1 Hz and may be neglected. Moreover, with increasing ξ_M, this fall in DFO further lessens. Thus, the DFO within the alpha band stays fairly stable with fluctuations $\leqslant 1\,\mathrm{Hz}$.

2.3   Discussion

The simultaneous analysis of the bifurcation behaviour and DFO within the alpha band of the model have shown that the alpha rhythmic content is maximal at close proximity to the bifurcation point. Indeed, in a non-linear study of the ARm, Stam [57] suggests that the dynamics of normal EEG "reflect" that of the model behaviour close to the bifurcation between the point attractor and the limit cycle attractor mode. Intuitively, such an observation "makes sense" and explains the ease of switching from an awake to a sleeping state and vice versa in normal and healthy adult, as opposed to a normal state far away from the point of transition, which might be the case in neurological disorders, for example in sleep disturbances associated with AD [38]. Furthermore, such an observation is in agreement with the DFO plots corresponding to variation of C_2 and C_3 as the "control" parameters for bifurcation, where, at a certain value of ξ_M and when the model is in the point attractor mode, the decrease in DFO is only within a certain range of the control parameter; beyond this range, the DFO plot flattens, possibly indicating a dissipation of the alpha frequency component and a change in the "brain state".

The limit cycle mode characteristics of the model are representative of the epileptic seizure activity and have been the topic of several studies [25, 58, 64]. Thus, limit cycles are referred to as a sign of abnormality in the EEG [58]. At the same time, slowing of alpha activity is also known to be a prominent feature in EEG
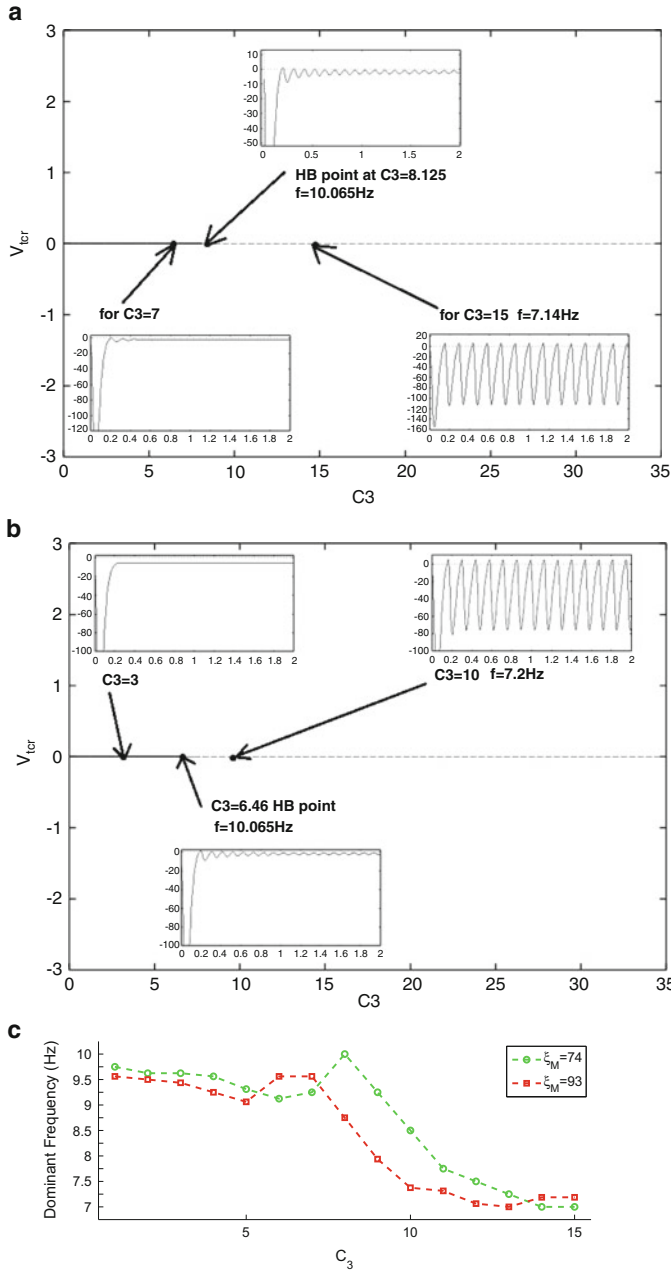
Fig. 3 Bifurcation analysis for varying values of the connectivity parameter C_3 when (**a**) $\xi_M = 74$ and (**b**) $\xi_M = 93$. The time-series plots corresponding to the points of bifurcation are shown as embedded figures. The dominant frequency of oscillation (DFO) of the model output corresponding to these points as generated in XPP*Aut* are also mentioned. (**c**) DFO with varying values of C_3. Two cases for $\xi_M = 74$ and 93 are shown

of AD under sleeping condition [2, 15, 50]. Moreover, the slow wave EEG during sleep bears a strong visual resemblance to limit cycle oscillations [43]. However, our study is aimed at investigating EEG of AD patients in an awake state (relaxed and eyes closed). Thus, we may say that the part of the DFO plots that correspond to the point attractor mode of the model is more appropriate for our current study. This observation is also supported by a report by Stam [57] on the local brain dynamics in AD which are "more noise like" than in healthy subjects. In support of this observation our results indicate a slowing in alpha rhythm for increasing values of the inhibitory connectivity parameter (C_2), albeit up to a certain range. This has been a consistent feature of our research while investigating alpha rhythmic slowing such as observed in AD using thalamocortical neural mass models [7–10]. Furthermore, the slowing "trend" for values immediately after and close to the bifurcation point matches with the above-mentioned speculative argument that the "normal" brain state is close to the bifurcation point; thus, any deviation from normal is bound to be in the close vicinity of this point. However, the slowing of the alpha rhythm in the point attractor mode with varying values of C_3 is observed to be minimal. This is unlike our earlier reports [8] of an increase in DFO with increasing values of the connectivity parameter C_3. One reason for this difference may be attributed to the use of different basal values of the input standard deviation ξ_D as well as the mean ξ_M (312 sps in [8], which in turn was as in [63]; 500 sps in [9], which was obtained by trial and error such that the model oscillated with dominant frequency of 10 Hz when all parameters were at their basal values).

One notable observation in the results is the classical slowing of alpha rhythms within the limit cycle mode for the cases pertaining to the control variables ξ_M and C_2, implied by a sharp fall in the DFO within the alpha band accompanied by an approximately linear increase in the corresponding power (not shown herein; interested readers may refer to [12]), thus conforming to the conventional definition of the alpha rhythm slowing in AD. Interestingly, slowing of the alpha rhythms has been reported as a more prominent feature in sleep conditions in AD [15] compared to waking conditions [2]. Thus, the larger drop in DFO within the limit cycle mode as compared to the point attractor mode may be speculated to imply a more prominent slowing of alpha rhythm during sleep in AD. Furthermore, the inhibitory pathway from the TRN to the TCR may be implicated in a transition from an awake to a sleeping brain state, as well as from normal to an abnormal brain state in awake conditions, both transitions occurring with decreasing levels of inhibition. The implication is that, while a decrease in inhibitory activity in the thalamocortical circuitry is desired for transition to a normal sleeping state, it is equally undesirable in awake conditions where the system can go into a state of "unconsciousness", an abnormal state seen for example during epileptic seizure. This observation finds strong support in a combined experimental and model-based research by Suffczyński et al. [64], where an induced decrease in inhibitory neurotransmitter in rats has generated conditions of epileptic seizure activity. On the other hand, while an increase in inhibitory activity is desirable to transit from sleeping to an awake state, too much increase of the same might lead to a decline in cognitive abilities associated with slowing of the alpha rhythm. Thus, our study

shows a critical role played by the inhibitory efferents from the TRN in maintaining normal thalamocortical activity in an awake and relaxed condition.

While the modARm is a useful starting point for model-based investigation of neurological disorders, it is an overly simplified representation of the thalamo-cortico-thalamic circuitry. To undertake a systematic approach towards understanding the underlying causes of abnormal brain rhythms in disease conditions, it is crucial that computational models are more biologically plausible. Towards such an endeavour, we have proposed an enhanced version of the thalamo-cortico-thalamic circuitry in [10]. We introduce this model in the following section.

3 An Enhanced Neural Mass Model of the Thalamo-Cortico-Thalamic (TCT) Circuitry

The TCT-model is an enhanced version of the modARm based on available experimental data on the structure and synaptic connectivity in the dorsal lateral geniculate nucleus (LGNd) of the thalamus, the basis of our choice of modelling the LGNd being: first, aberration of alpha rhythms in AD is dominant in the occipital area—the cortical region comprising of the visual cortex, which in turn is efferent mainly to the LGNd; second, the LGNd is the most widely studied thalamic nucleus to date [37, 54] and thus experimental data is more easily available; third, thalamocortical connectivity for most of the other thalamic nuclei is essentially similar to the LGN [54]. Thus we expect that insights from an LGNd model would also extend to some other nuclei (e.g. medial dorsal and anterior ventral nucleus [13, 51, 65]) which are more severely affected in early stage AD.

The term "functional connectivity" in connectionist networks is often used to indicate the total number of fibres from a given cell population converging on the dendrites of one cell in a neighbouring neural population. However, physiological data on exact fibre count of intra nuclei connectivity are scarce. On the other hand, there is substantial amount of data available on "synaptic connectivity", i.e. the total number of synapses by afferent fibres on the dendrites of TCR, IN and TRN cells in rats, cats and monkeys; in our work, we consider "synaptic connectivity".

In this section, we first discuss the biological basis of the enhancements made in the TCT-model (Sect. 3.1). The modelling (Sect. 3.2) and simulation (Sect. 3.3) methods are presented next. The results are presented and discussed in Sect. 3.4.

3.1 Biological Basis of the Synaptic Organisation in the TCT-Model

The synaptic organisation and structure of the model is as presented in [53] based on physiological data. A schematic of the model is presented in Fig. 4. It consists
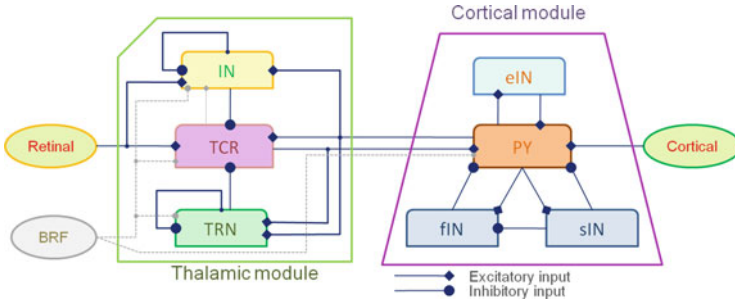
Fig. 4 The synaptic organisational structure in the thalamo-cortico-thalamic model presented in this work. The organisational structure of the thalamic module is as in [53]. The cortical module is as in [70, 75]. The thalamic module consists of the Thalamic Relay Cells (TCR), the Inter-Neurons (IN) and the Thalamic Reticular Nucleus (TRN). The cortical module consists of the Pyramidal Cells (PY), the Excitatory Inter-Neurons (eIN), the slow inhibitory Inter-Neurons (sIN) and the fast inhibitory Inter-Neurons (fIN). The *diamond shaped arrow heads* indicate excitatory input while the *round shaped arrowheads* indicate inhibitory input. The cholinergic inputs from the Brain Reticular Formation (BRF) is ignored. Also, feedback from the TCR to the IN is not considered due to ambiguous information from available experimental data [54]. The TCR cell population receive extrinsic input from the Retinal cells while the PY cell population receive extrinsic input from neighbouring cortical regions representing cortico-cortical connections

of two modules viz. the thalamic and the cortical module. The cortical module is as presented in [70, 75] and consists of four cell populations: Pyramid cells (PY), Excitatory Inter-Neurons (eIN), slow inhibitory Inter-Neurons (sIN) and fast inhibitory Inter-Neurons (fIN). Almost all thalamic nuclei (a few exceptions reported in the rat thalamus in [37]) is populated with the TCR and the IN cells. The location of the TRN is such that all thalamocortical and corticothalamic axons (except those of the olfactory pathway) pass through it and make excitatory synapses on its cells.

Both the retinal and the cortical cells make excitatory (Glutamatergic) synapses with the TCR and IN population. The other major source of input to the TCR, IN and TRN cell population comes from the brain reticular formation (BRF). The BRF makes excitatory synapses on the TCR cells and inhibitory synapses on the IN and TRN cell population, the synaptic nature for both types being mainly Cholinergic. In the TCT-model, we ignore inputs from the BRF and consider inputs only from the retina and the cortex. The IN cells make inhibitory (GABAergic) synapses on the TCR cells mainly through dendro-dendritic connections. The type and quantity of synapses made by the TCR population on the IN cells are however ambiguous [52]. We therefore ignore this connection for the current work. The TRN cells send inhibitory (GABAergic) feedback to the TCR cells. Each post-synaptic dendrite is divided broadly into three regions: proximal, intermediate and distal. Each of the three regions has different quantitative connectivity with the pre-synaptic terminals. Only the proximal dendrite connectivity is considered in the current model for

brevity. The pre-synaptic terminals are classified into four types depending on the shape and size of the synaptic vesicle [37]:

- *RL:* Round shaped large sized vesicles, characteristic of the excitatory axonal terminals of the retinal cells.
- *RS:* Round shaped small sized vesicles, characteristic of the excitatory axonal terminals of the cortical cells as well as those of the BRF and other parts of the brain.
- *PSD:* Inhibitory Pre-Synaptic-Dendritic terminals of the IN cells.
- *F:* Inhibitory pre-synaptic axonal terminals of the TRN cells.

The PSD and F terminals are similar in size and shape and thus pose a difficulty in estimating the exact proportion of each during quantitative data collection. However, it is still possible to estimate the major contributing thalamic area (viz. the TRN or the IN) towards the total number of inhibitory synapses. Thus, it is common in literature to collectively refer to the inhibitory synaptic terminals as "PSD/F". We summarise below the quantitative data available in current literature on synaptic connectivities reported in the TCR, IN and TRN population in context to data used in our current work.

- *TCR afferents:* For the proximal dendritic synapses of the TCR cells in monkey LGNd, 30–34 % of the synapses are from RL terminals, 40–45 % are from PSD/F terminals, while 25–30 % are from the RS terminals; quantitative data of similar synapses in the cat LGNd are 15–20 %, 25–30 % and 40–50 % respectively. However, reporting a later and more recent finding from the cat LGNd, the sources suggest the following figures: 7–12 % from RL, 24–31 % from PSD/F and 58–62 % from RS terminals.
- *IN afferents:* A study made on the cat LGNd in 1991 suggest that the IN cells receive around 25 % synapses from RL, 37 % from PSD/F, while 26 % synapses from RS terminals. However, according to a more recent study in 2000, these figures are reported as 48.7 %, 24.4 %, 26.9 %, respectively, in [37] and as 47 %, 24 %, 29 %, respectively, in [54]. On the other hand, data from the LGNd of a squirrel monkey (primate) indicates that the IN cells receive equal proportion of each of the three categories of synaptic terminals.
- *TRN afferents:* The most recent available data on synaptic afferents to the TRN are based on studies on the rat LGNd in 1999 as reported in [37]. The TRN can be divided into several sectors each of which correspond to a specific cortical region [54]. Both thalamocortical and corticothalamic synapses on the TRN sector corresponding to the rat visual cortex are excitatory (Glutamatergic) in nature and constitute 30–40 % and 50 % of the total synapses, respectively. The remaining and up to 25 % of the synapses are from inhibitory (GABAergic) sources from other parts of the brain as well as via dendro-dendritic connections among neighbouring cells within the TRN. Another study by Liu et al. in 1995 (as reported in [37]) gives the respective synaptic figures as 20 %, 60–65 % and 15 % respectively.

Table 2 Basal value for the connectivity parameters used in (6)–(14)

Module	Afferent (to)	Efferent (from)	Connectivity parameter	Value
Thalamic module	TCR	Retinal	C_{tre}	7.1
		IN	C_{tii}	15.45
		TRN	C_{tni}	15.45
		PY	C_{tpe}	62
	IN	Retina	C_{ire}	47.4
		IN	C_{isi}	23.6
		PY	C_{ipe}	29
	TRN	TCR	C_{nte}	35
		TRN	C_{nsi}	15
		PY	C_{npe}	50
Cortical module	PY	Cortical	C_{pce}	1
		TCR	C_{pte}	80
		eIN	C_{pxe}	108
		sIN	C_{pli}	33.75
		fIN	C_{pfi}	108
	eIN	PY	C_{xpe}	135
	sIN	PY	C_{lpe}	33.75
	fIN	PY	C_{fpe}	40.5
		sIN	C_{fli}	13.5

Each connectivity parameter in the thalamic module is a percentage of the total number of synapses T convergent on the respective cell population and is based on physiological data [29]. The thalamo-cortical connectivity parameter is a constant and is sourced from [56]. The connectivity parameters in the cortical module are as in [33, 75]

The basal values of the connectivity parameters used in this work are detailed in Table 2.

3.2   The Thalamo-Cortico-Thalamic Model

The model is defined by a set of second order ordinary differential equations defined in (6)–(14). The extrinsic input $P_1(t)$ to the thalamic module is from the retinal spiking neurons in an awake relaxed condition with eyes closed; extrinsic input $P_2(t)$ to the cortical module is from the neighbouring cortical regions.

TCR

$$\dot{x}_{\text{tcr1}} = x_{\text{tcr2}}$$

$$\dot{x}_{tcr2} = \frac{H_e}{\tau_e} S \left(C_{tre} x_{ret1} + C_{tpe} x_{tpy1} - C_{tii} x_{in1} - C_{tni} x_{trn1} \right)$$

$$- \frac{2}{\tau_e} x_{tcr2} - \frac{1}{\tau_e^2} x_{tcr1} \tag{6}$$

IN

$$\dot{x}_{in1} = x_{in2}$$

$$\dot{x}_{in2} = \frac{H_i}{\tau_i} S \left(C_{ire} x_{ret1} + C_{ipe} x_{tpy1} - C_{isi} x_{in1} \right)$$

$$- \frac{2}{\tau_i} x_{in2} - \frac{1}{\tau_i^2} x_{in1} \tag{7}$$

TRN

$$\dot{x}_{trn1} = x_{trn2}$$

$$\dot{x}_{trn2} = \frac{H_i}{\tau_i} S \left(C_{nte} x_{tcr1} + C_{npe} x_{py1} - C_{nsi} x_{trn1} \right)$$

$$- \frac{2}{\tau_i} x_{trn2} - \frac{1}{\tau_i^2} x_{trn1} \tag{8}$$

PY

$$\dot{x}_{py1} = x_{py2}$$

$$\dot{x}_{py2} = \frac{H_e}{\tau_e} S(C_{pce} x_{cc1} + C_{pte} x_{tcr1} + C_{pxe} x_{ein1}$$

$$- C_{pli} x_{sin1} - C_{pfi} x_{fin1})$$

$$- \frac{2}{\tau_e} x_{py2} - \frac{1}{\tau_e^2} x_{py1} \tag{9}$$

eIN

$$\dot{x}_{eIN1} = x_{eIN2}$$

$$\dot{x}_{eIN2} = \frac{H_e}{\tau_e} S \left(C_{xpe} x_{py1} \right)$$

$$- \frac{2}{\tau_e} x_{eIN2} - \frac{1}{\tau_e^2} x_{eIN1} \tag{10}$$

sIN

$$\dot{x}_{sIN1} = x_{sIN2}$$

$$\dot{x}_{sIN2} = \frac{H_{il}}{\tau_{il}} S\left(C_{lpe} x_{py1}\right)$$

$$- \frac{2}{\tau_{il}} x_{sIN2} - \frac{1}{\tau_{il}^2} x_{sIN1} \qquad (11)$$

fIN

$$\dot{x}_{fIN1} = x_{fIN2}$$

$$\dot{x}_{fIN2} = \frac{H_{if}}{\tau_{if}} S\left(C_{fpe} x_{py1} - C_{fli} x_{sIN1}\right)$$

$$- \frac{2}{\tau_{if}} x_{fIN2} - \frac{1}{\tau_{if}^2} x_{fIN1} \qquad (12)$$

Retinal

$$\dot{x}_{ret1} = x_{ret2}$$

$$\dot{x}_{ret2} = \frac{H_e}{\tau_e} P_1(t) - \frac{2}{\tau_e} x_{ret2} - \frac{1}{\tau_e^2} x_{ret1} \qquad (13)$$

Cortico-cortical

$$\dot{x}_{cc1} = x_{cc2}$$

$$\dot{x}_{cc2} = \frac{H_e}{\tau_e} P_2(t) - \frac{2}{\tau_e} x_{cc2} - \frac{1}{\tau_e^2} x_{cc1} \qquad (14)$$

All parameters in the above equation are as in [10]: $H_{e/i}$ is the strength of the excitatory (e) or inhibitory (i) PSP; a_1 is the inverse of the time constant of the excitatory PSP; b_i is the inverse of the time constant of an inhibitory PSP in the thalamic module; $b_{il/f}$ is the inverse of the time constant of the inhibitory PSP generated by the sIN (l) or fIN (f) cell population respectively; C_{xyz} are connectivity parameters with x representing the afferent population, y representing the efferent population and z representing either an excitatory or an inhibitory synapse; $P_1(t)$ is simulated by a Gaussian white noise with mean μ_r and standard deviation φ_r; $P_2(t)$ is simulated by a Gaussian white noise with mean μ_c and standard deviation φ_c; and $S(\cdot)$ is a sigmoid function which is as defined in (4). The connectivity parameters (C_{xyz}) are defined in Table 2; all other parameters are defined in Table 3. The connectivity parameters in the thalamic module are expressed as a percentage of the total number of synapses T convergent on an afferent thalamic cell population; T is defined in Table 3.

Table 3 Basal values of the parameters used in (6)–(14)

Parameter	Unit	Module	Value
ν	mV^{-1}		0.56
e_0	s^{-1}		2.5
s_0	mV		6
μ_r	Spikes per second (sps)	Thalamic	11
φ_r	sps^2	"	5
μ_c	sps	Cortical	30
φ_c	sps^2	"	5
H_e	mV	Thalamic	3.25
		Cortical	2.7
H_i	mV	Thalamic	22
H_{il}	mV	Cortical	4.5
H_{if}	mV	"	39
a_1	s^{-1}	Thalamic	100
		Cortical	40
b_1	s^{-1}	Thalamic	40
b_{il}	s^{-1}	Cortical	20
b_{if}	s^{-1}	"	300
T	Normalised	Thalamic	100

A blank field under the "Module" heading implies a common parameter value in both modules. The parameter T represents the total number of synapses convergent on a thalamic cell population

3.3 Simulation Methods

Model simulation is implemented using the 4th/5th order Runge–Kutta ODE solver within the Simulink® environment in Matlab (The ModelDB [28] accession number for the model simulated in Simulink® is 138970). The total simulation time is 60 s at a sampling rate of 250 Hz. As is done in experimental EEG studies [67], an epoch of the model output from the start of the 3rd to the end of the 59th second is abstracted for each simulation and averaged over 20 trials generated with random seeds. Each output vector obtained is bandpass filtered with a butterworth filter of order 10 with lower and upper cut-off frequencies of 0.5 and 50 Hz, respectively. The power spectral density vector is computed in Matlab using a Welch periodogram with Hamming window of segment length $\frac{1}{2}$ the size of the sampling frequency and overlap of 50 %. As in previous studies of alpha rhythm in AD [16, 45, 46], the alpha band in this work is further divided into the lower-alpha (α_{low} : 7.5−10 Hz) and upper-alpha (α_{high} : 10.5−13 Hz) bands. Our analysis is based on the relative band power, an analytic method recommended in dementia studies [45]. The relative power density is computed by dividing the absolute power at each frequency by the mean of the total power spectra. The relative power within each alpha sub-band is calculated by averaging the relative power density within the band.

For brevity, we consider a basal value of $T = 100$ for all the thalamic cell populations. In this study, T is varied from 5 to 200 at intervals of 5. It needs to
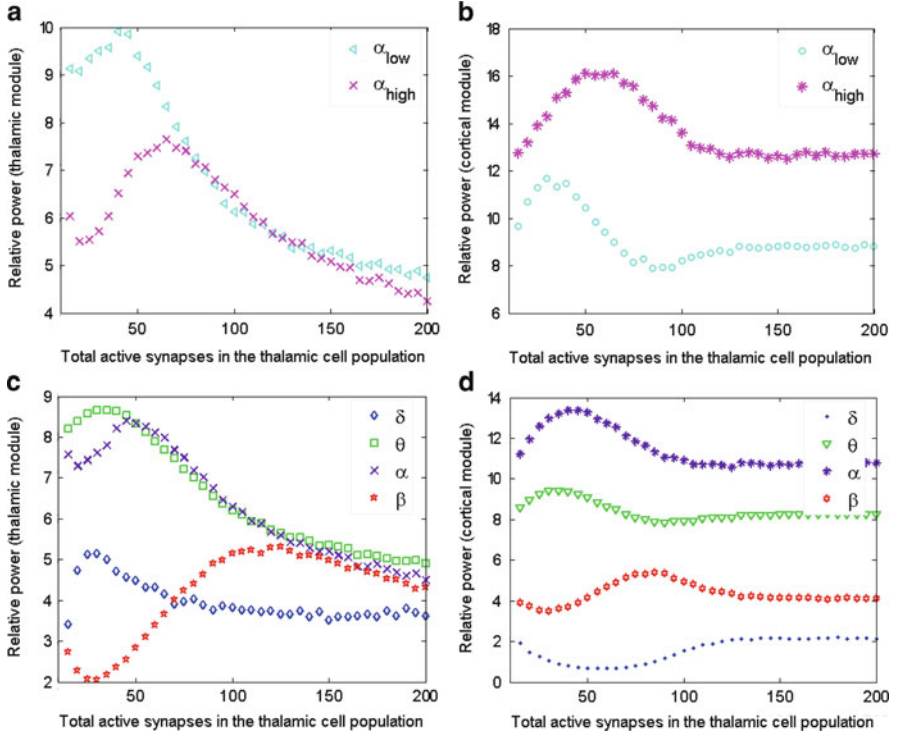
Fig. 5 (**a** and **b**) Relative power within the upper (α_{high}:10.5–13 Hz) and lower alpha (α_{low}:7.5–10 Hz) bands in the (**a**) thalamic and (**b**) cortical modules. (**c**) and (**d**) Relative power within the delta (δ:0.5–3 Hz), theta (θ:3.5–7 Hz), alpha (α:7.5–13 Hz) and beta (β:13.5–20 Hz) frequency bands in the (**a**) thalamic and (**b**) cortical modules. The results are obtained by variation of the parameter T about its basal value of 100. Thus, the synaptic connectivity in the thalamic module is varied about their basal values. However, the synaptic connectivity parameters of the cortical module are not varied in this study and are fixed at their basal values as defined in Table 2

be emphasised here that the variation of T about this basal value will change only the thalamic connectivity parameters. The connectivity parameters of the cortical module are not varied in this study and are kept at their basal values as defined in Table 2. Furthermore, the thalamo-cortical connectivity parameter C_{pte} is a constant [56] and is not affected with the variation of T. Thus, the output of both thalamic and cortical modules are analysed corresponding to variation of synaptic connectivity in the thalamic module only.

3.4 Results and Discussion

Power spectra analysis of the output of both the thalamic and the cortical modules with varying synaptic connectivity in the thalamic module only (i.e. the synaptic connectivity in the cortical module is not varied) is presented in Fig. 5. Normalised

power within the lower and upper alpha bands of the thalamic and cortical modules, respectively, with variation of the parameter T is shown in Fig. 5a, b. Normalised power within the delta, theta, alpha and beta bands in thalamic and cortical modules with variation of T is shown in Fig. 5c, d, respectively. We make the following observations:

1. The lower alpha band (α_{low}) peaks at values of $T < 50$ in both thalamic and cortical modules as observed in Fig. 5a, b, respectively. As T increases further, the power within this band decreases with an increase in power in the upper alpha band (α_{high}), which peaks for $T > 50$. This is consistent with the EEG alpha rhythm slowing in AD as well as with the results obtained in the study using the modARm (Sect. 2.2).
2. Diminished synapses correspond to increased power in delta and theta bands and decreased power in alpha and beta bands as shown in Fig. 5c, d. This conforms to the overall EEG slowing (increase in power within the lower frequency bands viz. delta and theta and decrease in power within the upper frequency bands viz. alpha and beta) reported in AD [36].
3. Power in delta and theta bands peak at values of $T < 50$ in both modules (Fig. 5c, d), followed by alpha at $T \approx 50$ and then by beta for $T > 50$. The observed phenomenon is like a ripple effect in the power of successive bands with increasing values of synaptic activity (T). It is now well known that the oscillatory activity of the brain changes with various states of sleep and wakefulness [60]. For example, the beta activity is more prominent when performing a physical or mental task, which might be linked with increased synaptic activity in a cell population; the delta rhythm is more prominent in states of sleep and correspond to reduced synaptic activity. Thus, our result conforms to the change in dominant frequency of oscillation with changing states of the brain.
4. Within each band in Fig. 5c, d, an increase in synaptic connectivity above a threshold does not affect the power, while a decrease below the threshold causes a sharp change in power. For the thalamic module (Fig. 5c), the threshold is at $T \approx 200$, while for the cortical module (Fig. 5d), this is at $T \approx 150$. The results imply a condition of redundancy—an all-pervading feature in biological information processing. We speculate that redundancy in synaptic connectivity within the thalamo-cortico-thalamic circuitry prevents disruption of brain oscillatory activity up to a certain threshold of synaptic depletion. This is consistent with the idea of cognitive reserve in AD compensation [1], a condition which prevents the clinical diagnosis of AD until at an advanced stage when severe damage has already been done to neurons and their synaptic pathways.

Additionally, the results emphasise that the synaptic parameter variation in the thalamic module only has a significant effect on the cortical module behaviour. This is a remarkable observation and is in agreement with several studies that emphasise the vital role of the thalamus in modulating cortical oscillations [3, 59], so much so that it has been speculated to be the "rosetta stone" in several neurological disorders [42].

4   Conclusion

We have been using neural mass models to improve understanding of neuronal correlates of abnormal electroencephalogram (EEG) in Alzheimer's disease (AD). While power spectra analysis is the more common trend in analysing EEG, it has failed to underpin a means of early detection of AD, a crucial aspect in being able to treat the condition. In this context, multi-modal techniques such as combination of power spectra and non-linear dynamical analysis have been proposed to gain understanding of the EEG signatures in AD. In the first part of this study, we have presented a combined study of power spectral and non-linear dynamical analysis on a simple thalamocortical model. Our study has shown that the model essentially behaves in two specific modes, viz. a point-attractor mode and a limit-cycle mode. The point of bifurcation is an area where the model oscillates with the maximum frequency of oscillation within the alpha band. Furthermore, with increasing values of the inhibitory connectivity from the TRN cell population to the TCR cell population, there is a decrease in the peak frequency of oscillation within the alpha band—a hallmark of EEG in AD. These observations emphasise the importance of such multi-modal techniques in model-based studies towards playing a meaningful role as a supporting tool in AD research.

However, the model used in the above-mentioned study is a simplistic representation of the intricate feed-forward and feed-back connectivities in the thalamocortical circuitry of the brain; more biologically plausible models would be desirable, especially when investigating biological phenomena such as EEG signatures. Towards this, we have presented an enhanced thalamo-cortico-thalamic model in the second part of the study. The synaptic connectivity and layout in the model is based on experimental biological data. A power spectra analysis on the model output and within the delta (1–3 Hz), theta (4–7 Hz), alpha (7–13 Hz) and beta (14–30 Hz) bands is presented. An important observation is the slowing of alpha rhythms with increasing values of the inhibitory synaptic connectivity from the TRN to the TCR cell population—this is consistent with the studies on the simple thalamocortical model mentioned above. Furthermore an overall slowing in EEG (i.e. an increase in power within the lower frequency bands and a decrease in power within the higher frequency bands) is observed using the enhanced model, once again conforming to experimental observations on EEG characteristics in AD. It may be speculated that a combined study of power spectra and bifurcation analysis of this enhanced model will add to the observations made on the model in previous research [10].

References

1. Abuhassan, K., Coyle, D., Maguire, L.P.: Investigating the neural correlates of pathological cortical networks in Alzheimer's disease using heterogeneous neuronal models. IEEE Trans. Biomed. Eng. **59**, 890–896 (2012)

2. Adeli, H., Ghosh-Dastidar, S., Dadmehr, N.: Alzheimer's disease: models of computation and analysis of EEGs. Clin. EEG Neurosci. **36**(3), 131–136 (2005)
3. Amsallem, B., Pollin, B.: Possible role of the nucleus reticularis thalami (nRT) in the control of specific, non-specific thalamic nuclei and cortex activity. Pain **18**, S-283 (1984)
4. Aradi, I., Erdi, P.: Computational neuropharmacology: dynamical approaches in drug discovery. Trends Pharmacol. Sci. **27**(5), 240–243 (2006)
5. Babloyantz, A., Destexhe, A.: Low-dimensional chaos in an instance of epilepsy. Proc. Natl. Acad. Sci. USA **83**, 3513–3517 (1986)
6. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: A computational modelling approach to investigate alpha rhythm slowing associated with Alzheimer's Disease. In: Proceedings of the Brain Inspired Cognitive Systems (BICS), Madrid, pp. 382–392 (2010)
7. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: Thalamocortical circuitry and alpha rhythm slowing: an empirical study based on a classic computational model. In: Proceedings of the International Journal of Neural Networks (IJCNN), Barcelona, pp. 3912–3918 (2010)
8. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: Alpha and theta rhythm abnormality in Alzheimer's disease: a study using a computational model. In: Hernandez, C., Gomez, J., Sanz, R., Alexander, I., Smith, L., Hussain, A., Chella, A. (eds.) Advances in Experimental Medicine and Biology, vol. 718, pp. 57–73. Springer, New York (2011)
9. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: Assessing retino-geniculo-cortical connectivities in Alzheimer's disease with a neural mass model. In: Proceedings of the IEEE Symposium Series in Computational Intelligence (SSCI), Paris, pp. 159–163 (2011)
10. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: A thalamo-cortico-thalamic neural mass model to study alpha rhythms in Alzheimer's disease. Neural Netw. **24**, 631–645 (2011)
11. Bhattacharya, B.S., Coyle, D., Maguire, L.P.: Assessing alpha band event-related synchronisation/desynchronisation with a mutually coupled thalamo-cortical circuitry model. J. Univers. Comput. Sci. **18**, 1888–1904 (2012)
12. Bhattacharya, B.S., Cakir, Y., Serap-Sengor, N., Maguire, L., Coyle, D.: Model-based bifurcation and power spectral analyses of thalamocortical alpha rhythm slowing in Alzheimer's disease. Neurocomputing **115**, 11–22 (2013). http://dx.doi.org/10.1016/j.neucom.2012.10.023
13. Braak, H., Braak, E.: Alzheimer's Disease affects limbic nuclei of the thalamus. Acta Neuropathol. **81**, 261–268 (1991)
14. Brenner, R.P., Ulrich, R.F., Spiker, D.G., Sclabassi, R.J., Reynolds, C.F., Marin, R.S., Boller, F.: Computerized EEG spectral analysis in elderly normal, demented and depressed subjects. Electroencephalogr. Clin. Neurophysiol. **64**, 483–492 (1986)
15. Cantero, J.L., Atienza, M., Salas, R.M.: Human alpha oscillations in wakefulness, drowsiness period, and REM sleep: different electroencephalographic phenomena within the alpha band. Clin. Neurophysiol. **32**, 54–71 (2002)
16. Cantero, J.L., Atienza, M., Cruz-Vadell, A., Suarez-Gonzalez, A., Gil-Neciga, E.: Increased synchronization and decreased neural complexity underlie thalamocortical oscillatory dynamics in mild cognitive impairment. Neuroimage **46**, 938–948 (2009)
17. da Silva, F.H.L., van Lierop, T.H.M.T., Schrijer, C.F., van Leeuwen, W.S.: Essential differences between alpha rhythms and barbiturate spindles: spectra and thalamo-cortical coherences. Electroencephalogr. Clin. Neurophysiol. **35**, 641–645 (1973)
18. Dauwels, J., Vialatte, F., Cichocki, A.: Diagnosis of Alzheimer's disease from EEG signals: Where are we standing? Neuroimage **49**, 668–693 (2010). doi:10.1016/j.neuroimage.2009.06.056
19. de Jong, L.W., van der Hiele, K., Veer, I.M., Houwing, J.J., Westendorp, R.G.J., Bollen, E.L.E.M., de Bruin, P.W., Middelkoop, H.A.M., van Buchem, M.A., van der Grond, J.: Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain **131**, 3277–3285 (2008)
20. Erdi, P., Kiss, T., Toth, J., Ujfalussy, B., Zalanyi, L.: From systems biology to dynamical neuropharmacology: proposal for a new methodology. IEEE Proc. Syst. Biol. **153**(4), 299–308 (2006)

21. Ermentrout, B.: Simulating, Analyzing, and Animating Dynamical Systems: A Guide to Xppaut for Researchers and Students (Software, Environments, Tools). Society for Industrial Mathematics, Philadelphia (2002)
22. Ermentrout, B.: Xppaut. Scholarpedia **1**(10), 1399 (2006)
23. Freeman, W.J.: Linear analysis of the dynamics of neural masses. Annu. Rev. Biophys. Bioeng. **1**, 225–256 (1972)
24. Freeman, W.J.: Mass Action in the Nervous System, 1st edn. Academic, New York (1975)
25. Freyer, F., Roberts, J.A., Becker, R., Robinson, P.A., Ritter, P., Breakspear, M.: Dynamic mechanisms of multistability in the human alpha rhythm. J. Neurosci. **31**, 6353–6361 (2011)
26. Gasser, U.S., Rousson, V., Hentschel, F., Sattel, H., Gasser, T.: Alzheimer disease versus mixed dementias: an EEG perspective. Clin. Neurophysiol. **119**, 2255–2259 (2008)
27. Grimbert, F., Faugeras, O.: Bifurcation analysis of Jansen's neural mass model. Neural Comput. **18**, 3052–3068 (2006)
28. Hines, M.L., Morse, T., Migliore, M., Carnevale, N.T., Shepherd, G.M.: ModelDB: a database to support computational neuroscience. J. Comput. Neurosci. **17**(1), 7–11 (2004)
29. Horn, S.C.V., Erisir, A., Sherman, S.M.: Relative distribution of synapses in the A-laminae of the lateral geniculate nucleus of the cat. J. Comp. Neurol. **416**, 509–520 (2000)
30. Hornero, R., Abasolo, D., Escudero, J., Gomez, C.: Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. Philos. Trans. R. Soc. A **367**, 317–336 (2009)
31. Hughes, S.W., Crunelli, V.: Thalamocortical mechanisms in EEG alpha rhythms and their pathological implications. Neuroscientist **11**(4), 357–372 (2005)
32. Izhikevich, E.M.: Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting. MIT, Cambridge (2007)
33. Jansen, B.H., Rit, V.G.: Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol. Cybern. **73**, 357–366 (1995)
34. Jelles, B., Scheltens, P., van der Flier, W.M., Jonkman, E.J., da Silva, F.H.L., Stam, C.J.: Global dynamical analysis of the EEG in Alzheimer's Disease: frequency-specific changes of functional interactions. Clin. Neurophysiol. **119**, 837–841 (2008)
35. Jeong, J.: Nonlinear dynamics of EEG in Alzheimer's disease. Drug Dev. Res. **56**, 57–66 (2002)
36. Jeong, J.: EEG dynamics in patients with Alzheimer's disease. Clin. Neurophysiol. **115**, 1490–1505 (2004)
37. Jones, E.G.: The Thalamus, vols. I and II, 1st edn. Cambridge University Press, Cambridge (2007)
38. Jonkman, E.J.: The role of the electroencephalogram in the diagnosis of dementia of the Alzheimer type: an attempt at technology assessment. Clin. Neurophysiol. **27**, 211–219 (1997)
39. Li, X., Coyle, D., Maguire, L.P., Watson, D., McGinnity, T.: Gray matter concentration and effective connectivity changes in Alzheimer's disease: a longitudinal structural MRI study. Neuroradiology **53**, 733–748 (2010)
40. Marten, F., Rodrigues, S., Suffczynski, P., Richardson, M.P., Terry, J.R.: Derivation and analysis of an ordinary differential equation mean-field model for studying clinically recorded epilepsy dynamics. Phys. Rev. E **79**(021911), 1–7 (2009)
41. MATLAB: Version 7.10.0 (R2010a). The MathWorks Inc., Natick (2010)
42. McCormick, D.A.: Are thalamocortical rhythms the Rosetta Stone of a subset of neurological disorders? Nat. Med. **5**(12), 1349–1351 (1999)
43. McCormick, D.A., Bal, T.: Sleep and arousal: thalamocortical mechanisms. Annu. Rev. Neurosci. **20**, 185–215 (1997)
44. Moran, R., Kiebel, S., Stephan, K., Reilly, R., Daunizeau, J., Friston, K.: A neural mass model of spectral responses in electrophysiology. Neuroimage **37**, 706–720 (2007)
45. Moretti, D.V., Babiloni, C., Binetti, G., Cassetta, E., Forno, G.D., Ferreric, F., Ferri, R., Lanuzza, B., Miniussi, C., Nobili, F., Rodriguez, G., Salinari, S., Rossini, P.M.: Individual analysis of EEG frequency and band power in mild Alzheimer's disease. Clin. Neurophysiol. **115**, 299–308 (2004)

46. Pons, A.J., Cantero, J.L., Atienza, M., Garcia-Ojalvo, J.: Relating structural and functional anomalous connectivity in the ageing brain via neural mass modelling. Neuroimage **52**(3), 848–861 (2010)
47. Prinz, P.N., Vitiello, M.V.: Dominant occipital (alpha) rhythm frequency in early stage Alzheimer's Disease and depression. Electroencephalogr. Clin. Neurophysiol. **73**, 427–432 (1989)
48. Pritchard, W.S., Duke, D.W.: Measuring chaos in the brain: a tutorial review of nonlinear dynamical EEG analysis. Int. J. Neurosci. **67**, 31–80 (1992)
49. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C.: Estimation of multiscale neurophysiologic parameters by electroencephalographic means. Hum. Brain Mapp. **23**, 53–72 (2004)
50. Roth, C., Achermann, P., Borbely, A.A.: Alpha activity in the human REM sleep EEG: topography and effect of REM sleep deprivation. Clin. Neurophysiol. **110**, 632–635 (1999)
51. Scheibel, A.B.: Thalamus. In: Encyclopedia of the Neurological Sciences, pp. 501–508. Elsevier Science, Amsterdam (2003)
52. Sherman, S.M.: Interneurons and triadic circuitry of the thalamus. Trends Neurosci. **27**(11), 670–675 (2004)
53. Sherman, S.M.: Thalamus. Scholarpedia **1**(9), 1583 (2006)
54. Sherman, S.M., Guillery, R.W.: Exploring the Thalamus, 1st edn. Academic, New York (2001)
55. Soininen, H., Reinikainen, K., Partanen, J., Helkala, E.L., Paljarvi, L., Riekkinen, P.: Slowing of electroencephalogram and choline acetyltransferase activity in post mortem frontal cortex in definite Alzheimer's Disease. Neuroscience **49**(3), 529–535 (1992)
56. Sotero, R.C., Tujillo-Barreto, N.J., Iturria-Medina, Y.: Realistically coupled neural mass models can generate EEG rhythms. Neural Comput. **19**, 479–512 (2007)
57. Stam, C.: Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin. Neurophysiol. **116**, 2266–2301 (2005)
58. Stam, C., Pijn, J., Suffczyński, P., da Silva, F.L.: Dynamics of the human alpha rhythm: evidence for non-linearity? Clin. Neurophysiol. **110**, 1801–1813 (1999)
59. Steriade, M., Deschenes, M.: The thalamus as a neuronal oscillator. Brain Res. Rev. **8**, 1–63 (1984)
60. Steriade, M.M., McCarley, R.: Brain Control of Wakefulness and Sleep, 2nd edn. Kluwer, New York (2005)
61. Steriade, M., Gloor, P., Llinas, R.R., da Silva, F.H.L., Mesulam, M.M.: Basic mechanisms of cerebral rhythmic activities. Electroencephalogr. Clin. Neurophysiol. **76**, 481–508 (1991)
62. Stoller, A.: Slowing of the alpha-rhythm of the electroencephalogram and its association with mental deterioration and epilepsy. J. Mental Sci. **95**, 972–984 (1949). doi:10.1192/bjp.95.401.972
63. Suffczyński, P.: Neural dynamics underlying brain thalamic oscillations investigated with computational models. Ph.D. Thesis, Institute of Experimental Physics, University of Warsaw (2000)
64. Suffczyński, P., Kalitzin, S., Silva, F.H.L.D.: Dynamics of non-convulsive epileptic phenomena modelled by a bistable neuronal network. Neuroscience **126**, 467–484 (2004)
65. Tai, Y., Yi, H., Ilinsky, I.A., Kultas-Ilinsky, K.: Nucleus reticularis thalami connections with the mediodorsal thalamic nucleus: a light and electron microscopic study in the monkey. Brain Res. Bull. **38**(5), 475–488 (1995)
66. Theiler, J.: On the evidence for low-dimensional chaos in an epileptic electroencephalogram. Phys. Lett. A **196**, 335–341 (1995)
67. Tognoli, E., Lagarde, J., DeGuzman, G.C., Kelso, J.A.S.: The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. USA **104**(19), 8190–8195 (2006)
68. Tombol, T.: Short neurons and their synaptic relations in the specific thalamic nuclei. Brain Res. **3**, 307–326 (1967)
69. Ursino, M., Cona, F., Zavaglia, M.: The generation of rhythms within a cortical region: analysis of a neural mass model. Neuroimage **52**(3), 1080–1094 (2010)

70. Wendling, F., Bartolomei, F., Bellanger, J.J., Chauvel, P.: Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. Eur. J. Neurosci. **15**, 1499–1508 (2002)
71. Wilson, H.: Spikes, Decisions and Actions. Oxford University Press, Oxford (1999)
72. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interaction in localized populations of model neurons. J. Biophys. **12**, 1–23 (1972)
73. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik **13**, 55–80 (1973)
74. Xuereb, J.H., Perry, R.H., Candy, J.M., Perry, E.K., Marshall, E., Bonham, J.R.: Nerve cell loss in the thalamus in Alzheimer's disease and Parkinson's disease. Brain **114**, 1363–1379 (1991)
75. Zavaglia, M., Astolfi, L., Babiloni, F., Ursino, M.: A neural mass model for the simulation of cortical activity estimated from high resolution EEG during cognitive or motor tasks. J. Neurosci. Methods **157**, 317–329 (2006)
76. Zou, X., Coyle, D., Wong-Lin, K., Maguire, L.P.: Computational study of hippocampal-septal theta rhythm changes due to beta-amyloid-altered ionic channels. PLoS One **6**, e21579 (2011)
77. Zou, X., Coyle, D., Wong-Lin, K., Maguire, L.P.: Beta-amyloid induced changes in a-type k^+ current can alter hippocampo-septal network dynamics. J. Comput. Neurosci. **32**(3), 465–477 (2012)

A Meta-learning Approach for Protein Function Prediction

Dariusz Plewczynski and Subhadip Basu

Abstract One of the major challenges in the post-genomic era is to accurately model the interactions taking place in most cellular processes. Detailed characterization of such interactions is critical for understanding the principles of living cell molecular machinery on the system biology level. This book chapter contains a review of the multiscale protein biological function prediction algorithms that are founded on protein sequence analysis, three-dimensional structure comparison, biological function annotation, and finally molecular interactions. We include diverse computational methods used to predict the biological function for a given biomolecule using multiscale features, and more generally to model a meta-learning prediction system to analyze the impact of micro-dynamics on global behavior for selected biological systems, with important roles in chemistry, biology, and medicine.

1 Introduction

Many fundamentally important biological processes are inherently multiscale in nature. These include biophysical phenomena like protein-folding, protein–protein, protein–peptide interactions, protein–ligand docking, and posttranslational modifications (PTMs), to name a few examples. They are intimately coupled to molecular level micro-dynamics, yet their long-time behavior or spatially distant characteristics are related to mesoscale processes. Diverse mathematical methodologies have

D. Plewczynski (✉)
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Pawinskiego 5a Street, 02-106 Warsaw, Poland
e-mail: D.Plewczynski@icm.edu.pl

S. Basu
Department of Computer Science and Engineering, Jadavpur University, Kolkata 32, India
e-mail: subhadip@cse.jdvu.ac.in

been developed recently for modeling such nonlocal biological processes using varying resolution, or granularity of a model. This book chapter contains a review of the multiscale protein biological function prediction algorithms that are founded on protein sequence analysis, three-dimensional structure comparison, biological function annotation, and finally molecular interactions. The presented approaches span different scales of description of biomolecules, starting from single atoms, ending with a network biology context of biomolecules. We provide a variety of different applications of the proposed methodology to biomedical and biophysical problems. We include diverse computational methods used to predict the biological function for a given biomolecule using multiscale features, and more generally to model a meta-learning prediction system to analyze the impact of micro-dynamics on the global behavior for selected biological systems, with important roles in chemistry, biology, and medicine.

Modern molecular biology provides a vast amount of experimentally verified information, mostly coming from high-throughput studies, such as the Human Genome Project [1, 2], where large-scale sequencing is taking place. In typical drug design procedure pharmaceutical companies perform high-throughput screening studies, where biological activity on the selected protein targets is performed for hundreds of thousands of small chemical molecules. Typical microarray experiments store relative expression profiles for thousands of genes in selected time points. Collecting and verifying currently available experimental data is therefore an important goal of bioinformatics, a rapidly developing computational discipline, which focuses on completing the functional annotation of proteins, inhibitors, DNA/RNA molecules, genes, or more generally all types of biomolecules that can be found in living cells.

Yet, proteins or other molecules are not single, acting-alone entities, they are often working together by forming either stable, permanent complexes with others (like in the case of ribosome), DNA/RNA strains, or ligands (natural metabolites), or interacting transiently with each other, with RNA/DNA molecules, and with small chemical entities. Therefore, a more general term, namely *Interactome*, was introduced that describes the whole universe of molecular interactions in cells, including the protein–protein, protein–ligand, and protein–DNA/RNA interactions. Systems biology addresses those issues, by experimental and theoretical analysis of interactions between biomolecules. Whole proteomes experimental high-throughput techniques rapidly populate available databases with a large amount of average quality data. Such data has to be further validated to qualify each data source contributing to the completion of the interactome. Those collaborative approaches are significantly limited by both time and total cost required for obtaining an accurate and complete map of protein interactions. Therefore, computational methods are beginning to be used to automate the procedure of the careful selection of high quality subsets of available data.

On the other hand, the gap between the number of known protein sequences and the number of crystallized structures is growing rapidly. A three-dimensional structure determines the protein's function; therefore, computational techniques have to be used in order to narrow the gap by predicting structures using only

sequence information. However, despite recent progress in the protein structure prediction community, greatly facilitated by the Critical Assessment of Protein Structure Prediction (CASP) experiment [3], the prediction of the three-dimensional protein structures from their amino acids sequence remains one of the major challenges in modern molecular biology. Therefore, the computational methods have to be designed in order to facilitate the functional annotation of proteins based only on partial biological information describing biomolecules (such as sequences, in the case of proteins). Such *bio-algorithms* are typically tested on a smaller number of examples that are very well characterized by various complementary experimental methods. In the case of proteins, the subset of known proteins with crystallized three-dimensional structures, where both protein sequence and structure is directly linked to performed protein biological function. The ability to learn knowledge on a smaller subset of available data, and apply it on diluted and very noisy biological information, performing data mining during large-scale experimental studies, is the core idea behind machine learning techniques applied in bioinformatics.

Those automatic algorithms nevertheless are focused on characterizing specific, pre-defined features of biomolecules, even if they are trying to study and predict interactions between those components of biological systems. Systems biology is trying to understand how these interactions lead to the function and behavior of that system (for example, the enzymes and metabolites in a metabolic pathway). The living cell is not just the collection of the above-mentioned *life* building blocks, described by their features, or functional annotations. The complex biological systems have to be understood on a higher level, where the organization of those single modules, or biomolecules, is emerges from their individual behaviors, or characteristics. In order to do so, first the interacting agents have to be described, as is done usually in typical low-throughput molecular biology experiments. Secondly, the interactions between those bio-agents have to be characterized, for example, by collecting data in high-throughput studies. Then, computational methods have to be applied in order to combine both levels of description and link agents and their interactions in order to better understand the biological complexity of *life*.

Summarizing, the proteins, genes, small chemical molecules, inhibitors, metabolites are linked in cells with varying scales of molecular interactions, such as protein–protein, protein–ligand, and protein–DNA/RNA. Those interactions are taking place in most cellular processes; therefore, detailed characterization of the interaction repertoire is critical for understanding the principles of living cell molecular machinery on the system biology level. The major challenge in the post-genomic era is to accurately model the organization of those genetic networks, signaling and metabolic pathways, and details of molecular interactions between proteins and their natural or artificial inhibitors, and to understand how they contribute to cellular and organism phenotypes.

1.1 The Goal of the Research

The main goal of the research work presented here is to undertake the challenge of automatically acquiring, storing, organizing, refining, analyzing, and finally building useful working hypotheses for those enormous bioinformatics datasets, especially in the context of multiscale protein function prediction. The integration of those experimental results with previously stored biological knowledge has to be done efficiently, allowing for detection of false or erratic information; both in previously acquired data objects and in newly processed ones. Therefore, new theoretical algorithms have to be introduced, and a new design of the previously used approaches is needed to handle those challenges on the whole genome scale. Due to limitations of traditional machine learning algorithms, various methods of meta-learning or computational intelligence are used to address those bioinformatics problems with very promising results. The importance and successes of those approaches over a diverse range of bioinformatics applications should encourage other scientists to apply these methods to their research.

Computational intelligence methods including traditional machine learning approaches, ensemble methods, artificial neural networks (ANNs), evolutionary algorithms, fuzzy systems, or cognitive computing have been developed during several decades of development. The recent development of CI and the ensemble learning research field is now extensively performed via hybridization and extensions of those algorithms, also in the context of bioinformatics and biomedical applications. This work is aimed at unification of both theoretical algorithms from computer sciences and their applications in biology, with meta-learning based bio-algorithms as the bridge between those two worlds.

2 Methods

This work is focused on various interdisciplinary applications of machine learning and data-mining techniques in bioinformatics. The computational methods are used to predict the biological functions for a given biomolecule in various scales, and more generally to model and analyze the selected biological systems. This section includes a description of selected machine learning, clustering, and computational intelligence algorithms, together with their applications in protein function annotation, protein structure prediction, the identification of proteins interactions with small chemical molecules, and finally the analysis of interactions in living cells. The computational methods presented here are based on various machine learning algorithms, dynamic programming methods, several clustering techniques, or similarity searchers. Novel consensus techniques are used to extend classical machine learning algorithms in many applications. Such an approach is inspired by meta-server applications in the protein structure prediction field. Those novel algorithms are now starting to be popular both in biomedical and bioinformatics applications under the general name of meta-approaches or meta-learning.

2.1 Meta-learning

The *Meta-learning* term is proposed in the context of computational intelligence as the successful combination of distributed intelligence approaches with traditional machine learning. It is useful in processing of metadata during typical machine learning experiments, when multiple algorithms are used to perform large-scale data mining. The metadata information can include various properties of the learning problem, performance measures of each learning algorithm, the structure of training data, or different patterns that can be deduced from the data. Therefore, the goal is to improve classical learning algorithms by combining different learning algorithms to effectively solve a given learning problem. Such meta-learning approaches are able to better resemble real world problems, allow to significantly improving the overall performance of learning algorithms. The flexibility of each machine learning algorithm is crucial in order to effectively store, organize, and process the acquired data. Each algorithm has its own inductive bias; often it is based on a preselected set of training data features, descriptors, also a set of assumptions about the nature of data. Therefore, it is able to build classification model only if its internal characteristics match the nature of external training data. This means that a learning algorithm may perform very well on one learning problem, but very badly on the next one.

The design of meta-model can be performed at least on three different scales. The microscale meta-learning approach combines several different machine learning algorithms, and builds the consensus between them. In the mesoscale solutions the larger number of independent methods is coupled with different representations of training data by performing feature selection process. In the macro-scale the whole semi-infinite ensemble of learning methods is trained on available data. The construction of meta-learning solution can therefore be drawn in three different layers:

(a) *Microensembles* are constructed using several classification methods combined into single consensus system, for example, by weighted voting procedure [4]. Such method was extensively used by us in applications' papers. We have reported 10 % improvement of error rate over the mean results for the wide range of various interdisciplinary applications.
(b) *Mesoscopic ensembles* may be implemented using standard software, where different machine learning methods are combined with features selection procedure [5]. The decrease of prediction error for the population of learning agents is linked with the distribution of quality of single methods, or statistical influence of selected features on the global meta-result.
(c) *Macroscopic meta-learning* solution can be approximated as semi-infinite learning ensemble, where mean-field theory can be used to get analytical stationary solutions for such system [6, 7].

The applications considered in this work focus on several examples from computational molecular biology that has become increasingly popular in modern research. It provides an excellent overview of machine learning approaches in the context of the complex bioinformatics problems with enormous amount of

Fig. 1 Structure of the HLA-DR1 peptide-binding site is illustrated as a sample application of meta-learning prediction system. Top view of the peptide-binding site, with HLA-DR1 residues in contact with the peptide indicated in *yellow* and *red* for the alpha and beta chain, respectively. The peptide residues are colored by atom type, with oxygen in *red*, nitrogen in *blue*, and carbon atoms in *green*. The alpha and beta chains of HLA-DR1 are indicated in *pale-blue* and *pink*, respectively
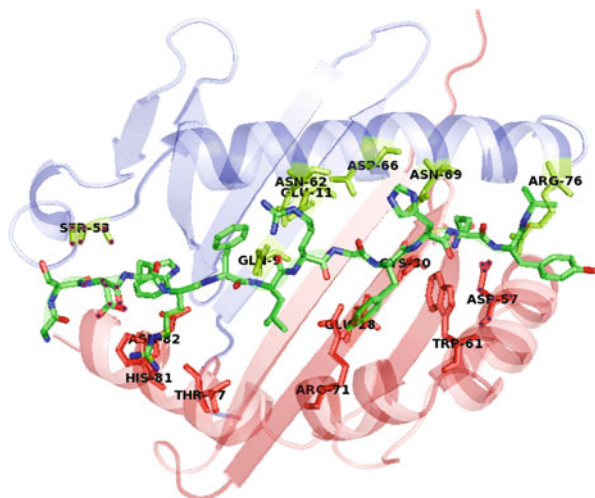
heterogeneous biological data. Apart from the theoretical foundations of meta-learning, the bioinformatics defines practice of this theoretical framework in the context of many biomedical engineering applications (for an example, see Fig. 1).

One can distinguish three major components of typical bioinformatics workflow, namely the biological data processing, features sampling and preparation, extraction of features (data acquisition), the clustering of relevant features or objects, extraction of similarities between various types of their descriptors (data clustering), and finally the construction of the ensemble of either different machine learning (ML) methods or similar/identical ML methods trained on different subsets of training data, and then the construction of consensus solution that is able to boost theoretical methods quality and precision (meta-learning). The actual treatment of the input training data in the meta-learning procedure is different from previous approaches. The ultimate goal of learning is to discover the relationships between the variables of a system (input, output, and hidden) from direct samples of the system. Most methods assume single representation of training data. Here, during data clustering one builds the set of multiple hypotheses by manipulating the training examples, input data points, target output (the class labels) of the training data, and by introducing randomness into the training data representation. Such approach provides the solid background for more advanced statistical analysis, the background noise extraction, and most informative features selection.

Although quite a few identifiers have been developed in this regard through various approaches, such as clustering, support vector machine (SVM), ANN, or K-nearest neighbor (KNN), and many other classifiers the way they operate the identification is basically individual. Yet, the proper approach usually takes into account the opinions from several experts rather than rely on only one when they are making critical decisions. Likewise, a sophisticated identifier should be trained by several different modes. A consensus of different classifiers often outperforms a single classifier: a learning algorithm searches the hypothesis space to find the best

possible hypothesis. When the size of training set is small, a number of hypotheses may appear to be optimal. An ensemble will average the hypotheses reducing the risk of choosing the wrong one. In addition, most classifiers perform a local search often getting stuck in local optima; multiple starting points provide a better approximation to the unknown function. Finally a single classifier may not be able to represent the true unknown function. A combination of hypotheses, however, may be able to represent this function. This is the core idea of *brainstorming*, which is the core procedure of the Meta-Learning approach. Our consensus approach is similar to other ensemble methods, yet differently from bagging, or boosting, it focuses of the use of heterogeneous set of algorithms in order to capture even remote, weak similarity of the predicted sample to the training cases.

In the following, we present the theoretical framework of the *brainstorming* consensus strategy followed by various applications of different computational intelligence, machine learning, or consensus learning techniques in several practical problems from chemo- and bio-informatics. The foundations of the brainstorming approach, namely the consensus between different types of machine learning algorithms, are described in the context of practical applications to prediction of PTMs of proteins and biological activity of small chemical molecules.

2.2   Theoretical Framework of Brainstorming

In general, we define *brainstorming* as a $n-$star quality consensus scheme as C_n^N, where N is the number of individual prediction routines or classifiers participating in the specific consensus strategy, and n ($1 \leq n \leq N$) is the quality of prediction. More specifically, $1-$star prediction says that any one of the possible N classifiers predicts the test sequence to be *positive* for the class type under consideration, and $N-$star represents that *all* classifiers agreed to the decision. Along this principle, we can consider the example of a neural network classifier, and we define the $10-$star quality consensus prediction C_n^{10} as the consensus over ten variations of hidden neurons (neurons in the single hidden layer are varied from 2 to 20 in steps of 2) for a specific performance measure. We designed three different performance measures in our work, one based on the area under the receiver operating characteristic (ROC) curve or the AUC measure, the others being the optimum recall (R) and precision (P) measures. C_n^{10} is defined over the optimum AUC performance. Similarly, we define C_n^{20} that combines 20 network predictions from AUC and R, and C_n^{30} that combines 30 network predictions from AUC, R and P optimized networks. In the following we first discuss the C_n^{10} consensus algorithm and then describe the other variations.

Let n_k^A, n_k^R, n_k^P be the MLP networks with K neurons in the hidden layer, designed to generate optimum AUC score (A), recall (R), and precision (P) scores, respectively, over the test dataset. Let p_k^A, p_k^R, p_k^P be the prediction results corresponding to the networks n_k^A, n_k^R, n_k^P for any unknown test pattern, where:

$$p_k^A = \begin{cases} 1; \text{ test pattern is classified as positive by } n_k^A \\ 0; \text{ otherwise} \end{cases}$$

Similarly, p_k^R, p_k^P also generate binary prediction decisions based on the classification confidence of the corresponding MLP classifiers n_k^R and n_k^P, respectively. Now the general $n-$star consensus is designed as C_n^N, where $n =$ minimum number of networks advocating for a test fragment to be positive. The sum of prediction scores is defined as S_p^N. For example, in case of C_n^{10} if $S_p^{10} = \sum_k p_k^A$; $k = 2$ to 20 in steps of 2, a test pattern is said to be predicted with $n-$star quality if $n \leq S_p^{10}$. Similarly, for C_n^{20}, we estimate $S_p^{20} = \sum_k p_k^A + \sum_k p_k^R$ and for C_n^{30}, $S_p^{30} = \sum_k p_k^A + \sum_k p_k^R + \sum_k p_k^P$, where $k = 2$ to 20 in steps of 2 in all cases.

In another variation of the brainstorming meta-learning strategy, more consensus prediction models are designed. C_n^3 is defined as the consensus among three best A, R, P networks. C_n^9 and C_n^{12} are defined as the consensus over the best networks across different feature sets.

For C_n^3 we first define a function Max_AUC_over_Testdata (MAT) to select the best performing network in any given optimization category. The performance is evaluated in terms of maximum AUC score over the unbiased test dataset. Therefore, we first compute $n_{\mathrm{MAT}}^A = \mathrm{MAT}(n_k^A)$; $k = 2$ to 20 in steps of 2. Similarly, we compute $n_{\mathrm{MAT}}^R = \mathrm{MAT}(n_k^R)$ and $n_{\mathrm{MAT}}^P = \mathrm{MAT}(n_k^P)$. The corresponding prediction scores are for the three selected networks are defined as p_{MAT}^A, p_{MAT}^R and p_{MAT}^P, respectively, and the sum of prediction scores as, $S_p^3 = p_{\mathrm{MAT}}^A + p_{\mathrm{MAT}}^R + p_{\mathrm{MAT}}^P$.

In the case of C_n^9 we use the MAT function separately for the three different feature sets, viz., HQI-8, HQI-24, and HQI-40 [5, 8]. We define the function $\mathrm{MAT}-\mathrm{HQI}-8$ to generate three best performing nets as $n_{\mathrm{MAT-HQI-8}}^A = \mathrm{MAT}-\mathrm{HQI}-8(n_k^A)$; $k = 2$ to 20 in steps of 2, and likewise $n_{\mathrm{MAT-HQI-8}}^R$ and $n_{\mathrm{MAT-HQI-8}}^P$. In the same way three best networks are generated by each of the functions $\mathrm{MAT}-\mathrm{HQI}-24$ and $\mathrm{MAT}-\mathrm{HQI}-40$. The sum of the corresponding prediction scores is then defined as:

$$S_p^9 = p_{\mathrm{MAT-HQI-8}}^A + p_{\mathrm{MAT-HQI-8}}^R + p_{\mathrm{MAT-HQI-8}}^P + p_{\mathrm{MAT-HQI-24}}^A + p_{\mathrm{MAT-HQI-24}}^R$$
$$+ p_{\mathrm{MAT-HQI-24}}^P + p_{\mathrm{MAT-HQI-40}}^A + p_{\mathrm{MAT-HQI-40}}^R + p_{\mathrm{MAT-HQI-40}}^P.$$

Similarly, for C_n^{12} we use four different MAT function separately for the four different feature sets, viz., $\mathrm{MAT}-\mathrm{HQI}\ 8$, $\mathrm{MAT}-\mathrm{HQI}-24$, $\mathrm{MAT}-\mathrm{HQI}-40$, and $\mathrm{MAT}-\mathrm{AMS3}-10$. The sum of the corresponding prediction scores is then defined as:

$$S_p^{12} = p_{\mathrm{MAT-HQI-8}}^A + p_{\mathrm{MAT-HQI-8}}^R + p_{\mathrm{MAT-HQI-8}}^P + p_{\mathrm{MAT-HQI-24}}^A + p_{\mathrm{MAT-HQI-24}}^R$$
$$+ p_{\mathrm{MAT-HQI-24}}^P + p_{\mathrm{MAT-HQI-40}}^A + p_{\mathrm{MAT-HQI-40}}^R + p_{\mathrm{MAT-HQI-40}}^P$$
$$+ p_{\mathrm{MAT-AMS3-10}}^A + p_{\mathrm{MAT-AMS3-10}}^R + p_{\mathrm{MAT-AMS3-10}}^P.$$

As discussed before, n-star quality result is obtained for any specific class type between the ANNs in any of the six ways, viz., C_n^{10}, C_n^{20}, C_n^{30}, C_n^3, C_n^9, or C_n^{12}. We

assign the statistical significance based on "how many ANNs agree that selected fragment is predicted as *Positive* for a given class type."

3 Case Studies

3.1 Prediction of Proteins Biological Function Using Sequence and Structural Similarity Searches

The sequence and structural similarity provide the important tool to infer the biological function of a protein. The structural similarity is able to recover twice as more distant relationships than sequence based methods, at the same error rate. Therefore, in [9] work we analyzed the ability of EC (enzyme classification) number prediction that describes the biological function for a given protein [9, 10]. The 3D-Hit structure comparison software developed in [11] provides a unique opportunity to perform fast comparison for a query protein structure in order to find its structural homologs. In another work [12], the Meta-BASIC protein sequence similarity method was modified and applied in order to find sequence homologs for proteins of medicinal relevance. The remote homology detection by Meta-BASIC uses both sequences and predicted secondary structure for similarity assignment. We provided the estimation of the upper limit of the number of molecular targets in the human genome that represent an opportunity for further therapeutic treatment. The druggability is defined here as the ability to bind small-molecule drug, therefore, being likely protein target for screening studies.

The sequence similarity searches were able to find around ∼6,300 human proteins that were similar to sequences of known protein targets as collected from DrugBank database. Therefore, our bioinformatics method estimates the size of druggable human genome to be around 20 % of human proteome. Those predicted protein targets from whole human genome present the unique opportunity for the virtual or experimental high-throughput screening. Figure 2 shows the three-dimensional structure of human glutathione S-transferase in complex with inhibitor.

3.2 Consensus Prediction of Protein Secondary Structures Using Two-Stage Multiclass SVMs

Secondary structure prediction is also crucial for understanding the variety of protein structures and performed biological functions. Prediction of secondary structures for new proteins using their amino acid sequences is of fundamental importance in bioinformatics. We proposed [13] a novel technique to predict protein secondary structures based on position-specific scoring matrices (PSSMs)

Fig. 2 Three-dimensional structure of human glutathione S-transferase in complex with inhibitor

and physicochemical properties of amino acids [14]. It is a two-stage approach involving multiclass SVMs as classifiers for three different structural conformations, viz., helix, sheet, and coil. In the first stage, PSSMs obtained from PSI-BLAST [15] and five specially selected physico-chemical properties of amino acids are fed into SVMs as features for sequence-to-structure prediction. Confidence values for forming helix, sheet, and coil that are obtained from the first stage SVM are then used in the second stage SVM for performing structure-to-structure prediction. The two-stage cascaded classifiers (PSP_MCSVM) are trained with proteins from RS126 dataset. The classifiers are finally tested on target proteins of critical assessment of protein structure prediction experiment-9 (CASP9) [3]. PSP_MCSVM with brainstorming consensus procedure performs better than the prediction servers like Predator [16], DSC [17], SIMPA96 [18], for randomly selected proteins from CASP9 targets. The overall performance is found to be comparable with the current state-of-the art tools.

3.3 Prediction of PTMs Using Local Sequence Motifs

We developed a tool that predicts the position of various PTM sites in proteins using only sequence information. Initial study published in [19] was focused on phosphorylation sites identification, and the extended work covering all known types of posttranslational single residue modifications was presented in [20]. Both versions of the method use the ensemble of SVM machine learning algorithms trained using different representations of training objects, i.e. 9-amino acid long fragments of protein sequences dissected around posttranslationally modified sites. Those fragments were extracted from Swiss-Prot database as experimentally confirmed to be phosphorylated by any kinase.

The different representations are build as vectors in a multidimensional abstract space of short sequence fragments linked with different physico-chemical features of amino acids, the sequence profile dissected from the whole protein homology alignments, local structural conformation, etc. Those multiple representations are combined with SVM algorithms (linear, polynomial, and radial kernel functions) to provide statistical models for PTM site predictors.

3.4 Automotif Server Version 3.0

In [4] we replaced the SVM method with ANNs (multilayer perceptron), also simplifying the selection of representations. We presented the recent update of AMS algorithm for identification of PTM sites in proteins based only on sequence information, using ANN method. The query protein sequence is dissected into overlapping short sequence segments. Ten different physico-chemical features describe each amino acid; therefore, nine residues long segment is represented as a point in a 90-dimensional space. The database of sequence segments with experimentally confirmed PTM sites are used for training a set of ANNs. The efficiency of the classification for each type of modification and the prediction power of the method is estimated here using recall (sensitivity), precision values, the area under ROC curves, and leave-one-out cross validation (LOOCV) tests. The significant differences in the performance for differently optimized neural networks are observed, yet the AMS 3.0 tool (http://ams3.bioinfo.pl) integrates those heterogeneous classification schemes into the single consensus scheme, and it is able to boost the precision and recall values independent of a PTM type in comparison with the currently available state-of-the-art methods.

3.5 Automotif Server Version 4.0

The 2011 update of the Auto-Motif Service (AMS 4.0) predicts a wide selection of 88 different types of the single amino acid PTMs in protein sequences [5]. The selection of experimentally confirmed modifications is acquired from the latest UniProt and Phospho.ELM databases for training. The sequence vicinity of each modified residue is represented using amino acids physicochemical features encoded using high quality indices (HQI) obtained by automatic clustering of known indices extracted from AA index database [8]. For each type of the numerical representation, the method builds the ensemble of multilayer perceptron (MLP) pattern classifiers, each optimizing different objectives during the training (for example, the recall, precision, or area under the ROC curve (AUC)). The consensus is built using brainstorming technology, which combines multi-objective instances of machine learning algorithm, and the data fusion of different training objects representations, in order to boost the overall prediction accuracy of conserved short

sequence motifs. The performance of AMS 4.0 is compared with the accuracy of previous versions, which were constructed using single machine learning methods (ANNs, SVM). Our software improves the average AUC score of the earlier version by close to 7 % as calculated on the test datasets of all 88 PTM types. Moreover, for the selected most-difficult sequence motifs types it is able to improve the prediction performance by almost 32 %, when compared with previously used single machine learning methods. Summarizing, the brainstorming consensus meta-learning methodology on the average boosts the AUC score up to around 89 %, averaged over all 88 PTM types. Detailed results for single machine learning methods and the consensus methodology are also provided, together with the comparison to previously published methods and state-of-the-art software tools.

3.6 Protein Alignment Method Using Sequence-Structure Motifs

Defining blocks forming the global protein structure on the basis of local structural regularity is a very fruitful idea, extensively used in description, and prediction of structure from only sequence information. Over many years the secondary structure elements were used as available building blocks with great success. Specially prepared sets of possible structural motifs can be used to describe similarity between very distant, nonhomologous proteins. The reason for utilizing the structural information in the description of proteins is straightforward. Structural comparison is able to detect approximately twice as many distant relationships as sequence comparison at the same error rate.

In our previous paper [21] we provided a new fragment library for local structure segment (LSS) prediction called FRAGlib which is integrated with a previously described segment alignment algorithm SEA. A joined FRAGlib/SEA server provides easy access to both algorithms, allowing a one stop alignment service using a novel approach to protein sequence alignment based on a network matching approach. The FRAGlib used as secondary structure prediction achieves only 73 % accuracy in Q3 measure, but when combined with the SEA alignment, it achieves a significant improvement in pairwise sequence alignment quality, as compared to previous SEA implementation and other public alignment algorithms. The FRAGlib algorithm takes 2 min to search over FRAGlib database for a typical query protein with 500 residues. The SEA service aligns two typical proteins within around 5 min. The joined FRAGlib/SEA server will be a valuable tool both for molecular biologists working on protein sequence analysis and for bioinformaticians developing computational methods of structure prediction and alignment of proteins.

3.7 Prediction of Protein–Protein Interaction Prediction Using Domain–Domain Affinities and Frequency Tables

Protein–protein interactions control most of the biological processes in a living cell. In order to fully understand protein functions, knowledge of protein–protein interactions is necessary. Prediction of PPI is challenging, especially when the three-dimensional structure of interacting partners is not known. We proposed a novel knowledge-based prediction method [22], which predicts interactions between two protein sequences by exploiting their domain information. We trained a two-class SVM on the benchmarking set of pairs of interacting proteins extracted from the Database of Interacting Proteins [23]. The method considers all possible combinations of constituent domains between two protein sequences, unlike most of the existing approaches. Moreover, it deals with both single-domain proteins and multi-domain proteins; therefore, it can be applied to the whole proteome in high-throughput studies. Our machine learning classifier, following a brainstorming consensus approach, achieves accuracy of 86 %, with specificity of 95 %, and sensitivity of 75 %, which are better results than most previous methods that sacrifice recall values in order to boost the overall precision. Our method has on average better sensitivity combined with good selectivity on the benchmarking dataset.

4 Summary and Future Directions

First, we performed an extensive review of different physico-chemical features of proteins that have a dramatic impact on their role in a cell. In addition, several computational methods are used for extending the acquired experimental knowledge to unknown or newly sequenced proteins, where no structural information is available. The work sheds some light on details of the fundamental link between a protein's sequence, structure, and its function performed in the living cell. This systems biology approach is further validated on known high quality experimental examples.

In the second step of our analysis, we limited the discussion to those cases where the three-dimensional structures of both a protein and its metabolites, or interaction partners, are known. We evaluated the various algorithms that are trying to predict the correct structure of the protein–protein or protein–ligand complexes, and estimate their binding affinity value by scoring the strength of interactions. The results were later used as the foundation of the effective protocol for docking, within the optimization technique based on multiple linear regression (MLR).

The ultimate goal of those computational approaches is to provide the methodology for automatic characterization of interaction partners, using either amino acids sequences (or full atom representations) and/or three-dimensional structures (if some structural information is known). We focused here on practical applications

of the theoretical methods, also elucidating the software or web interfaces needed to run them in the high-throughput predictions on the whole genome scale. The proposed general validation of theoretical methods on real life, experimental data provides the best estimation of their accuracy. We compared here various clustering, machine learning or statistical methods for bioinformatics knowledge acquisition, processing and mining. In addition, we presented a novel computational intelligence algorithm, namely the brainstorming meta-learning technique applied to various problems from bioinformatics and chemoinformatics. It covered the analysis of interactions of proteins, functional links between protein function, structure and sequence, and other applications in the context of the life sciences.

In this book chapter, we are more focused on applications than on the theoretical foundation of meta-learning. This extensive review of applications of meta-learning is focused on bioinformatics, an enormously rich application field for mathematical methods. The complexity of scientific problems and the large amount of hetero-geneous biological data provide an excellent test-ground for machine learning approaches in a real-life context. In return, bioinformatics, while using different theoretical methods, can also offer serious advances in theoretical computational intelligence. Most computational approaches are based on comparative molecular similarity analysis of proteins with known and unknown characteristics.

Eventually, the broader goal of this project is to develop a multiscale com-putational model of the entire life cycle of living organisms. In this process, both the bottom-up and the top-down approaches need to converge to delineate the underlying functionalities of cellular processes. The real-time analysis of incoming time-dependent non-equilibrium data performed by modern large-scale data-mining techniques, or computational intelligence algorithms, provides the theoretical framework, which in turn allows for better understanding the flexibility of learning mechanisms observed in real biological systems. Recent advances in theoretical neuroscience allow for better understanding of the brain structure, dynamics and performed basic cognitive functions on the biology level (see Fig. 3). We hypothesize that those two presently distant areas of science, namely cognitive science and computational intelligence, will eventually merge into a single research area, where complex and time-dependent meta-learning systems, inspired by the real mammalian brain structure, can be used as the implementation of cognitive systems.

The crucial point here is that in addition to learning algorithms the coupling to sensors has to be provided in order to allow the artificial system (or modeled biological one) to perform the effective cognitive process when discovering the world that surrounds it. The ultimate goal of artificial intelligence studies, i.e., a constantly evolving meta-learner that in real time accumulates acquired information in the form of processed knowledge, is still long way from the present state of research. Both theoretical algorithms and hardware resources (computers or specialized accelerators) have to be improved in order to perform instant, rapid learning using different algorithms, when new input is presented to the system. Only then will the "intelligent" system be able to answer most of our expectations focusing on computational intelligence.
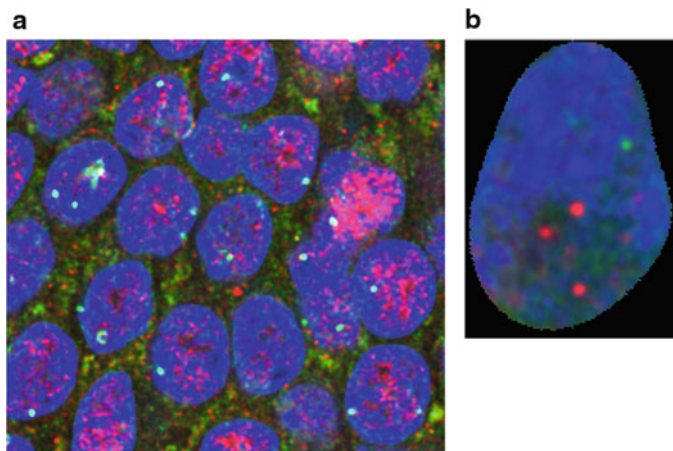
Fig. 3 (**a**) The confocal images of nuclei in living cells with selected active genes marked by fluorescent proteins. (**b**) Segmented view of a single nucleus. This type of data modality links the molecular view described in our chapter with imaging techniques, applied to living systems. *Data courtesy*: Nencki Institute of Experimental Biology, Warsaw, Poland

References

1. Watson, J.D.: The human genome project: past, present, and future. Science **248**, 44–49 (1990)
2. Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F.: Complementary DNA sequencing: expressed sequence tags and human genome project. Science **252**, 1651–1656 (1991)
3. Moult, J., Fidelis, K., Kryshtafovych, A., Tramontano, A.: Critical assessment of methods of protein structure prediction (CASP)—round IX. Proteins **79**, 1–5 (2011)
4. Basu, S., Plewczynski, D.: AMS 3.0: prediction of post-translational modifications. BMC Bioinformatics **11**, 210 (2010)
5. Plewczynski, D., Basu, S., Saha, I.: AMS 4.0: consensus prediction of post-translational modifications in protein sequences. Amino Acids **43**(2), 573–582 (2012)
6. Plewczynski, D.: Mean-field theory of meta-learning. J. Stat. Mech. **11**, P11003 (2009)
7. Plewczynski, D.: Landau theory of meta-learning. In: Security and Intelligent Information Systems, vol. 7053, pp. 142–153. Springer, Heidelberg (2012)
8. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Fuzzy clustering of physicochemical and biochemical properties of amino acids. Amino Acids **43**, 583–594 (2012)
9. von Grotthuss, M., Plewczynski, D., Ginalski, K., Rychlewski, L., Shakhnovich, E.: PDB-UF: database of predicted enzymatic functions for unannotated protein structures from structural genomics. BMC Bioinformatics **7**, 53 (2006)
10. von Grotthuss, M., Plewczynski, D., Vriend, G., Rychlewski, L.: 3D-Fun: predicting enzyme function from structure. Nucleic Acids Res. **36**, W303–W307 (2008)
11. Plewczyński, D., Paś, J., von Grotthuss, M., Rychlewski, L.: 3D-Hit: fast structural comparison of proteins. Appl. Bioinformatics **1**, 223 (2002)
12. Plewczynski, D., Rychlewski, L.: Meta-basic estimates the size of druggable human genome. J. Mol. Model. **15**, 695–699 (2009)
13. Chatterjee, P., Basu, S., Kundu, M., Nasipuri, M., Plewczynski, D.: PSP_MCSVM: brainstorming consensus prediction of protein secondary structures using two-stage multiclass support vector machines. J. Mol. Model. **17**, 2191–2201 (2011)

14. Kawashima, S., Kanehisa, M.: AAindex: amino acid index database. Nucleic Acids Res. **28**(374) (2000)
15. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402 (1997)
16. Frishman, D., Argos, P.: Seventy-five percent accuracy in protein secondary structure prediction. Proteins **27**, 329–335 (1997)
17. King, R.D., Sternberg, M.J.E.: Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci. **5**, 2298–2310 (1996)
18. Levin, J.M.: Exploring the limits of nearest neighbour secondary structure prediction. Protein Eng. **10**, 771–776 (1997)
19. Plewczynski, D., Tkacz, A., Wyrwicz, L.S., Rychlewski, L.: AutoMotif server: prediction of single residue post-translational modifications in proteins. Bioinformatics **21**, 2525–2527 (2005)
20. Plewczynski, D., Tkacz, A., Wyrwicz, L.S., Rychlewski, L., Ginalski, K.: AutoMotif Server for prediction of phosphorylation sites in proteins using support vector machine: 2007 update. J. Mol. Model. **14**, 69–76 (2008)
21. Plewczynski, D., Rychlewski, L., Ye, Y., Jaroszewski, L., Godzik, A.: Integrated web service for improving alignment quality based on segments comparison. BMC Bioinformatics **5**, 98 (2004)
22. Chatterjee, P., Basu, S., Kundu, M.M., Nasipuri, M., Plewczynski, D.: PPI_SVM: prediction of protein-protein interactions using machine learning, domain-domain affinities and frequency tables. Cell. Mol. Biol. Lett. **16**, 264–278 (2011)
23. Xenarios, I., Salwinski, L., Duan, X.J., Higney, P., Kim, S.-M., Eisenberg, D.: DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res. **30**, 303–305 (2002)

Part II
Biomedical Applications

Segmentation of the Carotid Arteries from 3D Ultrasound Images

Eranga Ukwatta and Aaron Fenster

Abstract Ultrasound (US) Doppler flow-velocity imaging has been used extensively in the diagnosis and management of carotid atherosclerosis. Doppler ultrasound-based measurement is a well-established screening tool for the assessment of stenosis severity. However, this method of measurement does not provide information on carotid plaque morphology, plaque vulnerability, or composition. In addition to stenosis severity, US-based phenotypes of carotid atherosclerosis include intima–media thickness (IMT), and total plaque area (TPA). More recently, vessel-wall-volume (VWV) and total-plaque-volume (TPV) have emerged as sensitive and useful US phenotypes of carotid atherosclerosis that measure plaque burden in 3D images. In order to accelerate the translation of these 3D US-based carotid atherosclerosis measurements to clinical practice, semiautomated methods of measurement are required to enable multiple observers to be trained in shorter time periods and with decreased inter-observer variability. This has stimulated investigators to develop accurate and robust segmentation algorithms allowing efficient quantification of carotid atherosclerosis. In this chapter, we demonstrate that 3D US is a viable technique for quantifying the progression and regression of carotid atherosclerosis and describe algorithms for segmentation of carotid vessels.

E. Ukwatta (✉)
Imaging Research Laboratories, Robarts Research Institute, Biomedical Engineering Graduate Program, University of Western Ontario, London, ON, Canada N6A 3K7
e-mail: eukwatta@robarts.ca

A. Fenster
Imaging Research Laboratories, Robarts Research Institute, Biomedical Engineering Graduate Program, Medical Biophysics, University of Western Ontario, London, ON, Canada N6A 3K7

1   Introduction

Despite more than five decades of active research since the Framingham study first identified risk factors for atherosclerosis, it is still the second leading cause of death in industrialized nations. Because most measurements of atherosclerosis currently in use provide indirect measurements of the disease, they may be inadequate for use in modern patient care and management and to test new treatments. New sensitive and specific measurements may provide critically needed and more specific risk management tools and more sensitive ways to assess new patient treatments.

Atherosclerosis and its sequelae, stroke, and heart attack are still the leading causes of mortality and morbidity in the developed world, and are expected to take a growing toll in the developing world. Stroke is the most common, serious neurological problem globally and the third leading cause of death among North American adults. Clearly, stroke represents a staggering mortality, morbidity, and economic cost. Improved methods to identify patients at increased risk for stroke, and better techniques to treat and monitor them will have an enormous impact.

Many new molecular therapies and phenotypic therapeutic targets for atherosclerosis are currently being investigated and developed. Parallel to the development of new therapies, an equally important requirement is ***the development of treatment-specific measurement tools or biomarkers***. As the numbers of new treatments for atherosclerosis are expected to increase rapidly, it is critical to continue to develop and validate cost-effective, noninvasive, and direct measurements or phenotypes of carotid atherosclerosis to use as efficacy targets for treatment and for monitoring therapy.

The utility of ultrasound (US) imaging in diagnosis and monitoring of carotid atherosclerosis is well established. Because it is noninvasive and technological advances are continually improving image quality, US imaging is achieving a greater role by providing direct views of the vessel wall and plaque. Thus, improved and well-characterized US imaging has the potential to be used for quantification of plaque surface morphology, volume, and composition, and therefore become the primary carotid atherosclerosis monitoring modality in patients at risk of stroke, who are being treated with plaque stabilization strategies.

2   3D Carotid Ultrasound Imaging

2.1   *Benefits of 3D Ultrasound Imaging*

Conventional 2D US imaging systems are used extensively to image the carotid arteries and diagnose carotid disease. These US systems are highly flexible, allowing users to manipulate hand-held US transducers freely over the neck to generate real-time 2D images of the carotid arteries and blood flow within them. However,

conventional 2D US systems suffer from the following disadvantages for use in carotid disease monitoring:

- The use of conventional 2D US imaging requires users to mentally integrate many 2D images to form an impression of three-dimensional (3D) pathology in the carotid arteries. Although this approach can be effective, at times, it leads to longer procedures and variability in diagnosis and monitoring changes in the vessels.
- Monitoring progression and regression of carotid atherosclerosis in response to nonsurgical therapy requires that the same view of the carotid arteries be images over time periods that can be as long as a few years. Since the 2D US imaging transducer is held and controlled manually, it is difficult to relocate of the 2D US image at the exact location and orientation in the body when imaging a patient. Thus, 2D US imaging is suboptimal as an accurate carotid disease monitoring technique.

Over the past two decades, many researchers and commercial companies have been developing efficient 3D US imaging systems for use in a wide variety of applications [1–7]. Because 3D US systems require significant computational speed for reconstructing and viewing of 3D US information in real- or near real-time on inexpensive systems, progress has been slow. However, advances in low-cost computer and visualization technology have now made 3D US imaging a viable technology that can be used in a wide range of applications. Thus, 3D US imaging capability is now widely available using commercial systems and ones developed in research laboratories. The challenge now is to develop efficient and informative 3D image analysis tools for use in monitoring carotid atherosclerosis in the large number of patients with the disease.

Before describing methods for segmentation and analysis of carotid atherosclerosis, we briefly review methods used for acquiring 3D carotid US images and the tools required for visualizing and measuring carotid disease. Further technical and computational aspects of 3D US can be found in recent review articles and books on the subject [8–15].

Because analysis of the carotid arteries requires at least a 4 cm length, use of real-time 3D systems (i.e., 4D systems with 2D arrays) with small transducer apertures cannot be used effectively. Thus, all current 3D US systems used to acquire images of the carotid arteries make use of conventional 1D transducer arrays that produce 2D US images. To produce a 3D image of the carotid arteries, the conventional US transducer must be moved over arteries to collect 2D images necessary to reconstruct the 3D US image. Thus, a method to track the position and orientation of the transducer must be used so that the acquired 2D images can be placed in their correct position and orientation relative to each other in the 3D image. Over the past decade, two methods have been developed to image the carotid arteries: mechanical linear scanners and magnetically tracked free-hand scanners.
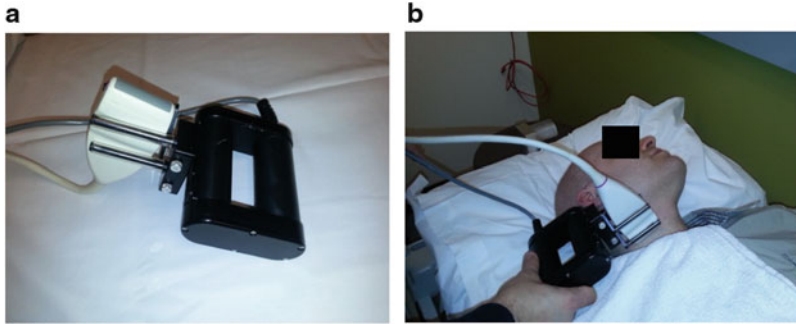
Fig. 1 (**a**) Photograph of a mechanical linear scanning mechanism used to acquire 3D carotid US images. The transducer is translated along the arteries, while conventional 2D US images are acquired by a computer and reconstructed into a 3D image in real-time. (**b**) Photograph of the system being used to scan the carotid arteries

2.2 Methods Used to Acquire 3D Carotid US Images

2.2.1 Mechanical Linear 3D Carotid Ultrasound Imaging

Linear mechanical scanners use a motorized mechanism to translate a conventional transducer linearly along the neck over the carotid arteries, as shown in Fig. 1, and two examples of linearly scanned 3D ultrasound images of carotid arteries with complex plaques are shown in Fig. 2. As the transducer is moved over the carotid arteries, transverse 2D images of the carotid arteries are acquired at regular spatial intervals. Since a motor is used to move the transducer linearly at a constant velocity, each acquired 2D image is spaced equally and all images are parallel to each other. This acquisition geometry makes 3D reconstruction easy and possible in real-time. Since mechanical mechanisms can be constructed with different lengths, the length of the scan can be made to be 4–6 cm. Typically, we acquire 2D US images every 0.2 mm. If the 2D US images are acquired at 30 fps, a 4 cm length will require 200 2D US images, which can be collected in 6.7 s without cardiac gating.

Because the 3D carotid US image is produced from a series of equally spaced conventional 2D US images, the resolution in the 3D image will not be isotropic. The spatial resolution of the image transverse to the vessel axis will depend on the resolution in the acquired 2D image plane. However, the spatial resolution in the 3D scanning direction (i.e., along the artery) will depend on the elevational resolution of the transducer as well as the spacing between the acquired images. The spacing between acquired image planes can be optimized by varying the translating speed and 2D US acquisition rate. Thus, the acquisition rate can be matched to the frame rate of the ultrasound machine, and the velocity can be adjusted so that the sampling interval is no more than half the elevational resolution of the transducer [16]. In this way, the spacing between acquired images is typically not limiting the spatial resolution. Since the elevational resolution is worse than the in-plane resolution of
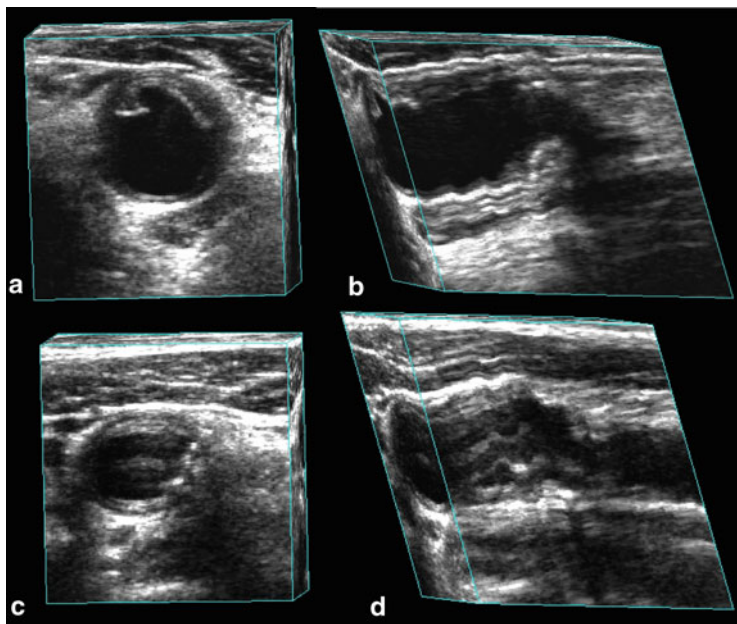
Fig. 2 Two 3D carotid ultrasound views of two different patients with complex and ulcerated carotid plaques. For each patient, a transverse (**a**) and (**c**) and longitudinal (**b**) and (**d**) views are shown side-by-side

the 2D US images, the resolution of the 3D US image will be the poorest in the 3D scanning direction. Therefore, a transducer with good elevational resolution should be used to obtain optimal results [13].

2.2.2 Free-Hand Scanning with Position Sensing

Since mechanical 3D US scanning mechanisms used for imaging the carotid arteries tend to be bulkier than conventional 2D US transducers, 3D US scanning techniques have been developed that do not require mechanical scanning devices. To use a conventional 2D US transducer for 3D US imaging, the position and orientation of the transducer must be tracked as it is moved over the carotid arteries. This can be accomplished by mounting a sensor on the transducer to allow measurement of the transducer's position and orientation as the transducer is moved over the body. In this way, the 2D US images and their relative location and orientation recorded and used to reconstruct the 3D US image [17]. Since the locations and orientations of the acquired 2D images are not predefined, an operator must move the transducer over the anatomy at an appropriate speed to ensure that the spatial sampling is proper and to ensure that the set of 2D images does not have any significant gaps. Several approaches for free-hand scanning have been developed,

but the approach most commonly used for carotid imaging makes use of magnetic field sensing technology: Ascension–Bird sensor [18]; Polhemus–Fastrak sensor [19]; and Northern Digital–Aurora sensor [20].

The magnetic field sensors approach makes use of a time-varying 3D magnetic field transmitter, which is placed near the patient, and a small receiver containing three orthogonal coils (with six degrees-of-freedom) mounted on the probe to sense the strength of the magnetic field in three orthogonal direction. The position and orientation of the transducer are calculated by measuring the strength of the three components of the local magnetic field as each 2D US image is acquired. This information is then used to reconstruct the 3D US image.

The magnetic field sensor 3D scanning approach allows free-hand scanning with a small and unobtrusive device that allows the transducer to be tracked. However, electromagnetic interference can compromise the accuracy of the tracking, and geometric distortions in the 3D US image can occur if ferrous (or highly conductive) metals are located nearby. Current magnetic field sensors produce excellent images and are less susceptible to sources of error, particularly ones that use a magnetic transmitter placed between the bed and the patient. Nonetheless, to minimize these potential sources of error, the position of the sensor relative to the transmitter must be calibrated accurately and precisely [21–29].

3   Carotid US Imaging Phenotypes

In this subsection, we briefly review the carotid US imaging phenotypes that have been used recently. As discussed in Sect. 1, carotid imaging techniques are increasingly becoming popular as tools to quantify surrogate end points of carotid atherosclerosis due to their capability of direct measurement of plaque as opposed to indirect measurements such as blood serum levels [30]. Such carotid imaging phenotypes allow assessment of atherosclerotic disease progression, and reduce the need for studying the effect of interventions on hard clinical end points, such as stroke or myocardial infarction (MI) [31].

There have been numerous US phenotypes of carotid atherosclerosis used in the past, which have been investigated as surrogate end points of vascular outcomes for monitoring the progression and regression of atherosclerotic plaque burden in patients during medical interventions. For example, Doppler ultrasound-based measurements of stenosis [32, 33], intima–media thickness (IMT) [34, 35], total plaque area (TPA) [36], total-plaque-volume (TPV) [37–39], vessel-wall-volume (VWV) [40, 41], and vessel-wall-plus-plaque thickness (VWT) maps [42–44]. Although Doppler US is a widely used tool for assessing stenosis severity [45, 46], it does not provide information on plaque morphology, plaque vulnerability, and artery remodeling. 2D US IMT measurement, a one-dimensional measurement of arterial thickness, is a reproducible measurement that has been shown to correlate significantly with the risk factors of stroke [47]. However, it is not sensitive to the changes in plaque burden, which is a stronger predictor of cardiovascular events [48,

49]. In addition, change in IMT occurs at a rate of fraction of a millimeter per year [50], resulting in long follow-up times in the monitoring study designs as well as algorithms with high precision to detect significant changes in IMT.

3.1 3D Carotid US Imaging Phenotypes

Recently, there has been an increased interest in the use of area and volumetric measurements of carotid plaque burden, because they are more reflective and sensitive to the changes of the plaque burden [36, 51–53]. With the availability of such measurements, it may be possible to show an effect of therapy with fewer numbers of patients in shorter clinical trials with imaging end points as opposed to longer clinical trials with hard end points involving large number of subjects. For example, the results from the 6-year follow-up of the Tromsø study [54] involving more than 6,000 subjects showed that in a general population, carotid TPA was a stronger predictor of myocardial infarction than IMT. With a 10-year follow-up, the study also reported that the TPA is a strong predictor of first-ever ischemic stroke than IMT [55].

With the availability of the 3D carotid ultrasound images, several 3D volumetric measurements have been used in clinical studies such as, 3D US TPV [37], 3D US VWV [40], and 3D US VWT maps [42, 53].

3D US TPV is a direct volumetric measurement of plaque burden in the left and right carotid arteries [38]. TPV measurements require an expert observer to delineate plaque–lumen and plaque–outer wall boundaries as shown in Fig. 3. TPV measurements encompass more of the complex plaque morphologies and geometries than TPA measurements that are performed only in longitudinal views. Therefore, volumetric measurements are expected to be even more sensitive than the two-dimensional measurements, because carotid plaque progression occurs in three dimensions. Thus, measurements of plaque volume have the potential to be more sensitive to change than do measurements of plaque area, IMT, and carotid stenosis [38]. Therefore, sample sizes required to test the effects of new therapies might be smaller for measurements of plaque volume than for 1D and 2D measurements [38].

The 3D US VWV is the volume enclosed between the outer wall and the lumen boundaries for the common carotid artery (CCA), internal carotid artery (ICA), and less commonly for external carotid artery (ECA). The measurement of 3D US VWV requires an observer to manually outline the lumen–intima/plaque and media–adventitia boundaries as shown in Fig. 4. The VWV measurements have a higher reproducibility than TPV measurements [40], due to the fact that lumen–intima/plaque and media–adventitia boundaries are more straightforward to interpret than plaque–lumen and wall boundaries in 3D US images. In addition, VWV boundaries measurements are more regular and circular, which may simplify the development of semiautomated segmentation techniques.
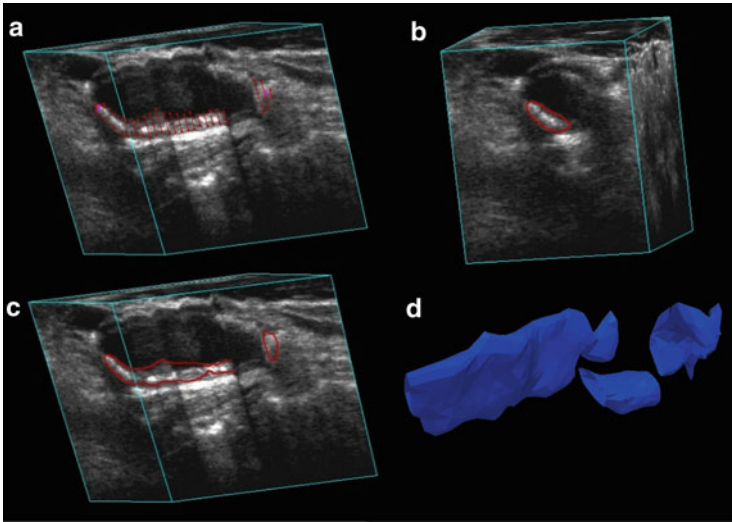
Fig. 3 Transverse and longitudinal views of the 3D US image are shown with overlaid manual segmentations for the generation of 3D US VWV. (**a**) Transverse and longitudinal view of a 3D carotid US image. (**b**) Manual delineations of the lumen and wall boundaries that are performed on transverse slices that are 1 mm apart. (**c**) and (**d**) The surfaces generated by interpolation of the 2D contours for the lumen and vessel wall boundaries using the manual segmentations

The 3D US phenotypes discussed so far are global measurements of plaque progression and/or wall thickening, which provide a single value of measurement such as an area or volume. While such phenotypes provide valuable quantitative information about global plaque burden, it does not identify the locations in the vessel, where volumetric changes are occurring. However, identification of the locations of change in plaque burden will enable more precise evaluation of changes in plaque burden during medical interventions. Chiu et al. [44] proposed 3D US VWT maps (see Fig. 5) and 3D US VWT-change maps (see Fig. 6) to quantify and visualize local changes in plaque morphology on a point-by-point basis using the segmented surfaces for the lumen and outer wall of the carotid arteries. To facilitate the visualization and interpretation of these maps for clinicians, Chiu et al. [43] proposed a technique to flatten the 3D VWT maps and VWT-change maps to 2D (see Fig. 6).

4   Segmentation Algorithms for 3D Carotid US Images

Image segmentation is a way of partitioning an image into two or more regions, where often the segmented region is an object of interest. To generate the volumetric measurements of carotid plaque burden, the outer wall, lumen, and plaque
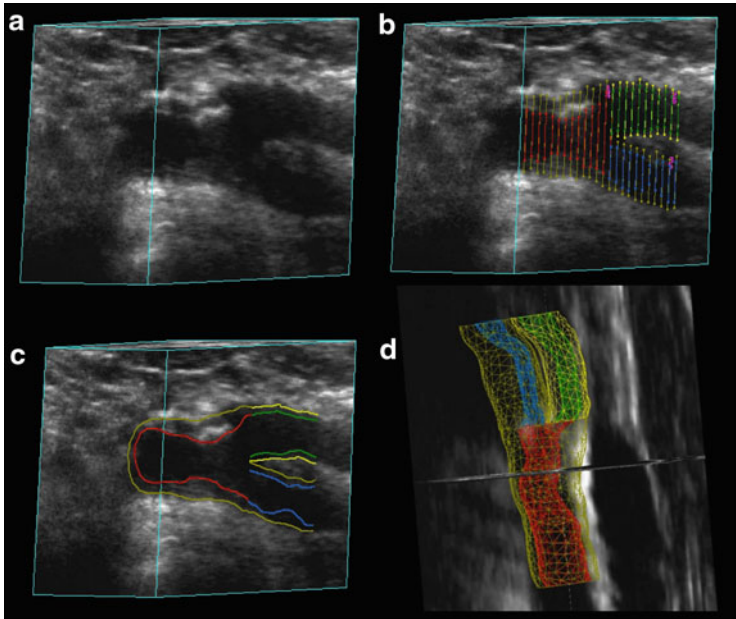
Fig. 4 Transverse and longitudinal views of the 3D US image are shown with overlaid manual segmentations for the generation of 3D US TPV. (**a**) Manual delineations of plaque boundaries are shown in *red*, which were performed on transverse views. (**b**) Transverse cross section of plaque surface. (**c**) Longitudinal and transverse view of the plaque surface. (**d**) The 3D plaque surface generated from interpolating the 2D transverse contours that are outlined manually
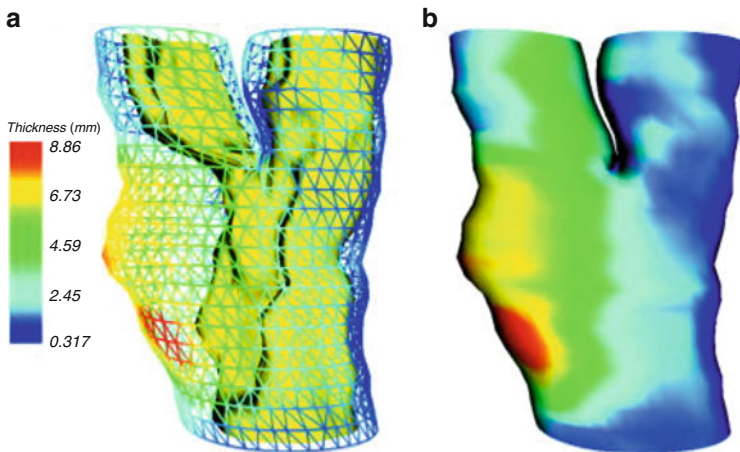
Fig. 5 (**a**) and (**b**) The 3D US VWT map color-coded and superimposed on the vessel wall, and the lumen of one patient. The *red region* on the VWT map (i.e., the region with maximum VWT) corresponds to a deep depression on the lumen surface (i.e., the location of a large plaque) [44]
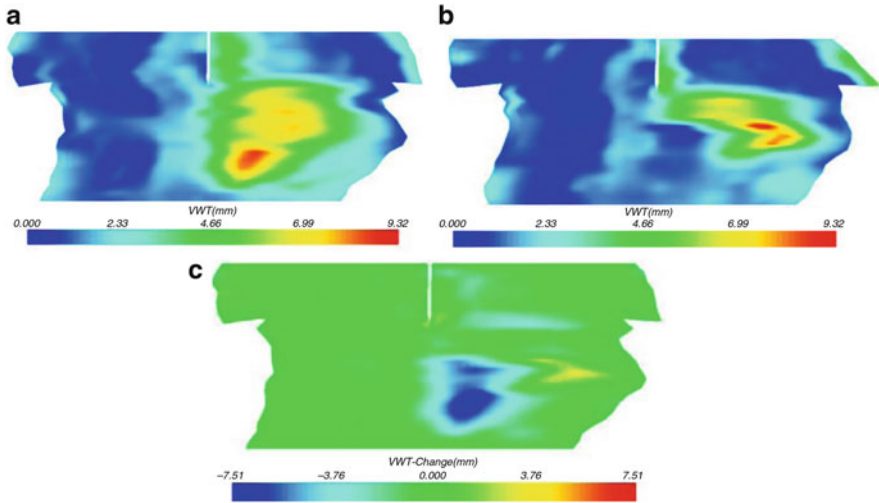
Fig. 6 The VWT maps at the first and second scan, and the VWT-change maps for a single patient. (**a**) The VWT map at baseline, (**b**) the VWT map at 3 months after baseline, and (**c**) the VWT-change map for the same patient [43]

boundaries are required to be segmented from 3D US images. In particular, for VWV and VWT measurements, the lumen and wall boundaries need to be segmented, whereas for TPV measurement the plaque boundaries need to be segmented. However, obtaining the lumen and outer wall boundaries prior to plaque segmentation may help to locate the plaque boundaries more efficiently.

4.1   Requirement for Segmentation Algorithms

Unfortunately, although 3D US VWV, TPV, and VWT maps provide many advantages compared to measurements of arterial wall thickening or TPA alone, they are still not widely used in research or clinical practice because of the inordinate amount of time required to train observers and to generate those 3D US phenotypes. In this regard, semiautomated methods for segmentation of the carotid outer wall, lumen, and plaque boundaries would greatly improve the time to train observers and for them to generate 3D US phenotypes with high reproducibility. In the following sections, the segmentation algorithms for carotid wall, lumen, and plaque boundaries will be described.

4.2 Segmentation of Carotid Wall and Lumen Boundaries

Currently, carotid outer wall and lumen segmentations are achieved either via manual segmentation or semiautomated segmentation techniques. Although, fully automated segmentation techniques are effective in some segmentation tasks of medical imaging, there are no techniques yet developed for 3D US carotid imaging. So far, manual segmentation is by far the most widely used technique for segmenting the carotid wall and lumen boundaries from 3D carotid US images [40, 41, 56–59]. However, there has been recent progress in the development of semiautomated techniques for 3D US segmentation, which are described in Sect. 4.2.2.

4.2.1 Manual Segmentation

Figure 4 shows manual segmentations of the outer wall and the lumen boundaries overlaid on a 3D US image and the surfaces of the outer wall and lumen for computing 3D US VWV. Typically, 15–20 mm portion of the artery is segmented including 10/15 mm section of the CCA from the carotid bifurcation and 5 mm section of the ICA from the carotid bifurcation. Manual segmentations of the carotid arteries from 3D US images are performed in transverse views. The observer first defines the bifurcation point and the axis of the CCA and ICA. This defined axis is used to obtain approximately orthogonal transverse slices to the artery for the segmentation. The observer is then shown the transverse and the longitudinal views of the artery simultaneously using multi-planar reformatting [13]. The observer first outlines the outer wall in transverse slices that are 1 mm apart, along the section of the artery that is to be outlined. The observer then outlines the lumen in the transverse slices. When the segmentation of both boundaries is completed for the CCA, the same segmentation steps are repeated for the ICA.

4.2.2 Semiautomated Segmentation Algorithms

Although, manual segmentation is still widely used for the carotid wall and lumen segmentation, semiautomated segmentation algorithms are preferred over the manual segmentation because manual techniques are tedious, time-consuming, and may be subject to high operator variability. In addition, manual segmentation demands more training and expertise from the observers. Therefore, developing semiautomated techniques for segmentation and analysis is important for translating the use of 3D carotid US phenotypes to clinical research and ultimately to clinical care. However, due to the low image contrast of the boundaries and less distinguishable intensity histograms, it is challenging for algorithms to segment 3D carotid US images robustly (Fig. 7).

A summary of the segmentation techniques for delineating the carotid outer wall and lumen from 3D US images is shown in Table 1. Most recent segmentation tasks are performed using optimization-based approaches, such as, active contours
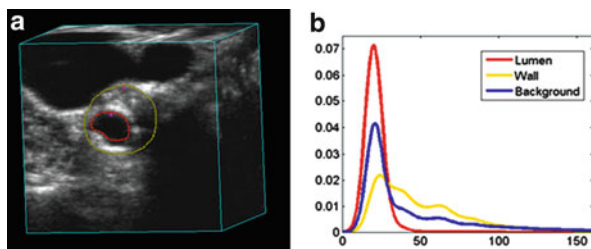
Fig. 7 This figure illustrates the degree of overlap in the intensity histograms for the lumen, wall, and the background. (**a**) Transverse view of a 3D US image with manual delineations of the lumen and the outer wall. (**b**) The corresponding intensity histogram for the lumen, outer wall, and the background region

Table 1 Previous papers describing carotid lumen and/or wall segmentations from 3D US B-mode images

Paper	Year	Dimension	Lumen	Wall	Time (min)	# images (type)
Ukwatta et al.	2011	2D	Level set	Level set	2.8 ± 0.4	21 (3D US)
Ukwatta et al.	2012	3D	Level set	Level set	1.72	21 (3D US)
Solovey	2010	3D	Level set	–	–	2 (3D US)
Gill et al.	2000	3D	Dynamic balloon model	–	30 s	1 (3D US)
Abolmaesumi et al.	2000	2D	Star algorithm	–	–	1 (B-mode)
Zahalka and Fenster	2001	2D	Active contours	–	25 s/slice	2 (3D US)
Mao et al.	2000	2D	Active contours	–	0.8 s/slice	7 (B-mode)

[60], level sets [61], and active shape models [62]. In such approaches, an objective function is defined, which describes the goodness of the segmentation. The objective function is then minimized to locate the correct boundary. Such objective functions typically contain two types of energy terms: external energy term, derived from the (pre-processed) image data, and internal energy terms, derived from high-level knowledge about the expected shape of the boundaries. A parameter is used to determine the relative importance of each term for the segmentation. Depending on the strength of the data term, this parameter(s) is be tuned for the segmentation task in hand. For example, if the data term is strong, more weight can be assigned to the data term.

As shown in Table 1, the algorithms developed for the segmentation of the carotid artery lumen and wall can be broadly categorized into two classes, which are based on the measurement of interest: methods that segment only the lumen boundary and ones that segment the lumen and the outer wall boundary. In this chapter, we briefly review all such algorithms, but focus more on the segmentation methods that are specifically designed for segmenting both boundaries for the purpose of generating volumetric phenotypes of plaque burden.

2D and 3D Methods that Segment Only a Single Boundary

Some methods have been developed to delineate only the lumen boundary [63–67]. For example, Gill et al. [64] proposed a semiautomated method based on dynamic balloon model to segment only the lumen boundary from 3D US images, but validated using a single image. After placing a seed mesh within the lumen, it is inflated to find the approximate boundary using an inflation force. The mesh is then refined using edge-based forces proximal to the lumen boundary. The algorithm yielded a submillimeter value for the mean separation between the manual and algorithm surface. However, the algorithm needed to be validated using multiple images to adequately evaluate its robustness and usefulness in practice. Zahalka and Fenster [67] proposed a lumen only segmentation method using a deformable model with a stopping term based on the image gradient, which was validated using only with a phantom. Solovey [66] proposed a level set-based method to segment the lumen by incorporating a region-based term that minimizes the probability distribution overlap and a weak geometric prior, which encourages convexity of the boundary in a level set framework.

2D Methods that Segment Both Boundaries

Ukwatta et al. [68, 69] proposed a combined 2D segmentation method to segment the outer wall and the lumen boundaries of the carotid arteries from 3D US images using a modified sparse field level set method [70]. Their method is the first semiautomated technique that was proposed to segment both the lumen and outer wall boundaries from 3D US images for the purpose of computing volumetric measurements of plaque burden. Therefore, the rest of this section is devoted to describing their segmentation method.

The segmentation pipeline of their method is shown in Fig. 8. Their method initializes the boundary close to the actual boundary by choosing four or more anchor points $\mathbf{x}_A^i$, where $i = 1 \ldots N$ on each boundary on each transverse slice to be segmented. After initialization, two level set segmentations with different energies were used to segment the outer wall and lumen boundaries. The outer wall was segmented initially, where the segmented outer wall contour was used to constrain the lumen segmentation from leaking in regions with low image contrast.

Here, the level set equations [61] are reviewed before describing the particulars of the carotid segmentation methods [68]. Let the image be defined as $I : \Omega \to \mathbb{R}$, and $\mathbf{x} \in \Omega$ is a spatial variable representing a single point within the image domain $\Omega \subset \mathbb{R}^2$. The level set is an implicit representation of a boundary defined by the function $\phi(\mathbf{x}) : \Omega \to \mathbb{R}$. The level set evolution equation is then given by:

$$\frac{\partial \phi(\mathbf{x})}{\partial t} + \nabla E \ |\nabla \phi(\mathbf{x})| = 0, \tag{1}$$
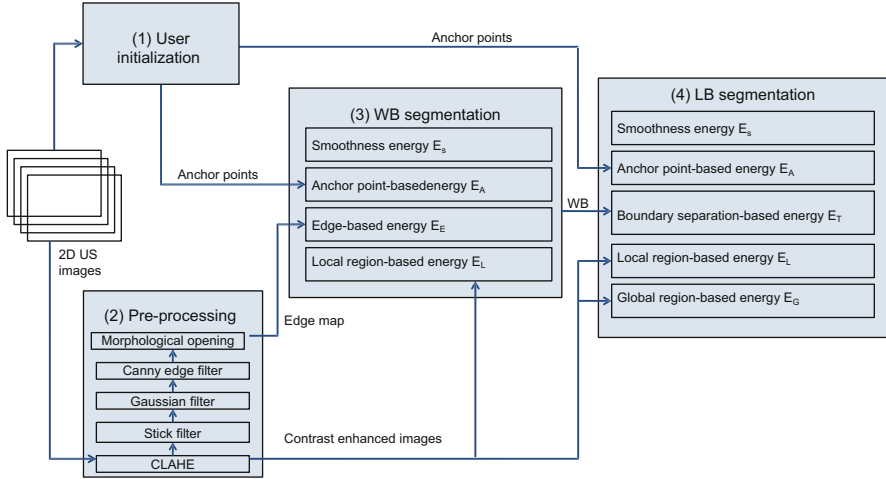
Fig. 8 Block diagram of the segmentation pipeline [68]

where E is the objective function that defines the goodness of the segmentation and ∇E denotes the speed of the evolution. A typical Chan–Vese (CV) energy term [71] that models an object and background by their mean image intensities is given by:

$$
E = \underbrace{\int_{\Omega} |\nabla H(\phi(\mathbf{x}))| \; d\mathbf{x}}_{\text{Smoothness term } (E_S)} + \underbrace{\lambda \int_{\Omega} H(\phi(\mathbf{x})) \; (I(\mathbf{x}) - u)^2 + (1 - H(\phi(\mathbf{x}))) \; (I(\mathbf{x}) - v)^2 d\mathbf{x}}_{\text{Global CV term } (E_G)},
$$
(2)

where $H(\mathbf{x})$ is the regularized Heaviside function that represent the interior region of the boundary, the exterior region of the boundary is given by $(1 - H(\phi(\mathbf{x}))$ and u and v are the mean image intensities of the inside and outside, respectively.

The speed equation is then obtained from the first variation of the energy function E as follows [71]:

$$
\frac{\partial \phi(\mathbf{x})}{\partial t} = \delta(\phi(\mathbf{x})) \left\{ \mathrm{div} \left(\frac{\nabla \phi(\mathbf{x})}{|\nabla \phi(\mathbf{x})|} \right) + \lambda \left((I(\mathbf{x}) - u)^2 - (I(\mathbf{x}) - v)^2 \right) \right\}
$$
(3)

where $\delta(\phi(\mathbf{x}))$ is the Dirac function, which is the derivative of the Heaviside function. $\lambda > 0$ is a parameter that determines the relative importance of the data or the smoothness term.

Outer wall segmentation: Since the outer wall has neither a strong edge nor consistent regional features for segmentation, a combination of various criteria was used for the outer wall segmentation. For example, the outer wall has high image contrast at orthogonal locations to the US beam direction, whereas the parallel locations of the boundary would generally have low image contrast and the image intensities of the interior and exterior of the outer wall are heterogeneous. Therefore,

an objective function combining local region-based, edge-based energies, and user interaction are used for the outer wall segmentation. Edge-based forces drive the segmentation successfully in the presence of strong image gradients, whereas the local region-based forces localize the boundary in the local neighborhood.

Let $\widehat{\mathbf{x}} \in \Omega$ be another independent spatial variable. The objective function E_{WB} for the outer wall segmentation is defined as:

$$
\begin{aligned}
E_{\text{WB}} = {} & \underbrace{\alpha_1 \int_\Omega \delta\left(\phi\left(\mathbf{x}\right)\right) \int_\Omega B_{\text{L}}\left(\mathbf{x},\widehat{\mathbf{x}}\right) \left[H\left(\phi\left(\widehat{\mathbf{x}}\right)\right) \left(I\left(\widehat{\mathbf{x}}\right) - u_{\text{x}}\right)^2 + \left(1 - H\left(\phi\left(\widehat{\mathbf{x}}\right)\right)\right) \left(I\left(\widehat{\mathbf{x}}\right) - v_{\text{x}}\right)^2 \right] d\widehat{\mathbf{x}}\, d\mathbf{x}}_{\text{Local region-based energy } (E_{\text{L}})} \\
& + \underbrace{\int_\Omega \left|\nabla H\left(\phi\left(\mathbf{x}\right)\right)\right|\, d\mathbf{x}}_{\text{Smoothness term } (E_{\text{S}})} + \underbrace{\alpha_2 \sum_{i=1}^{N_{\text{P}}} \int_\Omega \delta\left(\phi\left(\mathbf{x}\right)\right) B_{\text{A}}^i\left(\mathbf{x}\right) \left(\phi\left(\mathbf{x}\right) - \phi\left(\mathbf{x}_{\text{A}}^i\right)\right)^2 d\mathbf{x}}_{\text{Anchor-point based energy } (E_{\text{A}})} \\
& - \underbrace{\alpha_3 \int_\Omega \delta\left(\phi\left(\mathbf{x}\right)\right) \frac{B_{\text{E}}\left(\mathbf{x}\right)}{\varepsilon + \left|\phi\left(\mathbf{x}\right) - \phi\left(X_{\text{E}}\left(\mathbf{x}\right)\right)\right|}\, d\mathbf{x},}_{\text{Edge-based energy } (E_{\text{E}})}
\end{aligned}
\tag{4}
$$

where the parameters $\alpha_{1-3} > 0$. The local region-based energy (E_{L}), which was originally proposed by Lankton and Tannenbaum [72] assumes that at each point on the object boundary, the local interior and exterior neighborhoods can be modeled by their mean intensities $u_{\mathbf{x}}$ and $v_{\mathbf{x}}$ within the circular region with radius r_{L} defined by $B_{\text{L}}\left(\mathbf{x},\widehat{\mathbf{x}}\right) = \begin{cases} 1 \text{ if } \|\mathbf{x} - \widehat{\mathbf{x}}\| < r_{\text{L}} \\ 0 \text{ otherwise} \end{cases}$.

The anchor-point energy encourages the contour to pass through each anchor point $\mathbf{x}_A^i$, placed by the observer, when segmentation is within a distance r_{A} to an anchor point, which is defined by $B_{\text{A}}^i\left(\mathbf{x}\right) = \begin{cases} 1 \text{ if } \|\mathbf{x} - \mathbf{x}_{\text{A}}^i\| < r_{\text{A}} \\ 0 \text{ otherwise} \end{cases}$. When the contour points are away from the anchor point by more than r_{A} there is no influence by this energy term.

Similar to edge-based geometric active contours proposed by Caselles et al. [73] and Malladi et al. [74], an edge-based energy is used, which utilizes the edge map $M(\mathbf{x})$ generated during the pre-processing stage, where $B_{\text{E}}\left(\mathbf{x}\right) = \begin{cases} 1 \text{ if } 0 < \|\mathbf{x} - X_{\text{E}}\left(\mathbf{x}\right)\| < r_{\text{E}} \\ 0 \text{ otherwise} \end{cases}$, is the circular-shaped region around the point $\mathbf{x}$ with a radius r_{E} and ε is a small positive constant. $X_{\text{E}}\left(\mathbf{x}\right) = \underset{\widehat{\mathbf{x}} \in s}{\arg\min} \|\mathbf{x} - \widehat{\mathbf{x}}\|$, $s = \{\widehat{\mathbf{x}}; M\left(\widehat{\mathbf{x}}\right) = 1\}$ is the nearest edge point from point $\mathbf{x}$. The edge force is nonzero when the contour is within a distance r_{E} to the closest edge point.

By taking the first variation of the E_{WB} (4) with respect to ϕ we obtain the evolving equation

$$\frac{\partial \phi(\mathbf{x})}{\partial t} = \delta\left(\phi\left(\mathbf{x}\right)\right) \left\{ \alpha_{\mathrm{L}} \int_{\Omega} B_{\mathrm{L}}\left(\mathbf{x}, \widehat{\mathbf{x}}\right) \delta\left(\phi\left(\widehat{\mathbf{x}}\right)\right) \left[\left(I\left(\widehat{\mathbf{x}}\right) - u_{\mathrm{x}}\right)^2 - \left(I\left(\widehat{\mathbf{x}}\right) - v_{\mathrm{x}}\right)^2\right] d\widehat{\mathbf{x}} \right.$$

$$+ \alpha_{\mathrm{S}} \; \mathrm{div}\left(\frac{\nabla \phi(\mathbf{x})}{|\nabla \phi(\mathbf{x})|}\right) + \alpha_{\mathrm{A}} \sum_{i=1}^{N_{\mathrm{P}}} B_{\mathrm{A}}^{i}\left(\mathbf{x}\right) \left(\phi\left(\mathbf{x}\right) - \phi\left(\mathbf{x}_{\mathrm{A}}^{i}\right)\right)$$

$$\left. + \alpha_{\mathrm{E}} \; B_{\mathrm{E}}\left(\mathbf{x}\right) \frac{\mathrm{sgn}(\phi(\mathbf{x}) - \phi(X_{\mathrm{E}}(\mathbf{x})))}{(\varepsilon + |\phi(\mathbf{x}) - \phi(X_{\mathrm{E}}(\mathbf{x}))|)^2} \right\},$$

(5)

for the outer wall segmentation, where $\mathrm{sgn}(\phi(\mathbf{x}))$ is the signum function.

Lumen segmentation: Similar to the outer wall segmentation, a combination of numerous criteria is used for the lumen segmentation to utilize its characteristics. For example, the lumen boundary has an irregular shape due to the presence of plaque and does not typically contain sharp, high contrast edges. In addition, the intensity distribution of the interior of the lumen is fairly homogeneous. Therefore, a combination of local region-based and global region-based forces is employed to segment the lumen. To stop the leaking at the low-contrast boundaries, the previously segmented outer wall boundary is used to constrain the current lumen segmentation. The objective function E_{LB} used to segment the LIB is given as

$$E_{\mathrm{LB}} = \underbrace{\lambda_1 \int_{\Omega} \delta\left(\phi\left(\mathbf{x}\right)\right) \int_{\Omega} B_{\mathrm{L}}\left(\mathbf{x}, \widehat{\mathbf{x}}\right) \left[H\left(\phi\left(\widehat{\mathbf{x}}\right)\right) \left(I\left(\widehat{\mathbf{x}}\right) - u_{\mathrm{x}}\right)^2 + \left(1 - H\left(\phi\left(\widehat{\mathbf{x}}\right)\right)\right) \left(I\left(\widehat{\mathbf{x}}\right) - v_{\mathrm{x}}\right)^2\right] d\widehat{\mathbf{x}} \, d\mathbf{x}}_{\text{Local CV term } (E_{\mathrm{L}})}$$

$$+ \underbrace{\int_{\Omega} |\nabla H\left(\phi\left(\mathbf{x}\right)\right)| \, d\mathbf{x}}_{\text{Smoothness term } (E_{\mathrm{S}})} + \underbrace{\lambda_2 \sum_{i=1}^{N_{\mathrm{P}}} \int_{\Omega} \delta\left(\phi\left(\mathbf{x}\right)\right) B_{\mathrm{A}}^{i}\left(\mathbf{x}\right) \left(\phi\left(\mathbf{x}\right) - \phi\left(\mathbf{x}_{\mathrm{A}}^{i}\right)\right)^2 \, d\mathbf{x}}_{\text{Anchor-point based energy } (E_{\mathrm{A}})}$$

$$+ \underbrace{\lambda_3 \int_{\Omega} H\left(\phi\left(\mathbf{x}\right)\right) \left(I\left(\mathbf{x}\right) - u\right)^2 + \left(1 - H\left(\phi\left(\mathbf{x}\right)\right)\right) \left(I\left(\mathbf{x}\right) - v\right)^2 \, d\mathbf{x}}_{\text{Global CV term } (E_{\mathrm{G}})} + \underbrace{\lambda_4 \int_{\Omega} \delta\left(\phi\left(\mathbf{x}\right)\right) B_{\mathrm{T}}\left(\mathbf{x}\right) \, d\mathbf{x}}_{\text{Boundary separation term } (E_{\mathrm{T}})},$$

(6)

where the parameters $\lambda_{1-4} > 0$. Similarly to the outer wall segmentation, the smoothness energy, local region-based energy, and the anchor point-based energies are used in the lumen segmentation. The global CV term (E_{G}) (2) is used here to utilize the homogeneous intensity of the lumen region.

Because the outer wall and lumen are separated by the carotid media layer, they are enforced to have a greater or equal separation distance (d_{T}) of 0.5 mm from each other [75]. The boundary separation-based energy E_{T} is given by where $B_{\mathrm{T}}\left(\mathbf{x}\right) = \begin{cases} 1 \text{ if } \min\limits_{\widehat{\mathbf{x}}} D\left(\mathbf{x}, \widehat{\mathbf{x}}\right) < d_{\mathrm{T}} \\ 0 \text{ otherwise} \end{cases}$, and $D\left(\mathbf{x}, \widehat{\mathbf{x}}\right) = ||\mathbf{x} - \widehat{\mathbf{x}}||$, $\nabla H\left(\phi_{\mathrm{MAB}}\left(\widehat{\mathbf{x}}\right)\right) = 1$. Therefore, the term E_{T} avoids the intersection of the outer wall and lumen.

By taking the first variation of the E_{LB} (6) with respect to ϕ we obtain the evolving equation for the lumen segmentation.

$$
\begin{aligned}
\frac{\partial \phi(\mathbf{x})}{\partial t} = \delta\left(\phi\left(\mathbf{x}\right)\right) &\left\{ \lambda_{\mathrm{L}}\ \alpha_{\mathrm{L}} \int_{\Omega} B_{\mathrm{L}}\left(\mathbf{x}, \widehat{\mathbf{x}}\right)\ \delta\left(\phi\left(\widehat{\mathbf{x}}\right)\right) \left[\left(I\left(\widehat{\mathbf{x}}\right) - u_{\mathrm{x}}\right)^2 - \left(I\left(\widehat{\mathbf{x}}\right) - v_{\mathrm{x}}\right)^2\right]\ d\widehat{\mathbf{x}} \right. \\
&+ \lambda_{\mathrm{S}}\ \mathrm{div}\left(\frac{\nabla\phi(\mathbf{x})}{|\nabla\phi(\mathbf{x})|}\right) + \lambda_{\mathrm{G}}\ \left(\left(I\left(\mathbf{x}\right) - u\right)^2 - \left(I\left(\mathbf{x}\right) - v\right)^2\right) \\
&+ \lambda_{\mathrm{A}} \sum_{i=1}^{N_{\mathrm{P}}} \mathrm{sgn}\ \left(\phi\left(x_{\mathrm{A}}^i\right)\right)\ B_{\mathrm{A}}^i\left(\mathbf{x}\right)\ \left(\phi\left(\mathbf{x}\right) - \phi\left(x_{\mathrm{A}}^i\right)\right) + \left. \lambda_{\mathrm{T}} B_{\mathrm{T}}\left(\mathbf{x}\right) \right\}
\end{aligned}
\tag{7}
$$

This algorithm was evaluated with 231 2D carotid US slices extracted from 21 3D US images. The algorithm segmentations were compared to the manual segmentations on a slice-by-slice basis as shown in Fig. 9. The algorithm yielded high Dice coefficients (DC) (>92 %), submillimeter errors for the mean absolute distance (MAD) and maximum absolute distance (MAXD) (see Table 2 for more details), and a comparable coefficient of variation (COV) (5.1 vs. 3.9 %) to manual segmentations for the outer wall and lumen boundaries, respectively. The minimum detectable difference of the algorithm in computing the VWV was comparable (64.2 vs. 50.3 mm^3) to manual segmentations. Using this method, the total segmentation time is reduced by 5.5 min in comparison to the manual segmentation (2.8 vs. 8.3 min). Out of the mean time of 2.8 min, 1.2 min is required for the computation and the 1.6 min for the initialization by the observer. The authors also reported on a coupled level set segmentation method to segment both boundaries simultaneously [76] for a more robust segmentation.

This study was validated using the global metrics such as DC, MAD, MAXD, and volume errors, which are sufficient for evaluating for 3D US VWV. However, when the algorithm is required to be used for generating VWT maps, the algorithm must be validated using localized techniques. Chiu et al. [77] conducted a study, which evaluated the algorithm using standardized carotid maps for the population.

3D Methods that Segment Both Boundaries

One of the main drawbacks of the previous method proposed by Ukwatta et al. [68] is that it still requires a considerable amount of user interaction on each slice to obtain a good segmentation. If the user decides to segment each and every slice of the 3D image, a 2D segmentation technique would take a long time.

Ukwatta et al. [78] described a direct 3D segmentation algorithm to delineate the outer wall and lumen of the carotid CCA from 3D US images. The segmentation pipeline of the algorithm is shown in Fig. 10. The algorithm uses a similar initialization approach used previously for the 2D segmentation, but requires initialization only on a smaller subset of transverse slices than used in the 2D method [68]. A direct 3D segmentation method has the potential for reducing user interaction over a 2D method, while increasing the robustness of the segmentation by integrating out-of-plane image information. The 3D method used very similar energy functions to the 2D approach and an extension to the 2D sparse field level set method.
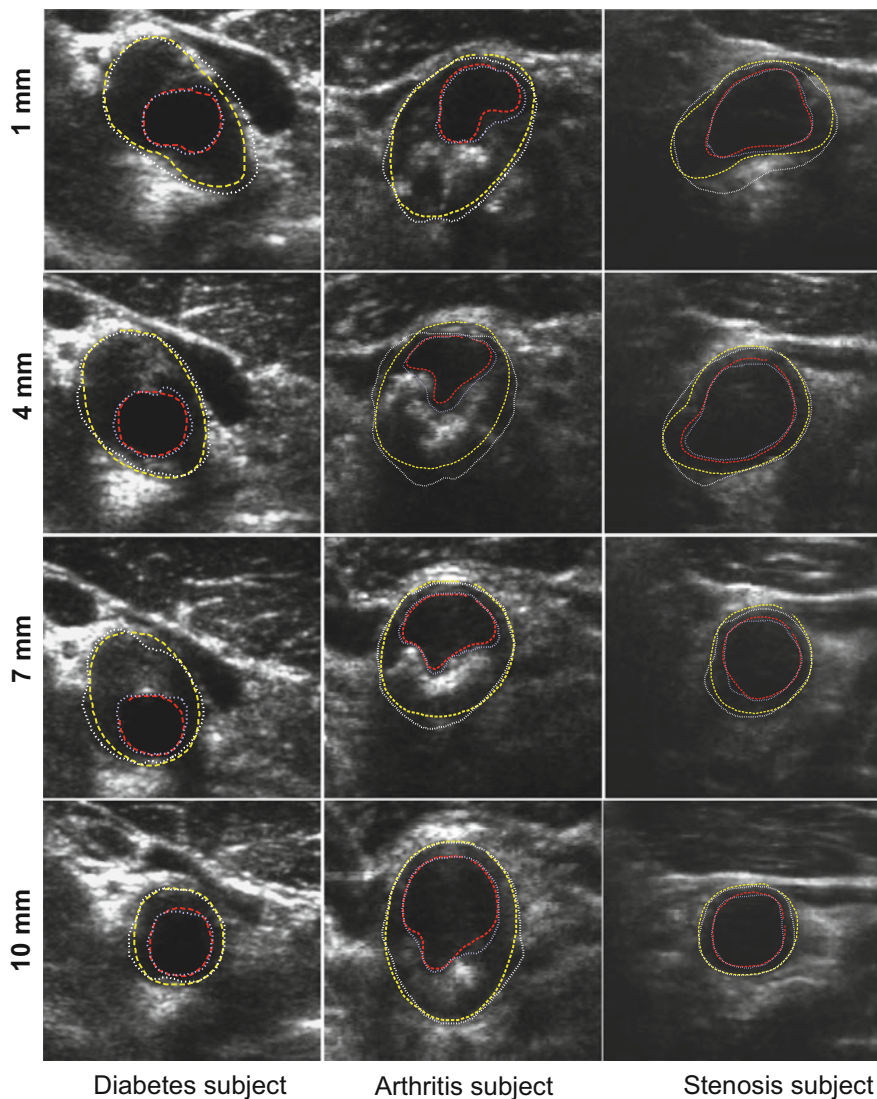
Fig. 9 Sample results of the lumen and wall segmentations for three patients using four anchor points chosen from the initial eight anchor points (*dashed outer contour* = manual wall, *dashed inner contour* = manual lumen, *dotted outer contour* = algorithm-generated wall, *dotted inner contour* = algorithm-generated lumen). 1, 5, 7, and 10 mm indicate the distance to the BF from each slice [68]

The algorithm-generated accuracy and intra-observer variability results are comparable to the previous methods, but with fewer user interactions. For example, for the ISD of 3 mm, the algorithm yielded an average Dice coefficient of $94.4 \pm 2.2\%$

Table 2 Overall CCA segmentation results for the segmentation algorithm [68] for 231 2D US images extracted from 21 3D US images

Metric	Outer wall segmentation	Lumen segmentation
Volume error (%)	2.4 ± 1.9	5.8 ± 2.7
Dice coefficient (%)	95.6 ± 1.5	92.8 ± 3.1
MAD (mm)	0.2 ± 0.1	0.3 ± 0.1
MAXD (mm)	0.6 ± 0.3	0.6 ± 0.4

MAD is the mean absolute distance error and MAXD is the maximum absolute distance error
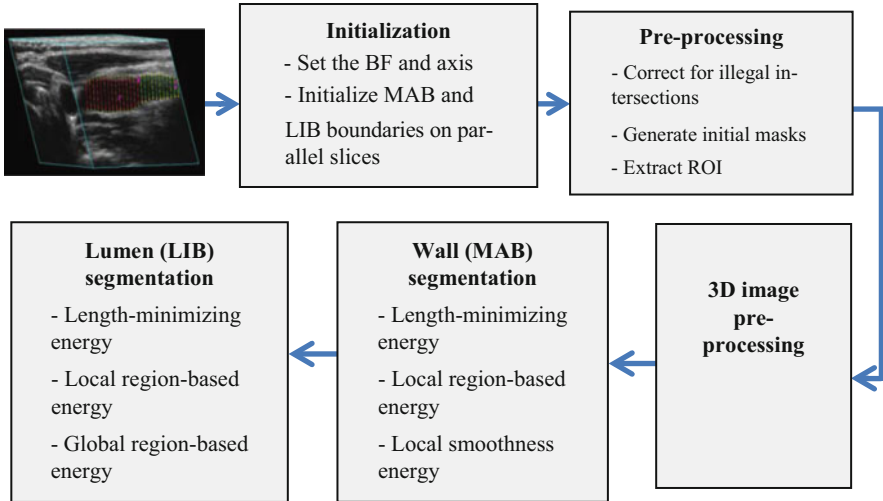
Fig. 10 Steps involved in 3D segmentation algorithm pipeline [78], which is similar to the previous 2D segmentation method proposed by Ukwatta et al. [68]

and 90.6 ± 5.0 % for the MAB and LIB and the COV of 6.8 % for computing the VWV of the CCA, while requiring only 1.72 min (vs. 8.3 min for manual segmentation) for a 3D US image.

4.3 Segmentation Algorithms of Carotid Plaque from 3D US Images

The segmentation of plaque boundaries is an even more challenging task than the lumen and outer wall segmentations due to the fuzzy boundaries of the plaque. Unlike the measurement of 3D US TPV, which requires observers to distinguish plaque–lumen and plaque–outer vessel wall boundaries, the measurement of 3D US VWV requires an observer to manually outline the lumen–intima/plaque and media–adventitia boundaries—similarly to the measurement of IMT. These boundaries are more straightforward to interpret than plaque–lumen and plaque–wall

boundaries in 3D US images. Furthermore, the measurement of TPV from 3D US images requires trained observers who are experts in 3D US image interpretation and in distinguishing vessel wall from plaque in 3D US images.

4.3.1 Manual Segmentation

Currently, TPV measurements are obtained by manual segmentation of the plaque [38, 39, 79–81]. Initially, the observer defines the medial axis of the artery in longitudinal view. After familiarizing with the orientation and geometry of the plaque using a multi-planer reformatting software, the observer outlines the plaque boundary on transverse slices with inter-slice distance (ISD) of 1 mm. After outlining the complete plaque, the software generates a 3D surface from these contours. As an illustration, Fig. 3 shows some example manual segmentations of plaque overlaid on a 3D US image, and Fig. 3 shows the reconstructed surface of plaque for the computation of TPV. Limitations of this approach include image interpretation and measurement differences within and between observers, long training times for observers, and long durations to perform manual segmentations.

4.3.2 Semiautomated Segmentation Algorithms

Although, there are several methods that have been developed for 2D segmentation of plaque from either 2D or 3D US images, there exists only one study that describes a semiautomated tool to generate the TPV measurement. Buchanan et al. [82] reported on a semiautomated plaque estimated method to outline the plaque boundaries from 3D US images thereby computing TPVs. The workflow of their method is shown in Fig. 11. For the initialization, the observer defines the beginning and end points of the plaque in the long axis view, followed by two contours in the transverse view to identify the regions of greatest change in shape. The algorithm was evaluated with 22 plaques from 10 3D US images. An example comparison of the manually and algorithm-generated plaque surfaces are shown in Fig. 12. The algorithm plaque volume and the manual plaque volumes were significantly correlated ($r = 0.99$, $p < 0.001$) while maintaining a mean COV of $\approx 12.0\ \%$. One of the main drawbacks of this method is that it currently does not use any image information by the algorithm. After the algorithm segmentation is obtained, the ability to adjust the segmented boundary by the user is highly desirable.

5 Future Direction

As discussed in this chapter, 3D US imaging is increasingly being used as a surrogate for monitoring carotid plaque burden in patients in clinical trials during medical interventions. For example, Shai et al. [59] used 3D US VWV measurements to
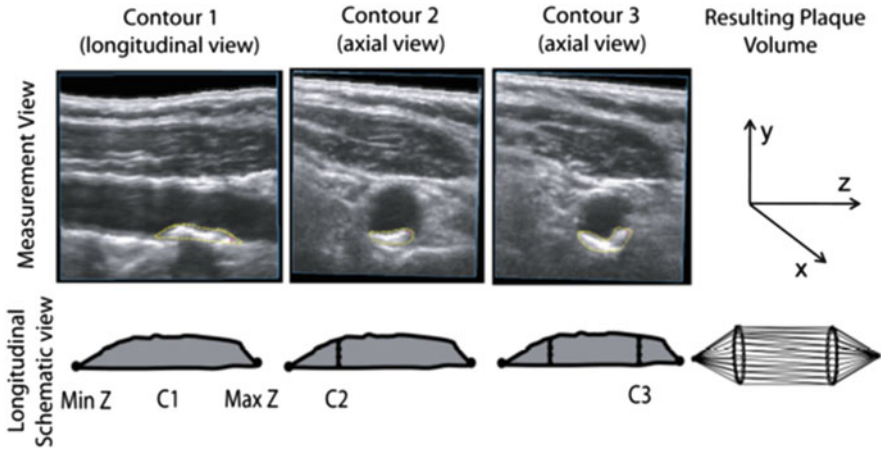
Fig. 11 User input for semiautomated TPV measurement. In the longitudinal view (and with assistance from the axial view, not shown) the user identifies the maximum and minimum z-values representing the end points of the plaque (Min Z and Max Z). The user identifies the mid-point of the plaque (C1) and finally C2 and C3 are identified and generated in the axial view. Uniform plaque geometry between C2 and C3 is assumed and a final volume is generated [82]
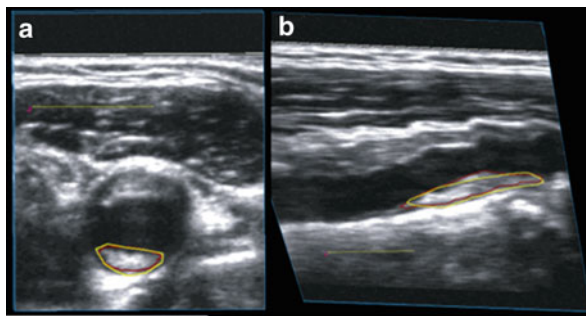
Fig. 12 Representative manual and semiautomated segmentation boundaries. (**a**) Axial view, (**b**) longitudinal view. Manual (*red*) algorithm contours (*yellow*). Yellow bar = 10 mm [82]

assess the effect of dietary interventions for reducing the carotid plaque burden. The study found that diet-mediated weight loss over a 2-year period induced significant regression of 3D US carotid VWV, even though the effect is similar in low-fat, Mediterranean, or low-carbohydrate diets. Krasinski et al. [41] used 3D US VWV measurement to assess the effect of atorvastatin in reducing the carotid plaque burden. The clinical trial had 35 subjects with carotid stenosis >60 %, where subjects were randomly assigned to either the statin drug or placebo in a three month trial. The results showed an increase of 70 ± 140 mm^3 in 3D US VWV in the placebo group and reduction of 30 ± 130 mm^3 in the atorvastatin group. Similarly, Ainsworth et al. [83] conducted a study to evaluate the effect of atorvastatin using 3D US TPV. They also observed significant difference of the two groups, such that

placebo group had a progression of 16.81 ± 74.10 mm^3 and the atorvastatin group had a regression of 90.25 ± 85.12 mm^3. Since the number of patients suffering from carotid atherosclerosis and subsequently the number of novel treatments for carotid atherosclerosis is expected to grow rapidly, it is important to continue to develop and validate noninvasive, cost-effective, and direct measurements of atherosclerosis for more accurately stratifying risk of stroke.

With improved strategies to treat atherosclerosis nonsurgically and cost effectively, sensitive and noninvasive 3D US techniques that allow quantification of plaque burden are becoming more important in the management of patients who are suffering from carotid atherosclerosis. However, for translating 3D US imaging to clinical research and subsequently to clinical care, development and validation of semiautomated algorithms for segmentation that reduce the overall observer time and observer variability are vital. Especially for large multicenter clinical trials, the inter-observer variabilities of the algorithm in computing 3D US phenotypes should be small enough to detect a significant change.

Although, algorithms have been developed that are either fully automated or that require minimal user interaction for the segmentation of carotid computed tomography angiography (CTA) [84], carotid black-blood MR images [85, 86], segmentation of carotid arteries from 3D US images is a more challenging problem, that still requires a considerable amount of user interaction. State of the art global segmentation algorithms, such as continuous max-flow-based algorithms that have recently been introduced to medical imaging, have been successfully applied for segmentation of the lumen and the outer wall from black-blood MR images. Such global optimization-based techniques that can be easily parallelized may address the limitations of the existing segmentation methods.

Even though 3D US volumetric measurements described in this chapter are directly correlated with the risk of stroke, they are not suggestive of plaque vulnerability for rupture. However, plaque composition measurements and inflammation are strong indicators of plaque vulnerability [87]. A plaque could remodel from a more vulnerable plaque to a stable plaque, without essential change in size during medical interventions [88]. Therefore, volumetric measurements alone are not adequate to stratify plaque vulnerability to rupture and cause ischemic stroke. Although the ability of MR imaging for plaque composition [89, 90] and inflammation analysis [91] have been validated, the ability of 3D US to identify individual plaque components such as fibrous, necrotic core, loose connective tissue, and calcifications is still unclear. Therefore, future work lies in investigating the capability of 3D US imaging to identify plaque components, which could be validated using digital histopathology images of endarterectomy specimens and in vivo MR imaging.

References

1. Fenster, A., Downey, D.B.: 3-Dimensional ultrasound imaging: a review. IEEE Eng. Med. Biol. **15**, 41–51 (1996)
2. Greenleaf, J.F., Belohlavek, M., Gerber, T.C., Foley, D.A., Seward, J.B.: Multidimensional visualization in echocardiography: an introduction [see comments]. Mayo Clin. Proc. **68**, 213–220 (1993)
3. King, D.L., Gopal, A.S., Sapin, P.M., Schroder, K.M., Demaria, A.N.: Three-dimensional echocardiography. Am. J. Card. Imaging **7**, 209–220 (1993)
4. Nelson, T.R., Pretorius, D.H.: Three-dimensional ultrasound of fetal surface features. Ultrasound Obstet. Gynecol. **2**, 166–174 (1992)
5. Peralta, C.F., Cavoretto, P., Csapo, B., Falcon, O., Nicolaides, K.H.: Lung and heart volumes by three-dimensional ultrasound in normal fetuses at 12–32 weeks' gestation. Ultrasound Obstet. Gynecol. **27**, 128–133 (2006)
6. Rankin, R.N., Fenster, A., Downey, D.B., Munk, P.L., Levin, M.F., Vellet, A.D.: Three-dimensional sonographic reconstruction: techniques and diagnostic applications. AJR Am. J. Roentgenol. **161**, 695–702 (1993)
7. Sklansky, M.: New dimensions and directions in fetal cardiology. Curr. Opin. Pediatr. **15**, 463–471 (2003)
8. Baba, K., Jurkovic, D.: Three-Dimensional Ultrasound in Obstetrics and Gynecology. The Parthenon Publishing Group, New York (1997)
9. Downey, D.B., Fenster, A.: Three-dimensional ultrasound: a maturing technology. Ultrasound Q. **14**, 25–40 (1998)
10. Fenster, A., Downey, D.: Three-dimensional ultrasound imaging. In: Beutel, J., Kundel, H., Van Metter, R. (eds.) Physics and Psychophysics. Handbook of Medical Imaging, vol. 1. SPIE Press, Bellingham (2000)
11. Fenster, A., Downey, D.: Basic principles and applications of 3-D ultrasound imaging. In: Stergiopoulos, S. (ed.) An Advanced Signal Processing Handbook. CRC, Boca Raton (2001)
12. Fenster, A., Downey, D.B.: Three-dimensional ultrasound imaging. Annu. Rev. Biomed. Eng. **02**, 457–475 (2000)
13. Fenster, A., Downey, D.B., Cardinal, H.N.: Three-dimensional ultrasound imaging. Phys. Med. Biol. **46**, R67–R99 (2001)
14. Nelson, T.R., Downey, D.B., Pretorius, D.H., Fenster, A.: Three-Dimensional Ultrasound. Lippincott–Raven, Philadelphia (1999)
15. Nelson, T.R., Pretorius, D.H.: Three-dimensional ultrasound imaging. Ultrasound Med. Biol. **24**, 1243–1270 (1998)
16. Smith, W.L., Fenster, A.: Optimum scan spacing for three-dimensional ultrasound by speckle statistics. Ultrasound Med. Biol. **26**, 551–562 (2000)
17. Pagoulatos, N., Haynor, D.R., Kim, Y.: A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. Ultrasound Med. Biol. **27**, 1219–1229 (2001)
18. Boctor, E.M., Choti, M.A., Burdette, E.C., Webster III, R.J.: Three-dimensional ultrasound-guided robotic needle placement: an experimental evaluation. Int. J. Med. Robot. **4**, 180–191 (2008)
19. Treece, G., Prager, R., Gee, A., Berman, L.: 3D ultrasound measurement of large organ volume. Med. Image Anal. **5**, 41–54 (2001)
20. Hummel, J., Figl, M., Bax, M., Bergmann, H., Birkfellner, W.: 2D/3D registration of endoscopic ultrasound to CT volume data. Phys. Med. Biol. **53**, 4303–4316 (2008)
21. Dandekar, S., Li, Y., Molloy, J., Hossack, J.: A phantom with reduced complexity for spatial 3-D ultrasound calibration. Ultrasound Med. Biol. **31**, 1083–1093 (2005)
22. Gee, A.H., Houghton, N.E., Treece, G.M., Prager, R.W.: A mechanical instrument for 3D ultrasound probe calibration. Ultrasound Med. Biol. **31**, 505–518 (2005)
23. Gooding, M.J., Kennedy, S.H., Noble, J.A.: Temporal calibration of freehand three-dimensional ultrasound using image alignment. Ultrasound Med. Biol. **31**, 919–927 (2005)

24. Hsu, P.W., Prager, R.W., Gee, A.H., Treece, G.M.: Real-time freehand 3D ultrasound calibration. Ultrasound Med. Biol. **34**, 239–251 (2008)
25. Leotta, D.F.: An efficient calibration method for freehand 3-D ultrasound imaging systems. Ultrasound Med. Biol. **30**, 999–1008 (2004)
26. Lindseth, F., Tangen, G.A., Lango, T., Bang, J.: Probe calibration for freehand 3-D ultrasound. Ultrasound Med. Biol. **29**, 1607–1623 (2003)
27. Mercier, L., Lango, T., Lindseth, F., Collins, D.L.: A review of calibration techniques for freehand 3-D ultrasound systems. Ultrasound Med. Biol. **31**, 449–471 (2005)
28. Poon, T.C., Rohling, R.N.: Comparison of calibration methods for spatial tracking of a 3-D ultrasound probe. Ultrasound Med. Biol. **31**, 1095–1108 (2005)
29. Rousseau, F., Hellier, P., Barillot, C.: Confhusius: a robust and fully automatic calibration method for 3D freehand ultrasound. Med. Image Anal. **9**, 25–38 (2005)
30. Bots, M.L., Grobbee, D.E.: Intima media thickness as a surrogate marker for generalised atherosclerosis. Cardiovasc. Drugs Ther. **16**, 341–351 (2002)
31. Bots, M.L., Evans, G.W., Riley, W.A., Grobbee, D.E.: Carotid intima-media thickness measurements in intervention studies: design options, progression rates, and sample size considerations: a point of view. Stroke **34**, 2985–2994 (2003)
32. Picot, P.A., Rickey, D.W., Mitchell, R., Rankin, R.N., Fenster, A.: Three-dimensional colour Doppler imaging. Ultrasound Med. Biol. **19**, 95–104 (1993)
33. Pretorius, D.H., Nelson, T.R., Jaffe, J.S.: 3-Dimensional sonographic analysis based on color flow Doppler and gray scale image data: a preliminary report. J. Ultrasound Med. **11**, 225–232 (1992)
34. Espeland, M.A., O'leary, D.H., Terry, J.G., Morgan, T., Evans, G., Mudra, H.: Carotid intimal-media thickness as a surrogate for cardiovascular disease events in trials of HMG-CoA reductase inhibitors. Curr. Control. Trials Cardiovasc. Med. **6**, 3 (2005)
35. O'leary, D.H., Polak, J.F., Kronmal, R.A., Manolio, T.A., Burke, G.L., Wolfson Jr., S.K., Cardiovascular Health Study Collaborative Research Group: Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults. N. Engl. J. Med. **340**, 14–22 (1999)
36. Johnsen, S.H., Mathiesen, E.B.: Carotid plaque compared with intima-media thickness as a predictor of coronary and cerebrovascular disease. Curr. Cardiol. Rep. **11**, 21–27 (2009)
37. Ainsworth, C., Tamayo, A., Landry, A., Blake, C., Beletsky, V., Fenster, A.: 3D ultrasound measurement of change in carotid plaque volume: a tool for quickly measuring effects of treatment on atherosclerosis. Stroke **36**, 1904–1909 (2005)
38. Landry, A., Spence, J.D., Fenster, A.: Measurement of carotid plaque volume by 3-dimensional ultrasound. Stroke **35**, 864–869 (2004)
39. Landry, A., Spence, J.D., Fenster, A.: Quantification of carotid plaque volume measurements using 3D ultrasound imaging. Ultrasound Med. Biol. **31**, 751–762 (2005)
40. Egger, M., Spence, J.D., Fenster, A., Parraga, G.: Validation of 3D ultrasound vessel wall volume: an imaging phenotype of carotid atherosclerosis. Ultrasound Med. Biol. **33**, 905–914 (2007)
41. Krasinski, A., Chiu, B., Spence, J.D., Fenster, A., Parraga, G.: Three-dimensional ultrasound quantification of intensive statin treatment of carotid atherosclerosis. Ultrasound Med. Biol. **35**, 1763–1772 (2009)
42. Chiu, B., Beletsky, V., Spence, J.D., Parraga, G., Fenster, A.: Analysis of carotid lumen surface morphology using three-dimensional ultrasound imaging. Phys. Med. Biol. **54**, 1149–1167 (2009)
43. Chiu, B., Egger, M., Spence, D.J., Parraga, G., Fenster, A.: Area-preserving flattening maps of 3D ultrasound carotid arteries images. Med. Image Anal. **12**, 676–688 (2008)
44. Chiu, B., Egger, M., Spence, J.D., Parraga, G., Fenster, A.: Quantification of carotid vessel wall and plaque thickness change using 3D ultrasound images. Med. Phys. **35**, 3691–3710 (2008)
45. Dabrowski, W., Dunmore-Buyze, J., Cardinal, H.N., Fenster, A.: A real vessel phantom for flow imaging: 3-D Doppler ultrasound of steady flow. Ultrasound Med. Biol. **27**, 135–141 (2001)

46. Guo, Z., Fenster, A.: Three-dimensional power Doppler imaging: a phantom study to quantify vessel stenosis. Ultrasound Med. Biol. **22**, 1059–1069 (1996)
47. O'leary, D.H., Bots, M.L.: Imaging of atherosclerosis: carotid intima-media thickness. Eur. Heart J. **31**, 1682–1689 (2010)
48. Inaba, Y., Chen, J.A., Bergmann, S.R.: Carotid plaque, compared with carotid intima-media thickness, more accurately predicts coronary artery disease events: a meta-analysis. Atherosclerosis **220**(1), 128–133 (2011)
49. Spence, J.D.: Ultrasound measurement of carotid plaque as a surrogate outcome for coronary artery disease. Am. J. Cardiol. **89**, 10B–15B (2002). discussion 15B–16B
50. Mallett, C., House, A.A., Spence, J.D., Fenster, A., Parraga, G.: Longitudinal ultrasound evaluation of carotid atherosclerosis in one, two and three dimensions. Ultrasound Med. Biol. **35**, 367 (2009)
51. Makris, G.C., Lavida, A., Griffin, M., Geroulakos, G., Nicolaides, A.N.: Three-dimensional ultrasound imaging for the evaluation of carotid atherosclerosis. Atherosclerosis **219**, 377–383 (2011)
52. Parraga, G., Fenster, A., Krasinski, A., Chiu, B., Egger, M., Spence, J.D.: 3D Carotid ultrasound imaging. In: Suri, J.S. (ed.) Atherosclerosis Disease Management. Springer, New York (2011)
53. Ukwatta, E., Buchanan, D., Parraga, G., Fenster, A.: Three-dimensional ultrasound imaging of carotid atherosclerosis. In: IEEE International Conference on Intelligent Computation and Bio-Medical Instrumentation (ICBMI), pp. 81–84, Wuhan (2011)
54. Johnsen, S.H., Mathiesen, E.B., Joakimsen, O., Stensland, E., Wilsgaard, T., Løchen, M.-L., Njølstad, I., Arnesen, E.: Carotid atherosclerosis is a stronger predictor of myocardial infarction in women than in men. Stroke **38**, 2873–2880 (2007)
55. Mathiesen, E.B., Johnsen, S.H., Wilsgaard, T., Bønaa, K.H., Løchen, M.-L., Njølstad, I.: Carotid plaque area and intima-media thickness in prediction of first-ever ischemic stroke. Stroke **42**, 972–978 (2011)
56. Chiu, B., Egger, M., Spence, J., Parraga, G., Fenster, A.: Development of 3D ultrasound techniques for carotid artery disease assessment and monitoring. Int. J. Comput. Assist. Radiol. Surg. **3**, 1–10 (2008)
57. Egger, M., Chiu, B., Spence, J., Fenster, A., Parraga, G.: Mapping spatial and temporal changes in carotid atherosclerosis from three-dimensional ultrasound images. Ultrasound Med. Biol. **34**, 64–72 (2008)
58. Krasinski, A., Chiu, B., Fenster, A., Parraga, G.: Magnetic resonance imaging and three-dimensional ultrasound of carotid atherosclerosis: mapping regional differences. J. Magn. Reson. Imaging **29**, 901–908 (2009)
59. Shai, I., Spence, J.D., Schwarzfuchs, D., Henkin, Y., Parraga, G., Rudich, A., Fenster, A., Mallett, C., Liel-Cohen, N., Tirosh, A., Bolotin, A., Thiery, J., Fiedler, G.M., Bluher, M., Stumvoll, M., Stampfer, M.J., Direct group: Dietary intervention to reverse carotid atherosclerosis. Circulation **121**, 1200–1208 (2010)
60. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **1**, 321–331 (1988)
61. Osher, S., Sethian, J.A.: Fronts propagating with curvature dependent speed: algorithms based on Hamilton–Jacobi formulations. J. Comput. Phys. **79**(1), 12–49 (1988)
62. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. Comput. Vis. Image. Underst. **61**, 38–59 (1995)
63. Abolmaesumi, P., Sirouspour, M.R., Salcudean, S.E.: Real-time extraction of carotid artery contours from ultrasound images. In: Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS), Houston, 2000
64. Gill, J.D., Ladak, H.M., Steinman, D.A., Fenster, A.: Accuracy and variability assessment of a semiautomatic technique for segmentation of the carotid arteries from three-dimensional ultrasound images. Med. Phys. **27**, 1333–1342 (2000)
65. Mao, F., Gill, J., Downey, D., Fenster, A.: Segmentation of carotid artery in ultrasound images: method development and evaluation technique. Med. Phys. **27**, 1961–1970 (2000)

66. Solovey, I.: Segmentation of 3D carotid ultrasound images using weak geometric priors. M.S. Thesis, University of Waterloo, Canada (2010)
67. Zahalka, A., Fenster, A.: An automated segmentation method for three-dimensional carotid ultrasound images. Phys. Med. Biol. **46**, 1321–1342 (2001)
68. Ukwatta, E., Awad, J., Ward, A.D., Buchanan, D., Samarabandu, J., Parraga, G., Fenster, A.: Three-dimensional ultrasound of carotid atherosclerosis: semiautomated segmentation using a level set-based method. Med. Phys. **38**, 2479 (2011)
69. Ukwatta, E., Awad, J., Ward, A.D., Samarabandu, J., Krasinski, A., Parraga, G., Fenster, A.: Segmentation of the lumen and media-adventitia boundaries of the common carotid artery from 3D ultrasound images. In SPIE Medical Imaging, International Society for Optics and Photonics, pp. 79630G–79630G (2011)
70. Whitaker, R.T.: A level-set approach to 3D reconstruction from range data. Int. J. Comput. Vis. **29**, 203–231 (1998)
71. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
72. Lankton, S., Tannenbaum, A.: Localizing region-based active contours. IEEE Trans. Image Process. **17**, 2029–2039 (2008)
73. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vis. **22**, 61–79 (1997)
74. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape modeling with front propagation: a level set approach. IEEE Trans. Pattern Anal. Mach. Intell. **17**(2), 158–175 (1995)
75. Touboul, P.J., Hennerici, M.G., Meairs, S., Adams, H., Amarenco, P., Bornstein, N., Csiba, L., Desvarieux, M., Ebrahim, S., Fatar, M., Hernandez Hernandez, R., Jaff, M., Kownator, S., Prati, P., Rundek, T., Sitzer, M., Schminke, U., Tardif, J.C., Taylor, A., Vicaut, E., Woo, K.S., Zannad, F., Zureik, M.: Mannheim carotid intima-media thickness consensus. Cerebrovasc. Dis. **23**, 75–80 (2007)
76. Ukwatta, E., Awad, J., Ward, A.D., Buchanan, D., Parraga, G.: Fenster, A.: Coupled level set approach to segment carotid arteries from 3D ultrasound images. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL (2011)
77. Chiu, B., Ukwatta, E., Shavakh, S., Fenster, A.: Quantification and visualization of carotid segmentation accuracy and precision using a 2D standardized carotid map. Phys. Med. Biol. **58**(11), 3671–3703 (2013)
78. Ukwatta, E., Awad, J., Buchanan, D., Parraga, G., Fenster, A.: Three-dimensional semi-automated segmentation of carotid atherosclerosis from three-dimensional ultrasound images. In: Proceedings of the SPIE, vol. 8315 (2012) (article id: 83150O-1)
79. Fenster, A., Blake, C., Gyacskov, I., Landry, A., Spence, J.D.: 3D ultrasound analysis of carotid plaque volume and surface morphology. Ultrasonics **44**, e153–e157 (2006)
80. Landry, A., Ainsworth, C., Blake, C., Spence, J.D., Fenster, A.: Manual planimetric measurement of carotid plaque volume using three-dimensional ultrasound imaging. Med. Phys. **34**, 1496–1505 (2007)
81. Spence, J.D., Blake, C., Landry, A., Fenster, A.: Measurement of carotid plaque and effect of vitamin therapy for total homocysteine. Clin. Chem. Lab. Med. **41**, 1498–1504 (2003)
82. Buchanan, D., Gyacskov, I., Ukwatta, E., Lindenmaiera, T., Fenster, A., Parraga, G.: Semi-automated Segmentation of Carotid Artery Total Plaque Volume from Three Dimensional Ultrasound Carotid Imaging. SPIE Medical Imaging, San Diego (2011)
83. Ainsworth, C.D., Blake, C.C., Tamayo, A., Beletsky, V., Fenster, A., Spence, J.D.: 3D ultrasound measurement of change in carotid plaque volume; a tool for rapid evaluation of new therapies. Stroke **35**, 1904–1909 (2005)
84. Vukadinovic, D., Van Walsum, T., Manniesing, R., Rozie, S., Van der Lugt, A., Niessen, W.J.: Region based level set segmentation of the outer wall of the carotid bifurcation in CTA. In SPIE Medical Imaging, International Society for Optics and Photonics, pp. 79623P–79623P (2011)

85. Ukwatta, E., Yuan, J., Rajchl, M., Fenster, A.: Efficient global optimization based 3D carotid AB-LIB MRI segmentation by simultaneously evolving coupled surfaces. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 377–384, Nice (2012)
86. Ukwatta, E., Yuan, J., Rajchl, M., Qiu, W., Tessier, D., Fenster, A.: 3D carotid multi-region MRI segmentation by globally optimal evolution of coupled surfaces. IEEE Trans. Med. Imaging **32**(4), 770–785 (2013)
87. Yuan, C., Oikawa, M., Miller, Z., Hatsukami, T.: MRI of carotid atherosclerosis. J. Nucl. Cardiol. **15**, 266–275 (2008)
88. Yuan, C., Kerwin, W.S., Yarnykh, V.L., Cai, J., Saam, T., Chu, B., Takaya, N., Ferguson, M.S., Underhill, H., Xu, D., Liu, F., Hatsukami, T.S.: MRI of atherosclerosis in clinical trials. NMR Biomed. **19**, 636–654 (2006)
89. Clarke, S.E., Beletsky, V., Hammond, R.R., Hegele, R.A., Rutt, B.K.: Validation of automatically classified magnetic resonance images for carotid plaque compositional analysis. Stroke **37**, 93–97 (2006)
90. Saam, T., Ferguson, M.S., Yarnykh, V.L., Takaya, N., Xu, D., Polissar, N.L., Hatsukami, T.S., Yuan, C.: Quantitative evaluation of carotid plaque composition by in vivo MRI. Arterioscler. Thromb. Vasc. Biol. **25**, 234–239 (2005)
91. Kerwin, W.S., O'Brien, K.D., Ferguson, M.S., Polissar, N., Hatsukami, T.S., Yuan, C.: Inflammation in carotid atherosclerotic plaque: a dynamic contrast-enhanced MR imaging study. Radiology **241**, 459 (2006)

Contemporary Problems in Quantitative Image Analysis in Structural Neuronal Plasticity

Błażej Ruszczycki, Monika Bijata, Agnieszka Walczak, Grzegorz Wilczyński, and Jakub Włodarczyk

Abstract In studies of the brain structure at the microscopic level the analysis of microscopic images is playing currently a dominant role. Insight into the processes of structural neuronal plasticity may be a key step towards understanding the bases of various neurodegenerative diseases. Almost in each experiment, there is a need for a quantitative analysis of brain tissue images. Such analysis often requires using elaborate computational tools, due to the rich structure present in the observed images. We review the most important problems encountered in the processes of analysis of fluorescent confocal microscopy images. Each of these problems suggests a dedicated computational approach.

1 Structural Neuronal Plasticity

The brain plasticity may be defined as reorganization of the neuronal network that allows for changes in responses to incoming environmental stimuli, in a context of the previous experience. This phenomenon underlies cognitive functions like learning and memory, and, in its aberrant form, plays an important pathogenic role in brain disorders, such as epilepsy, mental retardation, drug addiction, to name just

---

B. Ruszczycki (✉)
Nencki Institute of Experimental Biology, Laboratory of Imaging Tissue Structure and Function, Polish Academy of Sciences, Pasteura 3, Warszawa, Poland
e-mail: b.ruszczycki@nencki.gov.pl

M. Bijata · J. Włodarczyk (✉)
Nencki Institute of Experimental Biology, Laboratory of Cell Biophysics, Polish Academy of Sciences, Pasteura 3, Warszawa, Poland
e-mail: jwlodarc@nencki.gov.pl

A. Walczak · G. Wilczyński
Nencki Institute of Experimental Biology, Laboratory of Molecular and Systemic Neuromorphology, Polish Academy of Sciences, Pasteura 3, Warszawa, Poland
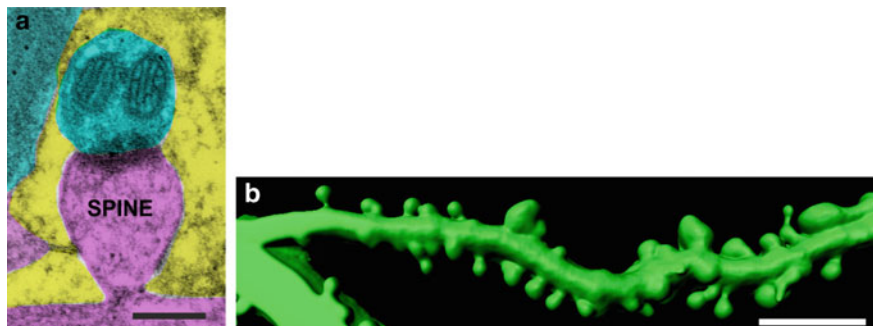
Fig. 1 Dendrites and spines. (**a**) Electron micrograph of the rat brain tissue showing an excitatory synapse formed by the dendritic spine. Pseudocolors indicate the cellular components: dendrite and dendritic spine (*purple*), presynaptic bouton(s) (*cyan*), perisynaptic astroglial processes (*yellow*). (**b**) The dendrite of the cultured rat neuron, visualized using confocal microscopy; see the multiple dendritic spines protruding from the shaft. Scale bars: (**a**) 500 nm, (**b**) 10 μm

a few. At the cellular level, the plasticity is apparently supported by modifications in neuronal connectivity and excitability that are driven by molecular changes in neurons and glia. Synapses are specialized membrane domains forming contact sites through which neurons communicate [1] (Fig. 1a). The typical synapse is composed of the pre- and postsynaptic domains, physically separated by a ∼20 nm wide gap, called the synaptic cleft [2]. The electrical signal is transmitted from pre- to post-synaptic domain by neurotransmitter molecules. The major excitatory and inhibitory neurotransmitters in mammalian brain are glutamate and gamma-aminobutyric acid [3], respectively. In mammals, more than 90 % of the excitatory synapses are located on small dendritic protrusions (0.2–2 μm in diameter) called the dendritic spines [4, 5] (Fig. 1b). Synapses are particularly prone to dynamic alterations and thus are believed to play the major role in the plasticity. Subpopulation of dendritic spines and axonal boutons from various cell types and in various brain areas continuously change in size or appear and disappear (undergoes spine turnover). The ability to change synaptic connections (called synaptic plasticity) comprises both alterations at the morphological level and changes at the functional level. It has been shown that changes in spine turnover and stabilization underlie synaptic plasticity and memory processes [6–8]. Short-term memory is considered to rely mainly on the strengthening and weakening of existing synapses by changes in postsynaptic receptor composition [9,10]. Long-term storage of information is thought to require structural reorganization such as increase in the number of synapses, the formation of new synapses and the loss of existing connections [10, 11]. Experimental paradigms mimicking some aspects of the plasticity have also been identified, e.g., long-term potentiation (LTP) and depression (LTD) [12,13].

Besides structural changes at the synaptic sites, long-lasting synaptic plasticity also involves dramatic changes in gene expression, occurring under the influence of specific signaling pathways and transcription factors [14]. The phenomenon

Fig. 2 Chromosomal domains. Three dimensional visualization (based on a confocal stack) of the chromosome 3 domains (*red*) in the nucleus of the rat hippocampal neuron; a growth factor *bdnf* embedded in its parent chromosomal domain is shown in *green*; nucleus outline is based on DNA-specific staining (*blue*)
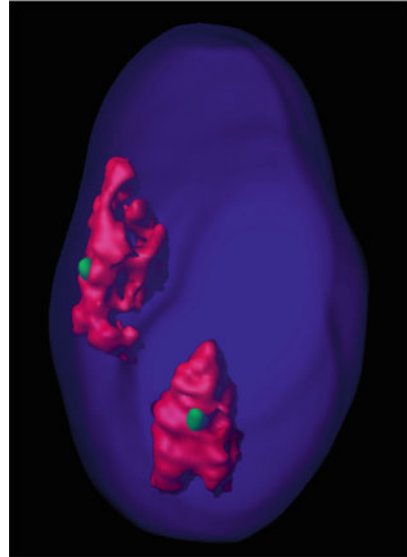
provides time- and space-correlated expression of proteins required to establish and/or maintain structural synaptic modifications. Recent data indicate that, in addition to *cis*- and *trans*-acting DNA sequence-dependent mechanisms, epigenetic mechanisms also play a role in these processes [15]. Numerous studies have shown that covalent modifications of chromatin, e.g., DNA methylation and histone methylation/acetylation, as well as non-coding RNAs, play key roles in neuronal plasticity underlying cognitive phenomena and the pathogenesis of brain diseases [15, 16]. Interestingly, recent studies in non-neuronal cell types indicate the existence of yet another class of epigenetic mechanisms, related to the structural remodeling of the cell nucleus [17,20]. It has been evident since early electron microscopic studies that the cell nucleus is not homogeneous, but contains a range of distinctive structural components [18]. However, the function and molecular composition of these structures, except chromatin and nucleolus, have remained largely obscure throughout decades. Only recently, with the advent of modern morphological approaches, widely utilizing molecular biology techniques, some new light has been shed on the significance of the nuclear substructure. Perhaps the most important discovery was the finding that interphase chromosomes are not intermingled but form distinctive, though irregular, chromosomal domains [19] (Fig. 2). Importantly, genes located at the periphery of different chromosomal domains can loop out and move toward one another, to form macromolecular factories in which transcription and mRNA processing occur in concert [20] . Such factories can utilize nuclear resources more efficiently under the specific requirements of the external environment. Another major structural element within the cell nucleus, shown to play an important role in the regulation of gene expression, is nuclear lamina, a fibrillar framework situated beneath the nuclear envelope. At least in vertebrates, lamina exerts a strong

inhibitory influence on the transcription of genes located in its vicinity [21]. The majority of data on the neuronal nuclear architecture comes from the studies on peripheral sympathetic and hypothalamic neurons [22], and references therein. In contrast, virtually nothing is known about the architectural dynamics of neuronal nuclei in the regions engaged in cognitive functions (or epileptogenesis) such as cerebral cortex and hippocampus. Similarly, the structural dynamics of neuronal nuclei has been hardly studied in the context of gene expression phenomena underlying synaptic plasticity. Despite the enormous progress in the understanding of neuronal plasticity that has occurred during the past decades, the molecular machinery that drives architectural changes at the synapses is far from complete understanding, and the architecture of the neuronal nucleus remains an unexplored world. Therefore, there is a need for constant improvement of the methods to study structural changes at these subcellular neuronal compartments. Here, we attempt to review our recent developments in the field of quantitative morphometric analysis of dendritic spines and neuronal nuclei, based on the images obtained using high-end confocal microscopy. These methodological aspects played a fundamental role in our work on the role of extracellular proteolysis [23], and of the nuclear architectural remodeling, in the neuronal plasticity [24].

2   Brain Tissue Imaging

The analysis of dendritic spine morphology in fluorescent confocal microscopy images has become a popular technique in contemporary neuroscience [7,25–45]. It is not uncommon that in a single experiment thousands (or even more) of dendritic spines or neuronal nuclei are analyzed. Manual analysis of such a large number of objects is very time consuming. Therefore, the use of a specialized software is required. As changes in the spine morphology which are looked for might be quite subtle, the conclusions drawn from the data may crucially depend on a performance of the computational tool applied. In consequence, the selection of an appropriate technique is a very important step which may crucially impact the final result of the experiment.

Fluorescence labeling of dendritic spines is currently the main tool for structural plasticity study by confocal microscopy. However, the use of various types of neurons, and different imaging and staining techniques, leads to images which appear superficially similar but once we look into the details we discover that they are actually quite different. This can present a major challenge for fully automatic algorithms which might be insufficiently flexible to cope with complicated structures. In turn, this leads to the necessity for manual analysis of images for meaningful interpretation of the data. An example of the discussed problem is illustrated in Fig. 3, where three different types of images are presented. We can recognize immediately the appearance of such artifacts as a halo around the dendrite, the complicated and often branching structure of dendritic spines, and inhomogeneity of spines and dendrite.
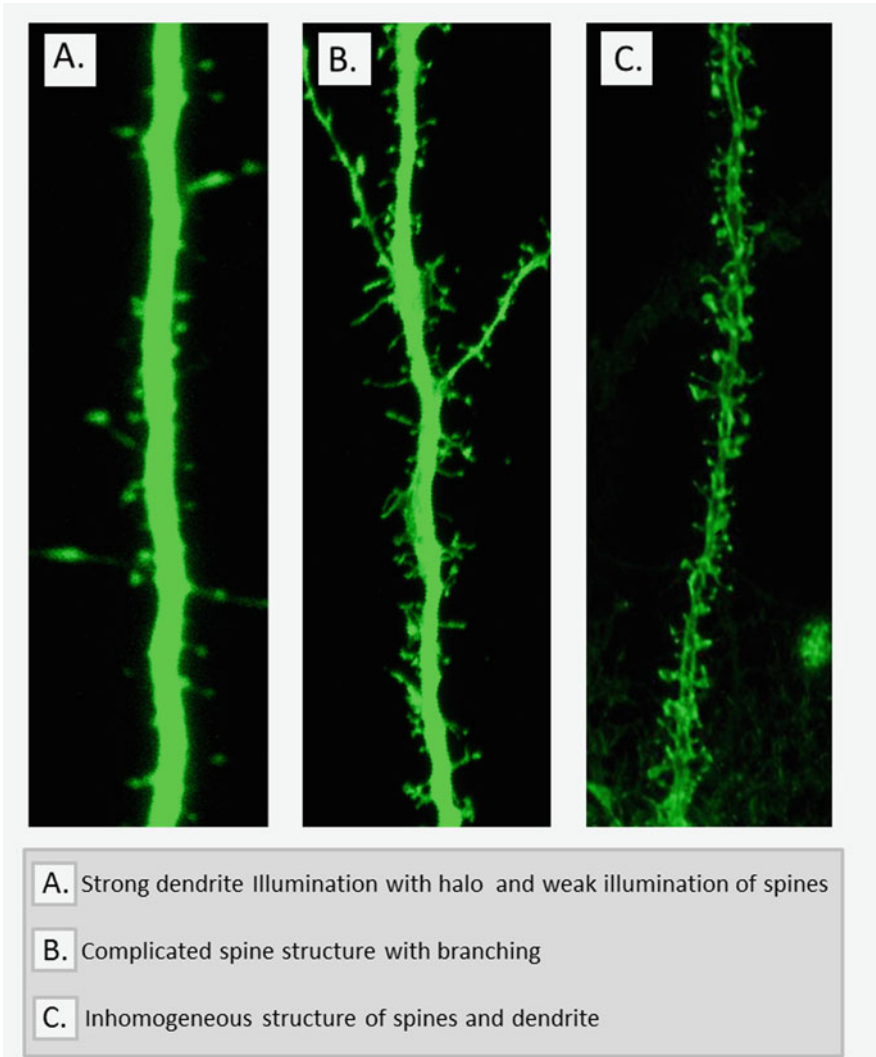
Fig. 3 The segment of a dendrite with spines; different types of neurons and staining techniques have been used. (**a**) Organotypic slices marked with RFP (the color was changed to *green* in postprocessing), (**b**) cell from dissociated cultured marked with GFP, (**c**) cell from thick slices marked with DiI. In spite of overall similarity of obtained images, the detailed structure of dendritic spines may be substantially different

There are several strategies to introduce fluorescence probe (i.e., fluorescence dye or fluorescence protein) into the sample, such as transfection, electroporation, microinjection, viral transformation, lipofection, and biolistic method. One of the methods for the fluorescence labeling employed in the studies described here is a gene gun method using a lipophilic carbocyanine dye (DiI) which diffuses
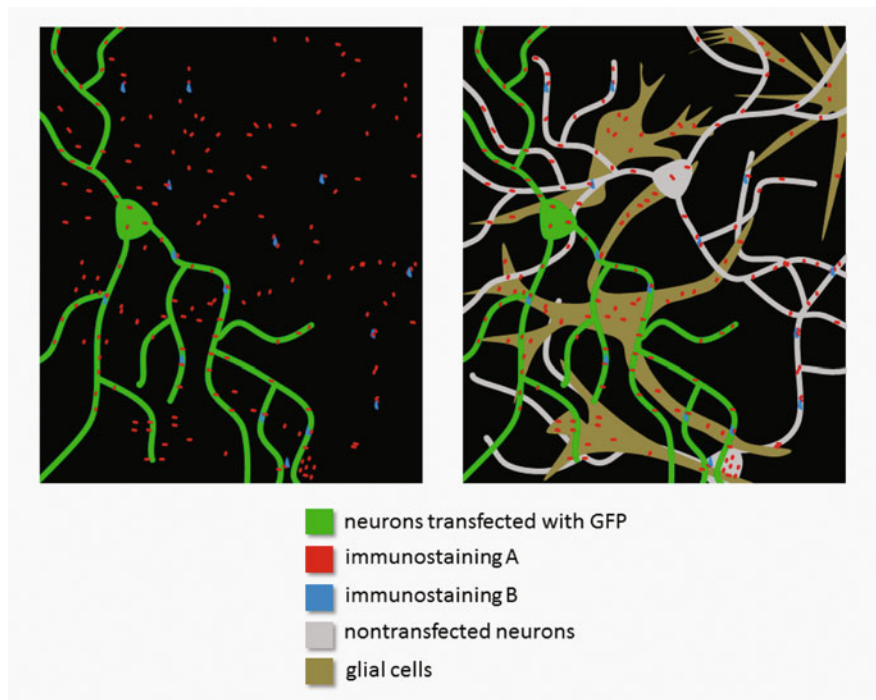
Fig. 4 Multichannel images combine observations of the transfected neuron with simultaneous detection of signals representing immunostaining

along the neuronal membrane labeling spine structures. Transfection using a gene gun (referred as dio- or biolistics method) is a procedure to introduce coated microparticles with dye or DNA respectively, by firing them into the sample. Another method employed here is transfection either by lipofection (using liposomes, vesicles made of phospholipid bilayers, which attaches to cell membrane) or biolistic method with DNA coding fluorescence protein (GFP, RFP) using neuron-specific promoter (Synapsin). In practice, the researchers do not restrain themselves to using a specific staining and imaging technique, therefore an appropriate computational tool has to possess an ability to cope with different types of images. Although all the fluorescent labeling methods are widely used when morphology of spines is studied, all of them faced with the problem of labeling efficiency and in practice transfected neuron is surrounded by non-transfected cells. This means that we see only few of the dendrites from the very complicated actual structure (see Fig. 4). This is obviously advantageous from the point of view of image recognition procedure, it is much easier to perform a segmentation and detection in an image embracing sparse neurons. However, from the perspective of data analysis this presents an additional challenge—the systematic effect can be "buried in the noise" when the analyzed population of neurons is small. Moreover, in case of

simultaneous analysis of immunostaining observed in the other channels we often compartmentalize the signal into the region of the interest, here the dendritic spines. However, outside the region of interest, we still observe the immunostaining signal from the dendritic spines within non-transfected neurons (and therefore not visible). In such cases we cannot easily define the background for signals on channels corresponding to immunostaining fluorescence. Moreover, the quantification of the immunostaining signal inside the dendritic spines may be strongly disturbed by signals from the other structures in the proximity to the dendritic spines, e.g. glial cells.

3   Reconstruction of the Spine Surface

The necessary operation in the image processing is the reconstruction of the spine surface and distinction of the objects inside the spine from the objects in the dendrite and in the spine vicinity (which includes extracellular space, and possibly axons, dendrites with spines belonging to other neurons, glial cells). In general, most of the reconstruction methods are either intensity or gradient based. If the spine surface in the image possesses a sharp boundary, with little internal details, the gradient-based method is a natural choice. Unfortunately, this is not often a case we encounter in confocal images. Lack of sharp boundary of the spine does not influence so much the intensity-based methods, which include image thresholding and binarization, often improved by various postprocessing steps. The binarized image is definitely easier to analyze, however in case of confocal images the global thresholding leads to the occurrence of different type of artifacts, presented in Fig. 5. We can clearly observe a phenomenon of the head detachment, when a small threshold value is used in the image binarization (red arrows). With increasing threshold, we may obtain continuity between the spine head and the dendrite; however we simultaneously lose details of spine structure close to the dendrite (blue arrows), where the observed fluorescence intensity is larger. Therefore the algorithms that perform the image binarization need additional postprocessing steps in order to correct the artifacts arising in the binarization process. For example, Koh et al. [51] perform the detection of detached spine heads and merge them back with the dendrite. This problem may be solved by adjusting the local value of a threshold (see Fig. 6) by using a sort of adaptive threshold techniques.

Several different algorithms dedicated to the recognition and segmentation of dendritic spines have been invented. The techniques employed in them include image binarization, skeletonization (backbone extraction), contour detection by means of active geodesics, and feature point extractions from the gradient flow methods.

Unfortunately, most of the algorithms described in literature are not supported by available implementations, what obstructs performing independent tests and evaluations. What may possibly happen is that the algorithm that performs quite well in recognition of the structures in a certain set of images fails to recognize the dendritic spines in the other types of neurons.
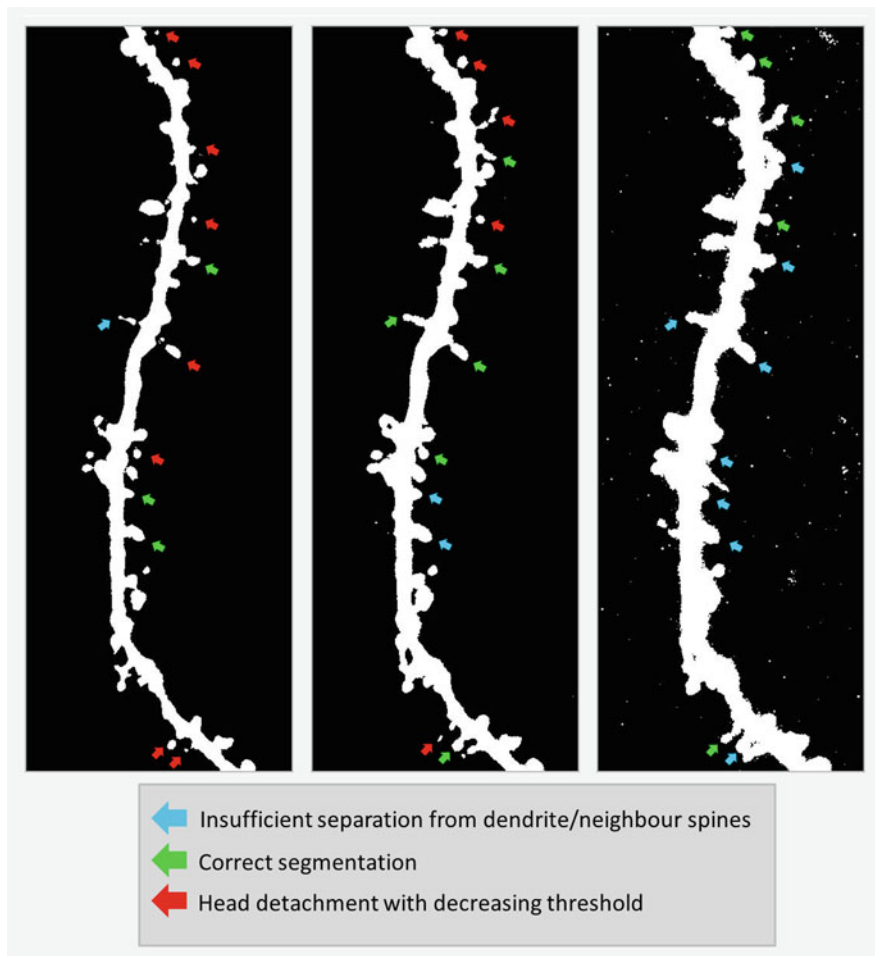
Fig. 5 Image binarization leads to the creation of artifacts. The image was thresholded using three different values of the threshold (9, 40, and 75 units). It is not possible to find the proper value of the global threshold that provides the correct spine segmentation

4   Limitation in Axial Resolution Leads to Uneven Directional Distribution of Observed Spines

The resolution of the conventional confocal microscope in the axial direction is approximately three times worse than in the lateral direction. Since the dimensions of the dendritic spines are comparable with the resolution of the microscope (e.g., the lateral and axial resolution limits could be around 0.25 and 0.7 μm, while the
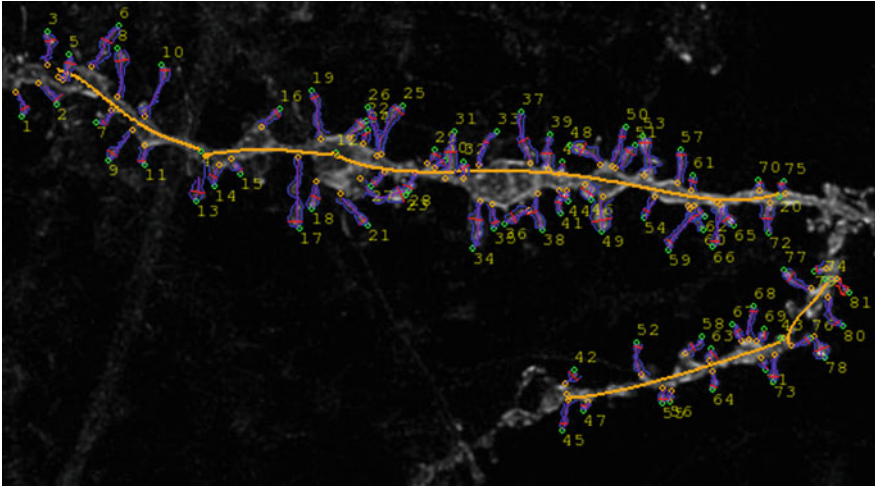
Fig. 6 Example of spine segmentation on a spiny dendrite using adaptive threshold technique. The positions of the top of the spine head (*green circles*) and the dendrite harboring (*orange circles*) were chosen manually. The applied algorithm reconstructed the spine skeleton and the boundary of the spine surface. The *red lines* show the size of the spine head width

head width of the spine is usually around $1\,\mu m$ and the neck width 0.2–$0.6\,\mu m$), the details of the spine shape are lost. Especially, along the axial direction, the resolution of the conventional confocal microscope is insufficient to acquire a high quality image. The situation might be improved by using superresolution techniques; however in many cases the application of these techniques is limited, for example due to the requirements of the high laser power which might be inadequate in live imaging application. The recent progress in application of superresolution techniques [46, 47] in live imaging of neurons may potentially allow to overcome the discussed limitations.

The visual inspection reveals that there is a large non-isotropy of directional distribution of the dendritic spines in the confocal images (most of the observed spines protrude in the lateral direction). One of the possible explanation of this effect is flattening of the cell by the growth on the cover glass, in case of visualization of dissociated culture cells. However, in case of visualization of thick organotypic slices, the effect of flattening is absent, yet the high non-isotropy of the spine distribution is still observed: Fig. 7 presents a projection of a fragment of dendrite along three different directions. Almost no spines that protrude towards axial direction are observed. In order to quantify this effect we performed an analysis of the spine distribution angle (the angle between the line connecting the point of spine harboring at the dendrite with the top of the spine head, and the axial direction of observation). The analysis revealed that this angle is highly centered around 90°. If in the analyzed images the effect of non-isotropic distribution persists, the full three-dimensional analysis will hardly give any improvement over two-dimensional
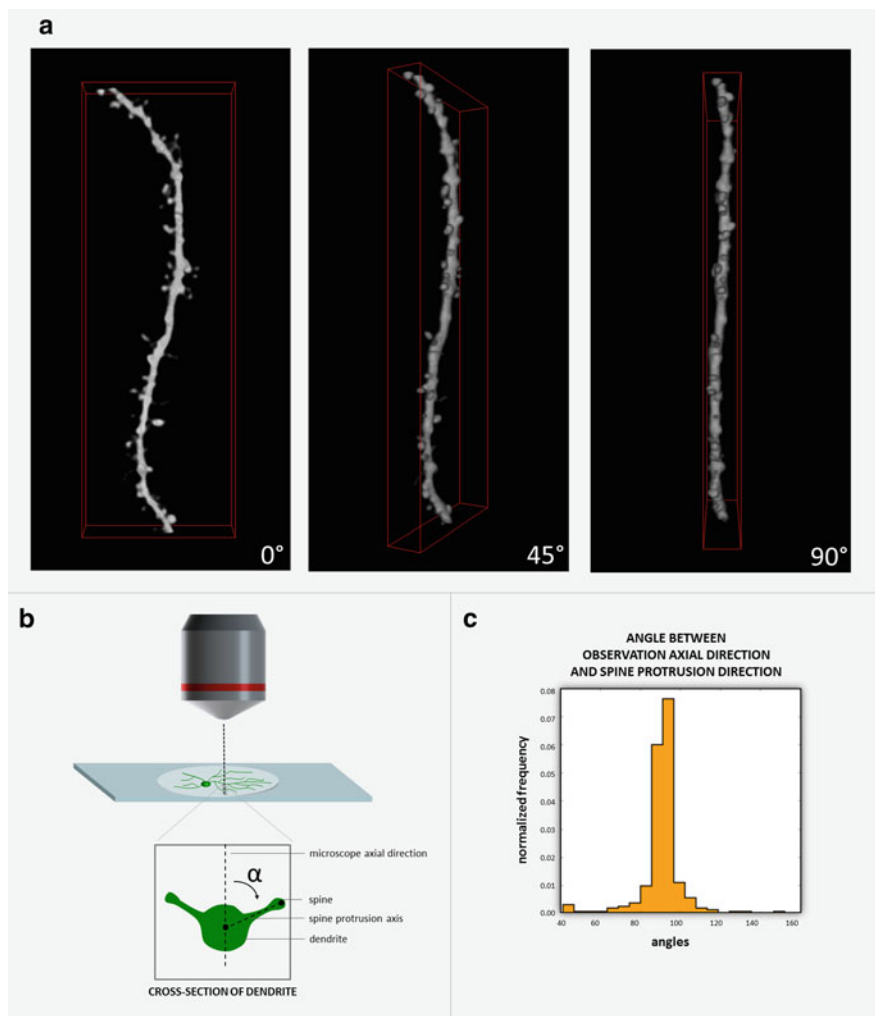
Fig. 7 The limitation in axial resolution of confocal microscopes leads to the large non-isotropy in the angular distribution of observed dendritic spines. (**a**) Fragment of the dendrite in a confocal stack projected onto three different planes (0° is the lateral plane). (**b**) The measurement of the protrusion angle. (**c**) The distribution of the protrusion angle

analysis. In literature, the two-dimensional analysis is commonly performed, presumably due to the discussed limitations. The phenomenon of anisotropy (however smaller) in the distribution of protruding angle was also reported by [48] . They observed that 65 % of spines protruded in "mostly horizontal" while 35 % in "mostly vertical" direction.

5   Choosing the Representative Set of Parameters

The most widely studied parameters describing the spine morphology are the spine head size (the head-width or the head area), the total spine volume or area, the spine length, and the neck width. The reason for the choice of these parameters is primarily their intuitive meaning and the feasibility of the measurements. However, the precise definition of the majority of morphological parameters is not straightforward. For example, by measuring the spine length as the distance along the straight line from the top to the bottom of the spine, we underestimate the length of the bent spines. The definition of the spine head size is usually ambiguous, especially in case of the highly irregular shapes (when the spine shape is not symmetric). Once we establish a certain number of precise definitions of what is measured, another issue appears, namely what is the appropriate number of parameters, which describe the spine shape.

An interesting approach has been suggested by Ceyhan et al. [52]. In order to detect the changes in shapes of dendritic spines they performed an analysis based on the *Large Deformation Diffeomorphic Metric Mapping* (LDDMM). The situation with which they deal is as follows: Let us consider two populations of spines, in control group and in diseased animals. The major source of the observed variation, while the comparative analysis is performed, origins from the spread of the spine overall sizes. The differences in the shapes of spines, between healthy and diseased animals, may be very subtle. By studying any of the parameters such as length, spine volume, head size, the differences might be hidden in the variation of the overall spine dimensions. It might be some combination of the parameters describing the spines shapes, which allows to detect the morphological deformation in diseased animals. In order to find an appropriate measure, the *Principal Component Analysis* was proposed to be performed. So far, there was no systematic study that exploited this potentially fruitful approach. With the recent advancements in the imaging and image segmentation techniques, the amount of available data describing the response of the dendritic spine morphology as a result of different treatments, stimulations and disease models has significantly increased, therefore making feasible the application of *Principal Component Analysis* techniques in studies of synaptic plasticity processes.

6   Choosing the Proper Size of Experimental Sample and Statistical Approach

Almost all experiments, in which the spine morphology is analyzed, employ a "comparative design", i.e. the morphology of dendritic spines in two or more groups of subjects is compared, e.g. in healthy versus diseased animals; in cells treated with stimulating chemicals versus cells treated with a pure solvent. In each case, the primary question is "Do we see the statistically significant changes in the value

of the studied variable?". Currently, there are no common standards regarding the value of the changes that still have a functional character. Taking into account the enormous complexity of the molecular machinery involved in the processes of synaptic plasticity, the determination of the magnitude of morphological changes, over which they represent a correlation with certain biochemical processes is established, is virtually impossible. By studying an extremely large number of samples, we increase our chances to find statistically significant changes; however this changes may represent some secondary effect, artifacts connected with the experimental procedure, or the chain of different effects. The sound interpretation of such tiny morphological changes is usually impossible. Difficulty in establishing the tentative values over which the magnitude of statistically significant changes may be interpreted unfortunately does not allow establishing the common standards.

Another important question regards the size of the experimental sample that has to be collected in order to provide a statistical significance, once a certain underlying effect is indeed present. For a certain magnitude of changes, the effect may be not observed in the data analysis if the sample size is too small; the shapes of dendritic spines form a very diverse spectrum, therefore the sampling variation is significant. In order to cope with this problem the Monte Carlo simulations were performed. The results of simulations provide the required number of samples and size of spine population that has to be studied, in order to obtain a statistically significant detection of the simulated effect.

The problem of choosing the proper size of experimental sample is universal and occurs in every experiment in which the statistical effect is sought in the sample of subjects. In case of dendritic spines this issue has an primary importance, due to the extremely large diversity of shapes of dendritic spines. Lack of common criteria may lead to the situation, when the outcome of an experiment (positive or negative) reflects mostly the size of the studied sample. For example two researchers may study an effect of two different chemical stimulations of neurons. Let us assume that both lead to 10 % increase of the head-width. If the researcher A uses 15 samples per group they will detect the studied effect with 95 % probability, using the t-test with p-value 0.005. The researcher B who studies only seven samples per group has more than 50 % chance of missing the studied effect, using the same statistical approach.

So far, most of the morphological changes reported in literature concerned the size of the spine head, the secondary variable was the spine length. This might be a result of the fact (observed in the Monte Carlo simulations, see [49]) that the changes in the spine head width are the easiest to be find, due to the tight distribution (no tails) of this variable. The simulation results show also that the detection of changes occurring in spine subpopulation depends significantly on the type of the test employed: The changes occurring in the numerous small spines are much easier to detect by comparing the spine distributions using the Kolgomorov–Smirnov test, rather than by using the t-test or the u-test, which take as arguments the mean values of the studied variable calculated for each animal or cell in the experiment. The opposite situation occurs once the changes take place in the subpopulation of the few large spines.

There is a popular classification scheme of the dendritic spines into the few subclasses, such as filopodia, stubby, mushroom, and thin. The morphological spectrum of spines might be described by the fraction of spines in these subclasses; however the Monte Carlo simulations show that the detection of changes in the values of morphological variables is more sensitive than the detection of changes by comparison of the size of subclasses between the studied groups.

7   Live Imaging

The main advantage of live imaging approach, from the perspective of statistics, is the elimination of the variation resulting from the spine morphology diversity. By comparing the changes in the morphology of the same spines observed at different time points, we expect the improvement in detection sensitivity. We need however to pay the price of additional technical difficulties. The most immediate challenge is to identify the same spine at two different snapshots. This may be an easy task if there is a very short time-lapse between consecutive images (of order of minutes) and there are no drastic effects of performed neuron stimulation. However, when the time between the observations is large, some spines disappear, the new spines are formed, and the existing spines are changing shape and orientation. The identification of the same spine might be therefore not a trivial task (see Fig. 8). Moreover, the change of spine orientation may lead to further artifacts, if we observe the two-dimensional projections we cannot distinguish whether the actual change of length has taken place, or whether the spine has changed its orientation. With a full three-dimensional image, such a distinction is possible; however the difference in the lateral and axial resolutions of the microscope comes in play. This difference leads to the observation of superficial changes in morphology, if the spine changes its orientation without changing its shape. Therefore, with the full three-dimensional live imaging morphometric measurements the influence of unequal microscope resolution in axial and lateral dimension should be evaluated and accounted in the data analysis.

In the article by Son et al. [50] the method of tracing the changes of dendritic spines in a series of time-lapse two-dimensional images has been proposed. This method is mainly based on detecting the spine boundaries by using a geodesic active contour model, supplied by the watershed algorithm. Once the spines are detected an optical flow method is used to track the changes between two consecutive images in the time-lapse series.

8   Quantitative Analysis of Neuronal Nuclei Architecture

The main challenge in the analysis of neuronal architecture is the proper segmentation of the nuclei, which is the necessary step in order to reconstruct the nuclear envelope, and to isolate objects inside the nucleus. In the brain tissue
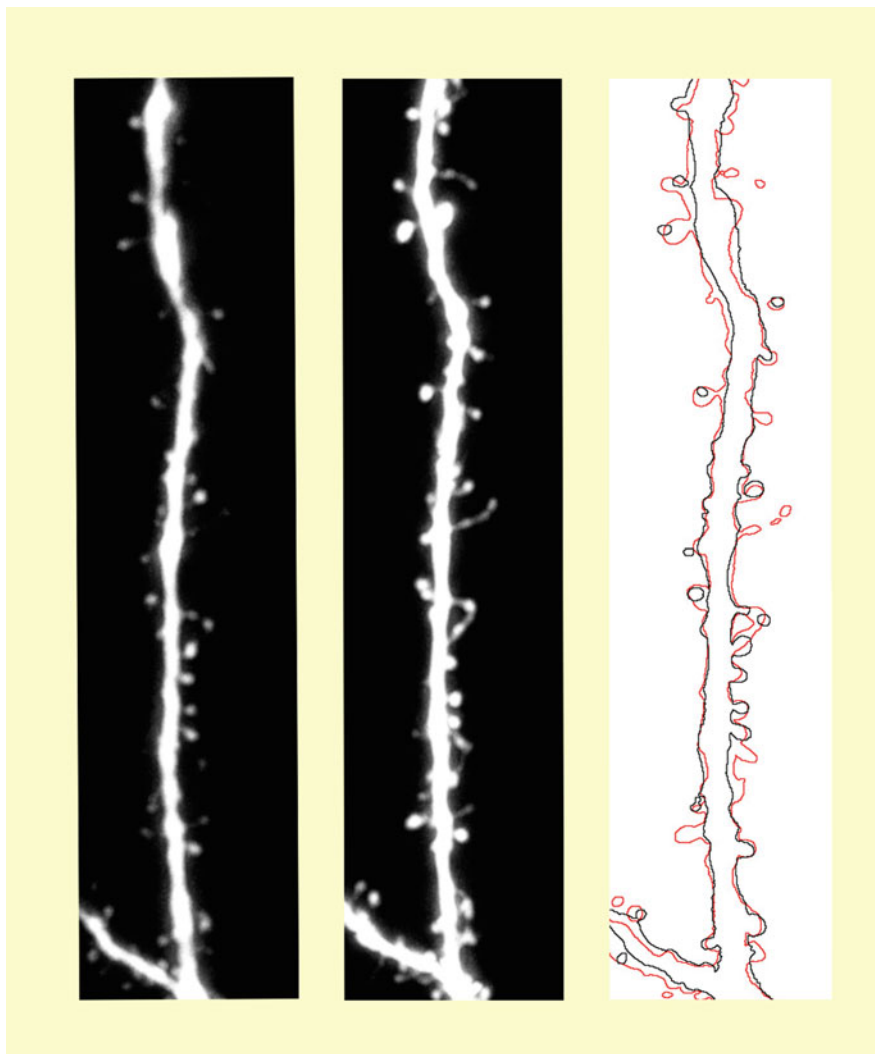
Fig. 8 Time-lapse imaging of the apical dendrite. *Left panel*: The observation before the chemical LTP stimulation. *Middle panel*: The observation 40 min after the chemical LTP stimulation. *Right panel*: Comparison of the contours before/after stimulation

images the nuclei are very densely packed, and often appear to overlap, which introduces a substantial difficulty into the segmentation task. The traditionally used watershed algorithm fails to perform the correct segmentation, as the number of undersegmented nuclei may be very high. In order to cope with the aforementioned problems we constructed an algorithm which starts the segmentation process at the two-dimensional plane, where there is a high chance that the analyzed nucleus is

well separated from the adjacent nuclei. By moving to the neighbor planes we can exploit the fact that the nucleus surface is continuous. The continuity constraint is helpful in segmenting the sections where the nucleus is no longer well separated from its neighbors. An example of correctly reconstructed nucleus is presented in Fig. 2

References

1. Yamada, S., Nelson, W.J.: Synapses: sites of cell recognition, adhesion, and functional specification. Annu. Rev. Biochem. **76**, 267–294 (2007)
2. Harris, K.M., Weinberg, R.J.: Ultrastructure of synapses in the mammalian brain. Cold Spring Harb. Perspect. Biol. **4**, 1–30 (2012)
3. Alexander, M.J.H.: Glutamate. In: Squire, L. (ed.) Encyclopedia of Neuroscience. Elsevier, Oxford (2009)
4. Nimchinsky, E.A., Sabatini, B.L., Svoboda, K.: Structure and function of dendritic spines. Annu. Rev. Physiol. **64**, 313–353 (2002)
5. McKinney, R.A., Thompson, S.M.: Glutamate regulation of dendritic spine form and function. In: Squire, L. (ed.) Encyclopedia of Neuroscience. Elsevier, Oxford (2009)
6. Holtmaat, A., Svoboda, K.: Experience-dependent structural synaptic plasticity in the mammalian brain. Nat. Rev. Neurosci. **10**, 647–658 (2009)
7. Xu, T., Xinzhu, Yu., Perlik, A.J., Tobin, W.F., Zweig, J.A., Tennant, K., Jones, T., Zuo, Y.: Rapid formation and selective stabilization of synapses for enduring motor memories. Nature **462**, 915–919 (2009)
8. Yang, G., Pan, F., Gan, W.B.: Stably maintained dendritic spines are associated with lifelong memories. Nature **462**, 920–924 (2009)
9. Neves, G., Cooke, S.F., Bliss, T.V.: Synaptic plasticity, memory and the hippocampus: a neural network approach to causality Nat. Rev. Neurosci. **9**, 65–75 (2008)
10. Hofer, S.B., Bonhoeffer, T.: Dendritic spines: the stuff that memories are made of? Curr. Biol. **20**, 157–159 (2010)
11. Bailey, C.H., Kandel, E.R.: Structural changes accompanying memory storage. Annu. Rev. Physiol. **55**, 397–426 (1993)
12. Luscher, C., Frerking, M.: Long-term depression (LTD): metabotropic glutamate receptor (mGluR) and NMDAR-dependent forms. In: Squire, L. (ed.) Encyclopedia of Neuroscience. Elsevier, Oxford (2009)
13. Sweatt, J.D.: Long-term potentiation (LTP). In: Squire, L. (ed.) Encyclopedia of Neuroscience. Elsevier, Oxford (2009)
14. Cohen, S., Greenberg, M.E.: Communication between the synapse and the nucleus in neuronal development, plasticity, and disease. Annu. Rev. Cell Dev. Biol. **24**, 183–209 (2008)
15. Sweatt, J.D.: Experience-dependent epigenetic modifications in the central nervous system. Biol. Psychiatry **65**, 191–197 (2009)
16. Urdinguio, R.G., Sanchez-Mut, J.V., Esteller, M.: Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. Lancet Neurol. **8**, 1056–1072 (2009)
17. Dundr, M., Ospina, J.K., Sung, M.H., John, S., Upender, M., Ried, T., Hager, G.L., Matera, A.G.: Actin-dependent intranuclear repositioning of an active gene locus in vivo. J. Cell Biol. **179**, 1095–1103 (2007)

18. Monneron, A., Bernhard, W.: Fine structural organization of the interphase nucleus in some mammalian cells. J. Ultrastruct. Res. **27**, 266–288 (1969)
19. Cremer, T., Cremer, M.: Chromosome territories. Cold Spring Harb. Perspect. Biol. **2**, a003889 (2010)
20. Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., Iborra, F.J., Buckle, V.J.: Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J. Cell Biol. **182**, 1083–1097 (2008)
21. Andrulis, E.D., Neiman, A.M., Zappulla, D.C., Sternglanz, R.: Perinuclear localization of chromatin facilitates transcriptional silencing. Nature **394**, 592–595 (1998)
22. Villagra, N.T., Bengoechea, R., Vaque, J.P., Llorca, J., Berciano, M.T., Lafarga, M.: Nuclear compartmentalization and dynamics of the poly(A)-binding protein nuclear 1 (PABPN1) inclusions in supraoptic neurons under physiological and osmotic stress conditions. Mol. Cell. Neurosci. **37**, 622–633 (2008)
23. Michaluk, P., Wawrzyniak, M., Alot, P., Szczot, M., Wyrembek, P., Mercik, K., Medvedev, N., Wilczek, E., De Roo, M., Zuschratter, W., Muller, D., Wilczynski, G.M., Mozrzymas, J.W., Stewart, M.G., Kaczmarek, L., Wlodarczyk, J.: Influence of matrix metalloproteinase MMP-9 on dendritic spine morphology. J. Cell Sci. **124**, 3369–3380 (2011)
24. Walczak, A., Szczepankiewicz, A.A., Ruszczycki, B., Magalska, A., Zamlynska, K., Dzwonek, J., Wilczek, E., Zybura-Broda, K., Rylski, M., Malinowska, M., Dabrowski, M., Szczepinska, T., Pawlowski, K., Pyskaty, M., Wlodarczyk, J., Szczerbal, I., Switonski, M., Cremer, M., Wilczynski, G.M.: Novel higher-order epigenetic regulation of the BDNF gene upon seizures. J. Neurosci. **33**, 2507–2511 (2013)
25. Yuste, R., Denk, W.: Dendritic spines as basic functional units of neuronal integration. Nature **375**:682–684 (1995)
26. Segal, M.: Release of calcium form stores alters the morphology of dendritic spines in cultured hippocampal neurons. Prog. Brain Res. **138**, 53–59 (2002)
27. Segal, M.: Dendritic spines, synaptic plasticity and neuronal survival activity shapes dendritic spines to enhance neuronal viability. Eur. J. Neurosci. **31**, 2178–2184 (2010)
28. Yuste, R., Bonhoeffer, T.: Morphological changes in dendritic spines associated with long-term synaptic plasticity. Annu. Rev. Neurosci. **24**, 1071–1089 (2001)
29. Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., Kasai, H.: Principles of long-term dynamics of dendritic spines. J. Neurosci. **28**, 13592–13608 (2008)
30. Matsuzaki, M., Honkura, N., Ellis-Davies, G.C.R., Kasai, H.: Structural basis of long-term potentiation in single dendritic spines. Nature **429**, 761–766 (2004)
31. Holtmaat, A.J., Trachtenberg, J.T., Wilbrecht, L., Shepherd, G.M., Zhang, X., Knott, G.W., Svoboda, K.: Transient and persistent dendritic spines in the neurocortex in vivo. Neuron **45**, 279–291 (2005)
32. Ziv, N., Smith, S.: Evidence for a role of dendritic filopodia in synaptogenesis and spine formation. Neuron **17**, 91–102 (1996)
33. Korkotian, E., Segal, M.: Release of calcium form stores alters the morphology of dendritic spines in cultured hippocampal neurons. Proc. Natl. Acad. Sci. USA **96**, 12068–12072 (1999)
34. Sala, C., Piëch, V., Wilson, N.R., Passafaro, M., Liu, G., Sheng, M.: Regulation of dendritic spine morphology and synaptic function by Shank and Homer. Neuron **31**, 115–130 (2001)
35. Kasai, H., Fukuda, M., Watanabe, S., Hayashi-Takagi, A., Noguch, J.: Structural dynamics of dendritic spines in memory and cognition. Trends Neurosci. **33**, 121–129 (2003)
36. Hung, A.Y., Futai, K., Sala, C., Valtschanoff, J.G., Ryu, J., Woodworth, M.A., Kidd, F.L., Sung, C.C., Miyakawa, T., Bearand, M.F., Weinberg, R.J., Sheng, M.: Smaller dendritic spines, weaker synaptic transmission, but enhanced spatial learning in mice lacking Shank. J. Neurosci. **38**, 1697–1708 (2008)
37. Oray, S., Majewska, A., Sur, M.: Effects of synaptic activity on dendritic spine motility of developing cortical layer V pyramidal neurons. Cereb. Cortex **16**, 730–741 (2006)
38. Hosokawa, T., Rusakov, D., Bliss, T., Fine, A.: Repeated confocal imaging of individual dendritic spines in the living hippocampal slice: evidence for changes in length and orientation associated with chemically induced LTP. J. Neurosci. **15**(8), 5560–6673 (1995)

39. Korkotian, E., Segal, M.: Regulation of dendritic spine motility in cultured hippocampal neurons. J. Neurosci. **21**, 6115–6124 (2001)
40. Hering, H., Sheng, M.: Dendritic spines: structure, dynamics and regulation. Nat. Rev. Neurosci. **12**, 880–886 (2001).
41. Fiala, J.C., Feinberg, M., Popov, V., Harris, K.M.: Synaptogenesis via dendritic filopodia in developing hippocampal area CA1. J. Neurosci. **18**, 8900–8911 (1998)
42. Sorra, K., Harris, K.: Overview on the structure, composition, function, development, and plasticity of hippocampal dendritic spines. Hippocampus **10**, 501–511 (2000)
43. Grutzendler, J., Kasthuri, N.: WBGan: long-term dendritic spine stability in the adult cortex. Nature **420**, 812–816 (2002)
44. Dailey, M.E., Smith, S.J.: The dynamics of dendritic structure in developing hippocampal slices. J. Neurosci. **16**, 2983–2994 (1996)
45. Kasai, H., Matsuzaki, M., Noguchi, J., Yasumatsu, N., Nakahara, H.: Structure-stability-function relationship of dendritic spines. Trends Neurosci. **26**, 360–368 (2003)
46. Izeddin, I., Specht, C.G., Lelek, M., Darzacq, X., Triller, A., Zimmer. C., Dahan, M.: Super-resolution dynamic imaging of dendritic spines using a low-affinity photoconvertible actin probe. PLoS One **6**(1), e15611 (2011)
47. Nägerl, U.V., Willig, K.I., Hein, B., Hell, S.W., Bonhoeffer. T.: Live-cell imaging of dendritic spines by STED microscopy. Proc. Natl. Acad. Sci. USA **105**, 18982–18987 (2008)
48. Rodriguez, A., Ehlenberger, D., Dickstein, D., Hof, P., Wearne, S.: Automated three-dimensional detection and shape classification of dendritic spines from fluorescence microscopy images. PLoS One **3**(4), e1997 (2008)
49. Ruszczycki, B., Szepesi, Z., Wilczynski, G.M., Bijata, M., Kalita, K., Kaczmarek, L., Wlodarczyk, J.: Sampling issues in quantitative analysis of dendritic spines morphology. BMC Bioinformatics **13**, 213 (2012)
50. Son, I., Song, S., Lee, S., Chang, S., Kim, M.: Morphological change tracking of dendritic spines based on structural features. J. Microsc. **241**, 261–272 (2011)
51. Koh, I., Lindquist, W.B., Zito, K., Nimchinsky, E.A., Svoboda, K.: An image analysis algorithm for dendritic spines. Neural Comput. **14**, 1283–1310 (2002)
52. Ceyhan, E.: Modeling metric distances of dendrite spines of mice based on morphometric measures. In: The Proceeding of International Symposium on Health Informatics and Bioinformatics (2007)

Advanced MRI of Cartilage and Subchondral Bone in Osteoarthritis

Gregory Chang and Ravinder R. Regatte

Abstract Osteoarthritis is a degenerative joint disorder that predominantly affects cartilage and subchondral bone. Magnetic resonance imaging (MRI) provides a noninvasive means to detect pathologic alterations in these two tissues. In this chapter, we provide an overview of MRI techniques to evaluate cartilage and subchondral bone macrostructure, and cartilage biochemical composition [T1rho mapping, T2 mapping, 23Na MRI, glycosaminoglycan chemical exchange saturation transfer, diffusion tensor imaging (DTI)]. The ability to detect early and short-term changes in the knee joint in vivo will allow new insight into the pathogenesis of osteoarthritis and may permit early diagnosis of osteoarthritis in at-risk subjects. This knowledge and capability should ultimately accelerate the discovery and testing of novel therapies to treat osteoarthritis, a disease which represents an enormous socioeconomic and health burden on society.

1 Osteoarthritis Epidemiology

According to the Centers for Disease Control, osteoarthritis (OA) affects 46 million Americans [1]. Joint deformity, pain, and decreased mobility are the main morbidities, resulting in an estimated $127 billion in annual costs [2]. The risk of disability attributable to knee OA alone is as great as that due to cardiac disease and greater than that due to any other medical condition in elderly persons [2]. Despite great efforts by the pharmaceutical industry and academia, no therapy exists to prevent or

G. Chang (✉) • R.R. Regatte

Quantitative Multinuclear Musculoskeletal Imaging Group (QMMIG), Center for Biomedical Imaging, Department of Radiology, New York University Langone Medical Center, 660 1st Avenue, 4th Floor, New York, NY 10016, USA
e-mail: gregory.chang@nyumc.org

reverse the structural damage of osteoarthritis, and interventions are limited to pain control and, ultimately, joint replacement.

2 Pathological Alterations in Cartilage and Subchondral Bone

In the healthy synovial joint, articular cartilage is the primary weight-bearing surface (2–4 mm in thickness) and is composed of an extracellular, fluid-filled matrix of type 2 collagen fibrils and proteoglycans (aggrecan) in which chondrocytes are sparsely distributed [3]. Proteoglycans cross-link collagen fibrils and serve to provide tensile and compressive strength to the matrix [4]. Deep to articular cartilage is a layer of corticalized bone, the subchondral bone plate, which itself is buttressed by subchondral trabecular bone, an intricately connected 3D-network of trabecular plates and rods (50–200 μm in dimension) [5]. The microarchitecture of this trabecular network can be characterized quantitatively in terms of trabecular number, thickness, separation, connectivity, and plate-to-rod ratio, and it is altered in states of high bone turnover or in response to mechanical loading at the joint [6].

Though originally conceptualized as a disorder of articular cartilage, osteoarthritis is now recognized as a disorder involving abnormal mechanical loading at the joint and damage to nearly all articular structures (menisci, synovium, ligaments), but especially articular cartilage and subchondral bone [7, 8]. Degeneration in articular cartilage is initially manifested by loss of proteoglycans and disruption of the type 2 collagen network [3, 9]. This results in loss of tensile and compressive strength of the cartilage matrix, and is followed by morphologic defects such as ulceration of the superficial layer of cartilage and decreased cartilage thickness [3, 9].

In parallel, subchondral bone structure is markedly altered. Beyond the radiographic demonstration of subchondral bone sclerosis, ex vivo histomorphometric and high-resolution micro-computed tomography (CT) studies reveal thickening of the subchondral bone plate and remodeling of trabecular bone microarchitecture, including increased trabecular thickness and number, and decreased trabecular separation in OA subjects compared to controls [5, 10–13]. Such trabecular bone micro-architectural alterations are believed to affect the biomechanical competence of bone in ways that cannot be predicted by bone density or bone volume fraction measurements alone [6, 10].

The ability to magnetic resonance imaging (MRI) to noninvasively monitor pathologic alterations in both cartilage and bone in OA subjects in vivo (and without the use of ionizing radiation such as in computed tomography studies) provides great insight into the pathogenesis of OA.

Fig. 1 High-resolution axial 7 T MR image (3-D FLASH, TR/TE = 26/4 ms, 0.234 × 0.234 × 1 mm) shows up to full thickness cartilage loss (*arrow*) at the median ridge of the patella in a patient with osteoarthritis

3 Morphological MRI of Osteoarthritis

3.1 Cartilage

On conventional MRI, numerous sequences exist to evaluate articular cartilage morphology [9]. The most commonly used sequences include fat-suppressed three-dimensional gradient-recalled echo techniques [14, 15] and fast-spin echo techniques without or with fat-suppression [16, 17]. These techniques have reported sensitivities and specificities for the detection of morphologic cartilage defects near or greater than 90 % in the knee when arthroscopy is used as the gold standard for comparison [14, 16].

Fat-suppressed T1-weighted gradient-echo images depict subchondral bone as dark, cartilage as hyperintense, and synovial fluid as lower in signal intensity (Fig. 1). Because a 3-D dataset can be obtained, they are considered ideal for quantitative assessment of cartilage thickness and volume [9]. One drawback is that they are time-consuming (6–15 min) especially when high-resolution images (0.3–0.4 mm in-plane, 1–2 mm slice thickness) are obtained. In addition, gradient-echo images are not ideal for the evaluation of other joint tissues, such as ligaments and menisci.

Fast-spin echo images (Fig. 2) are typically performed with proton density weighting and depict cartilage as intermediate in signal intensity, fluid as hyperintense, and subchondral bone as hyper- or hypointense depending on whether fat-suppression is used. Because fast-spin echo images are 2-D, they are not used for quantitation of cartilage thickness and volume, but they can be used to assess other joint structures, such as ligaments and menisci.

Fig. 2 Coronal clinical 7 T
MR image (fast-spin echo,
TR/TE = 1,800/26 ms,
0.5 × 0.5 × 2 mm) shows
marked cartilage loss within
the medial compartment
(*arrow*) and osteophyte
formation (*arrowhead*)

3.2 Subchondral Bone

Because conventional MRI sequences only depict subchondral bone as hyperintense
or hypointense, the evaluation of subchondral bone is limited to assessment of the
size, spatial distribution, and pattern of subchondral bone marrow lesions (BMLs).
These BMLs can reflect edema, hemorrhage, vascularity, or fibrosis. BMLs are
typically evaluated on fluid-sensitive fat-suppressed sequences. They can be semi-
quantitatively assessed according to the whole organ MRI score [18]. Alternatively,
it is also possible to quantify the volume of BMLs [19].

BMLs are associated with overlying cartilage defects [20]. Baseline BML
volume has been shown to correlate with longitudinal changes in cartilage thickness
[21] and predict later cartilage volume loss [22]. Enlarging BMLs are especially
associated with cartilage loss [23]. Finally, BMLs in the knee are associated with
knee pain [24, 25].

4 Advanced MRI of Osteoarthritis

4.1 Cartilage

4.1.1 T2 Mapping

Cartilage T2 relaxation value mapping is a potential tool to quantitatively assess
cartilage water content and collagen architecture in vivo (Fig. 3) [9, 26–28]. T2
values reflect the dipolar interaction of water protons, which have limited mobility in
the highly anisotropic cartilage extracellular matrix. In early cartilage degeneration,

Fig. 3 Sagittal T2 map of the medial knee compartment shows elevated cartilage T2 values (*arrow*) in a patient with osteoarthritis
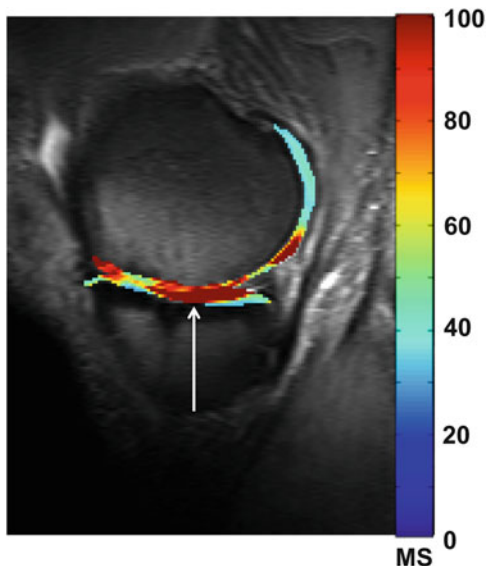
there is evidence for loss of the normal collagen framework [29] and increased water content [30, 31], both of which lead to increased cartilage T2 values [32, 33]. The potential value of T2 mapping as a biomarker for early cartilage degeneration is highlighted by its inclusion in the MRI protocol for the osteoarthritis initiative [34], a nationwide National Institutes of Health-funded observational study that aims to better understand how to prevent and treat osteoarthritis.

The T2 map is generated by: (1) obtaining a series of multiecho spin-echo images, (2) fitting the signal intensity of each pixel to a mono-exponential decay to compute the T2 value for each pixel, and (3) representing each pixel's T2 value with a color scale to create a "map," which is typically overlaid onto the anatomic image of the joint. At 3 Tesla, T2 values of knee cartilage ranges from approximately 30 to 50 ms. There is a well-known spatial variation in cartilage T2 values with higher T2 values closer to the articular surface [26, 35]. In addition, collagen orientation [36] and mechanical loading can also affect T2 values [37, 38], the latter likely through altered water content. Some recent promising studies have provided evidence that elevated knee cartilage T2 values at baseline may predict cartilage morphologic abnormalities at 2 years [39] and 3 years [40].

4.1.2 T1rho Mapping

Cartilage T1rho relaxation value mapping is a potential tool to quantitatively assess cartilage proteoglycan content in vivo (Fig. 4) [41–43]. The T1rho value reflects low frequency spin interactions (100 Hz–10 kHz) and is also known as T1 or spin lattice relaxation in the rotating frame. While cartilage T1 and T2 values are intrinsic

Fig. 4 Sagittal T1rho map of the medial knee compartment shows spatial variation in cartilage T1rho values (*arrow*) in a patient with osteoarthritis
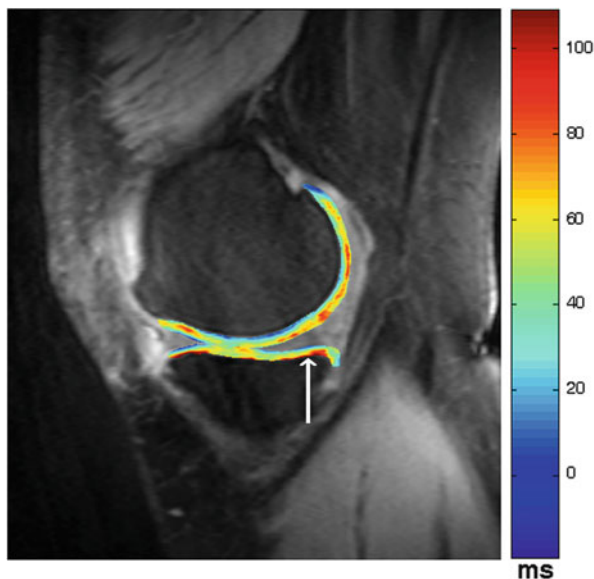
properties of a tissue, cartilage T1rho values are determined by both the tissue properties and MRI sequence parameters, in particular, the amplitude of the spin-locking pulse.

To perform T1rho mapping, the transverse magnetization is prepared using a long duration, lower power radiofrequency pulse known as the spin-lock pulse. The duration of this pulse is known as the time of spin-lock. It is during the time of spin-lock that the transverse magnetization undergoes relaxation with a rate of 1/T1rho. Multiple different types of molecular interactions (J-coupling, chemical shift, dipole–dipole, chemical exchange) contribute to the T1rho relaxation [41]. Similar to the T2 map, the T1rho map is generated by representing each pixel's T1rho value using a color look-up table and generation of a color map, which is overlaid onto an anatomic image.

In bovine cartilage specimens, T1rho values change linearly with changes in proteoglycan content [43], but there is also some evidence that collagen content may affect T1rho values as well [44]. In human cartilage specimens with normal T2 values, T1rho values correlate with cartilage proteoglycan content [45]. The high dynamic range of cartilage T1rho values in osteoarthritic human cartilage specimens suggest that it may be well-suited to monitor small changes in cartilage molecular composition and thus disease progression in humans [42]. In humans in vivo, elevated T1rho values correlate with cartilage degeneration [33], and altered T1rho values may serve as a marker for arthroscopically confirmed outerbridge grade 1 and 2 cartilage lesions [46]. Finally, one recent study has provided evidence that elevated knee cartilage T1rho values at baseline may predict cartilage morphologic abnormalities at 2 years [39].

Disadvantages of T1rho mapping include long scan times to acquire images with different spin-lock times and possible issues with specific absorption rate/energy deposition at higher field strength.

4.1.3 Delayed Gadolinium Enhanced MRI of Cartilage

Delayed gadolinium enhanced MRI of cartilage (dGEMRIC) is a method to assess cartilage proteoglycan content. Ions within the interstitial fluid of the cartilage extracellular matrix are distributed in relation to matrix levels of negatively charged glycosaminoglycans. Therefore, after administration of the anionic contrast agent Gd-DTPA^{2-}, the degree of penetration of Gd-DTPA^{2-} into the cartilage extracellular matrix is inversely related to matrix glycosaminoglycan content [3, 9]. Since Gd has a T1 shortening effect, cartilage T1 relaxation values can be used to estimate Gd concentration and thus cartilage glycosaminoglycan content.

To perform dGEMRIC, the patient is first injected intravenously with a dose of Gd-DTPA^{2-} (0.2 mM/kg, twice the recommended clinical dose) and is asked to exercise (e.g., walk on a treadmill) for approximately 10 min. The MRI and T1 mapping is performed 60–90 min after the contrast injection. Together, joint motion and the delay of imaging facilitates the entry of contrast into the joint. The T1 map is computed from a series of images with different T1 weighting. The recent development of a B1 insensitive two-dimensional (D) T1 mapping sequence [47] and a method to correct for B1 inhomogeneity in a 3-D variable flip-angle gradient-echo sequences [48] may allow for more accurate calculation of T1 values. The T1 maps are reported as the dGEMRIC index, with lower values corresponding to lower glycosaminoglycan levels.

In vitro and in vivo studies have shown that dGEMRIC measurements correlate well with reference standard measurements of glycosaminoglycans [49, 50]. In one cross-sectional study of patients with hip dysplasia, the dGEMRIC index correlated significantly with pain and severity of the hip dysplasia and differed significantly between mild, moderate, and severe dysplasia groups [51]. Other studies have provided evidence that a low dGEMRIC index at baseline may predict the development of osteoarthritis on radiographs at 6 years [52] or the later development of joint failure in patients undergoing periacetabular osteotomies [53]. Because of the wide intra- and inter-subject variation in dGEMRIC indices, a standardized dGEMRIC score (z-score based on number of standard deviations from the mean T1 value in the region of interest) has been proposed to increase the sensitivity for detection of abnormalities [54].

Disadvantages of dGEMRIC include the need for administration of an intravenous contrast agent, which is contraindicated in patients with renal dysfunction due to the increased risk of nephrogenic systemic sclerosis and the long examination time due to the delay after contrast injection.

Fig. 5 Sagittal sodium concentration map of the medial knee compartment shows markedly decreased cartilage sodium content (*arrow*) in a patient with osteoarthritis
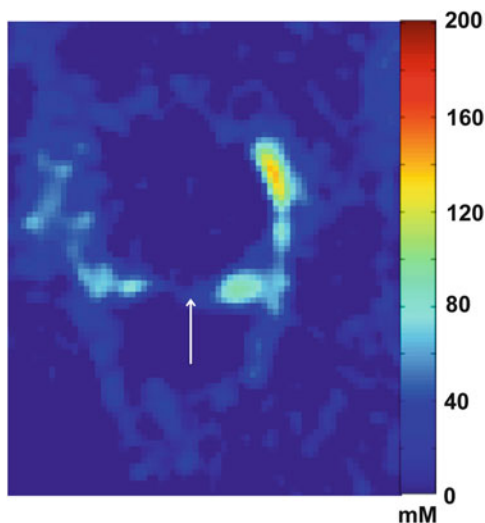
4.1.4 Sodium MRI

Sodium MRI is another method to evaluate cartilage proteoglycan content (Fig. 5). Because of the negative fixed charged density of glycosaminoglycans in the cartilage extracellular matrix, positively charged sodium ions penetrate into the interstitial fluid of the cartilage matrix in relation to the levels of glycosaminoglycans [4, 41]. That is, the lower cartilage glycosaminoglycan content results in lower sodium signal on MR images.

The sodium signal is approximately 4,000 times less than the proton signal [55]. This is because: (1) the sensitivity of MRI for sodium ions is only 9.2 % that of the MR sensitivity for protons and (2) the concentration of sodium ions in vivo is approximately 366 times lower than the water proton concentration. In addition, it should be noted that in intermediate regimes such as biological tissues, the nonspherical sodium nucleus demonstrates a quadrupolar interaction with surrounding electric field gradients. This results in multiple quantum coherences and biexponential T2 relaxation rates, with 60 % of the sodium signal relaxing rapidly (T2 value = 1–2 ms) and 40 % of the sodium signal relaxing more slowly (T2 value = 8–15 ms).

To perform sodium MRI, a dedicated coil tuned to the sodium Larmor frequency is required. Furthermore, because of sodium's rapid transverse relaxation and the already low baseline low signal-to-noise ratio of sodium MRI, imaging is facilitated by the implementation of ultrashort echo time (UTE) sequences and by the performance of imaging at high field (3 Tesla) and ultra high field (7 Tesla).

In trypsin-treated bovine cartilage specimens, cartilage proteoglycan content correlates strongly with sodium signal on MR images [56]. Because of the negative fixed charged density of proteoglycans in cartilage, it is possible to quantify cartilage

sodium concentration [57]. And in vivo, osteoarthritis subjects have demonstrated lower cartilage sodium concentration compared to healthy subjects [4, 58]. The recent development of a fluid-suppressed sodium inversion recovery sequence [59, 60] may improve the quantitation of sodium concentration in cartilage, since synovial fluid also contains sodium and can potentially confound results (Fig. 5).

Disadvantages of sodium MRI include the need for dedicated radiofrequency coils tuned to the sodium Larmor frequency and the lower spatial resolution/longer imaging times that result from low SNR.

4.1.5 Glycosaminoglycan Chemical Exchange Saturation Transfer

Glycosaminoglycan chemical exchange saturation transfer (gagCEST) has been proposed as a method to assess cartilage proteoglycan content [61, 62]. The basis of gagCEST is the detection of the magnetization associated with amide and hydroxyl groups of GAG molecules. Normally, the protons of these groups exchange with water protons. If this exchange is slow, then saturation of the exchange protons of the hydroxyl groups with a long RF pulse leads to attenuation of the water signal in proportion to the concentration of the hydroxyl groups and thus GAG molecules. A two-frequency irradiation approach has been proposed to discriminate CEST effects from magnetization transfer asymmetry effects [63].

Disadvantages of gagCEST include the need for high field strength, the critical dependence on B_o field homogeneity, and the possible challenges of isolating gagCEST effects from magnetization transfer effects.

4.1.6 Diffusion Tensor Imaging

Diffusion tensor imaging (DTI) represents a potential means to assess both cartilage proteoglycan and collagen content [64–67]. The rationale behind DTI is that proteoglycan and collagen restrict the motion of water molecules in different manners, which are reflected in diffusion parameters. Proteoglycans do not have anisotropy and restrict the displacement of water molecules in all directions, which can be reflected in the parameter mean diffusivity. Collagen fibers demonstrate anisotropy as they are oriented parallel to the articular surface superficially and perpendicular to subchondral bone in the deep cartilage layer. Water diffusion preferentially occurs along the direction of the collagen fibers, and this can be reflected by the parameter fractional anisotropy.

To perform DTI, multiple magnetic field gradients are applied in several directions to acquire a series of diffusion-weighted images. Each set of images provides information about diffusivity in that direction; their combination produces a map of the diffusion tensor. Eigenvalues and eigenvectors describe the magnitude and directions of the diffusion tensor, which can be represented by an ellipsoid. Mean diffusivity is calculated from the average of the three diffusion tensor eigenvalues.

Fig. 6 Axial high-resolution 7 T MR image (3-D FLASH, TR/TE = 26/5.1 ms, 0.234 × 0.234 × 1 mm) of trabecular bone microarchitecture of the knee shows osteophyte formation (*arrow*) and decreased trabecular number and connectivity (*circle*) in the medial femoral condyle of a patient with osteoarthritis
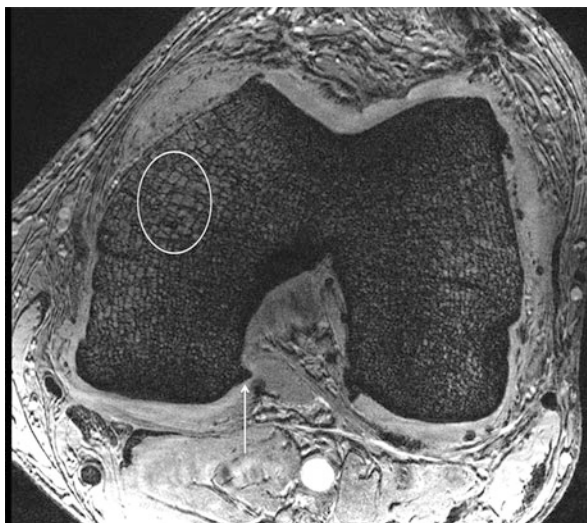
Fractional anisotropy also is calculated from the diffusion tensor eigenvalues, and it reflects the degree of anisotropy within a voxel.

DTI in cartilage demands high SNR. Thus imaging is facilitated by the move to higher field strength. Recently, in a small study involving ten patients with osteoarthritis, Raya et al. [64] showed that ADC and FA calculated from DTI may provide high sensitivity and specificity for the detection of patients with osteoarthritis.

4.2 Subchondral Bone

4.2.1 MRI of Bone Microarchitecture

The use of advanced MRI techniques combined with digital image analysis methods to quantify bone microarchitecture could provide additional insight into the role of subchondral bone in osteoarthritis pathogenesis. MRI of bone microarchitecture is indirect. The MRI scanner detects magnetization from mobile water and fat protons from within the marrow space and bony trabeculae, which contains solid-state protons that are devoid of signal, are outlined by the hyperintense surrounding marrow.

Because trabeculae are small (100–200 μm thick), MRI of bone microarchitecture requires imaging with high-resolution and voxel sizes that are on the order of the size of the trabeculae (Fig. 6). As a result, high-resolution MRI of bone microarchitecture is a high-SNR demand technique that is facilitated by imaging at higher field strength [68]. Different types of MRI pulse sequences have

been used for imaging bone microarchitecture, including fast large angle spin-echo, steady-state free precession, and gradient-recalled echo techniques [6]. Bone microarchitecture analysis can be performed with digital image analysis methods such as fuzzy distance transform [69], digital topological analysis [70], and finite element analysis [71], which allow quantification of morphologic, topologic, and mechanical properties of bone, respectively.

In the presence of knee malalignment, high-resolution MRI studies reveal elevated bone microarchitectural parameters (bone formation) associated with cartilage loss within the diseased compartment [72]. There is also bone loss within the contralateral compartment [72]. An inverse correlation has also been demonstrated between cartilage T1rho and T2 values and apparent trabecular number, separation, and bone volume fraction [73]. Finally, it is possible to detect two-year longitudinal changes in bone microarchitectural parameters in subjects with OA, implying that disease progression can be monitored [74].

5 Conclusions

MRI provides a noninvasive means to detect and monitor pathologic alterations in cartilage and subchondral bone on the macrostructural, microstructural, and even biochemical level in subjects with osteoarthritis. The ability to detect these changes in vivo will allow new insight into the pathogenesis of osteoarthritis and permit early diagnosis of osteoarthritis in subjects who are at risk. This knowledge and capability should ultimately accelerate the discovery and testing of novel therapies to treat osteoarthritis, a disease which represents an enormous socioeconomic and health burden on society.

References

1. Pleis, J.R., Lethbridge-Cejku, M.: Summary health statistics for U.S. adults: National Health Interview Survey, 2005. Vital and health statistics. Series 10, Data from the National Health Survey, 1 (2006)
2. Yelin, E., et al.: Medical care expenditures and earnings losses among persons with arthritis and other rheumatic conditions in 2003, and comparisons with 1997. Arthritis Rheum. **56**, 1397 (2007)
3. Burstein, D., Gray, M., Mosher, T., Dardzinski, B.: Measures of molecular composition and structure in osteoarthritis. Radiol. Clin. North Am. **47**, 675 (2009)
4. Wheaton, A.J., et al.: Proteoglycan loss in human knee cartilage: quantitation with sodium MR imaging – feasibility study. Radiology **231**, 900 (2004)

5. Burr, D.B.: Anatomy and physiology of the mineralized tissues: role in the pathogenesis of osteoarthrosis. Osteoarthritis Cartilage **12**(Suppl A), S20 (2004)
6. Wehrli, F.W.: Structural and functional assessment of trabecular and cortical bone by micro magnetic resonance imaging. J. Magn. Reson. Imaging **25**, 390 (2007)
7. Felson, D.T., Neogi, T.: Osteoarthritis: is it a disease of cartilage or of bone? Arthritis Rheum. **50**, 341 (2004)
8. Resnick, D., Kang, H.S., Pretterklieber, M.L.: Internal Derangements of Joints, 2nd edn. Saunders/Elsevier, Philadelphia (2006). pp. 2 v. (xvi, 2284, liv p.)
9. Eckstein, F., Burstein, D., Link, T.M.: Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis. NMR Biomed. **19**, 822 (2006)
10. Kamibayashi, L., Wyss, U.P., Cooke, T.D., Zee, B.: Trabecular microstructure in the medial condyle of the proximal tibia of patients with knee osteoarthritis. Bone **17**, 27 (1995)
11. Bobinac, D., Spanjol, J., Zoricic, S., Maric, I.: Changes in articular cartilage and subchondral bone histomorphometry in osteoarthritic knee joints in humans. Bone **32**, 284 (2003)
12. Layton, M.W., et al.: Examination of subchondral bone architecture in experimental osteoarthritis by microscopic computed axial tomography. Arthritis Rheum. **31**, 1400 (1988)
13. Chappard, C., et al.: Subchondral bone micro-architectural alterations in osteoarthritis: a synchrotron micro-computed tomography study. Osteoarthritis Cartilage **14**, 215 (2006)
14. Recht, M.P., Piraino, D.W., Paletta, G.A., Schils, J.P., Belhobek, G.H.: Accuracy of fat-suppressed three-dimensional spoiled gradient-echo FLASH MR imaging in the detection of patellofemoral articular cartilage abnormalities. Radiology **198**, 209 (1996)
15. Disler, D.G., McCauley, T.R., Wirth, C.R., Fuchs, M.D.: Detection of knee hyaline cartilage defects using fat-suppressed three-dimensional spoiled gradient-echo MR imaging: comparison with standard MR imaging and correlation with arthroscopy. AJR Am. J. Roentgenol. **165**, 377 (1995)
16. Potter, H.G., Linklater, J.M., Allen, A.A., Hannafin, J.A., Haas, S.B.: Magnetic resonance imaging of articular cartilage in the knee. An evaluation with use of fast-spin-echo imaging. J. Bone Joint Surg. Am. **80**, 1276 (1998)
17. Bredella, M.A., et al.: Accuracy of T2-weighted fast spin-echo MR imaging with fat saturation in detecting cartilage defects in the knee: comparison with arthroscopy in 130 patients. AJR Am. J. Roentgenol. **172**, 1073 (1999)
18. Peterfy, C.G., et al.: Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage **12**, 177 (2004)
19. Frobell, R.B.: Change in cartilage thickness, posttraumatic bone marrow lesions, and joint fluid volumes after acute ACL disruption: a two-year prospective MRI study of sixty-one subjects. J. Bone Joint Surg. Am. **93**, 1096 (2011)
20. Baranyay, F.J., et al.: Association of bone marrow lesions with knee structures and risk factors for bone marrow lesions in the knees of clinically healthy, community-based adults. Semin. Arthritis Rheum. **37**, 112 (2007)
21. Driban, J.B., et al.: Quantitative bone marrow lesion size in osteoarthritic knees correlates with cartilage damage and predicts longitudinal cartilage loss. BMC Musculoskelet. Disord. **12**, 217 (2011)
22. Wluka, A.E., et al.: Bone marrow lesions predict increase in knee cartilage defects and loss of cartilage volume in middle-aged women without knee pain over 2 years. Ann. Rheum. Dis. **68**, 850 (2009)
23. Hunter, D.J., et al.: Increase in bone marrow lesions associated with cartilage loss: a longitudinal magnetic resonance imaging study of knee osteoarthritis. Arthritis Rheum. **54**, 1529 (2006)
24. Torres, L., et al.: The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthritis Cartilage **14**, 1033 (2006)
25. Zhang, Y., et al.: Fluctuation of knee pain and changes in bone marrow lesions, effusions, and synovitis on magnetic resonance imaging. Arthritis Rheum. **63**, 691 (2011)
26. Dardzinski, B.J., Mosher, T.J., Li, S., Van Slyke, M.A., Smith, M.B.: Spatial variation of T2 in human articular cartilage. Radiology **205**, 546 (1997)

27. Mosher, T.J., Dardzinski, B.J.: Cartilage MRI T2 relaxation time mapping: overview and applications. Semin. Musculoskelet. Radiol. **8**, 355 (2004)
28. Nieminen, M.T., et al.: T2 relaxation reveals spatial collagen architecture in articular cartilage: a comparative quantitative MRI and polarized light microscopic study. Magn. Reson. Med. **46**, 487 (2001)
29. Maroudas, A.I.: Balance between swelling pressure and collagen tension in normal and degenerate cartilage. Nature **260**, 808 (1976)
30. Venn, M., Maroudas, A.: Chemical composition and swelling of normal and osteoarthrotic femoral head cartilage. I. Chemical composition. Ann. Rheum. Dis. **36**, 121 (1977)
31. Maroudas, A., Venn, M.: Chemical composition and swelling of normal and osteoarthrotic femoral head cartilage. II. Swelling. Ann. Rheum. Dis. **36**, 399 (1977)
32. Mosher, T.J., Dardzinski, B.J., Smith, M.B.: Human articular cartilage: influence of aging and early symptomatic degeneration on the spatial variation of T2 – preliminary findings at 3 T. Radiology **214**, 259 (2000)
33. Li, X., et al.: In vivo T(1rho) and T(2) mapping of articular cartilage in osteoarthritis of the knee using 3 T MRI. Osteoarthritis Cartilage **15**, 789 (2007)
34. Peterfy, C.G., Schneider, E., Nevitt, M.: The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage **16**, 1433 (2008)
35. Li, X., et al.: Spatial distribution and relationship of T1rho and T2 relaxation times in knee cartilage with osteoarthritis. Magn. Reson. Med. **61**, 1310 (2009)
36. Xia, Y., Moody, J.B., Alhadlaq, H.: Orientational dependence of T2 relaxation in articular cartilage: a microscopic MRI (microMRI) study. Magn. Reson. Med. **48**, 460 (2002)
37. Nishii, T., Kuroda, K., Matsuoka, Y., Sahara, T., Yoshikawa, H.: Change in knee cartilage T2 in response to mechanical loading. J. Magn. Reson. Imaging **28**, 175 (2008)
38. Shiomi, T., et al.: Loading and knee alignment have significant influence on cartilage MRI T2 in porcine knee joints. Osteoarthritis Cartilage **18**, 902 (2010)
39. Prasad, A.P., Nardo, L., Schooler, J., Joseph, G., Link, T.M.: T(1rho) and T(2) relaxation times predict progression of knee osteoarthritis. Osteoarthritis Cartilage **21**(1), 69–76 (2013)
40. Joseph, G.B., et al.: Baseline mean and heterogeneity of MR cartilage T2 are associated with morphologic degeneration of cartilage, meniscus, and bone marrow over 3 years – data from the osteoarthritis initiative. Osteoarthritis Cartilage **20**, 727 (2012)
41. Borthakur, A., et al.: Sodium and T1rho MRI for molecular and diagnostic imaging of articular cartilage. NMR Biomed. **19**, 781 (2006)
42. Regatte, R.R., Akella, S.V., Lonner, J.H., Kneeland, J.B., Reddy, R.: T1rho relaxation mapping in human osteoarthritis (OA) cartilage: comparison of T1rho with T2. J. Magn. Reson. Imaging **23**, 547 (2006)
43. Akella, S.V., et al.: Proteoglycan-induced changes in T1rho-relaxation of articular cartilage at 4T. Magn. Reson. Med. **46**, 419 (2001)
44. Menezes, N.M., Gray, M.L., Hartke, J.R., Burstein, D.: T2 and T1rho MRI in articular cartilage systems. Magn. Reson. Med. **51**, 503 (2004)
45. Keenan, K.E., et al.: Prediction of glycosaminoglycan content in human cartilage by age, T1rho and T2 MRI. Osteoarthritis Cartilage **19**, 171 (2011)
46. Witschey, W.R., et al.: T1rho MRI quantification of arthroscopically confirmed cartilage degeneration. Magn. Reson. Med. **63**, 1376 (2010)
47. Lattanzi, R., et al.: A B1-insensitive high resolution 2D T1 mapping pulse sequence for dGEMRIC of the HIP at 3 Tesla. Magn. Reson. Med. **66**, 348 (2011)
48. Siversson, C., et al.: Effects of B1 inhomogeneity correction for three-dimensional variable flip angle T1 measurements in hip dGEMRIC at 3 T and 1.5 T. Magn. Reson. Med. **67**, 1776 (2012)
49. Burstein, D., Gray, M.L., Hartman, A.L., Gipe, R., Foy, B.D.: Diffusion of small solutes in cartilage as measured by nuclear magnetic resonance (NMR) spectroscopy and imaging. J. Orthop. Res. **11**, 465 (1993)

50. Trattnig, S., et al.: MRI visualization of proteoglycan depletion in articular cartilage via intravenous administration of Gd-DTPA. Magn. Reson. Imaging **17**, 577 (1999)
51. Kim, Y.J., Jaramillo, D., Millis, M.B., Gray, M.L., Burstein, D.: Assessment of early osteoarthritis in hip dysplasia with delayed gadolinium-enhanced magnetic resonance imaging of cartilage. J. Bone Joint Surg. Am. **85-A**, 1987 (2003)
52. Owman, H., Tiderius, C.J., Neuman, P., Nyquist, F., Dahlberg, L.E.: Association between findings on delayed gadolinium-enhanced magnetic resonance imaging of cartilage and future knee osteoarthritis. Arthritis Rheum. **58**, 1727 (2008)
53. Kim, S.D., Jessel, R., Zurakowski, D., Millis, M.B., Kim, Y.J.: Anterior delayed gadolinium-enhanced MRI of cartilage values predict joint failure after periacetabular osteotomy. Clin. Orthop. Relat. Res. **470**(12), 3332–3341 (2012)
54. Lattanzi, R., et al.: A new method to analyze dGEMRIC measurements in femoroacetabular impingement: preliminary validation against arthroscopic findings. Osteoarthritis Cartilage **20**, 1127 (2012)
55. Regatte, R.R., Schweitzer, M.E.: Novel contrast mechanisms at 3 Tesla and 7 Tesla. Semin. Musculoskelet. Radiol. **12**, 266 (2008)
56. Borthakur, A., et al.: Sensitivity of MRI to proteoglycan depletion in cartilage: comparison of sodium and proton MRI. Osteoarthritis Cartilage **8**, 288 (2000)
57. Shapiro, E.M., et al.: Sodium visibility and quantitation in intact bovine articular cartilage using high field (23)Na MRI and MRS. J. Magn. Reson. **142**, 24 (2000)
58. Wang, L., et al.: Rapid isotropic 3D-sodium MRI of the knee joint in vivo at 7T. J. Magn. Reson. Imaging **30**, 606 (2009)
59. Madelin, G., Lee, J.S., Inati, S., Jerschow, A., Regatte, R.R.: Sodium inversion recovery MRI of the knee joint in vivo at 7T. J. Magn. Reson. **207**, 42 (2010)
60. Chang, G., et al.: Improved assessment of cartilage repair tissue using fluid-suppressed (2)(3)Na inversion recovery MRI at 7 Tesla: preliminary results. Eur. Radiol. **22**, 1341 (2012)
61. Ling, W., Regatte, R.R., Navon, G., Jerschow, A.: Assessment of glycosaminoglycan concentration in vivo by chemical exchange-dependent saturation transfer (gagCEST). Proc. Natl. Acad. Sci. U.S.A. **105**, 2266 (2008)
62. Singh, A., et al.: Chemical exchange saturation transfer magnetic resonance imaging of human knee cartilage at 3 T and 7 T. Magn. Reson. Med. **68**, 588 (2012)
63. Lee, J.S., Regatte, R.R., Jerschow, A.: Isolating chemical exchange saturation transfer contrast from magnetization transfer asymmetry under two-frequency RF irradiation. J. Magn. Reson. **215**, 56 (2012)
64. Raya, J.G., et al.: Articular cartilage: in vivo diffusion-tensor imaging. Radiology **262**(2), 550–559 (2012)
65. Filidoro, L., et al.: High-resolution diffusion tensor imaging of human patellar cartilage: feasibility and preliminary findings. Magn. Reson. Med. **53**, 993 (2005)
66. Meder, R., de Visser, S.K., Bowden, J.C., Bostrom, T., Pope, J.M.: Diffusion tensor imaging of articular cartilage as a measure of tissue microstructure. Osteoarthritis Cartilage **14**, 875 (2006)
67. Deng, X., Farley, M., Nieminen, M.T., Gray, M., Burstein, D.: Diffusion tensor imaging of native and degenerated human articular cartilage. Magn. Reson. Imaging **25**, 168 (2007)
68. Banerjee, S., et al.: Rapid in vivo musculoskeletal MR with parallel imaging at 7T. Magn. Reson. Med. **59**, 655 (2008)
69. Saha, P.K., Wehrli, F.W., Gomberg, B.R.: Fuzzy distance transform: theory, algorithms, and applications. Comput. Vis. Image Underst. **86**, 171 (2002)
70. Saha, P.K., Gomberg, B.R., Wehrli, F.W.: Three-dimensional digital topological characterization of cancellous bone architecture. Int. J. Imaging Syst. Technol. **11**, 81 (2000)
71. Rajapakse, C.S., et al.: Computational biomechanics of the distal tibia from high-resolution MR and micro-CT images. Bone **47**, 556 (2010)
72. Lindsey, C.T., et al.: Magnetic resonance evaluation of the interrelationship between articular cartilage and trabecular bone of the osteoarthritic knee. Osteoarthritis Cartilage **12**, 86 (2004)

73. Bolbos, R.I., et al.: Relationship between trabecular bone structure and articular cartilage morphology and relaxation times in early OA of the knee joint using parallel MRI at 3 T. Osteoarthritis Cartilage **16**, 1150 (2008)
74. Blumenkrantz, G., et al.: A pilot, two-year longitudinal study of the interrelationship between trabecular bone and articular cartilage in the osteoarthritic knee. Osteoarthritis Cartilage **12**, 997 (2004)

Computer Vision Based Hairline Mandibular Fracture Detection from Computed Tomography Images

Ananda S. Chowdhury, Anindita Mukherjee, Suchendra M. Bhandarkar, and Jack C. Yu

Abstract This chapter addresses the problem of detection of hairline mandibular fractures from a sequence of computed tomography (CT) images. It has been observed that such a fracture can be easily overlooked during manual detection due to the absence of sharp surface and contour discontinuities and the presence of intensity inhomogeneity in the CT images. In this work, the 2D CT image slices of a mandible with hairline fractures are first identified from the input sequence of a fractured craniofacial skeleton. Two intensity-based image retrieval schemes with different measures of similarity, namely the Jaccard index and the Kolmogorov–Smirnov distance, are applied for that purpose. In the second part, we detect a hairline fracture in the previously identified subset of images using the maximum flow-minimum cut algorithm. Since a hairline fracture is essentially a discontinuity in the bone contour, we model it as a minimum cut in an appropriately weighted flow network constructed using the geometry of the human mandible. The Ford–Fulkerson algorithm with Edmonds–Karp refinement is employed to obtain a minimum cut. Experimental results demonstrate the effectiveness of the proposed method.

A.S. Chowdhury (✉)
Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata 700032, India
e-mail: aschowdhury@etce.jdvu.ac.in

A. Mukherjee
Department of Computer Science, Dream Institute of Technology, Kolkata 700104, India

S.M. Bhandarkar
Department of Computer Science, University of Georgia, Augusta, GA 30912, USA

J.C. Yu
Department of Plastic Surgery, Georgia Health Sciences University, Athens, GA 30602, USA

193

Fig. 1 CT scan of an
unbroken mandible
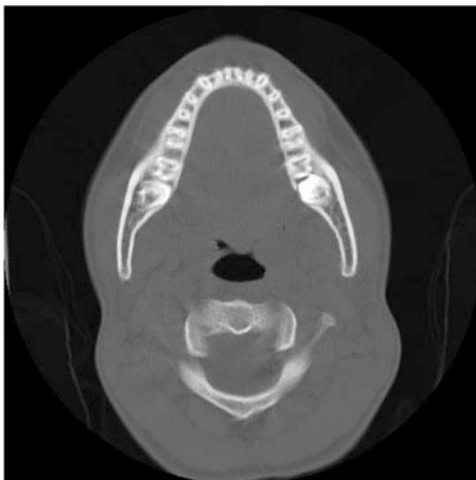
1   Introduction

Craniofacial fractures are very frequently encountered in modern society, the
principal reasons being gunshot wounds, motor vehicle accidents, and sports related
injuries [1]. Ogundare et al. [2] in their research illustrated that the fracture
patterns sometimes imply a single fracture, and, in some other cases, there can
be a combination of single fractures. The craniofacial skeleton consists of high
stress-bearing buttresses and low stress-bearing curve planes. When external loads
are applied, the craniofacial skeleton undergoes strain. When the strain exceeds the
ultimate strain limit, failure occurs [3]. The resulting loss of spatial continuity in the
craniofacial skeleton is clinically termed as fracture which causes pain, disfigure-
ment, and functional impairment due to the disruption of force transmission. In some
situations, the broken fragments may suffer negligible relative displacement. They
are termed as hairline/minor fractures. For the purpose of illustration, we first show
a 2D slice of an intact human mandible in Fig. 1. A typical hairline/minor fracture
is presented in Fig. 2.

The input to our problem is a sequence of computed tomography (CT) images
with scans of human craniofacial skeleton. A typical sequence consists of many
slices out of which only a handful is observed to contain the fractured mandible.
Figure 3 shows a typical input sequence.

Note that not all four slices in Fig. 3 contain the fractures. A careful scrutiny
reveals that only the two middle slices have these fractures. So, we decide to first
retrieve only the fracture containing slices from the entire input sequence. This
is achieved using two intensity-based image retrieval approaches [4] with Jaccard
index and Kolmogorov–Smirnov distance as measures of similarity. We next apply
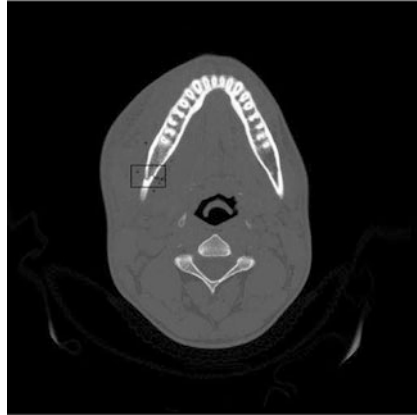
Fig. 2 CT scan of a mandible with a minor fracture (within the *black rectangle*)
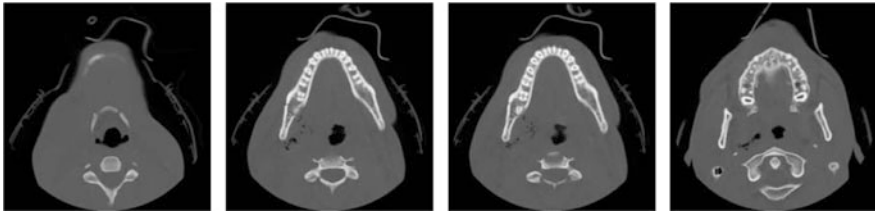
Fig. 3 An input sequence with four different slices

the max-flow min-cut algorithm to only the above fracture containing slices. A flow network with appropriate weights is constructed following the geometry of the mandible. A fracture is modeled as a minimum cut in such a network. The Ford–Fulkerson [5] algorithm is applied for the detection of a minimum cut. To expedite the Ford–Fulkerson algorithm, the augmenting paths are detected using breadth-first search, as indicated by Edmonds–Karp [6]. The overall goal is to develop an enabling technology that leverages the advances in computer vision to detect hairline mandibular fractures. The work presented here can be of good help for practicing radiologists and surgeons.

The rest of this chapter is organized in the following manner. In Sect. 2, we provide the state-of-the-art and highlight our contribution. Section 3 is devoted to the intensity-based image retrieval methods. In Sect. 4, we describe the max-flow min-cut based fracture detection. Experimental results are presented in Sect. 5. The chapter is concluded in Sect. 6 with an outline of future directions of research.

2   Literature Review

While there exists recent work published in the literature dealing with simulation of mandibular fractures [7], and simulation of dental implants [8]; only a little has been reported about computer-aided detection of craniofacial fractures. The relative paucity of existing literature on automated mandibular surgery has made us interested in carrying out this important research work.

Detection of well-displaced mandibular fractures, wherein the broken fragments have undergone noticeable relative displacement, using a Bayesian inference approach has been reported in a previous work by Chowdhury et al. [9]. A Markov random field (MRF)-based approach for detection of mandibular hairline fractures has been also presented by Chowdhury et al. [10]. However, the MRF-based scheme is computationally intensive. Furthermore, the MRF-based scheme is based on an underlying assumption that the bilateral symmetry of a mandible is preserved in the case of a hairline fracture, which may not necessarily hold in some cases.

In this paper, the 2D CT slices containing a fractured mandible are first identified within the entire CT image sequence of a fractured craniofacial skeleton using the Jaccard index and the Kolmogorov–Smirnov (KS) distance measures. Within each identified 2D CT slice, a hairline fracture is modeled as a *cut* in the flow of image intensities along the bone contours. The flow is conceived between two anatomical landmark points on the human mandible, called condyles [11]. The geometry of the mandibular contours is used to construct the flow network. In graph theory, this is equivalent to finding a minimum cut in the flow graph between the *source* and the *sink* vertices. It is important to note that unlike the MRF-based scheme in [10], the proposed method is computationally fast and does not rely on the assumption of preservation of bilateral symmetry for mandibles with hairline fractures. A part of the work described here has been published in [12].

3   Retrieval of Fracture Containing Slices

Content-based image retrieval (CBIR) consists of retrieving the most visually similar images to a given query/reference image from a database of images. The ever-increasing number of digital images in the medical domain has amplified the need for automated search and retrieval tools. For an excellent survey in this subject, see the work by Müller et al. [4]. Rahman et al. [13] proposed a CBIR framework for diverse collection of medical images of different imaging modalities and anatomic regions with different orientations. In another paper, Selvarani and Annadurai [14] discuss general Fourier descriptor (GFD) for shape representation for the purpose of medical image retrieval. Sedghi et al. [15] discuss the relevance criteria for searching of various medical images.

In the present problem, an intact mandible serves as the query/reference image and the candidate slices in any input sequence constitute the database of images. We propose two methods which are based on a general assumption that similar-looking

images should exhibit similar intensity patterns. The areas of high intensity and low intensity in an image can constitute important features because they often mark relevant image objects. As mandible is a bony structure, it typically exhibits considerably high intensity values compared to other components like soft tissues. We exploit these contrasting intensity patterns of the CT scans for retrieving slices containing the mandibles.

3.1 Image Pre-processing

We initially apply the following three image-processing tasks prior to intensity-based retrieval [16]:

1. Entropy thresholding
2. Median filtering
3. Histogram construction

3.1.1 Entropy Thresholding

In the CT data, the selection of the appropriate threshold is not obvious since the CT images typically contain objects or artifacts of different intensities (varying Hounsfield unit values). For example, a fractured mandibular fragment could contain cavities, dental fillings, crowns, and other dental prostheses. Thus, entropy-based thresholding [17] is chosen over simple thresholding. In the case of entropy-based thresholding, the threshold value is determined via maximization of the interclass entropy computed from the grayscale histogram of an image. For an image, the entropy S_c for each graylevel class c (consisting of several gray level values) can be computed using the grayscale histogram in the following manner:

$$S_c = -\sum_{k \epsilon G_c} p(k)\log_2\left(p(k)\right) \tag{1}$$

where $p(k)$ is the probability of a pixel having a grayscale value k and G_c is the set of graylevels for class c. In the context of binarization, the graylevel threshold T is chosen such that the total entropy $S = S_1 + S_2$ is maximized.

In Fig. 4, we show the image of the intact mandible. Result of entropy thresholding on that image is shown in Fig. 5.

3.1.2 Median Filtering

Median filter is a nonlinear digital filtering technique, often used to remove noise from images or other signals. Note that the filter could have been applied before the

Fig. 4 CT scan of an
unbroken mandible
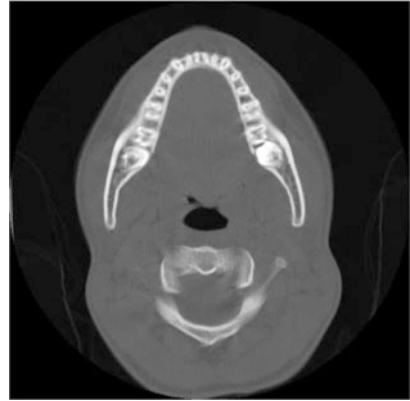
Fig. 5 Result of entropy
thresholding in Fig. 4

entropy thresholding also. Here, we have applied the same after the thresholding is
performed.

In Fig. 6, we show the result of median filtering on the image in Fig. 5.

3.1.3   Histogram Construction

In an image-processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. The histogram is a graph that shows the number of pixels in an image at each different intensity value found in that image (Fig. 7).

3.2   Set-Based Approach and Choice of Similarity Measure

In this section, the set-based representation of an image is discussed along with the choice of *Normalized Jaccard Index* (*NJI*) as the measure of similarity. An image

Fig. 6 Result of median filtering in Fig. 5
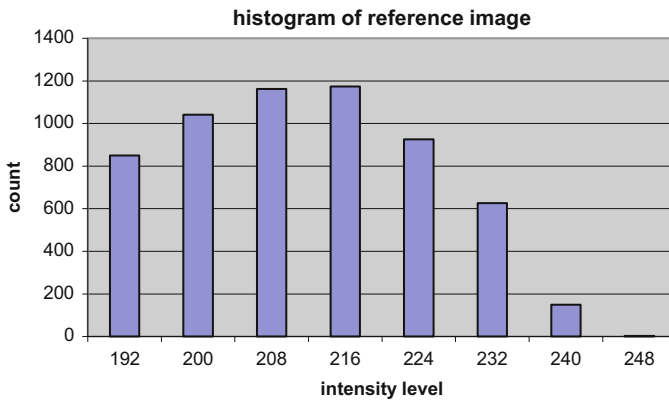
Fig. 7 Histogram of the image shown in Fig. 4. Only top 25 % intensities, i.e. values in the range 192–255 are considered

can be represented as a mathematical function over some finite domain. In general, we can represent that function f as $f:(x, y) \to z$ where x, y, $z \in R$ and R is the set of real numbers. The ordered pair (x, y) for an image represents spatial coordinates of a pixel with respect to some reference frame and z is the intensity value at that pixel location. For an n-bit grayscale image, $R = [0, 2^n - 1]$ and for a binary image $R = [0, 1]$. Now, we define union and intersection between two binary images $f(x, y)$ and $g(x, y)$.

Definition 1 If at a certain pixel location (x, y), either $f(x, y)$ is 1 or $g(x, y)$ is 1, then the **union** between the two sets $f(x, y) \cup g(x, y) = 1$, else $f(x, y) \cup g(x, y) = 0$.

Definition 2 If at a certain pixel location (x, y), both $f(x, y)$ and $g(x, y)$ are 1, then the **intersection** between the two sets $f(x, y) \cap g(x, y) = 1$, else $f(x, y) \cap g(x, y) = 0$.

Definition 3 **Jaccard index** is defined as follows:

$$J = \frac{|f(x, y) \cap g(x, y)|}{|f(x, y) \cup g(x, y)|} \qquad (2)$$

In Eq. (2), the symbol "| |" denotes cardinality of a set. Note that when $f(x, y)$ and $g(x, y)$ are two disjoint sets, which in this context means there is no overlap between the two binary images, $J = 0$. In contrast, when $f(x, y) = g(x, y)$, which in this context means two binary images completely overlap, $J = 1$. For partial overlap, J will assume appropriate intermediate values between 0 and 1. Among other related measures [18], both *Dice coefficient* and *Cosine coefficient* give more weight to the intersection part. We decide to give equal emphasis to intersection and union terms and hence choose *Jaccard index*.

3.3 Method-I: Image Retrieval Using Normalized Jaccard Index

Now, we describe our first method (method-I) for the image retrieval. Here, we compute the overlap between the reference image (i.e., unbroken mandible) and the candidate images (2D slices in any input sequence). The rationale behind the method is that two similar images exhibit a higher value of the overlap compared to two dissimilar images.

Figure 8 shows the flowchart for the first retrieval method. While computing the *Jaccard index*, we do not consider the entire images of dimensions $W \times H$. Here, W denotes the width of the image and H denotes the height of the image. We use a smaller area $w \times h$ instead, where $w < W$ and $h < H$. This selection of smaller area is based on our domain knowledge. After studying large number of slices, we came to conclusion that a mandible typically appears within a centrally located smaller region in the CT scans. Selection of a smaller area of interest increases the speed of *Jaccard index* computation (Fig. 9).

We normalize the value of the *Jaccard index* by dividing it with the maximum value for a given sequence. Normalization refers to the division of multiple sets of data by a common variable in order to negate that variable's effect on the data, thus allowing underlying characteristics of the datasets to be compared. We finally threshold the *NJI* values to retrieve appropriate slices. Slices possessing values above the chosen threshold are retrieved.

3.4 Basic Statistical Concepts and Choice of Distance Measure

In this section, we discuss some basic statistical concepts pertinent to this work. We next define the Kolmogorov–Smirnov distance, used to denote similarity between two images.

Fig. 8 Flowchart for method-I
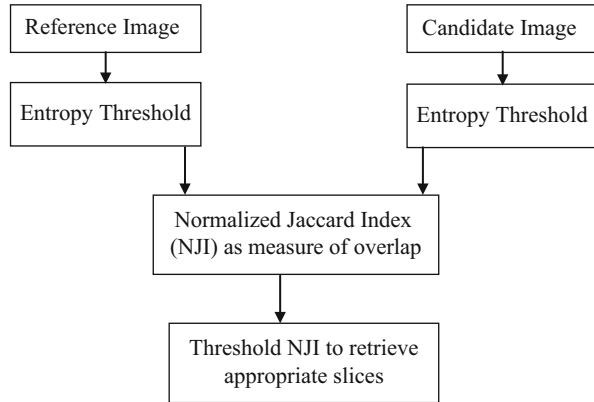
Fig. 9 Image in Fig. 4 with a selected area
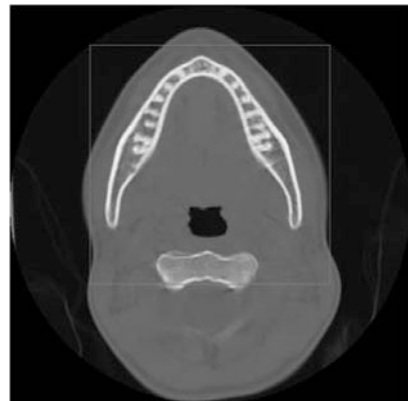
Definition 4 If X is a discrete random variable, the function given by $f(x) = \Pr(X = x)$ for each x within the range of X is called the **probability distribution function** (pdf) of X.

Definition 5 The function which denotes the probability that a discrete random variable X takes on a value less than or equal to x is called the **cumulative distribution function** (cdf) of X.

Definition 6 Kolmogorov–Smirnov (KS) distance is defined as the largest difference (supremum) between two cdfs. Let X^c and Y^c, respectively, denote the cdfs of the distributions X and Y. Then, we can write [19]:

$$KS\,(X^c, Y^c) = \sup_i |X^c(i) - Y^c(i)| \tag{3}$$

We have essentially modeled image histograms as pdfs. So, Kullback–Leibler (KL) divergence or Jenson–Shannon (JS) divergence can be also employed as measures of similarity. However, it is often observed that cdfs are less sensitive
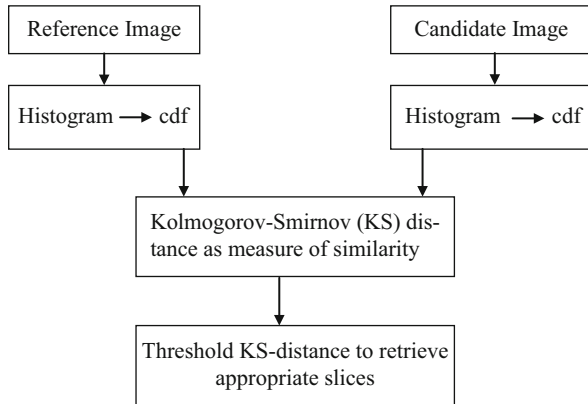
Fig. 10 Flowchart for method-II

to noise as compared to their pdf counterparts. So, we decide to use the cdf-based KS distance as a measure of similarity. It is important to note that the KS distance computation is based on nonparametric statistics, where we do not assume any particular form of a distribution [20].

3.5 Method-II: Image Retrieval Using Kolmogorov–Smirnov Distance

We now describe our second method (method-II) for the image retrieval. Here, we compute the similarity between the reference image (i.e., unbroken mandible) and the candidate images (2D slices in any input sequence) using *Kolmogorov–Smirnov* (KS) distance. The rationale behind the method is that two similar images exhibit a lower value of the KS distance as compared to two dissimilar images.

Figure 10 presents the flowchart for the second retrieval method. In this method, we first construct the histograms of the reference image and a candidate image. Since, a mandible has high intensity; we choose top 25 % of the intensity values, i.e. 192–255 for the construction of the histogram. We further consider the pixels only within a central region of interest and not in the entire image. We divide the above range of intensities to eight equal intervals of width eight pixels. Treating the above histograms as two *pdfs*, we construct the corresponding *cdfs*. Note that a very small interval for the histogram will make the *cdf* susceptible to noise. On the contrary, a very large interval will mask detailed intensity information. We choose an intermediate interval to address both these issues. This makes the *cdfs* less sensitive to noise without sacrificing detailed intensity information. From a pair of *cdfs*, *Kolmogorov–Smirnov* (*KS*) distance is computed. We finally threshold the *KS distances* to retrieve appropriate slices. Slices possessing values below the chosen threshold are retrieved.
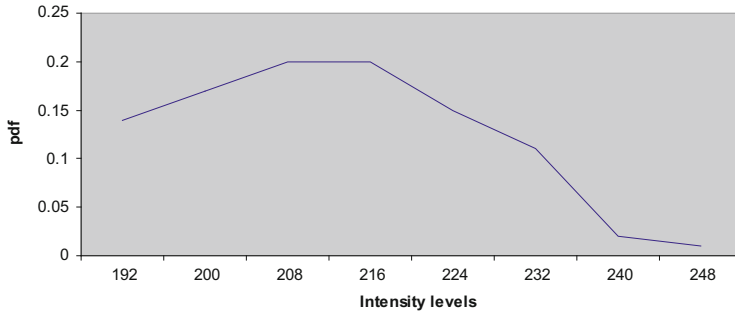
Fig. 11 Pdf of the histogram shown in Fig. 7
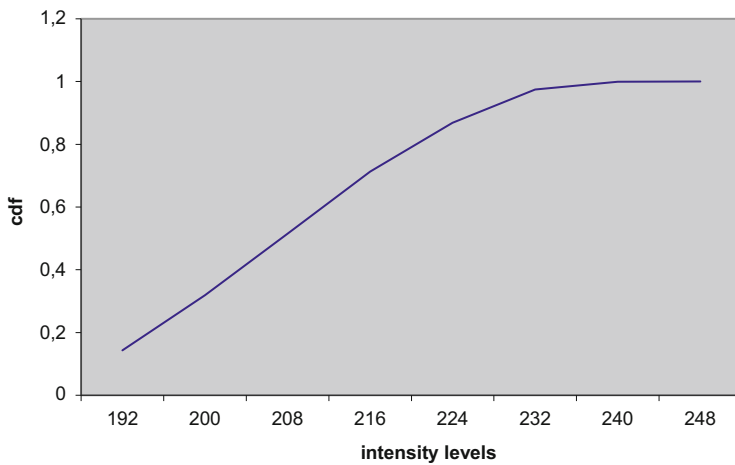
Fig. 12 Cdf of the pdf shown in Fig. 11

For the purpose of illustration, we now show the pdf (Figs. 11) and the corresponding cdf (Fig. 12) obtained from the histogram shown in Fig. 7. As mentioned above, we use an intensity range [192, 255] with eight equal intervals of eight pixels width.

4 Fracture Detection Using Max-Flow Min-Cut

Here, we first provide some basic concepts regarding flow networks. We next discuss the construction of the flow network for a 2D CT slice with suitable choices of vertices, edges, and capacity functions to be used as edge weights. We finally justify that the fracture detection problem can be modeled as identification of a minimum cut in the above flow network.

4.1 Network Flow Basics

We briefly state some useful definitions and an important theorem for the understanding of flow networks following [6].

Definition 7 A flow network $G = G(V, E)$ is a directed graph in which each edge $(u, v) \epsilon E$ has a nonnegative capacity $c(u, v) \geq 0$. Two vertices in the flow network are distinguished as a source vertex s and a sink vertex t.

Definition 8 A cut (S, T) of a flow network $G = (V, E)$ is a partition of V into S and $T = V - S$ such that $s \epsilon S$ and $t \epsilon T$.

Definition 9 The capacity of a cut (S, T) is the sum of the capacities of the edges from S to T.

We next state the all important **Max-Flow Min-Cut theorem** without proof. For proof, please see [5, 6].

Theorem For any graph the maximum flow value from source vertex s to sink vertex t is equal to the minimal cut capacity of all cuts separating s and t.

4.2 Construction of Our Flow Network

A typical hairline fracture appears along the bone contours. So, we choose the set of boundary pixels P on the two mandibular contours (i.e., the inner and the outer contour) as the vertices of the proposed flow network $G = G(V, E)$. The two condyles, which are anatomical landmark points at the two terminals of the jaw, serve as the natural choices for the source vertex s and the sink vertex t. Thus, we can write: $V = P \cup \{s, t\}$. Since the boundary pixels for a mandible lie essentially on an arc, we construct *tangential* (T) and *normal* (N) edges in the proposed flow network representation. For each boundary pixel on a specific mandibular contour, we create edge links with the immediately forward and backward neighboring pixels. These constitute the T edges. On the other hand, the N edges are established between any two normal (or near-normal) boundary points across the two contours q and r. In addition, the first boundary pixel of both the contours q_f and r_f is attached to the source s and the last boundary pixel of both the contours q_l and r_l is attached to the sink t. So, we can write: $E = T \cup N \cup \{(q_f, s),(r_f, s),(q_l, t),(r_l, t)\}$. A typical flow network is shown below (Fig. 13):

The rationale of having tangential and normal neighbors is guided by the geometry of the mandible as well as by the fracture pattern. A typical hairline fracture appears along and also across the two mandibular contours. The tangential edges are appropriate for capturing a fracture along the contour and the normal edges are suitable for identifying a fracture across the contour. Let j and k be any two consecutive points along a mandibular contour with coordinates (x_j, y_j) and (x_k, y_k), respectively. Then the equation of the line that is normal to the contour at point j is given by:

Fig. 13 A flow network with source s, sink t, tangential edges T and normal edges N
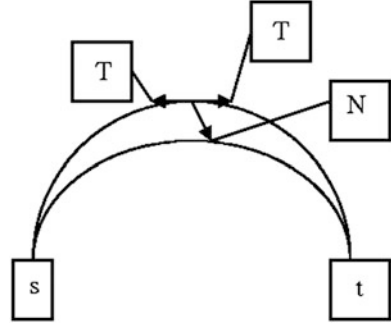
$$\left(x_k - x_j\right) x + \left(y_k - y_j\right) y + \left(x_j \left(x_j - x_k\right) + y_j \left(y_j - y_k\right)\right) = 0 \qquad (4)$$

To be a normal neighbor to the point j, a point on the other contour should ideally satisfy Eq. (4). However, it is not always possible to find the exact normal neighbor (primarily due to sampling error). So, we compute the distance d_{jm} of a set of competing candidate points m having coordinates (m_x, m_y) and choose the one which yields the minimum value of d_{jm}. From basic coordinate geometry, we obtain:

$$d_{jm} = \frac{Am_x + Bm_y + C}{\sqrt{A^2 + B^2}} \qquad (5)$$

where $A = (x_k - x_j), B = (y_k - y_j), C = (x_j(x_j - x_k) + y_j(y_j - y_k))$.

We choose a simple capacity function, with intensity and distance as its two parameters, as an edge weight between a pair of pixels. Let I_j and I_k be the intensities of two pixels j and k and let d_{jk} be the Euclidean distance between them. Then the capacity function c_{jk} is given by:

$$c_{jk} = \frac{I_j I_k}{d_{jk}} \qquad (6)$$

We now provide an intuitive justification for the correctness of the proposed modeling of fracture detection as a minimum cut in an appropriately constructed flow network. From the discussion in the previous section, it is evident that every cut C in the flow network G satisfies following two properties [21]:

1. C groups the vertices of G into two disjoint sets
2. One set will contain the source vertex s and the other set will contain the sink vertex t

Now we justify the following claim with the max-flow min-cut theorem stated earlier.

Claim: A minimum cut C^* correctly identifies a fracture in the proposed 2D flow network G.

Justification: The justification is based on the max-flow min-cut theorem and the construction of the flow network with the capacity function given by Eq. (6). We seek to determine the maximum flow between s and t. Applying the above theorem, we obtain the minimum cut C^*. The minimum cut will consist of the cut edges in our flow network G. Basically, the cut edges are edges with comparatively low capacity values. From Eq. (6), it is evident that the low capacity edges are edges with relatively lower pixel intensity values and relatively higher distance values. These are exactly the characteristics of an edge in the vicinity of a fracture site. Note that since the broken fragments in a hairline fracture are not visibly out of alignment, a nonzero flow exists between them albeit with a lower value. The drop in the flow value is due to the presence of the fracture, which is a flow bottleneck in the constructed flow network. Thus, identification of a minimum cut C^* corresponds to detection of a fracture in the 2D flow network G.□

The computation of the augmenting path is done using a *breadth-first-search*, as outlined in the *Edmonds–Karp* enhancement of the classical *Ford–Fulkerson* algorithm [6]. This algorithm has a worst-case time complexity $O(|V||E|^2)$, where $|V|$ denotes the cardinality of the vertex set V and $|E|$ denotes the cardinality of the edge set E in the flow network G.

5 Experimental Results

In this section, we present the experimental results with a detailed analysis. For implementation, we have used *ImageJ* [22], a public domain image-processing software, developed at National Institutes of Health (NIH), Bethesda, MD, USA. Five datasets with mandibles showing single hairline fracture, and one dataset with intact mandible are obtained from the Georgia Health Sciences University. The intact mandible is treated as the reference image. Each image in a dataset is of eight bit and has a dimension of 512×512 pixels. The resolution varies from 2.5 to 3.5 pixels/mm in various datasets. We first present the results for retrieval based on thresholding the NJI (Figs. 14 and 15) and KS distance (Figs. 16 and 17). We next show the results of actual detection of hairline fractures in some of the retrieved slices (Figs. 18 and 19).

For the purpose of illustration, four retrieved slices are shown for each dataset. For the NJI-based approach, retrieved images are those, whose binarized versions exhibit high similarity to the binarized version of the reference image. Similarity is computed from area of overlap measured using NJI. We have experimentally chosen the threshold value as 0.5. Slices with threshold values greater than 0.5 are retrieved. The threshold value, once chosen, is kept constant for all the datasets (Table 1). We show the mean and standard deviation of precision and recall values for this method in Table 2.
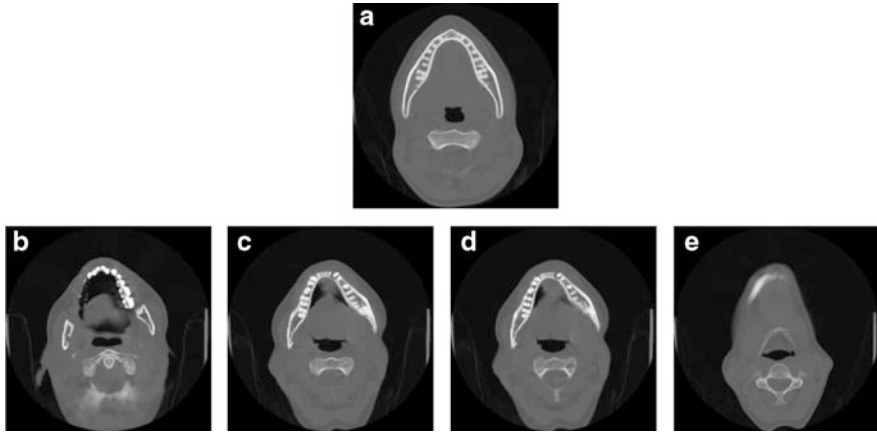
Fig. 14 Reference slice (**a**) and sample slices (**b**–**e**) from one dataset. Method-I retrieves (**c**) and (**d**)
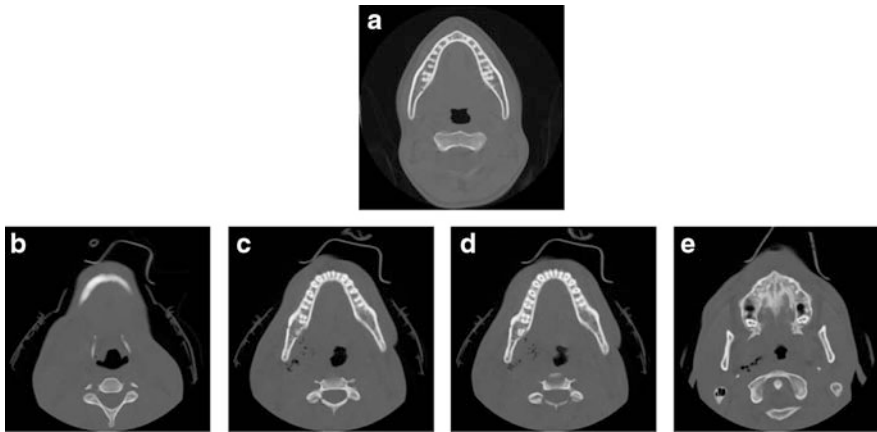
Fig. 15 Reference slice (**a**) and sample slices (**b**–**e**) from another dataset. Method-I retrieves (**c**) and (**d**)

For the KS distance-based approach also, the retrieved images are those, which exhibit high similarity with the reference image. However, similarity is computed in this case by estimating the distance between cdfs of the histograms with KS distance. We have experimentally chosen the threshold value to be 0.25. Slices with a threshold value smaller than 0.25 are retrieved. The threshold value, once chosen, is kept constant for all the datasets (Table 3). We show the mean and standard deviation of precision and recall values for this method in Table 4.

From Table 2, we see that method-I yields an average precision of 66.8 % and an average recall of 84.1 %. In contrast, Table 4 demonstrates that method-II has an average precision of 75.9 % and an average recall of 78.1 %. So, method-I
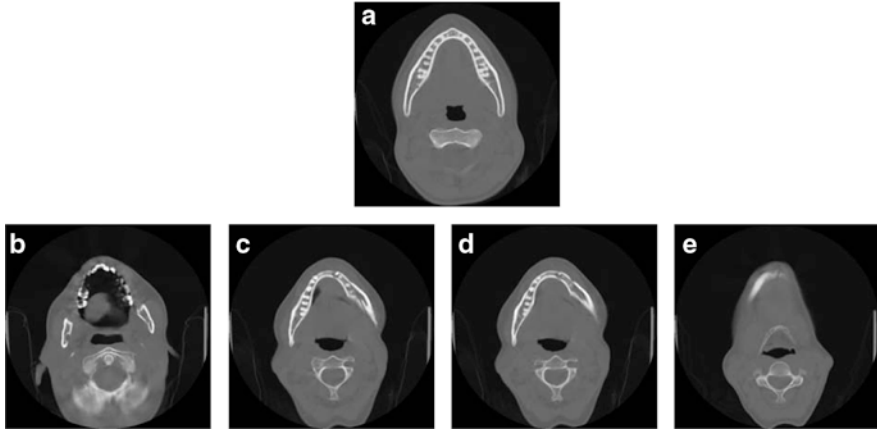
Fig. 16 Reference slice (**a**) and sample slices (**b–e**) from one dataset. Method-II retrieves (**c**) and (**d**)
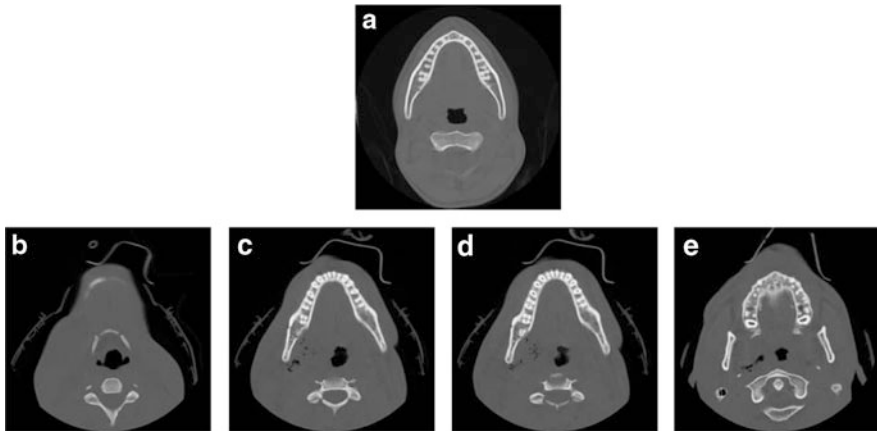
Fig. 17 Reference slice (**a**) and sample slices (**b–e**) from another dataset. Method-II retrieves (**c**) and (**d**)

has a lower value of precision but a higher value of recall compared to method-II. Method-I is less strict compared to method-II. So, in general, method-I retrieves a large number of slices. When more slices are retrieved, the chance of extracting a larger subset of the ground-truth, i.e., actual number of correct slices is higher. Thus, retrieval of more slices leads to an increase in the recall value. However, retrieval of large number of slices may contain more incorrect slices, which causes a decrease in the precision value. Method-II is more conservative than method-I, i.e., it retrieves lesser number of slices. So, chance of extracting a large number of incorrect slices is less, which in turn increases the precision value. However, extraction of lesser number of slices reduces the chance of retrieving a higher subset of ground-truth.
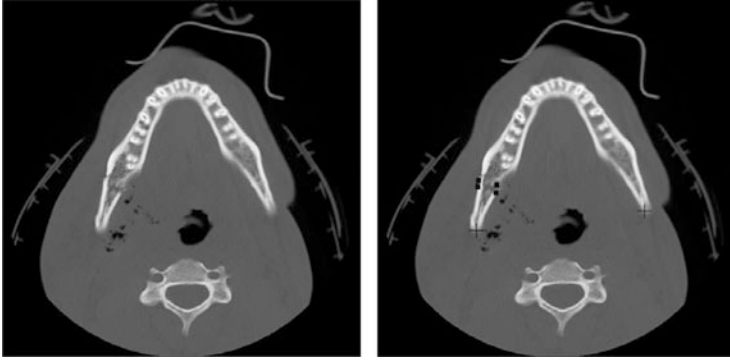
Fig. 18 Two-dimensional slice of a fractured mandible in the *left*. Fracture detection (*black squares*) with source and sink identification (*black crosses*) in the *right*
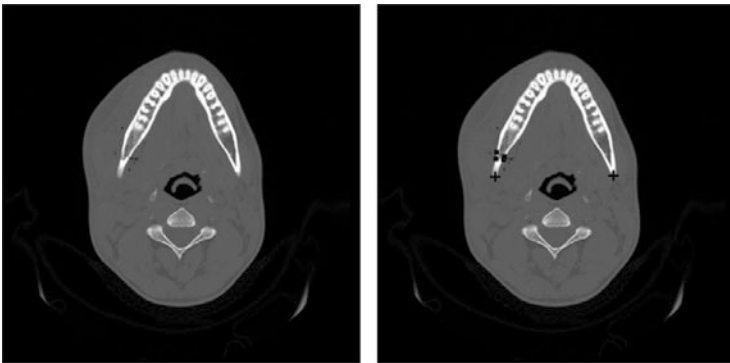
Fig. 19 Two-dimensional slice of a fractured mandible in the *left*. Fracture detection (*black squares*) with source and sink identification (*black crosses*) in the *right*

This leads to a fall in the value of recall. Now, we try to explain why method-I is less strict compared to method-II. Method-I estimates the overlap between the binarized versions of the reference image and a candidate image. Thus, it essentially involves comparison of just two intensities. But, method-II compares top 25 % of the intensity values between the reference image and a candidate image after modeling the respective histograms as cdfs.

Two different mandibles with the fracture sites identified are shown in Figs. 18 and 19. In each of these figures, the centers of the crosses mark the source vertex (in the left half of the image) and the sink vertex (in the right half of the image) and the fractures are indicated by dark squares. Upon execution of the max-flow min-cut algorithm, we obtain the edges in the cut set. Each such edge joins a vertex on the source-side with another vertex on the sink-side of the flow network. For

Table 1 Performance measure of method-I in terms of precision-recall values

Dataset	Retrieved slices (b)	Ground-truth slices (a)	Common slices (c)	Precision (%) (100 × c/b)	Recall (%) (100× c/a)
I	14–21, 24–26, 29–39	25–39	25–26, 29–39	59.1	86.7
II	3–16	6–20	6–16	78.6	73.3
III	9–29, 36–40	8–27	9–27	73.1	95
IV	12–14, 17–38	10–30	12–14, 17–30	65.4	81
V	15–25, 35–37, 40–44	15–27	15–25	57.9	84.6

Table 2 Overall precision-recall value of method-I

Precision	66.8 ± 8.9
Recall	84.1 ± 7.9

Table 3 Performance measure of method-II in terms of precision-recall values

Dataset	Retrieved slices (b)	Ground-truth slices (a)	Common slices (c)	Precision (%) (100 × c/b)	Recall (%) (100 × c/a)
I	12–18, 25–37	25–39	25–37	65	86.7
II	1, 9–20	6–20	9–20	92.3	80
III	8–9, 13–23, 31, 39, 40	8–27	8–9, 13–23	81.3	65
IV	10,12, 14–25, 35–43	10–30	10, 12, 14–25	60.8	66.6
V	15–26, 42–44	15–27	15–26	80	92.3

Table 4 Overall precision-recall value of method-II

Precision	75.9 ± 12.9
Recall	78.1 ± 12.1

proper visualization, each such vertex is represented by a black square. Note that the slice in Fig. 18 is the one retrieved in Fig. 15c using method-I and Fig. 17c using method-II. The slice in Fig. 19 is from a different dataset (for which the retrieved slices are not shown). It is evident from the figures that the fractures are identified accurately in both cases. The ground-truth is obtained via manual detection by trained radiologists. We achieve a sensitivity of 79 % and a specificity of 59 % for the max-flow min-cut based fracture detection algorithm. The above values of sensitivity and specificity are due to (1) use of a simple capacity function during the min-cut and (2) use of only the intensity feature in the image retrieval procedure. The execution time for the max-flow min-cut algorithm is only few seconds on a 1.73 GHz Intel© Pentium© M Processor.

6 Conclusion and Future Scope

In this work, we discuss an automated detection of hairline fracture in mandibular images. Hairline fracture detection is a challenging problem as in many cases it can be missed by a practicing radiologist. The input to our problem is a sequence of CT images with scans of human craniofacial skeleton. A typical sequence consists of many slices out of which only a handful is observed to contain the fractured mandible. We first apply intensity-based image retrieval to identify the slices containing the fractured mandible. CT scan of an unbroken mandible is used as the reference image. Two different measures of similarity, namely the NJI and the Kolmogorov–Smirnov distance are used in this connection. We then model a hairline fracture as a minimum cut in an appropriately weighted flow network. The flow network is based on the geometry of the human mandible and some prior knowledge of the fracture pattern. A simple capacity function is used to compute the edge weights. The Edmonds–Karp enhanced Ford–Fulkerson algorithm is employed on the 2D flow network to obtain a minimum cut. The present work serves as an important tool for the radiologists and the craniofacial surgeons.

One direction of future work is to incorporate the shape information of the mandible in the retrieval process. A combination of shape and intensity information [23] and/or some new techniques like multi-texton histogram [24] can enhance the retrieval performance. In the present work, we have experimentally chosen the threshold values. However, based on statistical properties of the values of similarity measures, the thresholds can be properly selected. So, we plan to apply sophisticated statistical techniques like bootstrap [25] for selection of threshold. Finally, we will try to make the capacity function more robust by incorporating anatomical knowledge. For example, we can use the information that tissue swelling and specific low intensity regions called emphysema [26] typically appear in the vicinity of a mandibular fracture. All the aforementioned measures will increase the sensitivity and specificity values for computer-aided detection of hairline mandibular fractures.

References

1. King, R.E., Scianna, J.M., Petruzzelli, G.J.: Mandible fracture patterns: a suburban trauma center experience. Am. J. Otolaryngol. **25**(5), 301–307 (2004)
2. Ogundare, B.O., Bonnick, A., Bayley, N.: Pattern of mandibular fractures in an urban major trauma center. J. Oral Maxillofac. Surg. **61**(6), 713–718 (2003)
3. Pashley, D.H., Borke, J.L., Yu, J.C.: Biomechanics and craniofacial morphogenesis. In: Lin, K.Y., Ogle, R.C., Jane, J.A. (eds.) Craniofacial Surgery – Science and Surgical Technique. Saunders, Philadelphia (2002)
4. Müller, H., Michoux, N., Bandon, D., Geissbuhler, A.: A review of content based image retrieval systems in medical applications – clinical benefits and future directions. Int. J. Med. Inform. **73**(1), 1–23 (2004)

5. Ford Jr., L.R., Fulkerson, D.R.: Maximum flow through a network. Can. J. Math. **8**, 399–404 (1956)
6. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms. MIT Press, Cambridge (2001)
7. Kurihara, T.: The fourth dimension in simulation surgery for craniofacial surgical procedures. Keio J. Med. **50**(2), 155–165 (2001)
8. Siessegger, M., Schneider, B.T., Mischkowski, R.A., Lazar, F., Krug, B., Klesper, B., Zoller, J.E.: Use of an image-guided navigation system in dental implant surgery in anatomically complex operation sites. J. Craniomaxillofac. Surg. **29**(5), 276–281 (2001)
9. Chowdhury, A.S., Bhandarkar, S.M., Datta, G., Yu, J.C.: Automated detection of stable fracture points in computed tomography image sequences. In: Proceedings of the Third IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1320–1323, Arlington, 6–9 April 2006
10. Chowdhury, A.S., Bhattacharya, A., Bhandarkar, S.M., Datta, G., Yu, J.C., Figueroa, R.: Hairline fracture detection using MRF and Gibbs sampling. In: Proceedings of the Eighth IEEE International Workshop on Applications of Computer Vision (WACV), p. 56, Austin (2007)
11. Hashimoto, T., Kuroda, S., Tanimoto, Y., Miyawaki, S., Takano-Yamamoto, T.: Correlation between craniofacial and condylar path asymmetry. J. Oral Maxillofac. Surg. **66**(10), 2020–2027 (2008)
12. Chowdhury, A.S., Bhandarkar, S.M., Robinson, R.W., Yu, J.C., Liu, T.: Detection of hairline mandibular fracture using max-flow min-cut and Kolmogorov–Smirnov distance. In: Proceedings of the Eighth IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1962–1965, Chicago (2011)
13. Rahman, M., Bhattacharya, P., Desai, B.C.: A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE Trans. Inf. Technol. Biomed. **11**(1), 58–69 (2007)
14. Selvarani, A.G., Annadurai, S.: Content based image retrieval for medical images using generic Fourier descriptor. J. Comput. Intell. Bioinform. **1**(1), 65–72 (2008)
15. Sedghi, S., Sanderson, M., Clough, P.D.: A study on the relevance criteria for medical images. Pattern Recognit. Lett. **29**(15), 2046–2057 (2009)
16. Gonzalez, R.C., Woods, R.E.: Digital image processing. Addison-Wesley, Noida (2001)
17. Sahoo, P.K., Soltani, S., Wong, K.C., Chen, Y.C.: A survey of thresholding techniques. Comput. Vis. Graph. Image Process. **41**, 233–260 (1988)
18. Crum, W.R., Camara, O., Hill, D.L.G.: Generalized overlap measures for evaluation and validation in medical image analysis. IEEE Trans. Med. Imaging **25**(11), 1451–1461 (2006)
19. Chowdhury, A.S., Arabnia, H.R., Bandyopadhyay, D.: Improved stereo correlation using Moravec operator and Kolmogorov–Smirnov test. In: Proceedings of the World Academy of Science International Conference on Computer Vision, pp. 24–30, Las Vegas (2005)
20. Hollander, M., Wolfe, D.A.: Nonparametric Statistical Methods. Wiley, New York (1999)
21. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: IEEE International Conference on Computer Vision (ICCV), pp. 105–112, Vancouver (2001)
22. Rasband, W.S.: ImageJ, U. S. National Institutes of Health, Bethesda, MD. http://imagej.nih.gov/ij/ (1997–2012)
23. Gevers, T., Smeulders, A.W.M.: PicToSeek: combining color and shape invariant features for image retrieval. IEEE Trans. Image Process. **9**(1), 102–119 (2000)
24. Liu, G.H., Zhang, L., Hou, Y.K., Li, Z.Y., Yang, J.Y.: Image retrieval based on multi-texton histogram. Pattern Recognit. **43**(7), 2380–2389 (2010)
25. Efron, B., Tibshirani, R.: An Introduction to the Bootstrap. Chapman & Hall, New York (1993)
26. Giannoudis, P.V., Dinopoulos, H.: Current concepts of the inflammatory response after major trauma: an update. Injury **36**(1), 229–230 (2005)

Index