

Nikos Paragios · James Duncan
Nicholas Ayache *Editors*

Handbook of Biomedical Imaging

Methodologies and Clinical Research

 Springer

Handbook of Biomedical Imaging

Nikos Paragios • James Duncan
Nicholas Ayache
Editors

Handbook of Biomedical Imaging

Methodologies and Clinical Research

 Springer

Editors

Nikos Paragios
Department of Applied Mathematics
École Centrale de Paris
Chatenay-Malabry, France

Nicholas Ayache
Inria Sophia Antipolis - Méditerranée
Sophia Antipolis Cedex, France

James Duncan
Department of Biomedical Engineering,
Diagnostic Radiology and Electrical
Engineering
Yale University
New Haven, CT, USA

ISBN 978-0-387-09748-0

ISBN 978-0-387-09749-7 (eBook)

DOI 10.1007/978-0-387-09749-7

Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014945340

© Springer Science+Business Media New York 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Part I Methodologies

Object Segmentation and Markov Random Fields	3
Y Boykov	
Fuzzy methods in medical imaging	25
I. Bloch	
Curve Propagation, Level Set Methods and Grouping	45
N. Paragios	
Kernel Methods in Medical Imaging	63
G. Charpiat, M. Hofmann, and B. Schölkopf	
Geometric Deformable Models	83
Y. Bai, X. Han, and J.L. Prince	
Active Shape and Appearance Models	105
T.F. Cootes, M.G. Roberts, K.O. Babalola, and C.J. Taylor	

Part II Statistical & Physiological Models

Statistical Atlases	125
C. Davatzikos, R. Verma, and D. Shen	
Statistical Computing on Non-Linear Spaces for Computational Anatomy	147
X. Pennec and P. Fillard	
Building Patient-Specific Physical and Physiological Computational Models from Medical Images	169
H. Delingette and N. Ayache	

Constructing a Patient-Specific Model Heart from CT Data	183
D.M. McQueen, T. O’Donnell, B.E. Griffith, and C.S. Peskin	
Image-based haemodynamics simulation in intracranial aneurysms	199
A.G. Radaelli, H. Bogunović, M.C. Villa Uriol, J.R. Cebal, and A.F. Frangi	
Part III Biomedical Perception	
Atlas-based Segmentation	221
M. Bach Cuadra, V. Duay, and J.-Ph. Thiran	
Integration of Topological Constraints in Medical Image Segmentation ..	245
F. Ségonne and B. Fischl	
Monte Carlo Sampling for the Segmentation of Tubular Structures	263
C. Florin, N. Paragios, and J. Williams	
Non-rigid registration using free-form deformations	277
D. Rueckert and P. Aljabar	
Image registration using mutual information	295
F. Maes, D. Loeckx, D. Vandermeulen, and P. Suetens	
Physical Model Based Recovery of Displacement and Deformations from 3D Medical Images	309
P. Yang, C. Delorenzo, X. Papademetris, and J.S. Duncan	
Graph-based Deformable Image Registration	331
A. Sotiras, Y. Ou, N. Paragios, and C. Davatzikos	
Part IV Clinical Biomarkers	
Cardiovascular Informatics	363
I.A. Kakadiaris, U. Kurkure, A. Bandekar, S. O’Malley, and M. Naghavi	
Rheumatoid Arthritis Quantification using Appearance Models	375
G. Langs, P. Peloschek, H. Bischof, and F. Kainberger	
Medical Image Processing for Analysis of Colon Motility	391
N. Navab, B. Glocker, O. Kutter, S.M. Kirchhoff, and M. Reiser	
Segmentation of Diseased Livers: A 3D Refinement Approach	403
R. Beichel, C. Bauer, A. Bornik, E. Sorantin, and H. Bischof	

Part V Emerging Modalities & Domains

Intra and inter subject analyses of brain functional Magnetic Resonance Images (fMRI) 415
 J.B. Poline, P. Ciuciu, A. Roche, and B. Thirion

Diffusion Tensor Estimation, Regularization and Classification..... 437
 R. Neji, N. Azzabou, G. Fleury, and N. Paragios

From Local Q-Ball Estimation to Fibre Crossing Tractography 455
 M. Descoteaux and R. Deriche

Segmentation of Clustered Cells in Microscopy Images by Geometric PDEs and Level Sets 475
 A. Kuijper, B. Heise, Y. Zhou, L. He, H. Wolinski, and S. Kohlwein

Atlas-based whole-body registration in mice..... 489
 M. Baiker, J. Dijkstra, J. Milles, C.W.G.M. Löwik, and B.P.F. Lelieveldt

Potential carotid atherosclerosis biomarkers based on ultrasound image analysis 501
 S. Golemati, J. Stoitsis, and K.S. Nikita

Part I

Methodologies

Object Segmentation and Markov Random Fields

Y Boykov

Abstract This chapter discusses relationships between *graph cut* approach to object delineation and other standard techniques optimizing segmentation boundaries. Graph cut method is presented in the context of globally optimal labeling of binary Markov Random Fields (MRFs). We review algorithms details and show several 2D and 3D examples.

1 Overview of object segmentation methods

In the last 20 years the computer vision and medical imaging communities have produced a number of useful algorithms for localizing object boundaries in images. The most basic methods like *thresholding* and *region growing* are simple heuristics that work well only for a fairly limited class of problems: either there should be 100 % consistent intensity edges on the object boundary, or there should be 0 % overlap between the object and background intensity histograms. In practice, there are relatively few applications where object appearance satisfies such assumptions (one good example is [24]). More robust segmentation techniques are needed in most applications that do not support such strong assumptions.

More advanced segmentation methods typically compute optimal contours/surfaces for specific energy functionals combining different boundary and/or region-based cues for an object of interest. Examples of such energy-based object extraction methods are *snakes* [36], *balloons* [18], other *active contour* models [30], geometric methods [14, 68], “shortest path” techniques [21, 50], *ratio cycles* [33], *ratio cuts* [65], *random walker* [27], *graph cuts* [4, 8, 9, 37, 41], continuous max-flow [2], total variation (TV) [15], and TV-based convex relaxation methods [16].

Y. Boykov, PhD (✉)

Computer Science Department, University of Western Ontario, Middlesex College,
355, Ontario N6A 5B7, Canada
e-mail: yuri@csd.uwo.ca

contour/surface representation	local optimization	global optimization
explicit	<i>snake/balloon</i> [36, 18] <i>DP snake</i> (2D) [1] <i>spline snake</i> [30]	<i>optimal path</i> (2D) [50, 21] <i>ratio cut/cycle</i> (2D) [33, 65] <i>graph cut on complex</i> [37]
implicit (region-based)	<i>level set</i> [14, 17, 52] <i>PDE cuts</i> [11]	<i>graph cut</i> [8, 9, 4, 39] <i>continuous max-flow</i> [2] <i>TV convex relaxation</i> [16]

Fig. 1 Object extraction methods using contour/surface energy functionals. Only a few representative references are shown for each method

There are many different ways to formulate object segmentation as an optimization problem. Yet, there are more similarities between existing methods than it may seem. While the primary focus of this chapter is *graph cut* approach to object extraction and its relation to binary MRF estimation, this section reviews connections between graph cuts and other segmentation methods optimizing boundary cost functionals (see Fig. 1).

Note that graph cut methods for object extraction, e.g. [4, 8, 9, 39], were preceded by a number of graph-based methods for image “clustering” that use either combinatorial optimization algorithms [22, 32, 63, 67] or spectral analysis techniques, e.g. *normalized cuts* [60]. Typically, the goal of these methods is to automatically divide an image into a number of “blobs” or “clusters” using only generic cues of coherence or affinity between pixels.¹ We focus on a fairly different group of image segmentation methods shown in Fig. 1 that integrate model-specific visual cues and contextual information to accurately delineate particular object of interest.

1.1 Implicit representation of contours/surfaces

To solve an image segmentation problem, one should first decide how to numerically represent contours or surfaces. Many existing segmentation methods (Fig. 1) use *explicit* representation of contours/surfaces based on meshes or splines defined by a set of moving or “active” control points forming a chain [1, 18, 25, 30, 36]. A 2D

¹ *Impossibility theorem* in [38] shows that no clustering algorithm can simultaneously satisfy three basic axioms on *scale-invariance*, *richness*, and *consistency*. Thus, any clustering method has some bias.

contour can be explicitly represented by a set of adjacent edges connecting nodes on a graph [21, 33, 50]. Both contours in 2D and surfaces in 3D can be explicitly represented by a set of adjacent facets on a complex [37].

Many recent segmentation methods use *implicit* surface representation, see Fig. 1. Implicit object representation is frequently referred to as *region-based* representation. Such methods store a *mask* describing object’s interior. Such a mask is a 2D (or 3D) array over pixels (or voxels) assigning different values to interior and exterior regions. For example, standard *level-sets* methods, e.g. [52, 58], maintain a real valued mask with negative values for interior pixels and positive values for exterior pixels (or vice versa). The object’s boundary is implicit, it is a zero-level set of this real-valued mask. Many level-sets techniques use a mask that closely approximates a *signed distance map* of the object’s boundary. Real-value precision of the mask also helps to reconstruct the boundary with sub-pixel accuracy, if needed.

Some other implicit techniques, e.g. [2, 15, 16, 27], use a regional mask storing real-values in some bounded interval, e.g. [0, 1]. While each method may provide a different interpretation for mask values, a surface is typically extracted as a $\frac{1}{2}$ -level set of the mask. If necessary, sub-pixel accuracy can be obtained using bilinear interpolation. Notably, [2] and [16] prove that (in a theoretical case of continuous-resolution) their masks should converge to binary (1/0) values. In particular, this shows that specific choice of a threshold is not essential in [2, 16].

Graph cut [4, 8, 9, 39] is another implicit approach to representing object boundaries. Graph cut methods compute a binary-valued mask assigning 1 to interior and 0 to exterior pixels. In this case, object boundary cannot be uniquely identified with sub-pixel accuracy since any curve/surface in a narrow space between 1-0 valued points in \mathcal{R}^n (pixels/voxels) is consistent with the corresponding binary mask. If necessary, one simple way to output a boundary explicitly is to draw a surface following pixels borders. This may work sufficiently well if pixels/voxels are small enough.² However, even tiny “pixelization” of the surface causes visually noticeable shading artifacts in applications using 3D surface rendering. Thus, fitting a smooth surface in a narrow band between 1 and 0 labeled points in \mathcal{R}^n could be a better solution in case of 3D rendering, e.g. [46, 66].

1.2 Local vs. Global optimization

Each surface representation approach may have some advantages and disadvantages. For example, methods using control points may suffer from mesh irregularities and from rigid topological structure of surfaces they represent. Discrete approaches based on graphs/trees and dynamic programming must address potential geometric metrification artifacts. Some types of representation are specific

²If necessary, one can build a graph with resolution finer than the pixel grid.

to contours in 2D (see Fig. 1) and do not generalize to surfaces in 3D. Moreover, each surface representation approach is typically tied with certain types of energy functionals and optimization algorithms.

The underlying optimization technique significantly affects numerical robustness, convergence, speed, and memory efficiency of the segmentation method. Moreover, it also determines quality guarantees for the solution. For example, some algorithms can find only a local minima near some provided initial guess, while others can compute a globally optimal solution in the whole domain of interest (e.g. image or volume).

The majority of existing techniques with explicit representation of surfaces via control points (meshes, splines) use either variational methods [18, 30, 36] or dynamic programming [1] to find a local optima positions for these control points near some initial guess. Dynamic programming can also be used for global minimization in some 2D segmentation problems. For example, segmentation methods using explicit representation of contours as “edge-paths” on graphs use variations of dynamic programming (e.g. Dijkstra/Viterbi algorithms, *ratio cycles*, *ratio cuts*) to obtain globally optimal solutions in 2D applications [21, 33, 50, 65]. Unfortunately, there are no straightforward generalizations of DP-based methods to surfaces in 3D. However, a dual approach explicitly representing contours/surfaces as a set of adjacent facets on *complexes* allows global optimization via combinatorial graph cut algorithms in either 2D or 3D cases [37].

Many segmentation methods using implicit representation of surfaces via level sets [14, 17, 52] obtain only local minima solutions since they apply gradient descent to non-convex spaces (e.g. of characteristic set functions). More recently, continuous max-flow/min-cut ideas developed by Strang [61] in the early eighties inspired convex relaxation techniques computing globally optimal surfaces [2, 16] using implicit surface representations.

Graph cuts methods for object segmentation [8, 9, 4, 39] represent object boundaries implicitly via binary pixel labeling. Globally optimal labeling is produced by combinatorial algorithms for discrete max-flow/min-cut problems on graphs known since the early sixties [23]. Interestingly, continuous versions of max-flow/min-cut problems formulated by Strang [61] were inspired by the corresponding discrete problems.

First application of combinatorial graph cut algorithms to image analysis dates back to the late eighties [28] (binary image restoration). But graph cuts have become popular in vision only 10 years later when high-quality globally optimal solutions made breakthroughs in stereo [13, 31, 44, 56] and multi-view reconstruction [45, 64, 47]. Before combinatorial graph cut approach to object extraction was first presented in [8], globally optimal object segmentation was possible only in 2D using dynamic programming.

Global optimization and numerically stable fast algorithms (e.g. [10]) are the key strengths of the graph cut framework. In general, global solutions are attractive due to potentially better robustness and predictability. For example, imperfections in a globally optimal solution directly relate to the cost functional rather than to numerical problems during minimization or bad initialization. Yet, global minima

could be unacceptable. For example, the trivial null solution is a global minima in some segmentation problems since length/area of the segmentation boundary is a typical regularizing component in most energy functionals. Some methods add a ballooning term to their energy or constrain solution space to avoid trivial/null solutions. Computing local minima could be another reasonable and efficient alternative if good initial solution is available. Interestingly, graph cut approach can also be used for local optimization when the search space is appropriately constrained near given initial solution [11].

1.3 Continuous vs. Discrete Optimization

Energy-based object segmentation methods can be also distinguished by the type of energy functional they use. The majority of methods optimizing segmentation boundary cost can be divided into two groups:

- (A) Optimization of a functional defined on continuous contour/surface
- (B) Optimization of an energy defined on a finite set of discrete variables

The standard methods in group (A) normally formulate segmentation as an optimization problem in an infinite-dimensional space \mathcal{R}^∞ (of continuous contours or functions). Most methods in this group rely on a variational approach and gradient descent. In contrast, segmentation methods from group (B) formulate some integer optimization problem in finite dimensional space \mathcal{Z}^n and use combinatorial optimization algorithms. Note that discrete energy functionals in group (B) are frequently designed to approximate some continuous optimization problem. Figure 1 contains examples of methods from both groups (A) and (B).

The methods in group (A) include *snakes* [36, 18], *region competition* [69], *geodesic active contours* [14], other methods based on *level-sets* [52, 59], and more recent global optimization methods using convex relaxation [2, 16].

Optimization methods in group (B) minimize an energy defined over a finite set of integer-valued variables. Such variables are usually associated with graph nodes corresponding to image pixels, super pixels, cells, or control points. Many methods in group (B) use discrete variables representing “direction” of a path along some graph relying on various versions of dynamic programming to compute optimal paths or cycles [1, 21, 33, 50, 65]. Graph cut methods also belong to group (B). Both implicit [4, 8, 9, 39] and explicit [37] techniques use binary variables representing (object/background) labels for either pixels, voxels, or *complex* cells. Numerically, graph cut methods rely on combinatorial max-flow/min-cut algorithms [10, 23, 26]. Once appropriate (*submodular*) discrete energy is formulated [6, 29, 53], these algorithms compute an exact global minima.

Continuous methods in (A) inherently rely on approximating numerical schemes (e.g. *finite differences* or *finite elements*) that must be carefully designed to insure robustness and stability. Convergence of such numerical methods is an important and non-trivial issue that has to be addressed. Segmentation results generated by two

variational techniques using the same energy may depend on their implementation details.

Discrete optimization methods in (B) are numerically robust and “repeatable”. For example, assuming the same energy function, one would always get identical segments even though one can choose from a number of different combinatorial min-cut/max-flow algorithms for computing minimum s - t cuts on graphs [10, 23, 26].

[10, 20, 35] studied practical efficiency of combinatorial min-cut/max-flow algorithms on applications in computer vision. It was shown that some max-flow techniques could solve 2D and 3D segmentation problems in close to real-time using regular PCs. Graph cut methods allow a user to change hard and soft constraints on-a-fly [4, 7, 8] enabling fast interactive editing of the segments. Significant speed-ups were demonstrated for dynamic segmentation problems using *flow-recycling* [39] and *cut-recycling* [34]. While straightforward implementation of graph cuts may require a lot of memory in 3D applications, [46, 49] showed that multi-level and adaptive banded techniques can alleviate the problem. Recent *region-push-relabel* algorithm [20]³ demonstrated good scalability to large 3D volumes under limited memory resources and significant speed-ups for multi-processor PCs.

1.4 Integrating Boundary and Regions

Continuous surface functionals in group (A) can incorporate various regional and boundary properties of segments motivated by ideas from geometry [14, 62] or physics [36]. It is more straightforward to incorporate regional properties of segments into implicit surface representation techniques like level-sets, which typically optimize geometric surface functionals related to continuous Mumford-Shah model [51].

All discrete *path*-based methods in group (B) can easily encode boundary-based segmentation cues. [33, 65] combine them with regional properties of segments using algorithms for optimizing some ratios of regional and boundary terms. Unfortunately, path-based methods are limited to 2D images since object boundaries in 3D cannot be represented by a path.

Graph cuts techniques in group (B) can optimize discrete energies that combine boundary regularization with regional consistency of segments in either 2D or 3D applications. The corresponding energies are consistent with the *weak membrane model* proposed by Blake and Zisserman [3, 5], which is a discrete counterpart of continuous Mumford-Shah functional [51]. As shown in [9, 40], graph cut methods can approximate the same geometric surface functionals that are widely used by level-set methods. This further bridges the gap between the *weak membrane* and Mumford-Shah models. Recently, [41] proposed *parametric max-flow* algorithm for

³[20] can be seen as a hierarchical version of standard *push-relabel* method [26].

optimizing ratios of some geometric functionals, extending some earlier ratio-based segmentation methods [33, 65] to applications in 3D.

1.5 Topological Constraints

Most segmentation methods using explicit boundary representation compute segmentation with fixed topological structure⁴, e.g. a single closed contour. Such methods typically have higher sensitivity to local minima. Moreover, object's topology may not be known *a priori*. In contrast, most implicit representation methods produce segments with flexible topological structure. Object segment may have multiple isolated blobs with holes. Sometimes, however, it may help to impose topological constraints reflecting certain high-level contextual information about the object of interest.

Region-based topological constraints naturally fit into implicit methods (graph cuts, level-sets). These constraints correspond to an infinite cost regional bias that enforces certain pixels to belong to object's interior or exterior. Similar ideas can also constrain the region of interest containing the object's boundary. Typically, optimal segmentation can be efficiently recomputed if some regional constraints (seeds) are added or removed. [12] discusses some non-regional topological constraints that potentially could be used in the graph cut algorithms for object extraction. It is also possible to simultaneously segment multiple surfaces topologically coupled to each other, e.g. when distance between them is constrained to be in some interval [48]. This approach can be extended to elastic interaction between coupled surfaces using results in [31].

Note that *intelligent scissors* [50] and *live wire* [21] use boundary-based hard constraints where the user can indicate certain pixels where the segmentation boundary should pass. In contrast, regional hard constraints used for graph cuts do not have to be precisely positioned. Moving the seeds around the object of interest (within some limits) does not normally change the segmentation results.

2 Related work on Markov Random Fields

In this sections we review relationship between graph cut methods for object extraction and estimation of Markov Random Fields. Greig et al. [28] were first to recognize that powerful max-flow/min-cut algorithms from combinatorial optimization can be applied to problems in computer vision. In particular, they

⁴[37] is a notable exception.

showed that graph cuts can be used for restoration of binary images.⁵ This problem can be formulated as Maximum *A Posterior* estimation of a Markov Random Field (MAP-MRF) that required minimization of posterior energy

$$E(I) = -\lambda \sum_{p \in \mathcal{P}} \ln \Pr(I_p | I) + \sum_{\{p,q\} \in \mathcal{N}} \delta_{I_p \neq I_q} \quad (1.1)$$

where

$$\delta_{I_p \neq I_q} = \begin{cases} 1 & \text{if } I_p \neq I_q \\ 0 & \text{if } I_p = I_q. \end{cases}$$

is a Kronecker delta representing interaction potential, $I = \{I_p | p \in \mathcal{P}\}$ is an unknown vector of original binary intensities $I_p \in \{0, 1\}$ of image pixels \mathcal{P} , vector I^o represents observed binary intensities corrupted by noise, and \mathcal{N} is a set of all pairs of neighboring pixels.

Greig et al. [28] recognized that if $I_p \in \{0, 1\}$ for all \mathcal{P} then posterior energy (1.1) is a typical example of energy of binary variables from *pseudo-boolean* optimization literature [29, 53]. In particular, [29, 53] show that minimization of such energies can be done by combinatorial algorithms for optimal graph partitioning. They show construction of a two terminal graph where the minimum cost cut divides the nodes (variables) into two disjoint sets corresponding to the globally optimal labeling of binary variables I_p .

Greig et al. [28] were first to compute globally optimal results for MAP-MRF estimation on 2D images. Previously, energies like (1.1) were approached with iterative sampling algorithms like simulated annealing. In fact, Greig et al. mainly used their results to show that in practice simulated annealing reaches solutions very far from the global minimum even in very simple binary image restoration examples.

Unfortunately, the graph cut approach to MAP-MRF estimation in [28] remained unnoticed for almost 10 years mainly because binary image restoration looked very limited as an application. In the late 90's new computer vision techniques appeared that showed how to use graph cut algorithms for more interesting non-binary problems. [57] was the first to use graph cuts to compute multi-camera stereo. Later, [31] showed that with the right edge weights on a graph similar to [57] one can globally minimize a non-binary case of (1.1) with any convex pairwise interaction terms, e.g. L_1 or L_2 metrics on the space of labels. Other algorithms for convex or even more general class of (multi-label) submodular interactions were also discovered. A different case of multi-label energies where the interaction penalty (*metric*) is not-necessarily convex was studied in [13]. In general, this is an NP-hard problem. The α -expansion algorithm proposed in [13] finds provably good approximate solutions by iteratively running min-cut/max-flow

⁵A typed or hand-written letter is an example of a binary image. Restoration of such an image may involve removal of a salt and pepper noise.

algorithms on appropriate graphs. The general case of *metric* includes “truncated” interaction potentials, also known as *discontinuity preserving*, which is practically very important for solving image analysis problems, e.g. [3, 5]). Later, it was shown that iterative α -expansion technique can be also used for non-metric interactions [55, 43] while further weakening optimality guarantees.

One of the insights in [8, 7] was that the problem of segmenting an object of interest from its background can be formulated as a binary image labeling problem and that an energy similar to (1.1) can be useful for object extraction. For example, binary image restoration energy (1.1) contains two terms representing “regional” and “boundary” properties. Such a combination looks appropriate for object segmentation. Moreover, this energy can be globally minimized in N-D images/volumes using combinatorial graph cut algorithms as suggested in pseudo-boolean optimization literature. [8, 7] showed how to integrate various basic cues and constraints in order to extract objects of interest.

3 Optimal object segmentation via graph cuts

This section describe basic MRF object segmentation framework in more detail. First, consider some terminology. A graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ is defined as a set of nodes or vertices \mathcal{V} and a set of edges \mathcal{E} connecting “neighboring” nodes. For simplicity, we mainly concentrate on undirected graphs where each pair of connected nodes is described by a single edge $e = \{p, q\} \in \mathcal{E}$ ⁶. A simple 2D example of an undirected graph that can be used for image segmentation is shown in Fig. 2(b).

The nodes of our graphs represent image pixels or voxels. There are also two specially designated terminal nodes S (*source*) and T (*sink*) that represent “object” and “background” labels. Typically, neighboring pixels are interconnected by edges in a regular grid-like fashion. Edges between pixels are called *n-links* where n stands for “neighbor”. Note that a neighborhood system can be arbitrary and may include diagonal or any other kind of n-links. Another type of edges, called *t-links*, are used to connect pixels to terminals. All graph edges $e \in \mathcal{E}$ including n-links and t-links are assigned some nonnegative weight (cost) w_e . In Fig. 2(b) edge costs are shown by the thickness of edges.

An *s-t* cut is a subset of edges $C \subset \mathcal{E}$ such that the terminals S and T become completely separated on the induced graph $\mathcal{G}(C) = \langle \mathcal{V}, \mathcal{E} \setminus C \rangle$. Note that a cut (see Fig. 2(c)) divides the nodes between the terminals. As illustrated in Fig. 2 (c-d), any cut corresponds to some binary partitioning of an underlying image into “object” and “background” segments. Note that in the simplistic example of Fig. 2 the image is divided into one “object” and one “background” regions. In general, cuts can generate binary segmentation with arbitrary topological properties.

⁶Each pair of connected nodes on a directed graph is linked by two distinct (directed) edges (p, q) and (q, p) . Directed edges can be useful in applications (see Sect. 3.3).

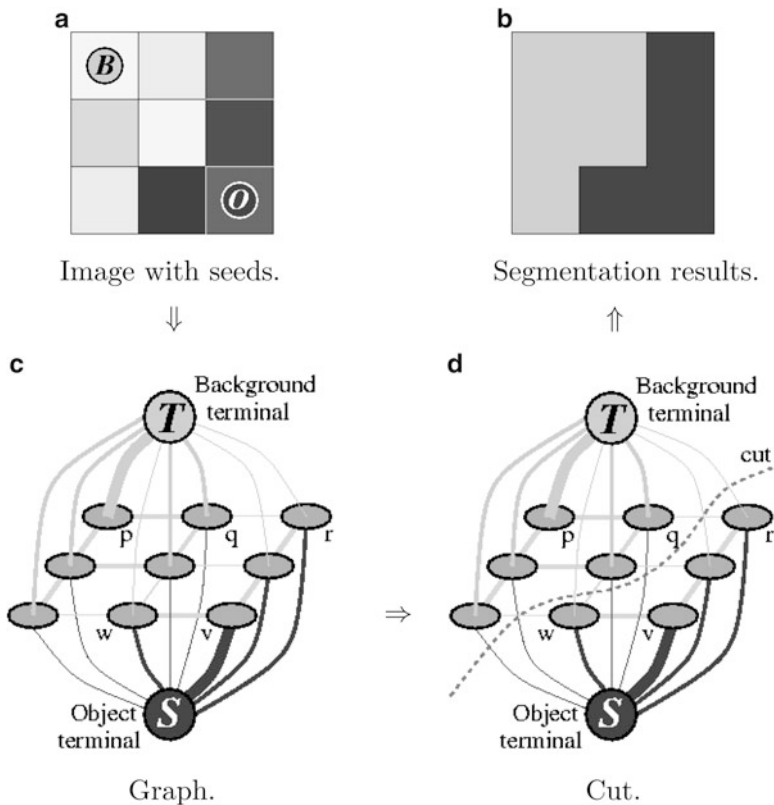


Fig. 2 A simple 2D segmentation example for a 3 X 3 image. The seeds are $\mathcal{O} = \{v\}$ and $\mathcal{B} = \{p\}$. The cost of each edge is reflected by the edge's thickness. The boundary term (1.4) defines the costs of n-links while the regional term (1.3) defines the costs of t-links. Inexpensive edges are attractive choices for the minimum cost cut. Hard constraints (seeds) (1.8,1.9) are implemented via infinity cost t-links. A globally optimal segmentation satisfying hard constraints can be computed efficiently in low-order polynomial time using max-flow /min-cut algorithms on graphs [23, 26, 19]

Examples in Sect. 4 illustrate that object and background segments may consist of several isolated connected blobs that also may have holes.

The goal is to compute the best cut that would give an “optimal” segmentation. In combinatorial optimization the cost of a cut is defined as the sum of the costs of edges that it severs

$$|C| = \sum_{e \in C} w_e$$

Note that severed n-links are located at the segmentation boundary. Thus, their total cost represents the cost of segmentation boundary. On the other hand, severed t-links can represent the regional properties of segments. Thus, a minimum cost cut

may correspond to a segmentation with a desirable balance of boundary and regional properties. Section 3.1 formulates a precise segmentation energy function that can be encoded via n-links and t-links. Note that infinity cost t-links can impose hard constraints on segments.

Globally minimum s - t cut can be computed efficiently in low-order polynomial time [23, 26, 10]. The corresponding algorithms work on any graphs. Therefore, graph cut segmentation method is not restricted to 2D images and computes globally optimal segmentation on volumes of any dimensions. Section 4 shows a number of 3D examples.

3.1 Segmentation energy

Consider an arbitrary set of data elements (pixels or voxels) \mathcal{P} and some neighborhood system represented by a set \mathcal{N} of all (unordered) pairs $\{p, q\}$ of neighboring elements in \mathcal{P} . For example, \mathcal{P} can contain pixels (or voxels) in a 2D (or 3D) grid and \mathcal{N} can contain all unordered pairs of neighboring pixels (voxels) under a standard 8 (or 26-) neighborhood system. Let $A = (A_1, \dots, A_p, \dots, A_{|\mathcal{P}|})$ be a binary vector whose components A_p specify *assignments* to pixels p in \mathcal{P} . Each A_p can be either “obj” or “bkg” (abbreviations of “object” and “background”). Vector A defines a segmentation. Then, the soft constraints that we impose on boundary and region properties of A are described by the cost function

$$E(A) = \lambda \cdot R(A) + B(A) \quad (1.2)$$

where

$$R(A) = \sum_{p \in \mathcal{P}} R_p(A_p) \quad (\text{regional term}) \quad (1.3)$$

$$B(A) = \sum_{\{p,q\} \in \mathcal{N}} B_{p,q} \cdot \delta_{A_p \neq A_q} \quad (\text{boundary term}) \quad (1.4)$$

and

$$\delta_{A_p \neq A_q} = \begin{cases} 1 & \text{if } A_p \neq A_q \\ 0 & \text{if } A_p = A_q. \end{cases}$$

The coefficient $\lambda \geq 0$ in (1.2) specifies a relative importance of the region properties term $R(A)$ versus the boundary properties term $B(A)$. The regional term $R(A)$ assumes that the individual penalties for assigning pixel p to “object” and “background”, correspondingly R_p (“obj”) and R_p (“bkg”), are given. For example,

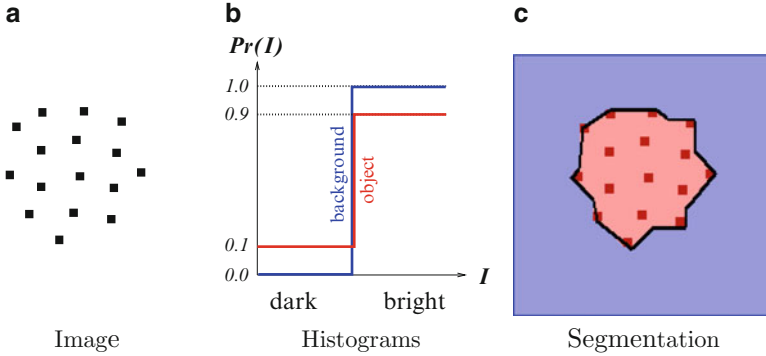


Fig. 3 Synthetic Gestalt example. The optimal object segment (red area in (c)) finds a balance between “region” and “boundary” terms in (1.2). The solution is computed using graph cuts. Some ruggedness of the segmentation boundary is due to metrication artifacts that can be realized by graph cuts in textureless regions. Such artifacts can be minimized [9]

$R_p(\cdot)$ may reflect on how the intensity of pixel p fits into given intensity models (e.g. histograms) of the object and background

$$R_p(\text{“obj”}) = -\ln \Pr(I_p | \text{“obj”}) \quad (1.5)$$

$$R_p(\text{“bkg”}) = -\ln \Pr(I_p | \text{“bkg”}) \quad (1.6)$$

This use of negative log-likelihoods is motivated by the MAP-MRF formulations in [13, 28].

The term $B(A)$ comprises the “boundary” properties of segmentation A . Coefficient $B_{p,q} \geq 0$ should be interpreted as a penalty for a discontinuity between p and q . Normally, $B_{p,q}$ is large when pixels p and q are similar (e.g. in their intensity) and $B_{p,q}$ is close to zero when the two are very different. The penalty $B_{p,q}$ can also decrease as a function of distance between p and q . Costs $B_{p,q}$ may be based on local intensity gradient, Laplacian zero-crossing, gradient direction, geometric [9, 40] or other criteria. Often, it is sufficient to set the boundary penalties from a simple function like

$$B_{p,q} \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{dist}(p, q)}. \quad (1.7)$$

This function penalizes a lot for discontinuities between pixels of similar intensities when $|I_p - I_q| < \sigma$. However, if pixels are very different, $|I_p - I_q| > \sigma$, then the penalty is small. Intuitively, this function corresponds to the distribution of noise among neighboring pixels of an image. Thus, σ can be estimated as “camera noise”.

A simple example of Fig. 3 illustrates some interesting properties of our cost function (1.2). The object of interest is a cluster of black dots in Fig. 3(a) that we

would like to segment as one blob. We combine boundary and region terms (1.3,1.4) taking $\lambda \geq 0$ in (1.2). The penalty for discontinuity in the boundary cost is

$$B_{p,q} = \begin{cases} 1 & \text{if } I_p \neq I_q \\ 0.2 & \text{if } I_p = I_q. \end{cases}$$

To describe regional properties of segments we use *a priori* known intensity histograms (Fig. 3(b)). Note that the background histogram concentrates exclusively on bright values while the object allows dark intensities observed in the dots. If these histograms are used in (1.5,1.6) then we get the following regional penalties R_p (A_p) for pixels of different intensities.

I_p	R_p (“obj”)	R_p (“bkg”)
dark	2.3	$+\infty$
bright	0.1	0

The optimal segmentation in Fig. 3(c) finds a balance between the regional and the boundary term of energy (1.2). Individually, bright pixels slightly prefer to stay with the background (see table above). However, spatial coherence term (1.4) forces some of them to agree with nearby dark dots which have a strong bias towards the object label (see Table).

3.2 Hard constraints and initialization

In the simple example of Fig. 3 the regional properties of the object of interest are distinct enough to segment it from the background. In real examples, however, objects may not have sufficiently distinct regional properties. In such cases it becomes necessary to further constraint the search space of possible solutions before computing an optimal one.

Assume that \mathcal{O} and \mathcal{B} denote the subsets of pixels *a priori* known to be a part of “object” and “background”, correspondingly. Naturally, the subsets $\mathcal{O} \subset \mathcal{P}$ and $\mathcal{B} \subset \mathcal{P}$ are such that $\mathcal{O} \cap \mathcal{B} = \emptyset$. For example, consider sets \mathcal{O} (red pixels) and \mathcal{B} (blue pixels) in Fig. 4(b). Our goal is to compute the global minimum of (1.2) among all segmentations A satisfying hard constraints

$$\forall p \in \mathcal{O} : A_p = \text{“obj”} \tag{1.8}$$

$$\forall p \in \mathcal{B} : A_p = \text{“bkg”} \tag{1.9}$$

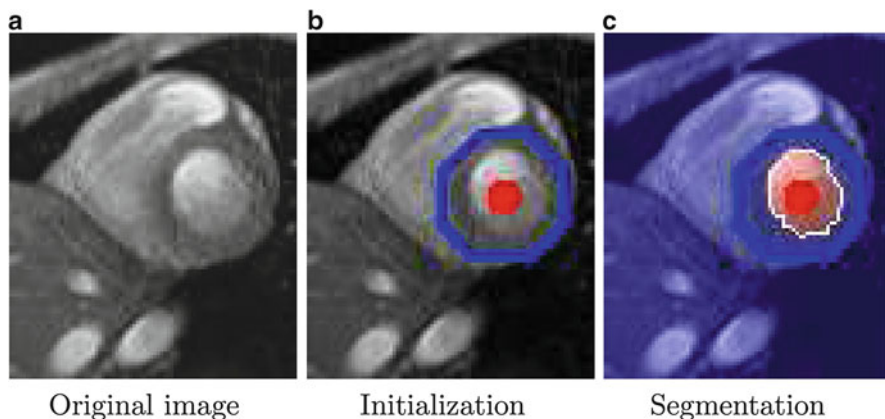


Fig. 4 Automatic segmentation of cardiac MR data. Initialization in (b) is based on hard constraints that can be placed automatically using simple template matching. Then, graph cuts accurately localize object boundaries in (c)

Figure 4(c) shows an example of an optimal segmentation satisfying the hard constraints in (b). Throughout this paper we use red tint to display object segments and blue tint for background.

Ability to incorporate hard constraints (1.8,1.9) is an interesting feature of graph cut methods. There is a lot of flexibility in how these hard constraints can be used to adjust the algorithm for different tasks. The hard constraints can be used to initialize the algorithm and to edit the results. The hard constraints can be set either automatically or manually depending on an application.

For example, consider example in Fig. 4(a). A simple template matching can roughly localize the left ventricle in the image, e.g. using its known circular shape. The hard constraints in Fig. 4(b) can be set automatically as soon as a rough position of the blood pool is known. Then, our graph cut technique can accurately localize the boundary of the blood pool in Fig. 4(c).

The band in Fig. 4 also restricts the area where the actual computation takes place. It is enough to build a graph only in the area of the band since max-flow/min-cut algorithms would not access any other nodes.

Note that it is possible to make a double use of the seeds. First of all, they can constrain the search space as discussed above. In addition, we can use intensities of pixels (voxels) marked as seeds to learn the histograms for “object” and “background” intensity distributions: $\Pr(I|“obj”)$ and $\Pr(I|“bkg”)$ in (1.5,1.6). Other ideas for initializing intensity distributions are studies in [4, 54].

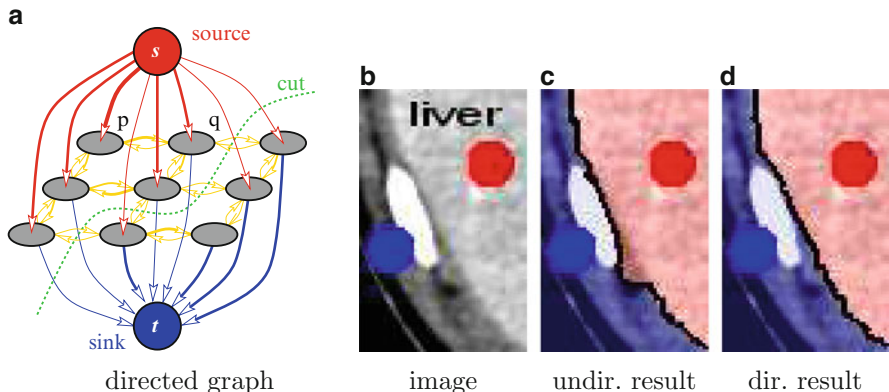


Fig. 5 Segmentation via cuts on a directed graph. Compare the results on an undirected graph (c) with the results on a directed graph in (d)

3.3 Using directed edges

For simplicity, we previously concentrated on the case of *undirected* graphs as in Fig. 2. In fact, the majority of *s-t* cut algorithms from combinatorial optimization work for *directed* graphs as well. Figure 5(a) gives one example of such a graph where each pair of neighboring nodes is connected by two directed edges (p, q) and (q, p) with distinct weights $w_{(p,q)}$ and $w_{(q,p)}$. If a cut separates two neighboring nodes p and q so that p is connected to the source while q is connected to the sink then the cost of the cut includes $w_{(p,q)}$ while $w_{(q,p)}$ is ignored. Vice versa, if q is connected to the source and p to the sink then the cost of the cut includes only $w_{(q,p)}$.

In certain cases one can take advantage of such directed costs to obtain more accurate object boundaries. For example, compare two segmentations in Fig. 5(c,d) obtained on a medical image in (b) using the same set of constraints. A relatively bright object of interest on the right (liver) is separated from a small bright blob on the left (bone) by a thin layer of a relatively dark muscle tissue. The contrast between the bone and the muscle is much better than the contrast between the muscle and the liver. Thus, according to (1.7) the standard “undirected” cost of edges between the bone and the muscle is much cheaper than the cost of edges between the muscle and the liver. Consequently, an optimal cut on an undirected graph produces segmentation in Fig. 5(c) that sticks to the bone instead of following the actual liver boundary.

Directed graphs can automatically distinguish between the incorrect boundary in Fig. 5(c) and the desirable one in (d). The key observation is that the weights of directed edges (p, q) can depend on a sign of intensity difference ($I_q - I_p$). In contrast, weights of undirected edges should be symmetric with respect to its end points and could depend only on the absolute value $|I_p - I_q|$ as in (1.7). Note that the object boundary that stuck to the bone in (c) separates darker tissue (muscle) in the

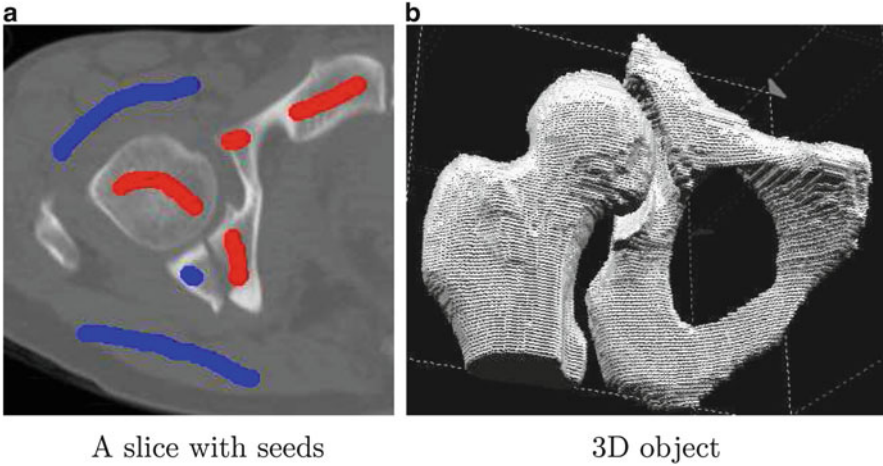


Fig. 6 Segmentation of bones in a CT volume [256x256x119]

object segment from brighter tissue (bone) in the background. On the other hand, the correct object boundary in (d) goes from brighter tissue (liver) in the object to darker tissue (muscle) in the background. Note that directed edge weights

$$w_{(p,q)} = \begin{cases} 1 & \text{if } I_p \leq I_q \\ \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) & \text{if } I_p > I_q \end{cases}$$

would specifically encourage cuts from brighter tissue in the object to darker tissue in the background. The results in Fig. 5(d) show optimal segmentation on a directed graph using such edge weights.

4 Some examples

We demonstrate a few examples of original image data and segments generated by graph cuts for a given set of hard constraints. User can enter hard constraints (seeds) via mouse operated brush of red (for object) or blue (for background) color. We used simple 4-neighborhood systems in 2D examples and 26-neighborhood system in 3D examples. All running times are given for 1.4GHz Pentium III. Our implementation is based on “max-flow” algorithm from [10].

Figures 6, and 7 show segmentation results that we obtained on a 3D medical volumes. Each object was segmented in 10 to 30 seconds. In the examples of Fig. 6 and 7 the objects were extracted from 3D volumes after entering seeds in only one slice shown in (a).

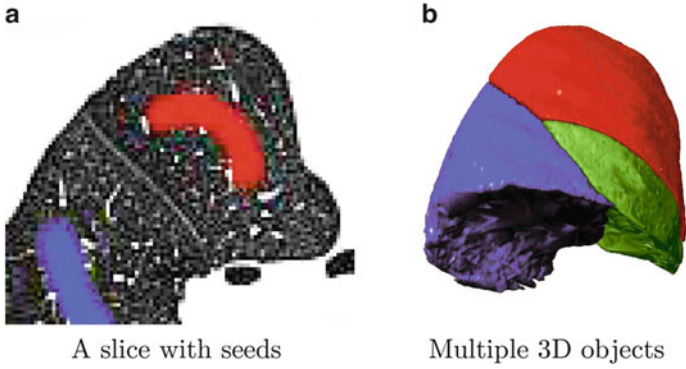


Fig. 7 Segmentation of lung lobes in a CT volume [205x165x253]

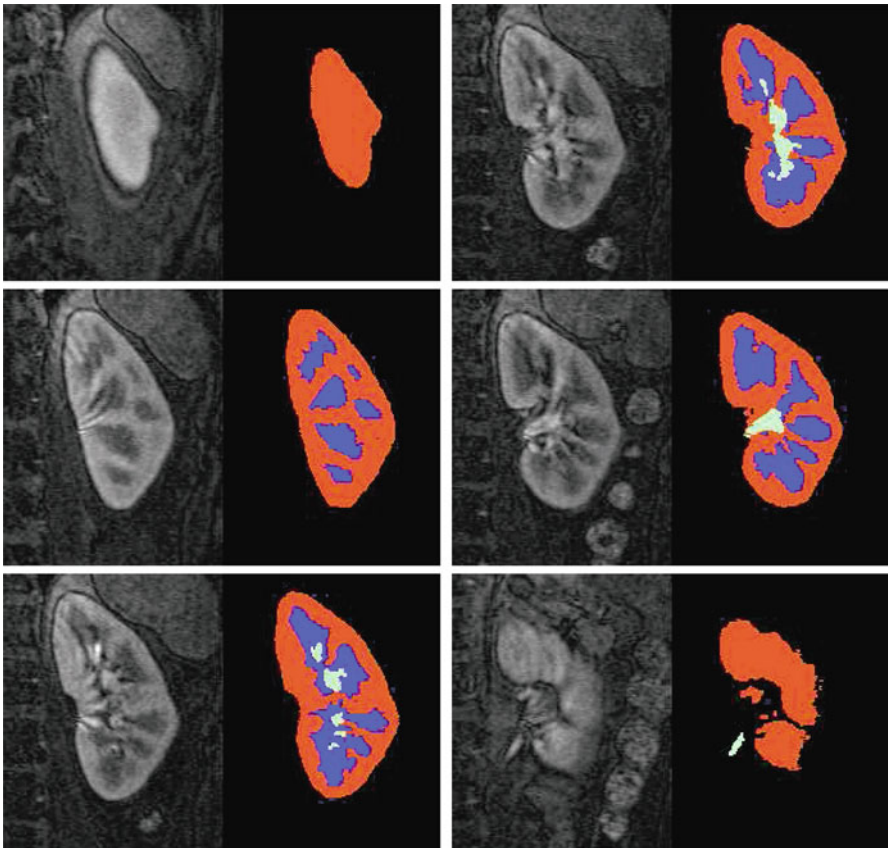


Fig. 8 Kidney in a 3D MRI angio data [55x80x32] segmented into cortex, medulla, and collecting system

We did not use regional term (1.3) for the experiments in Figs. 6, and 7. In some applications, however, the regional term may significantly simplify, if not completely automate, the segmentation process. In Fig. 8 we demonstrate segmentation on 3D kidney MR data that benefited from regional term (1.3). We segmented out cortex, medulla, and collecting system of a kidney in three consecutive steps. First, the whole kidney is separated from the background and the latter is cropped. The remaining pixels belong to three distinct types of kidney tissue (cortex, medulla, or collecting system) with identifiable regional properties. At this point it becomes useful to engage the regional term (1.3) of energy.

The results in Fig. 8 are shown without seeds since the process involved three different segmentations. Using regional bias allows to get 3D segmentation results by entering only a few seeds in one slice. Initial optimal segments are computed in 1-10 seconds and minor correction can be incorporated in less than a second. This example also demonstrates unrestricted topological properties of our segments. Fully automatic segmentation of kidney might be possible with more sophisticated models for regional.

References

1. A. A. Amini, T. E. Weymouth, and R. C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):855–867, September 1990.
2. B. Appleton and H. Talbot. Globally minimal surfaces by continuous maximal flows. *IEEE transactions on Pattern Analysis and Pattern Recognition (PAMI)*, 28(1):106–118, January 2006.
3. A. Blake. The least-disturbance principle and weak constraints. *Pattern Recognition Letters*, 1:393–399, 1983.
4. A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr. Interactive image segmentation using an adaptive gmmrf model. In *European Conference on Computer Vision (ECCV)*, Prague, Czech Republic, 2004.
5. A. Blake and A. Zisserman. *Visual Reconstruction*. Cambridge, 1987.
6. E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123(1-3):155–225, November 2002.
7. Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision (IJCV)*, 70(2):109–131, 2006.
8. Y. Boykov and M.-P. Jolly. *Interactive graph cuts* for optimal boundary & region segmentation of objects in N-D images. In *International Conference on Computer Vision*, volume I, pages 105–112, July 2001.
9. Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *International Conference on Computer Vision*, volume I, pages 26–33, 2003.
10. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, September 2004.
11. Y. Boykov, V. Kolmogorov, D. Cremers, and A. Delong. An integral solution to surface evolution PDEs via geo-cuts. In *European Conference on Computer Vision*, Graz, Austria, May 2006 (to appear).

12. Y. Boykov and O. Veksler. Graph cuts in vision and graphics: Theories and applications. In N. Paragios, Y. Chen, and O. Faugeras, editors, *Handbook of Mathematical Models in Computer Vision*, pages 79–96. Springer-Verlag, 2006.
13. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, November 2001.
14. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
15. A. Chambolle. Total variation minimization and a class of binary MRF models. pages 136–152, 2005.
16. T. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648, 2006.
17. T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.
18. L. D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2):211–218, 1991.
19. W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver. *Combinatorial Optimization*. John Wiley & Sons, 1998.
20. A. Delong. *A Scalable Max-Flow/Min-Cut Algorithm for Sparse Graphs*. MS thesis, University of Western Ontario, 2006.
21. A. X. Falcão, J. K. Udupa, S. Samarasekera, and S. Sharma. User-steered image segmentation paradigms: Live wire and live lane. *Graphical Models and Image Processing*, 60:233–260, 1998.
22. P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
23. L. Ford and D. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.
24. B. Geiger and R. Kikinis. Simulation of endoscopy. In *CVRMed*, pages 277–281, 1995.
25. D. Geiger, A. Gupta, L. A. Costa, and J. Vlontzos. Dynamic programming for detecting, tracking, and matching deformable contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(3):294–402, March 1995.
26. A. V. Goldberg and R. E. Tarjan. A new approach to the maximum flow problem. *Journal of the Association for Computing Machinery*, 35(4):921–940, October 1988.
27. L. Grady. Multilabel random walker segmentation using prior models. In *IEEE Conference of Computer Vision and Pattern Recognition*, volume 1, pages 763–770, San Diego, CA, June 2005.
28. D. Greig, B. Porteous, and A. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 51(2):271–279, 1989.
29. P. L. Hammer. Some network flow problems solved with pseudoboolean programming. *Operations Research*, 13:388–399, 1965.
30. M. Isard and A. Blake. *Active contours*. Springer-Verlag, 1998.
31. H. Ishikawa. Exact optimization for Markov Random Fields with convex priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1333–1336, 2003.
32. H. Ishikawa and D. Geiger. Segmentation by grouping junctions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 125–131, 1998.
33. I. H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *PAMI*, 23(10):1075–1088, October 2001.
34. O. Juan and Y. Boykov. Active graph cuts. In *IEEE Conference of Computer Vision and Pattern Recognition*, volume I, pages 1023–1029, 2006.
35. O. Juan and Y. Boykov. Accelerating graph cuts in vision via capacity scaling. In *International Conference on Computer Vision*, 2007.
36. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

37. D. Kirsanov and S. J. Gortler. A discrete global minimization algorithm for continuous variational problems. *Harvard Computer Science Technical Report*, TR-14-04, July 2004, (also submitted to a journal).
38. J. Kleinberg. An impossibility theorem for clustering. In *The 16th conference on Neural Information Processing Systems (NIPS)*, 2002.
39. P. Kohli and P. H. Torr. Efficiently solving dynamic markov random fields using graph cuts. In *International Conference on Computer Vision*, October 2005.
40. V. Kolmogorov and Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *International Conference on Computer Vision*, October 2005.
41. V. Kolmogorov, Y. Boykov, and C. Rother. Applications of parametric maxflow in computer vision. In *International Conference on Computer Vision (ICCV)*, Nov. 2007.
42. V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, and C. Rother. Bilayer segmentation of binocular stereo video. In *IEEE Conference of Computer Vision and Pattern Recognition*, San Diego, CA, 2005.
43. V. Kolmogorov and C. Rother. Minimizing non-submodular functions with graph cuts - a review. *PAMI*, 29(7), July 2007.
44. V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions via graph cuts. In *International Conference on Computer Vision*, July 2001.
45. V. Kolmogorov and R. Zabih. Multi-camera scene reconstruction via graph cuts. In *7th European Conference on Computer Vision*, volume III of *LNCS 2352*, pages 82–96, Copenhagen, Denmark, May 2002. Springer-Verlag.
46. V. Lempitsky and Y. Boykov. Global optimization for shape fitting. In *IEEE Conference of Computer Vision and Pattern Recognition*, June 2007.
47. V. Lempitsky, Y. Boykov, and D. Ivanov. Oriented visibility for multiview reconstruction. In *European Conference on Computer Vision*, Graz, Austria, May 2006 (to appear).
48. K. Li, X. Wu, D. Z. Chen, and M. Sonka. Optimal surface segmentation in volumetric images—a graph-theoretic approach. *IEEE transactions on Pattern Analysis and Pattern Recognition (PAMI)*, 28(1):119–134, January 2006.
49. H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *International Conference on Computer Vision*, October 2005.
50. E. N. Mortensen and W. A. Barrett. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60:349–384, 1998.
51. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
52. S. J. Osher and R. P. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer Verlag, 2002.
53. J. C. Picard and H. D. Ratliff. Minimum cuts and related problems. *Networks*, 5:357–370, 1975.
54. C. Rother, V. Kolmogorov, and A. Blake. Grabcut interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (SIGGRAPH)*, August 2004.
55. C. Rother, S. Kumar, V. Kolmogorov, and A. Blake. Digital tapestry. In *IEEE Conference of Computer Vision and Pattern Recognition*, San Diego, CA, June 2005.
56. S. Roy. Stereo without epipolar lines: A maximum-flow formulation. *International Journal of Computer Vision*, 34(2/3):147–162, August 1999.
57. S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. In *IEEE Proc. of Int. Conference on Computer Vision*, pages 492–499, 1998.
58. G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, 2001.
59. J. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
60. J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, August 2000.
61. G. Strang. Maximal flow through a domain. *Mathematical Programming*, 26:123–143, 1983.

62. A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *PAMI*, 24(12):1565–1578, December 2002.
63. O. Veksler. Image segmentation by nested cuts. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 339–344, 2000.
64. G. Vogiatzis, P. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *IEEE Conference of Computer Vision and Pattern Recognition*, pages 391–398, 2005.
65. S. Wang and J. M. Siskind. Image segmentation with ratio cut. 25(6):675–690, June 2003.
66. R. Whitaker. Reducing aliasing artifacts in iso-surfaces of binary volumes. pages 23–32, 2000.
67. Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1101–1113, November 1993.
68. A. Yezzi, Jr., S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum. A geometric snake model for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*, 16(2):199–209, 1997.
69. S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, September 1996.

Fuzzy methods in medical imaging

I. Bloch

Abstract Fuzzy sets theory is of great interest in medical image processing, for dealing with imprecise information and knowledge. It provides a consistent mathematical framework for knowledge representation, information modeling at different levels, fusion of heterogeneous information, reasoning and decision making. In this chapter, we provide an overview of the potential of this theory in medical imaging, in particular for classification, segmentation and recognition of anatomical and pathological structures.

1 Introduction

Imprecision is often inherent to images, and its causes can be found at several levels: observed phenomenon (imprecise limits between structures or objects), acquisition process (limited resolution, numerical reconstruction methods), image processing steps (imprecision induced by a filtering for instance). Fuzzy sets have several advantages for representing such imprecision. First, they are able to represent several types of imprecision in images, as for instance imprecision in spatial location of objects, or imprecision in membership of an object to a class. For instance, partial volume effect, which occurs frequently in medical imaging, finds a consistent representation in fuzzy sets (membership degrees of a voxel to tissues or classes directly represent partial membership to the different tissues mixed up in this voxel, leading to a consistent modeling with respect to reality). Second, image information can be represented at different levels with fuzzy sets (local, regional, or global), as well as under different forms (numerical, or symbolic). For instance, classification based only on grey levels involves very local information (at the pixel level);

I. Bloch (✉)
Signal and Image Processing, Telecom ParisTech - CNRS LTCI,
46 rue Barrault, Paris 75013, France
e-mail: isabelle.bloch@telecom-paristech.fr

introducing spatial coherence in the classification or relations between features involves regional information; and introducing relations between objects or regions for scene interpretation involves more global information and is related to the field of spatial reasoning. Third, the fuzzy set framework allows for the representation of very heterogeneous information, and is able to deal with information extracted directly from the images, as well as with information derived from some external knowledge, such as expert knowledge. This is exploited in particular in model-based pattern recognition, where fuzzy information extracted from the images is compared and matched to a model representing knowledge expressed in fuzzy terms.

Fuzzy set theory is of great interest to provide a consistent mathematical framework for all these aspects. It allows representing imprecision of objects, relations, knowledge and aims, at different levels of representation. It constitutes an unified framework for representing and processing both numerical and symbolic information, as well as structural information (constituted mainly by spatial relations in image processing). Therefore this theory can achieve tasks at several levels, from low level (e.g. grey-level based classification) to high level (e.g. model based structural recognition and scene interpretation). It provides a flexible framework for information fusion as well as powerful tools for reasoning and decision making.

In this chapter, we provide an overview of the potential of this theory in medical imaging, in particular for classification, segmentation and recognition of anatomical and pathological structures. The chapter is organized according to the level of information and processing. We assume that the basics of fuzzy sets theory are known (details can be found e.g. in [30]).

2 Low-level processing

The use of fuzzy sets in medical imaging at low level concerns mainly classification, often based on grey levels.

2.1 Representation

We denote by \mathcal{S} the spatial domain (\mathbb{R}^n in the continuous case or \mathbb{Z}^n in the discrete case). Fuzzy sets can be considered from two points of view. In the first one, a membership function is a function μ from the space \mathcal{S} on which the image is defined into $[0, 1]$. The value $\mu(x)$ is the membership degree of x ($x \in \mathcal{S}$) to a spatial fuzzy object. In the second one, a membership function is defined as a function μ' from a space of attributes \mathcal{A} into $[0, 1]$. At numerical level, such attributes are typically the grey levels. The value $\mu'(g)$ represents the degree to which a grey level g supports the membership to an object or a class. There is an obvious relation between μ and μ' in grey level based processing: $\mu(x) = \mu'[g(x)]$, where $g(x)$ denotes the grey level of x in the considered image.

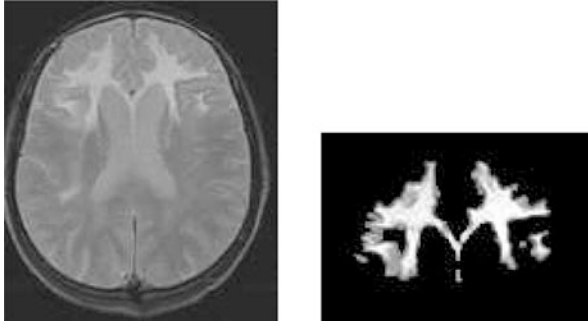


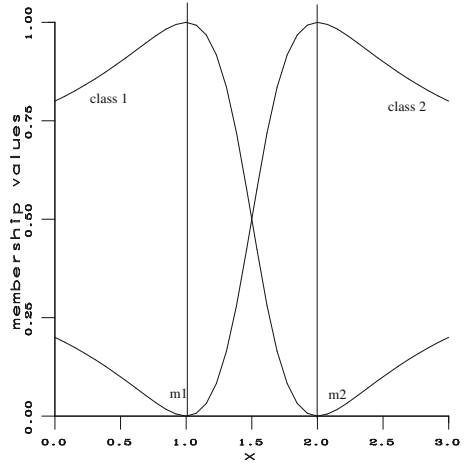
Fig. 1 MR image of the brain (left) (courtesy Prof. C. Adamsbaum, Saint-Vincent de Paul Hospital, Paris), and estimation of the partial membership to the pathology (right) in the pathological area (white means that there is only pathological tissue in the considering voxel, black means no pathological tissue, and intermediate values represent the partial volume effect, i.e. voxels that have also a non zero membership value to the white matter class)

Such models explicitly represent imprecision in the information provided by the images, as well as possible ambiguity between classes. For instance the problem of partial volume effect finds a consistent representation in this model. A pixel or voxel suffering from partial volume effect is characterized by its partial belonging to two (or more) different tissues or classes, i.e. by non zero membership values to several classes. Figure 1 shows an example of an MR image of the brain of a patient suffering from adrenoleukodystrophy, and where the slice thickness induces a high partial volume effect. The grey levels on the right figure represent the membership values to the pathology. The pathology is then considered as a fuzzy object, represented by a membership function defined on the spatial domain.

More generally, a spatial fuzzy object may represent different types of imprecision, either on the boundary of the objects (due for instance to partial volume effect, or to the spatial resolution), or on the individual variability of these structures, etc.

There is no definite answer to the question of how defining the membership functions. As mentioned above, they can be directly derived from the grey levels, but other characteristics can be used as well. For instance the contours of an object can be defined as a fuzzy set with a membership function depending on the gradient intensity. Based on a detection operator of some specific objects, the membership functions can be derived from the magnitude of the answer provided by this operator. Imprecision can also be introduced from a first crisp estimation of the objects, typically at their boundary as a function of the distance to the crisp object, to account for imprecision in this estimation. Finally, several approaches rely on fuzzy classification methods to derive membership functions.

Fig. 2 Membership values in a fuzzy C-means classification as a function of x in the case of 2 classes, with centroids in positions $x = 1$ and $x = 2$. In this example $m = 2$



2.2 Fuzzy classification

Learning of membership functions is a difficult task that still does not have a definite answer. Several methods have been proposed in the literature, often based on the minimization of some criteria. Among these methods, the most used is the fuzzy C-means algorithm (FCM) [5]. The idea is to define a membership function of a point to each class (which is then a fuzzy set), instead of deriving crisp assignments.

The FCM algorithm iteratively modifies a fuzzy partition so as to minimize an objective function defined as: $J_m = \sum_{j=1}^C \sum_i \mu_{ij}^m \|x_i - m_j\|^2$, under the constraint that $\forall i, \sum_{j=1}^C \mu_{ij} = 1$, where C denotes the number of class, N the number of points to be classified, μ_{ij} the membership function of point i to class j , and m is a parameter belonging to $]1, +\infty[$ called fuzzy factor, which controls the amount of “fuzziness” of the classification. The membership function is deduced from the cluster center position as: $\mu_{ij} = \frac{1}{\sum_{j=1}^C \left(\frac{\|x_i - m_j\|}{\|x_i - m_j\|} \right)^{\frac{2}{m-1}}}$, and the cluster center

position is obtained by: $m_j = \frac{\sum_i \mu_{ij}^m x_i}{\sum_i \mu_{ij}^m}$. From an initialization of cluster centers, the membership values and cluster centers are alternatively updated using these two equations, until convergence. Convergence towards a local minimum of the objective function has been proved. An example is provided in Fig. 2, in the case of a 1-dimensional 2-class problem. It also illustrates one of the main drawbacks of this approach: the membership functions are not decreasing with respect to the distance to the cluster center. This is due to the normalization constraint, and this phenomenon gets even worse with more classes.

An alternative solution to fuzzy C-means classification, which avoids the normalization drawbacks, is given by possibilistic C-means (PCM) [40]. The objective functional is defined as: $J = \sum_{j=1}^C \sum_{i=1}^N \mu_{ij}^m \cdot \|x_i - m_j\|^2 + \sum_{j=1}^C \eta_j \sum_{i=1}^N (1 - \mu_{ij})^m \cdot \|x_i - m_j\|^2$ and the obtained membership function is: $\mu_{ij} = \frac{1}{1 + \frac{\|x_i - m_j\|^2}{\eta_j} \frac{1}{m-1}}$. Now the membership functions are decreasing with respect

to the distance to the class centers. However this algorithm is very sensitive to initialization and sometimes coincident clusters may occur.

To address the problems of FCM and PCM a new fuzzy possibilistic C-mean (FPCM) algorithm was proposed in [49] by combining these two algorithms. The objective function involves both membership and typicality. FPCM solves the noise sensitivity defect of FCM and overcomes the problem of coincident clusters of PCM. Although FPCM is less prone to the problems of FCM and PCM, in the case of a large data set this algorithm does not work properly since the typicality values are very small in such cases, again due to a normalization constraint. This constraint has been suppressed in possibilistic fuzzy c-mean (PFM) [50]. Recently, approaches have been proposed by modifying the objective function to increase the robustness of FCM to noise [1, 33, 44, 45, 59, 64]. They also try to incorporate spatial information, by defining membership functions that depend on a local neighborhood around each point.

Another class of methods relies on probability-possibility transformations [29, 31, 39]. Other methods based on statistical information have been proposed, also by minimizing some criteria (e.g. [23, 25]). However, most criteria provide a function that depends on the shape of the histogram. Accounting for frequent situations where a pixel may belong completely and without any ambiguity to a class while having a grey-level with low occurrence frequency thus becomes difficult.

In [13] an original approach was proposed to obtain membership values from grey-level histogram. Two types of criteria are used simultaneously. The first type is based on a “resemblance” between the grey-level histogram and the membership function in the form of a distance between the two distributions. This type is very close to existing methods. The second type accounts for prior information on the expected shape of the membership function, in order to deal with problems mentioned above concerning low occurrence frequencies. This calls for a parametric representation of the functions. The combination of these two types of criteria leads to a simpler interpretation of the obtained functions that fits better the intuitive notion of membership. Membership functions are chosen as simple trapezoidal functions, whose parameters are estimated (simultaneously for all class membership functions) using simulated annealing in order to optimize the two criteria. The results obtained on a MR brain image are illustrated in Fig. 3. This method has been applied successfully to several problems like multi-image classification or segmentation of internal brain structures.

Finally, other types of classification methods, such as k -nearest neighbors, have also been extended to the fuzzy case.

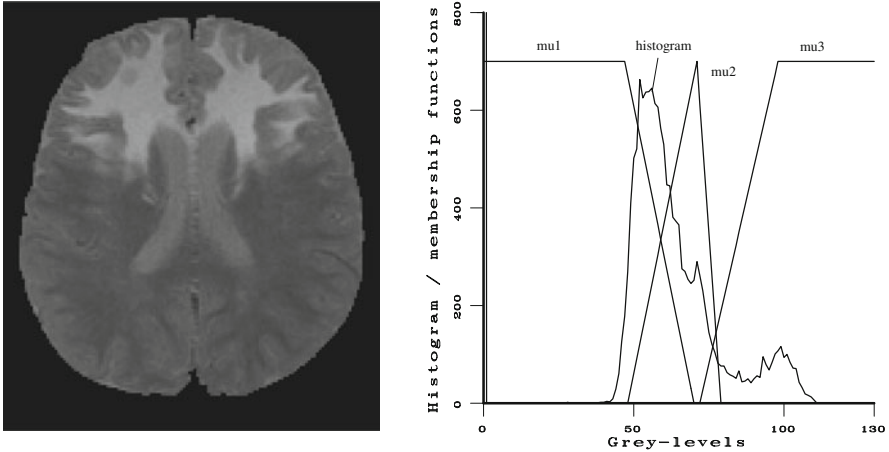


Fig. 3 MR image of the brain, showing three main classes: brain, ventricles and pathology (the white area on the left image), and result of the estimation of the three classes

Despite their drawbacks, these methods are quite widely used, mostly as an initialization for further more sophisticated processing. For instance, an adaptive C-means algorithm was used in [69] in order to take the partial volume effect into account in a deformable model approach. An original fuzzy classification method taking spatial context into account was also used as the initialization of a deformable model in [38] for segmenting brain tumors of different types, shapes and locations in 3D MRI. Some results are shown in Sect. 6.

2.3 Local operations for filtering or edge detection

In this section, we summarize the main techniques for local filtering in a broad sense, aiming at enhancing the contrast of an image, at suppressing noise, at extracting contours, etc. Note that these aims are different and often contradicting each other. However, the principles of the techniques are similar, and they can be grouped into two classes: techniques based on functional optimization on the one hand, and rule based techniques on the other hand. These aspects have been largely developed in the literature (see e.g. [2, 6, 42, 67]), and we provide here just the main lines.

Functional approaches consist in minimizing or maximizing a functional, which can be interpreted as an analytical representation of some objective. For instance, enhancing the contrast of an image according to this technique amounts to reduce the fuzziness of the image. This can be performed by a simple modification of membership functions (for instance using intensification operators), by minimizing a fuzziness index such as entropy, or even by determining an optimal threshold value

(for instance optimal in the sense of minimizing a fuzziness index) which provides an extreme enhancement (until binarization) [51, 52].

Other methods consist in modifying classical filters (median filter for instance) by incorporating fuzzy weighting functions [43].

Rule based techniques rely on ideal models (of filters, contours, etc.). These ideal cases being rare, variations and differences with respect to these models are permitted through fuzzy representations of the models, as fuzzy rules. For instance, a smoothing operator can be expressed by [62, 63]:

IF	a pixel is <i>darker</i> than its neighbors
THEN	<i>increase</i> its grey level
ELSE IF	the pixel is <i>lighter</i> than its neighbors
THEN	<i>decrease</i> its grey level
OTHERWISE	keep it unchanged

In this representation, the emphasized terms are defined by fuzzy sets or fuzzy operations. Typically, the grey level characteristics are defined by linguistic variables, the semantics of which are provided by fuzzy sets on the grey level interval. Actions are fuzzy functions applied on grey levels and on pixels. The implementation of these fuzzy rules follows the general principles of fuzzy logic [30].

More complex rules can be found, for instance in [41, 56], where a contour detector is expressed by a set of rules involving the gradient, the symmetry and the stiffness of the contour. Fuzzy rule based systems have also been proposed for contour linking, based on proximity and alignment criteria.

Note that rules are sometimes but a different representation of functional approaches. Their main advantage is that they are easy to design (in particular for adaptive operators) and to interpret, and they facilitate the communication with the user.

3 Intermediate level

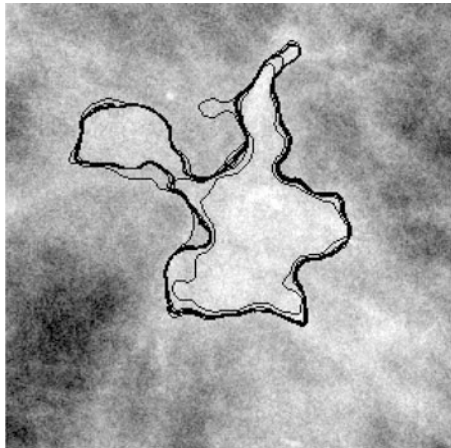
Several operations have been defined in the literature on fuzzy objects, in particular spatial fuzzy objects, since the early works of Zadeh [71] on set operations, and of Rosenfeld on geometrical operations [60].

Typical examples of geometrical operations are area and perimeter of a fuzzy object. They can be defined as crisp numbers, where the computation involves each point up to its degree of membership. But since objects are not well defined, it can also be convenient to consider that measures performed on them are imprecise too. This point of view leads to definitions as fuzzy numbers [30].

Such geometrical measures can typically be used in shape recognition, where geometrical attributes of the objects are taken into account.

As an example, fuzzy measures have been used in [58] for detecting masses in digital breast tomosynthesis. The measures are performed on detected fuzzy regions,

Fig. 4 Fuzzy particle (black lines represent the contours of the α -cuts of the fuzzy object) extracted from a digital mammography. Computing fuzzy attributes on this fuzzy object leads to a decision concerning this region. (From [57])



that are considered as candidate particles (Fig. 4). A decision concerning their recognition is performed by combining fuzzy attributes. Fuzzy decision trees can be used to this aim [20, 53].

It has been shown in [65, 66] that using fuzzy representations of digital objects allow deriving more robust measures than using crisp representations, and in particular dealing properly with the imprecision induced by the digitization process.

Such measures can also be used as descriptors for indexation and data mining applications.

Let us now consider topological features and the example of fuzzy connectivity. The degree of connectivity between two points x and y in a fuzzy object μ in a finite discrete space is defined as [60]: $c_\mu(x, y) = \max_{L_{xy}} \min_{t_i \in L_{xy}} \mu(t_i)$, where L_{xy} is any path from x to y . This definition was exploited in fuzzy connectedness notions [68], now widely used in medical image segmentation and incorporated in freely available softwares such as ITK¹.

Morphological operations have also been defined on fuzzy objects (see e.g. [17]). We give here general definitions, for fuzzy erosion and dilation, from which several other morphological operations can be derived:

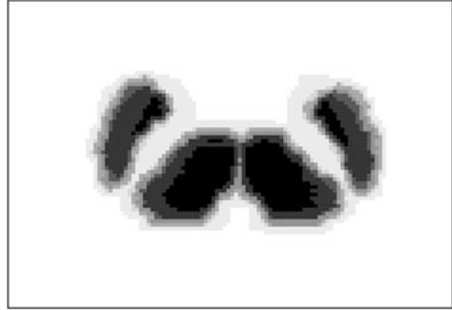
$$\forall x \in \mathcal{S}, E_v(\mu)(x) = \inf_{y \in \mathcal{S}} T[c(v(y-x)), \mu(y)], \quad (1)$$

$$\forall x \in \mathcal{S}, D_v(\mu)(x) = \sup_{y \in \mathcal{S}} t[v(x-y), \mu(y)]. \quad (2)$$

In these equations, μ denotes the fuzzy set to be dilated or eroded, v the fuzzy structuring element, t a conorm (fuzzy intersection), T the t-conorm (fuzzy union) associated to t with respect to the complementation c .

¹<http://www.itk.org/>

Fig. 5 Fuzzy median sets between the 18 instances of the IBSR database for four internal brain structures (thalamus and putamen in both hemispheres)



Such fuzzy morphological operations have been used in medical imaging for instance for taking into account the spatial imprecision on the location of vessel walls for 3D reconstruction of blood vessels by fusing angiographic and ultrasonic acquisitions [19]. They also constitute a good formal framework for defining fuzzy spatial relations, as will be seen in Sect. 4. Another application of fuzzy morphology is for defining median fuzzy sets and series of interpolating fuzzy sets [12], which can typically be used for representing variability based on several instances of an anatomical structures or for atlas construction. An example is illustrated in Fig. 5.

Some approaches using fuzzy rules can also be found at intermediate level. Let us just mention two examples. The first one [28] deals with the segmentation of osseous surface in ultrasound images. It uses fuzzy representations of image intensity and gradient, as well as their fusion, in rules that mimic the reasoning process of a medical expert and that include knowledge about the physics of ultrasound imaging. This approach was successfully tested on a large image data set.

The second example is completely different and fuzzy rules are used in [24] to tune the parameters of a deformable model for segmenting internal structures of the brain. This approach elegantly solves the difficult problem of parameter tuning in such segmentation methods, and proved to provide very good results on normal cases.

4 Higher level

The main information contained in the images consists of properties of the objects and of relations between objects, both being used for pattern recognition and scene interpretation purposes. Relations between objects are particularly important since they carry structural information about the scene, by specifying the spatial arrangements between objects. These relations highly support structural recognition based on models. This models can be of iconic type, as an anatomical atlas, or of symbolic type, as linguistic descriptions or ontologies. Although the use of iconic representations for normal structure recognition is well acknowledged, they remain difficult to exploit in pathological cases. Anatomical knowledge is also

available in textbooks or dedicated web sites, and is expressed mainly in linguistic form. These models involve concepts that correspond to anatomical objects, their characteristics, or the spatial relations between them. Human experts use intensively such concepts and knowledge to recognize visually anatomical structures in images. This motivates their use in computer aided image interpretation. Some attempts to formalize this knowledge has been recently performed, in particular in the form of ontologies (e.g. the Foundational Model of Anatomy [61]).

In our work, we concentrate mainly on spatial relations, which are strongly involved in linguistic descriptions. They constitute a very important information to guide the recognition of structures embedded in a complex environment, and are more stable and less prone to variability (even in pathological cases) than object characteristics such as shape or size. We proposed mathematical models of several spatial relations (adjacency, distances, directional relations, symmetry, between...) [8, 9, 10, 15, 18, 27], in the framework of fuzzy sets theory, which proved useful to recognize thoracic and brain structures [14, 16, 26]. These fuzzy representations can enrich anatomical ontologies and contribute to fill the semantic gap between symbolic concepts, as expressed in the ontology, and visual percepts, as extracted from the images. These ideas were used in particular in our segmentation and recognition methods [3, 36]: a concept of the ontology is used for guiding the recognition by expressing its semantics as a fuzzy set, for instance in the image domain or in an attribute domain, which can therefore be directly linked to image information.

The methods we develop in our group for segmentation and recognition of 3D structures in medical images can be seen as spatial reasoning processes. Two main components of this domain are spatial knowledge representation and reasoning. In particular spatial relations constitute an important part of the knowledge we have to handle, as explained before. Imprecision is often attached to spatial reasoning in images, and can occur at different levels, from knowledge to the type of question we want to answer. The reasoning component includes fusion of heterogeneous spatial knowledge, decision making, inference, recognition. Two types of questions are raised when dealing with spatial relations:

1. given two objects (possibly fuzzy), assess the degree to which a relation is satisfied;
2. given one reference object, define the area of space in which a relation to this reference is satisfied (to some degree).

In order to answer these questions and address both representation and reasoning issues, we rely on three different frameworks and their combination: (i) mathematical morphology, which is an algebraic theory that has extensions to fuzzy sets and to logical formulas, and can elegantly unify the representation of several types of relations; (ii) fuzzy set theory, which has powerful features to represent imprecision at different levels, to combine heterogeneous information and to make decisions; (iii) formal logics and the attached reasoning and inference power. The association of these three frameworks for spatial reasoning is an original contribution of our work [11].

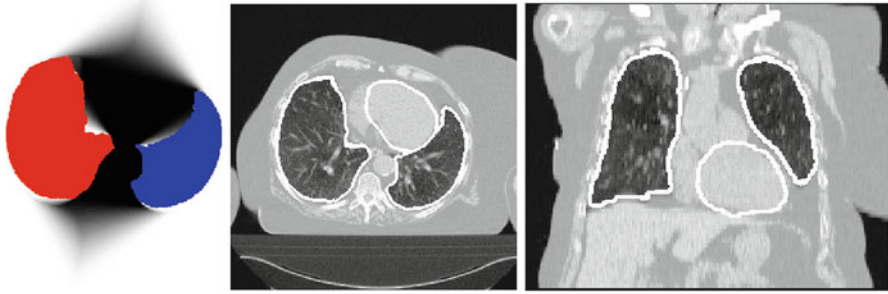


Fig. 6 Fuzzy region between the lungs, segmentation of the lungs and the heart on an axial slice and a coronal one

An example of using the second type of question was used for segmenting the heart in low resolution CT images [46], relying on the anatomical knowledge “the heart is between the lungs”. The translation of this knowledge uses an original definition of the concept “between” [15], that defines a fuzzy region of interest in which the heart can then be segmented using a deformable model integrating the spatial relation constraints, as in [26]. An example is shown in Fig. 6.

Further examples in brain imaging will be illustrated in Sect. 6.

5 Fusion

As seen in the previous sections, a lot of approaches, whatever their level, involve fusion steps.

Information fusion becomes increasingly important in medical imaging due to the multiplication of imaging techniques. The information to be combined can be issued from several images (like multi-echo MR images for instance), or from one image only, using for instance combination of several relations between objects or several features of the objects, or from images and a model, like an anatomical atlas, or knowledge expressed in linguistic form or as ontologies.

The advantages of fuzzy sets and possibilities rely in the variety of combination operators, offering a lot of flexibility in their choice, and which may deal with heterogeneous information [32, 70]. We proposed a classification of these operators with respect to their behavior (in terms of conjunctive, disjunctive, compromise [32]), the possible control of this behavior, their properties and their decisiveness, which proved to be useful for several applications in image processing [7]. It is of particular interest to note that, unlike other data fusion theories (like Bayesian or Dempster-Shafer combination), fuzzy sets provide a great flexibility in the choice of the operator, that can be adapted to any situation at hand. Indeed, image fusion has often to deal with situations where an image is reliable only for some classes, or does

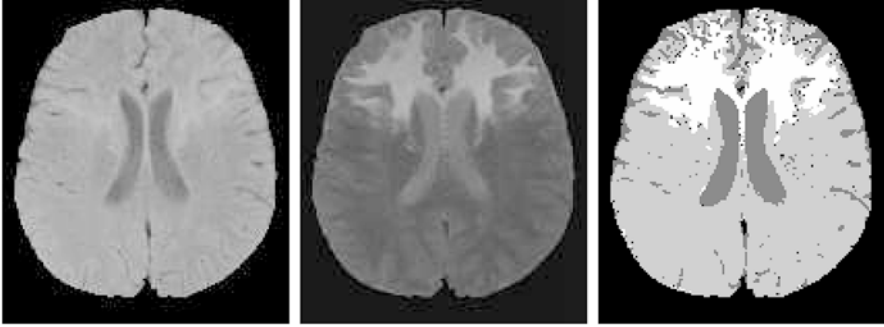


Fig. 7 Dual echo MR image of the brain, showing three main classes: brain, ventricles and pathology (the white area on the middle image). Right: final decision after fuzzy combination (note that the decision is taken at each pixel individually, without spatial regularization)

not provide any information about some class, or is not able to discriminate between two classes while another does. In this context, some operators are particularly powerful, like operators that behave differently depending on whether the values to be combined are of the same order of magnitude or not, whether they are small or high, and operators that depend on some global knowledge about source reliability about classes, or conflict between images (global or related to one particular class). The combination process can be done at several levels of information representation, from pixel level to higher level. A noticeable advantage of this approach is that it is able to combine heterogeneous information, like it is usually the case in multi-image fusion.

At a numerical level, the typical application is multi-source classification. We show an example of image fusion problem in brain imaging, where we combine dual-echo brain MR images in order to provide a classification of the brain into three classes: brain, ventricles and CSF, and pathology. These images are shown in Fig. 7. The membership functions for these classes have been estimated in a completely unsupervised way on both images, as described before. We then use these membership functions in a fuzzy fusion scheme [13]. Since both images provide similar information about the ventricles, we use a mean operator to combine the membership functions obtained in both images for this class. Brain and pathology cannot be distinguished in the first echo and we obtain only one class for this image, denoted by μ_c^1 . In the second image, we obtain two classes denoted by μ_c^2 and μ_{path}^2 respectively. We combine μ_c^1 and μ_c^2 using an arithmetical mean again. As for the pathology, we combine μ_c^1 and μ_{path}^2 using a symmetrical sum defined as: $\frac{ab}{1-a-b+2ab}$. This guarantees that no pathology is detected in the areas where $\mu_{path}^2 = 0$, and this reinforces the membership to that class otherwise, in order to include the partial volume effect areas in the pathology (this corresponds to what radiologists do). After the combination, the decision is made according to the maximum of membership values. The result is shown in Fig. 7 (right).

At a structural level, the operations defined on fuzzy objects as well as the relations between fuzzy objects can serve as a basis for structural recognition. An example will be provided in the next section.

A noticeable advantage of fuzzy fusion is that it is able to combine heterogeneous information, like is the case when dealing with higher level approaches, where several types of knowledge and information with different semantics have to be combined, and to avoid to define a more or less arbitrary and questionable metric between pieces of information.

Let us give a few examples. If we have different constraints about an object (for instance concerning the relations it should have with respect to another object) which have all to be satisfied, these constraints can be combined using a t-norm (a conjunction). If one object has to satisfy one relation or another one then a disjunction represented by a t-conorm has to be used. This occurs for instance when two symmetrical structures with respect to the reference object can be found (this situation often occurs in medical imaging). Mean operators can be used to combine several estimations and try to find a compromise between them. Associative symmetrical sums can be used for reinforcing the dynamics between high and low membership degrees. Importance of a constraint or reliabilities can be easily introduced in adaptive operators, and so on.

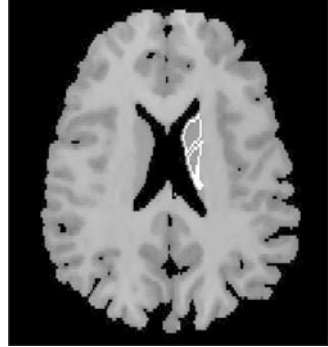
6 An application to the recognition of brain structures based on anatomical knowledge representation

Let us now illustrate how fuzzy spatial relations can be used for recognizing structures in a scene based on a model. The chosen example concerns the recognition of internal brain structures (ventricular system and grey nuclei) in 3D MRI. Two types of approaches have been developed, that correspond to the two types of questions raised in Sect. 4.

6.1 Global approach

In the first approach, which relies on the first type of question, spatial relations evaluated between spatial entities (typically objects or regions) are considered as attributes in a graph. The model is a graph derived from an anatomical atlas. Each node represents an anatomical structure, and edges represent spatial relations between these structures. A data graph is constructed from the MRI image where recognition has to be performed. Each node represents a region obtained from a segmentation method. Since it is difficult to segment directly the objects, usually the graph is based on an over-segmentation of the image, for instance based on watersheds. Attributes are computed as for the model. The use of fuzzy relations is particularly useful in order to be less sensitive to the segmentation step.

Fig. 8 Two regions of a 3D MR image, selected from an over-segmentation of the image as the ones having the best matching degree to the caudate nucleus in the atlas (only one slice is shown)



One important problem to be solved then is graph matching. Because of the schematic aspect of the model and the difficulty to segment the image into meaningful entities, no isomorphism can be expected between both graphs. In particular, several regions of the image can be assigned to the same node of the model graph. Such problems call for inexact graph matching. In general, it consists in finding a morphism, which furthermore optimizes an objective function based on similarities between attributes. Here the fusion applies not directly on the relations but on the similarities between them (see Sect. 5). A weighted mean operator allows us to give more importance to the edges, which show less variability between subjects and therefore constitute stronger anchors for guiding recognition. The morphism aims at preserving the graph structure, while the objective function privileges the association between nodes, respectively between edges, with similar attribute values. This approach can benefit from the huge literature on fuzzy comparison tools (see e.g. [21]) and from recent developments on fuzzy morphisms [54]. The optimization is not an easy task since the problem is NP-hard. Genetic algorithms, estimation of distribution algorithms and tree search methods have been developed towards this aim [4, 22, 55]. An example of recognition of the caudate nucleus is shown in Fig. 8.

Another approach consists in representing all knowledge on spatial relations between structures in a graph and expressing the joint segmentation and recognition problem as a constraint satisfaction problem [47, 48]. Propagators are defined for each spatial relation, and applied sequentially in order to progressively reduce the domain of each anatomical structure.

6.2 Sequential approach

In the second type of approach, relying on the second type of question, we use spatial representations of spatial knowledge [16, 26]. It consists in first recognizing simple structures (typically brain and lateral ventricles), and then progressively more and more difficult structures, based on relations between these structures and

previously recognized ones. The order in which structures can be recognized can be provided by the user, or estimated as suggested in [34, 35]. Each relation describing the structure to be recognized is translated into a spatial fuzzy set representing the area satisfying this relation, to some degree. The fuzzy sets representing all relations involved in the recognition process are combined using a numerical fusion operator. While we first used an atlas in [16], this constraint has been relaxed in our recent work [26, 35]. This presents two main advantages: the high computation time associated with the computation of a deformation field between the atlas and the image is left aside and the procedure is potentially more robust because it uses only knowledge expressed in symbolic form, which is generic instead of being built from a single individual as in the iconic atlas.

Finally, a refinement stage is introduced using a deformable model. This stage uses an initial classification (using a low level approach based on grey levels) as a starting point and has the potential to correct possible imperfections of the previous stage together with regularizing the contours of structures. This deformable model makes use of a fusion of heterogeneous knowledge: edge information derived from the image, regularization constraints and spatial relations contained in the linguistic description. All pieces of information are combined in the energy of a parametric deformable model. For instance the caudate nucleus can be recognized based on its grey level (roughly known depending on the type of acquisition), and, more importantly, on its relations to the lateral ventricles (exterior and close to them). Here, the primary role of spatial relations is to prevent the deformable model from progressing beyond the limit of structures with weak boundaries.

Figure 9 shows 3D views of some cerebral objects recognized in an MR image with our method. In particular, the importance of spatial relations is illustrated in the case of the caudate nucleus. The lower part of this structure has a very weakly defined boundary and the use of a spatial relation is essential to achieve a good segmentation.

One of the advantages of this approach is that it can be extended to pathological cases, since spatial relations remain quite stable in the presence of pathologies, unlike shapes and absolute locations. Moreover, it is possible to learn the parameters of the relations, and their stability according to the type of pathology [3, 37]. Two examples of segmentation and recognition results in pathological cases are shown in Fig. 10, based on a segmentation of the tumor (based on fuzzy classification) [38].

7 Conclusion

In this chapter, several examples illustrating the potential of fuzzy methods for medical imaging have been described. While low level methods are still the most widely used, recently several higher level approaches were developed, based on

Fig. 9 Segmentation and recognition results obtained for the lateral ventricles, third ventricle, caudate nuclei and thalami by integrating spatial relations in 3D deformable models. Illustration of the importance of spatial relations in the deformable model: in the case of caudate nucleus, the force derived from spatial relations prevents the model to grow below the lower limit of the structure (left: result obtained without this force, right: with this force)

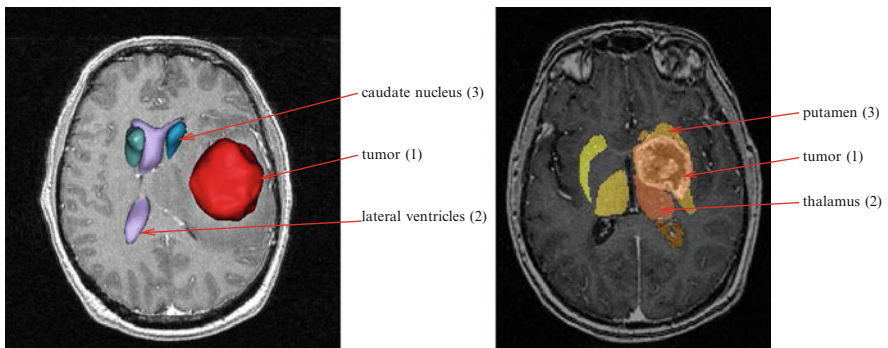
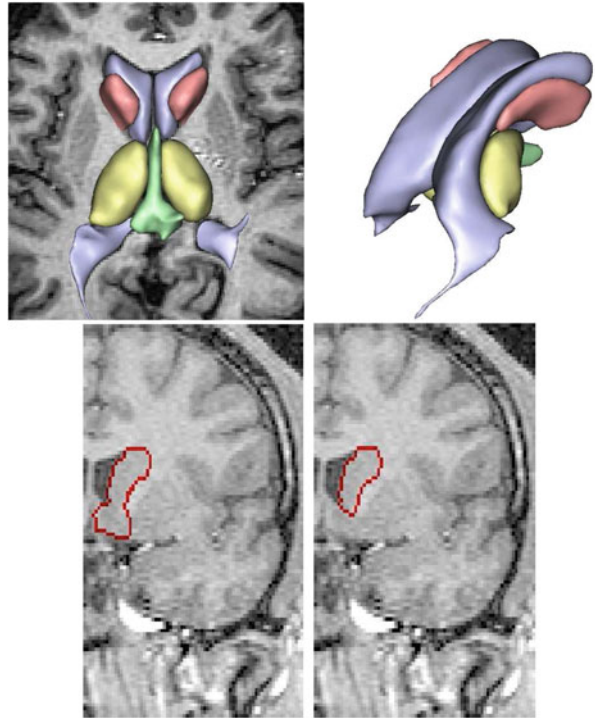


Fig. 10 Examples of segmentation and recognition in pathological cases

a rigorous and powerful mathematical basis. The association of a mathematical framework for modeling imprecision at different levels and of artificial intelligence methods for representing concepts and knowledge and for reasoning on them seems to be a very interesting current trend, where promising results are expected in a near future.

References

1. M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 21(3):193–199, March 2002.
2. K. Arakawa. Fuzzy Rule-Based Image Processing with Optimization. In E. E. Kerre and M. Nachtgaele, editors, *Fuzzy Techniques in Image Processing*, Studies in Fuzziness and Soft Computing, chapter 8, pages 222–247. Physica-Verlag, Springer, 2000.
3. J. Atif, C. Hudelot, G. Fouquier, I. Bloch, and E. Angelini. From Generic Knowledge to Specific Reasoning for Medical Image Interpretation using Graph-based Representations. In *International Joint Conference on Artificial Intelligence IJCAI'07*, pages 224–229, Hyderabad, India, jan 2007.
4. E. Bengoetxea, P. Larranaga, I. Bloch, A. Perchant, and C. Boeres. Inexact Graph Matching by Means of Estimation of Distribution Algorithms. *Pattern Recognition*, 35:2867–2880, 2002.
5. J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum, New-York, 1981.
6. J. C. Bezdek, J. Keller, R. Krishnapuram, and N. R. Pal. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Handbooks of Fuzzy Sets series. Kluwer Academic Publisher, Boston, 1999.
7. I. Bloch. Information Combination Operators for Data Fusion: A Comparative Review with Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.
8. I. Bloch. Fuzzy Relative Position between Objects in Image Processing: a Morphological Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):657–664, 1999.
9. I. Bloch. On Fuzzy Distances and their Use in Image Processing under Imprecision. *Pattern Recognition*, 32(11):1873–1895, 1999.
10. I. Bloch. Fuzzy Spatial Relationships for Image Processing and Interpretation: A Review. *Image and Vision Computing*, 23(2):89–110, 2005.
11. I. Bloch. Spatial Reasoning under Imprecision using Fuzzy Set Theory, Formal Logics and Mathematical Morphology. *International Journal of Approximate Reasoning*, 41:77–95, 2006.
12. I. Bloch. Fuzzy Skeleton by Influence Zones - Application to Interpolation between Fuzzy Sets. *Fuzzy Sets and Systems*, 159:1973–1990, 2008.
13. I. Bloch, L. Aurdal, D. Bijno, and J. Müller. Estimation of Class Membership Functions for Grey-Level Based Image Fusion. In *ICIP'97*, volume III, pages 268–271, Santa Barbara, CA, Oct. 1997.
14. I. Bloch, O. Colliot, O. Camara, and T. Géraud. Fusion of Spatial Relationships for Guiding Recognition. Example of Brain Structure Recognition in 3D MRI. *Pattern Recognition Letters*, 26:449–457, 2005.
15. I. Bloch, O. Colliot, and R. Cesar. On the Ternary Spatial Relation Between. *IEEE Transactions on Systems, Man, and Cybernetics SMC-B*, 36(2):312–327, apr 2006.
16. I. Bloch, T. Géraud, and H. Maître. Representation and Fusion of Heterogeneous Fuzzy Information in the 3D Space for Model-Based Structural Recognition - Application to 3D Brain Imaging. *Artificial Intelligence*, 148:141–175, 2003.
17. I. Bloch and H. Maître. Fuzzy Mathematical Morphologies: A Comparative Study. *Pattern Recognition*, 28(9):1341–1387, 1995.
18. I. Bloch, H. Maître, and M. Anvari. Fuzzy Adjacency between Image Objects. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(6):615–653, 1997.
19. I. Bloch, C. Pellot, F. Sureda, and A. Herment. Fuzzy Modelling and Fuzzy Mathematical Morphology applied to 3D Reconstruction of Blood Vessels by Multi-Modality Data Fusion. In D. D. R. Yager and H. Prade, editors, *Fuzzy Set Methods in Information Engineering: A Guided Tour of Applications*, chapter 5, pages 93–110. John Wiley and Sons, New-York, 1996.

20. S. Bothorel, B. Bouchon Meunier, and S. Muller. A fuzzy logic based approach for semi-ological analysis of microcalcifications in mammographic images. *International Journal of Intelligent Systems*, 12(11-12):819–848, 1997.
21. B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards General Measures of Comparison of Objects. *Fuzzy Sets and Systems*, 84(2):143–153, Sept. 1996.
22. R. Cesar, E. Bengoetxea, and I. Bloch. Inexact Graph Matching using Stochastic Optimization Techniques for Facial Feature Recognition. In *International Conference on Pattern Recognition ICPR 2002*, volume 2, pages 465–468, Québec, aug 2002.
23. H. D. Cheng and J. R. Chen. Automatically Determine the Membership Function based on the Maximum Entropy Principle. In *2nd Annual Joint Conf. on Information Sciences*, pages 127–130, Wrightsville Beach, NC, 1995.
24. C. Ciofalo and C. Barillot. Brain Segmentation with Competitive Level Sets and Fuzzy Control. In *19th International Conference on Information Processing in Medical Imaging, IPMI 2005*, Glenwood Springs, CO, USA, 2005.
25. M. R. Civanlar and H. J. Trussel. Constructing Membership Functions using Statistical Data. *Fuzzy Sets and Systems*, 18:1–13, 1986.
26. O. Colliot, O. Camara, and I. Bloch. Integration of Fuzzy Spatial Relations in Deformable Models - Application to Brain MRI Segmentation. *Pattern Recognition*, 39:1401–1414, 2006.
27. O. Colliot, A. Zuzikov, R. Cesar, and I. Bloch. Approximate Reflectional Symmetries of Fuzzy Objects with an Application in Model-Based Object Recognition. *Fuzzy Sets and Systems*, 147:141–163, 2004.
28. V. Daanen, J. Tonetti, and J. Troccaz. A Fully Automated Method for the Delineation of Osseous Interface in Ultrasound Images. In *MICCAI*, volume LNCS 3216, pages 549–557, 2004.
29. B. B. Devi and V. V. S. Sarma. Estimation of Fuzzy Memberships from Histograms. *Information Sciences*, 35:43–59, 1985.
30. D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New-York, 1980.
31. D. Dubois and H. Prade. Unfair Coins and Necessity Measures: Towards a Possibilistic Interpretation of Histograms. *Fuzzy Sets and Systems*, 10(1):15–20, 1983.
32. D. Dubois and H. Prade. A Review of Fuzzy Set Aggregation Connectives. *Information Sciences*, 36:85–121, 1985.
33. Y. Feng and W. Chen. Brain MR image segmentation using fuzzy clustering with spatial constraints based on markov random field theory. In *Second International Workshop on Medical Imaging and Augmented Reality (MIAR)*, volume 3150 of *Lecture Notes in Computer Science*, pages 188–195, 2004.
34. G. Fouquier, J. Atif, and I. Bloch. Local Reasoning in Fuzzy Attributes Graphs for Optimizing Sequential Segmentation. In *6th IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition, GbR'07*, volume LNCS 4538, pages 138–147, Alicante, Spain, jun 2007.
35. G. Fouquier, J. Atif, and I. Bloch. Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations. *Computer Vision and Image Understanding*, 116(1):146–165, Jan. 2012.
36. C. Hudelot, J. Atif, and I. Bloch. Fuzzy Spatial Relation Ontology for Image Interpretation. *Fuzzy Sets and Systems*, 159:1929–1951, 2008.
37. H. Khotanlou, J. Atif, E. Angelini, H. Duffau, and I. Bloch. Adaptive Segmentation of Internal Brain Structures in Pathological MR Images Depending on Tumor Types. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 588–591, Washington DC, USA, apr 2007.
38. H. Khotanlou, O. Colliot, J. Atif, and I. Bloch. 3D Brain Tumor Segmentation in MRI Using Fuzzy Classification, Symmetry Analysis and Spatially Constrained Deformable Models. *Fuzzy Sets and Systems*, 160:1457–1473, 2009.
39. G. J. Klir and B. Parviz. Probability-Possibility Transformations: A Comparison. *International Journal of General Systems*, 21:291–310, 1992.

40. R. Krishnapuram and J. M. Keller. A Possibilistic Approach to Clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98–110, 1993.
41. T. Law, H. Itoh, and H. Seki. Image Filtering, Edge Detection and Edge Tracing using Fuzzy Reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:481–491, 1996.
42. C. S. Lee and Y. H. Kuo. Adaptive Fuzzy Filter and its Applications to Image Enhancement. In E. E. Kerre and M. Nachtgeael, editors, *Fuzzy Techniques in Image Processing*, Studies in Fuzziness and Soft Computing, chapter 6, pages 172–193. Physica-Verlag, Springer, 2000.
43. C. S. Lee, Y. H. Kuo, and P. T. Yu. Weighted Fuzzy Mean Filters for Image Processing. *Fuzzy Sets and Systems*, 89:157–180, 1997.
44. A. W. C. Liew and H. H. Yan. An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation. *IEEE Transactions on Medical Imaging*, 22(9):1063–1075, September 2003.
45. L. Ma and R. C. Staunton. A modified fuzzy c-means image segmentation algorithm for use with uneven illumination patterns. *Pattern Recognition*, 40(11):3005–3011, 2007.
46. A. Moreno, C. M. Takemura, O. Colliot, O. Camara, and I. Bloch. Using Anatomical Knowledge Expressed as Fuzzy Constraints to Segment the Heart in CT images. *Pattern Recognition*, 41:2525–2540, 2008.
47. O. Nempont, J. Atif, E. Angelini, and I. Bloch. Structure Segmentation and Recognition in Images Guided by Structural Constraint Propagation. In *European Conference on Artificial Intelligence ECAI*, pages 621–625, Patras, Greece, jul 2008.
48. O. Nempont, J. Atif, and I. Bloch. A constraint propagation approach to structural model based image segmentation and recognition. *Information Sciences*, 246:1–27, 2013.
49. N. R. Pal, K. Pal, and J. C. Bezdek. A mixed c-Mean clustering model. In *IEEE International Conference on Fuzzy Systems*, volume 1, pages 11–21, July 1997.
50. N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek. A possibilistic fuzzy c-means clustering algorithm. *IEEE Transactions on Fuzzy Systems*, 13(4):517–530, Aug. 2005.
51. S. K. Pal, R. A. King, and A. A. Hashim. Automatic Grey-Level Thresholding through Index of Fuzziness and Entropy. *Pattern Recognition Letters*, 1:141–146, 1983.
52. S. K. Pal and A. Rosenfeld. Image Enhancement and Thresholding by Optimization of Fuzzy Compactness. *Pattern Recognition Letters*, 7:77–86, 1988.
53. G. Palma, G. Peters, S. Muller, and I. Bloch. Masses Classification using Fuzzy Active Contours and Fuzzy Decision Trees. In *SPIE Medical Imaging: Computer-Aided Diagnosis*, number 6915, San Diego, CA, USA, feb 2008.
54. A. Perchant and I. Bloch. Fuzzy Morphisms between Graphs. *Fuzzy Sets and Systems*, 128(2):149–168, 2002.
55. A. Perchant, C. Boeres, I. Bloch, M. Roux, and C. Ribeiro. Model-based Scene Recognition Using Graph Fuzzy Homomorphism Solved by Genetic Algorithm. In *Gbr'99 2nd International Workshop on Graph-Based Representations in Pattern Recognition*, pages 61–70, Castle of Haendorf, Austria, 1999.
56. O. Perez-Oramas. *Contribution à une méthodologie d'intégration de connaissances pour le traitement d'images. Application à la détection de contours par règles linguistiques floues*. PhD thesis, Université de Nancy, 2000.
57. G. Peters. *Computer-Aided Detection for Digital Breast Tomosynthesis*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, ENST2007E012, jun 2007.
58. G. Peters, S. Muller, S. Bernard, R. Iordache, and I. Bloch. Reconstruction-Independent 3D CAD for Mass Detection in Digital Breast Tomosynthesis using Fuzzy Particles. In *SPIE Medical Imaging*, volume 6147, San Diego, CA, USA, feb 2006.
59. D. L. Pham. Spatial models for fuzzy clustering. *Computer Vision and Image Understanding*, 84(2):285–297, November 2001.
60. A. Rosenfeld. The Fuzzy Geometry of Image Subsets. *Pattern Recognition Letters*, 2:311–317, 1984.
61. C. Rosse and J. L. V. Mejino. A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*, 36:478–500, 2003.
62. F. Russo and G. Ramponi. Introducing the Fuzzy Median Filter. In *Signal Processing VII: Theories and Applications*, pages 963–966, 1994.

63. F. Russo and G. Ramponi. An Image Enhancement Technique based on the FIRE Operator. In *IEEE Int. Conf. on Image Processing*, volume I, pages 155–158, Washington DC, 1995.
64. S. Shen, W. Sandham, M. Granat, and A. Sterr. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine*, 9(3):459–467, 2005.
65. N. Sladoje and J. Lindblad. Representation and Reconstruction of Fuzzy Disks by Moments. *Fuzzy Sets and Systems*, 158(5):517–534, 2007.
66. N. Sladoje, I. Nyström, and P. K. Saha. Perimeter and Area Estimations of Digitized Objects with Fuzzy Borders. In *DGCI 2003 LNCS 2886*, pages 368–377, Napoli, Italy, 2003.
67. H. R. Tizhoosh. Fuzzy Image Enhancement: An Overview. In E. E. Kerre and M. Nachtegael, editors, *Fuzzy Techniques in Image Processing*, Studies in Fuzziness and Soft Computing, chapter 5, pages 137–171. Physica-Verlag, Springer, 2000.
68. J. K. Udupa and S. Samarasekera. Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. *Graphical Models and Image Processing*, 58(3):246–261, 1996.
69. C. Xu, D. Pham, M. Rettmann, D. Yu, and J. Prince. Reconstruction of the human cerebral cortex from magnetic resonance images. *IEEE Transactions on Medical Imaging*, 18(6):467–480, June 1999.
70. R. R. Yager. Connectives and Quantifiers in Fuzzy Sets. *Fuzzy Sets and Systems*, 40:39–75, 1991.
71. L. A. Zadeh. The Concept of a Linguistic Variable and its Application to Approximate Reasoning. *Information Sciences*, 8:199–249, 1975.

Curve Propagation, Level Set Methods and Grouping

N. Paragios

Abstract Image segmentation and object extraction are among the most well addressed topics in computational vision. In this chapter we present a comprehensive tutorial of level sets towards a flexible frame partition paradigm that could integrate edge-drive, regional-based and prior knowledge to object extraction. The central idea behind such an approach is to perform image partition through the propagation planar curves/surfaces. To this end, an objective function that aims to account for the expected visual properties of the object, impose certain smoothness constraints and encode prior knowledge on the geometric form of the object to be recovered is presented. Promising experimental results demonstrate the potential of such a method.

1 Introduction

Image segmentation has been a long term research initiative in computational vision. Extraction of prominent edges [14] and discontinuities between in-homogeneous image regions was the first attempt to address segmentation. Statistical methods that aim to separate regions according to their visual characteristics was an attempt to better address the problem [11], while the snake/active contour model [16] was a breakthrough in the the domain.

Objects are represented using parametric curves and segmentation is obtained through the deformation of such a curve towards the lowest potential of an objective function. Data-driven as well as internal smoothness terms were the components of such a function. Such a model refers to certain limitations like, the initial conditions, the parameterisation of the curve, the ability to cope with structures with multiple components, and the estimation of curve geometric properties.

N. Paragios (✉)
Center for Visual Computing, Department of Applied Mathematics,
Ecole Centrale Paris, Paris, France
e-mail: nikos.paragios@ecp.fr

Balloon models [8] where a first attempt to make the snake independent with respect to the initial conditions, while the use of regional terms forcing visual homogeneity [45] was a step further towards this direction. Prior knowledge was also introduced at some later point [37] through a learning stage of the snake coefficients. Geometric alternatives to snakes [3] like the geodesic active contour model [4] were an attempt to eliminate the parameterisation issue.

Curves are represented in an implicit manner through the level set method [24]. Such an approach can handle changes of topology and provide sufficient support to the estimation of the interface geometric properties. Furthermore, the use of such a space as an optimisation framework [44], and the integration of visual cues of different nature [25] made these approaches quite attractive to numerous domains [23]. One can also point recent successful attempts to introduce prior knowledge [19, 32] within the level set framework leading to efficient object extraction and tracking methods [33].

To conclude, curve propagation is an established technique to perform object extraction and image segmentation. Level set methods refer to a geometric alternative of curve propagation and have proven to be a quite efficient optimisation space to address numerous problems of computational vision. In this chapter, first we present the notion of curve optimisation in computer vision, then establishes a connection with the level set method and conclude with the introduction of ways to perform segmentation using edge-driven, statistical clustering and prior knowledge terms.

2 On the Propagation of Curves

Let us consider a planar curve $\Gamma : [0, 1] \rightarrow \mathcal{R} \times \mathcal{R}$ defined at a plane Ω . The most general form of the snake model consists of:

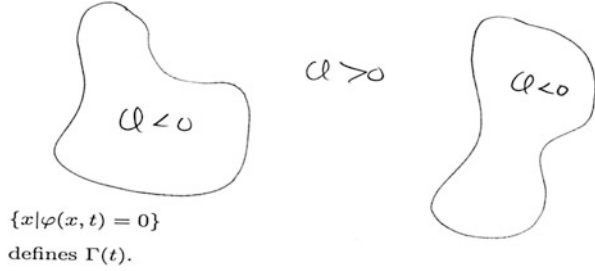
$$E(\Gamma) = \int_0^1 (\alpha E_{int}(\Gamma(p)) + \beta E_{img}(\mathcal{I}(\Gamma(p))) + \gamma E_{ext}(\Gamma(p))) dp \quad (1)$$

where \mathcal{I} is the input image, $E_{int} [=w_1|\Gamma'| + w_2|\Gamma''|]$ imposes smoothness constraints (smooth derivatives), $E_{img} [= -|\nabla \cdot \mathcal{I}|]$ makes the curve to be attracted from the image features (strong edges), E_{ext} encodes either user interaction or prior knowledge and α, β, γ are coefficients that balance the importance of these terms.

The calculus of variations can be used to optimise such a cost function. To this end, a certain number of control points are selected along the curve, and their positions are updated according to the partial differential equation that is recovered through the derivation of $E(\Gamma)$ at a given control point of Γ . In the most general case a flow of the following nature is recovered:

$$\Gamma(p; \tau) = \underbrace{(\alpha F_{gm}(\Gamma) + \beta F_{img}(\mathcal{I}) + \gamma F_{pr}(\Gamma))}_{F} \mathcal{N} \quad (2)$$

Fig. 1 Level set method and tracking moving interfaces; the construction of the (implicit) ϕ function [figure is courtesy of S. Osher]



where \mathcal{N} is the inward normal and F_{gm} depends on the spatial derivatives of the curve, the curvature, etc. On the other hand, F_{img} is the force that connects the propagation with the image domain and $F_{pr}(\Gamma)$ is a speed term that compares the evolving curve with a prior and enforces similarity with such a prior. The tangential component of this flow has been omitted since it affects the internal position of the control points and doesn't change the form of the curve itself.

Such an approach refers to numerous limitations. The number and the sampling rule used to determined the position of the control points can affect the final segmentation result. The estimation of the internal geometric properties of the curve is also problematic and depends on the sampling rule. Control points move according to different speed functions and therefore a frequent re-parameterisation of the contour is required. Last, but no least the evolving contour cannot change the topology and one cannot have objects that consist of multiple components that are not connected.

2.1 Level Set Method

The level set method was first introduced in [10] and re-invented in [24] to track moving interfaces in the community of fluid dynamics and then emerged in computer vision [3, 21]. The central idea behind these methods is to represent the (closed) evolving curve Γ with an implicit function ϕ that has been constructed as follows:

$$\phi(s) = \begin{cases} 0, & s \in \Gamma \\ -\epsilon, & s \in \Gamma_{in} \\ +\epsilon, & s \in \Gamma_{out} \end{cases}$$

where *epsilon* is a positive constant, Γ_{in} the area inside the curve and Γ_{out} the area outside the curve as shown in [Fig. (1)]. Given the partial differential equation that dictates the deformation of Γ one now can derive the one for ϕ using the chain rule according to the following manner:

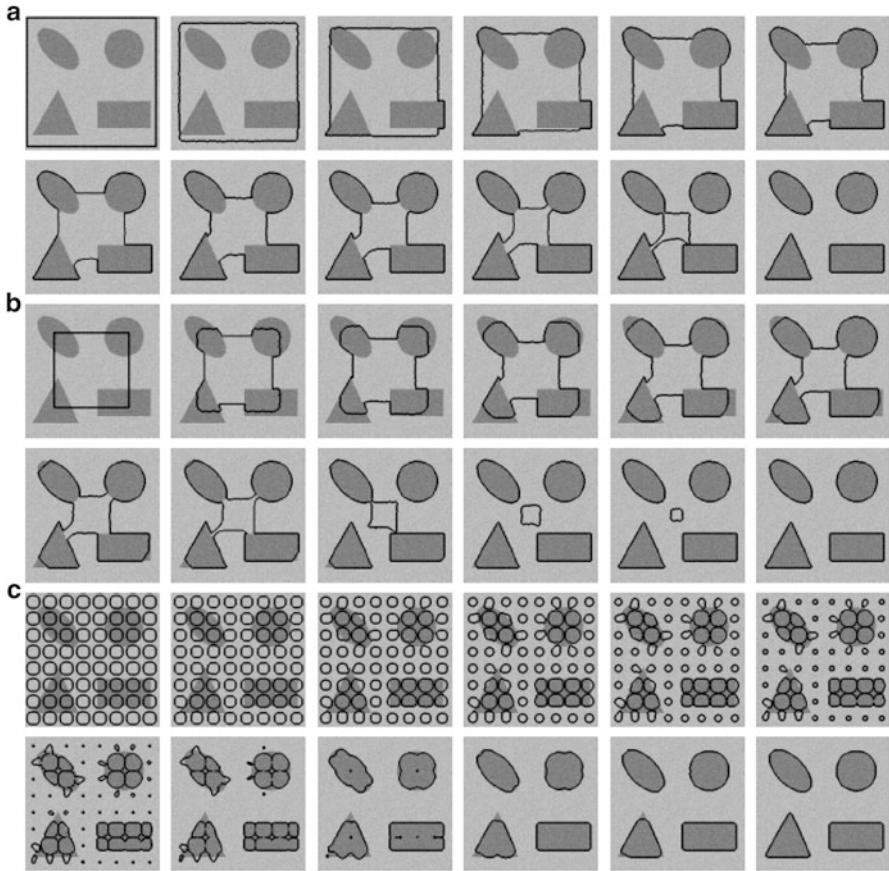


Fig. 2 Demonstration of curve propagation with the level set method; handling of topological changes is clearly illustrated through various initialization configurations (a,b,c)

$$\frac{\partial}{\partial \tau} \phi(\Gamma(p; \tau)) = \frac{\phi(\Gamma(p; \tau))}{\partial \Gamma} \underbrace{\frac{\partial \Gamma(p; \tau)}{\partial \tau}}_{F \mathcal{N}} + \frac{\partial \phi}{\partial \tau} = F(\nabla \phi \cdot \mathcal{N}) + \phi_\tau = 0 \quad (3)$$

Let us consider the arc-length parameterisation of the curve $\Gamma(c)$. The values of ϕ along the curve are 0 and therefore taking the derivative of ϕ along the curve Γ will lead to the following conditions:

$$\frac{\partial \phi(\Gamma(c))}{\partial c} = 0 \rightarrow \frac{\partial \phi}{\partial \Gamma}(\Gamma(c)) \cdot \frac{\partial \Gamma}{\partial c} = 0 \rightarrow \nabla \phi(c) \cdot \mathcal{T}(c) = 0 \quad (4)$$

where $\mathcal{T}(c)$ is the tangential vector to the contour. Therefore one can conclude that $\nabla\phi$ is orthogonal to the contour and can be used (upon normalisation) to replace the inward normal $\left[\mathcal{N} = -\frac{\nabla\phi}{|\nabla\phi|}\right]$ leading to the following condition on the deformation of ϕ :

$$-F|\phi| + \phi_\tau = 0 \rightarrow \phi_\tau = F|\phi| \quad (5)$$

Such a flow establishes a connection between the family of curves Γ that have been propagated according to the original flow and the ones recovered through the propagation of the implicit function ϕ . The resulting flow is parameter free, intrinsic, implicit and can change the topology of the evolving curve under certain smoothness assumptions on the speed function F . Last, but not least, the geometric properties of the curve like its normal and the curvature can also be determined from the level set function [24]. One can see a demonstration of such a flow in [Fig. (2)].

In practice, given a flow and an initial curve the level set function is constructed and updated according to the corresponding motion equation in all pixels of the image domain. In order to recover the actual position of the curve, the marching cubes algorithm [20] can be used that is seeking for zero-crossings. One should pay attention on the numerical implementation of such a method, in particular on the estimation of the first and second order derivatives of ϕ , where the ENO schema [24] is the one to be considered. One can refer to [36] for a comprehensive survey of the numerical approximation techniques.

In order to decrease computational complexity that is inherited through the deformation of the level set function in the image domain, the narrow band algorithm [7] was proposed. The central idea is update the level set function only within the evolving vicinity of the actual position of the curve. The fast marching algorithm [35, 40] is an alternative technique that can be used to evolve curves in one direction with known speed function. One can refer to earlier contribution in this book [Chap. 7] for a comprehensive presentation of this algorithm and its applications. Last, but not least semi-implicit formulations of the flow that guides the evolution of ϕ were proposed [12, 42] namely the additive operator splitting. Such an approach refers to a stable and fast evolution using a notable time step under certain conditions.

2.2 Optimisation and Level Set Methods

The implementation of curve propagation flows was the first attempt to use the level set method in computer vision. Geometric flows or flows recovered through the optimisation of snake-driven objective functions were considered in their implicit nature. Despite the numerous advantages of the level set variant of these flows, their added value can be seen as a better numerical implementation tool since the definition of the cost function or the original geometric flow is the core part of the

solution. If such a flow or function does not address the desired properties of the problem to be solved, its level set variant will fail. Therefore, a natural step forward for these methods was their consideration in the form of an optimisation space.

Such a framework was derived through the definition of simple indicator functions as proposed in [44] with the following behaviour

$$\delta(\phi) = \begin{cases} 0, & \phi \neq 0 \\ 1, & \phi = 0 \end{cases}, \quad \mathcal{H}(\phi) = \begin{cases} 1, & \phi > 0 \\ 0, & \phi = 0 \\ 0, & \phi < 0 \end{cases} \quad (6)$$

Once such indicator functions have been defined, an evolving interface Γ can be considered directly on the level set space as

$$\Gamma = \{s \in \Omega : \delta(\phi) = 1\} \quad (7)$$

while one can define a dual image partition using the \mathcal{H} indicator functions as:

$$\begin{aligned} \Gamma_{in} &= \{s \in \Omega : \mathcal{H}(-\phi) = 1\} \\ \Gamma_{out} &= \{s \in \Omega : \mathcal{H}(-\phi) = 0\}, \quad \Gamma_{in} \cup \Gamma_{out} = \Omega \end{aligned} \quad (8)$$

Towards continuous behaviour of the indicator function [\mathcal{H}], as well as well-defined derivatives [δ] in the entire domain a more appropriate selection was proposed in [44], namely the DIRAC and the HEAVISIDE distribution:

$$\delta_\alpha(\phi) = \begin{cases} 0 & , |\phi| > \alpha \\ \frac{1}{2\alpha} \left(1 + \cos\left(\frac{\pi\phi}{\alpha}\right)\right) & , |\phi| < \alpha \end{cases} \quad (9)$$

$$\mathcal{H}_\alpha(\phi) = \begin{cases} 1 & , \phi > \alpha \\ 0 & , \phi < -\alpha \\ \frac{1}{2} \left(1 + \frac{\phi}{\alpha} + \frac{1}{\pi} \sin\left(\frac{\pi\phi}{\alpha}\right)\right) & , |\phi| < \alpha \end{cases}$$

Such an indicator function has smooth, continuous derivatives and the following nice property:

$$\frac{\partial}{\partial \phi} \mathcal{H}_\alpha(\phi) = \delta_\alpha(\phi)$$

Last, but not least one consider the implicit function ϕ to be a signed distance transform $D(s, \Gamma)$,

$$\phi(s) = \begin{cases} 0, & s \in \Gamma \\ D(s, \Gamma), & s \in \Gamma_{in} \\ -D(s, \Gamma), & s \in \Omega - \Gamma_{in} = \Gamma_{out} \end{cases} \quad (10)$$

Such a selection is continuous and supports gradient descent minimisation techniques. On the other hand it has to be maintained, and therefore frequent re-initialisations using either the fast marching method [35] or PDE-based approaches [38] were considered. In [13] the problem was studied from a different perspective. The central idea was to derive the same speed function for all level lines - the one of the zero level set - an approach that will preserve the distance function constraint.

3 Data-driven Segmentation

The first attempt to address such task was made in [21] where a geometric flow was proposed to image segmentation. Such a flow was implemented in the level set space and aimed to evolve an initial curve towards strong edges constrained by the curvature effect. Within the last decade numerous advanced techniques have taken advantage of the level set method for object extraction.

3.1 Boundary-based Segmentation

The geodesic active contour model [4, 17] - a notable scientific contribution in the domain - consists of

$$E(\Gamma) = \int_0^1 g(|\nabla \mathcal{J}_\sigma(\Gamma(p))|) |\Gamma'(p)| dp \quad (11)$$

where \mathcal{J}_σ is the output of a convolution between the input image and a Gaussian kernel and g is a decreasing function of monotonic nature. Such a cost function seeks a minimal length geodesic curve that is attracted to the desired image features, and is equivalent with the original snake model once the second order smoothness component was removed. In [4] a gradient descent method was used to evolve an initial curve towards the lowest potential of this cost function and then was implemented using the level set method.

A more elegant approach is to consider the level set variant objective function of the geodesic active contour;

$$E(\phi) = \iint_{\Omega} \delta_\alpha(\phi(\omega)) g(|\nabla \mathcal{J}_\sigma(\omega)|) |\nabla \phi(\omega)| d\omega \quad (12)$$

where Γ is now represented in an implicit fashion with the zero-level set of ϕ . One can take the derivative of such a cost function according to ϕ :

$$\phi_\tau = \delta_\alpha(\phi) \operatorname{div} \left(g(\cdot) \frac{\nabla \phi}{|\nabla \phi|} \right) \quad (13)$$

where ω and $|\nabla \mathcal{I}_\sigma(\omega)|$ were omitted from the notation. Such a flow aims to shrink an initial curve towards strong edges. While the strength of image gradient is a solid indicator of object boundaries, initial conditions on the position of the curve can be issue. Knowing the direction of the propagation is a first drawback (the curve has either to shrink or expand), while having the initial curve either interior to the objects or exterior is the second limitation. Numerous provisions were proposed to address these limitations, some of them aimed to modify the boundary attraction term [29], while most of them on introducing global regional terms [45].

3.2 Region-based Segmentation

In [26] the first attempt to integrate edge-driven and region-based partition components in a level set approach was reported, namely the geodesic active region model. Within such an approach, the assumption of knowing the expected intensity properties (supervised segmentation) of the image classes was considered. Without loss of generality, let us assume an image partition in two classes, and let $r_{in}(\mathcal{I})$, $r_{out}(\mathcal{I})$ be regional descriptors that measure the fit between an observed intensity \mathcal{I} and the class interior $[r_{in}(\mathcal{I})]$ and exterior to $[r_{out}(\mathcal{I})]$ the curve. Under such an assumption one can derive a cost function that separates the image domain into two regions:

- according to a minimal length geodesic curve attracted by the regions boundaries,
- according to an optimal fit between the observed image and the expected properties of each class,

$$E(\phi) = w \iint_{\Omega} \delta_{\alpha}(\phi(\omega)) g(|\nabla \mathcal{I}_{\sigma}(\omega)|) |\nabla \phi(\omega)| d\omega + \iint_{\Omega} \mathcal{H}_{\alpha}(-\phi(\omega)) r_{in}(\mathcal{I}) d\omega + \iint_{\Omega} (1 - \mathcal{H}_{\alpha}(-\phi(\omega))) r_{out}(\mathcal{I}) d\omega \quad (14)$$

where w is a constant balancing the contributions of the two terms. One can see this framework as an integration of the geodesic active contour model [4] and the region-based growing segmentation approach proposed in [45]. The objective is to recover a minimal length geodesic curve positioned at the object boundaries that creates an image partition that is optimal according to some image descriptors. Taking the partial derivatives with respect to ϕ , one can recover the flow that is to be used towards such an optimal partition:

$$\phi_{\tau} = \delta_{\alpha}(\phi)(r_{in}(\mathcal{I}) - r_{out}(\mathcal{I})) + \omega \delta_{\alpha}(\phi) \operatorname{div} \left(g(\cdot) \frac{\nabla \phi}{|\nabla \phi|} \right) \quad (15)$$



Fig. 3 Multi-class image segmentation [27] through integration of edge-driven and region-based image metrics; The propagation with respect to the four different image classes as well as the final presentation result is presented

where the term $\delta_\alpha(-\phi)$ was replaced with $\delta_\alpha(\phi)$ since it has a symmetric behaviour. In [26] such descriptor function was considered to be the $-\log$ of the intensity conditional density $[p_{in}(\mathcal{I}), p_{out}(\mathcal{I})]$ for each class

$$r_{in}(\mathcal{I}) = -\log(p_{in}(\mathcal{I})), \quad r_{out}(\mathcal{I}) = -\log(p_{out}(\mathcal{I}))$$

In [34] the case of supervised image segmentation for more than two classes was considered using the frame partition concept introduced in [44]. One can also refer to other similar techniques [1]. Promising results were reported from such an approach for the case of image in [27] [Figure (3)] and for supervised texture segmentation in [28].

However, segmentation often refers to unconstrained domains of computational vision and therefore the assumption of known appearance properties for the objects to be recovered can be unrealistic. Several attempts were made to address this limitation. To this end, in [5, 43] an un-supervised region based segmentation

approach based on the Mumford-Shah [22] was proposed. The central idea behind these approaches of bi-modal [5] and tri-modal [43] segmentation was that image regions are piece-wise constant intensity-wise.

The level set variant of the Mumford-Shah [22] framework consists of minimising

$$E(\phi, \mu_{in}, \mu_{out}) = w \iint_{\Omega} \delta_{\alpha}(\phi(\omega)) |\nabla \phi(\omega)| d\omega + \iint_{\Omega} \mathcal{H}_{\alpha}(-\phi(\omega)) (\mathcal{I}(\omega) - \mu_{in})^2 d\omega + \iint_{\Omega} (1 - \mathcal{H}_{\alpha}(-\phi(\omega))) (\mathcal{I}(\omega) - \mu_{out})^2 d\omega \quad (16)$$

where both the image partition $[\phi]$ and the region descriptors $[\mu_{in}, \mu_{out}]$ for the inner and the outer region are to be recovered. The calculus of variations with respect to the curve position and the piece-wise constants can be consider to recover the lowest potential of such a function,

$$\mu_{in} = \frac{\iint_{\Omega} \mathcal{H}(-\phi)_{\mathcal{I}(\omega)} d\omega}{\iint_{\Omega} \mathcal{H}(-\phi) d\omega}, \quad \mu_{out} = \frac{\iint_{\Omega} (1 - \mathcal{H}(-\phi))_{\mathcal{I}(\omega)} d\omega}{\iint_{\Omega} (1 - \mathcal{H}(-\phi)) d\omega}, \quad (17)$$

$$\phi_{\tau} = \delta_{\alpha}(\phi) \left[((\mathcal{I}(\omega) - \mu_{in})^2 - (\mathcal{I}(\omega) - \mu_{out})^2) + w \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right]$$

Such a framework was the basis to numerous image segmentation level set approaches, while certain provisions were made to improve its performance. In [18] the simplistic Gaussian assumption of the image reconstruction term (piece-wise constant) was replaced with a non-parametric approximation density function while in [31] a vectorial unsupervised image/texture segmentation approach was proposed.

Last, but not least in [41] the same framework was extended to deal with multi-class segmentation. The most notable contribution of this approach is the significant reduction of the computational cost and the natural handling (opposite to [44]) of not forming neither vacuums nor overlapping regions. Such an approach can address the N -class partition problem, using $\log_2(N)$

level set functions.

4 Prior Knowledge

Computational vision tasks including image segmentation often refer to constrained environments. Medical imaging is an example where prior knowledge exists on the structure and the form of the objects to be recovered. One can claim that the level set method is among the most promising framework to model-free segmentation.

Introducing prior knowledge within such a framework is a natural extension that could make such level sets an adequate selection to numerous applications like object extraction, recognition, medical image segmentation, tracking, etc. In [19] a first attempt to perform knowledge-based segmentation was reported, while later numerous authors have proposed various alternatives [6, 9, 32, 39].

4.1 Average Models

Statistical representation of shapes is the first step of such an approach. Given a set of training examples, one would like to recover a representation of minimal length that can be used to reproduce the training set. To this end, all shapes of the training set should be registered to the same pose. Numerous methods can be found in the literature for shape registration, an adequate selection for building shape models in the space of implicit functions is the approach proposed in [15] where registration is addressed on this space. Without loss of generality we can assume that registration problem has been solved.

Let $S_{\mathcal{A}} = \{\phi_1, \phi_2, \dots, \phi_n\}$ be the implicit representations of n training samples according to a signed Euclidean distance transform. Simple averaging of the shape belonging to the training set can be used to determine a mean model

$$\phi_{\mathcal{M}} = \frac{1}{n} \sum_{i=1}^n \phi_i \quad (18)$$

that was considered in [9, 19, 39]. Such a model is a not an signed Euclidean implicit function, an important limitation. However, one can recover a mean model in the form of a planar curve $\Gamma_{\mathcal{M}}$ through the marching cubes algorithm [20]. Once such a model has been determined, one can impose shape prior knowledge through the constraint that the object to be recovered at the image plane Γ that is a clone of the average shape $\Gamma_{\mathcal{M}}$ according to some transformation:

$$\Gamma = \mathcal{A}(\Gamma_{\mathcal{M}}) \quad (19)$$

where \mathcal{A} can be a linear or non-linear transformation. In [6] prior knowledge has been considered in the form of a mean represented with a signed distance function. Once such a model was recovered, it was used [6] within the geodesic active contour model [4] to impose prior knowledge in the level set space:

$$E(\phi, \mathcal{A}) = \iint_{\Omega} \delta_{\alpha}(\phi) \left(g(|\nabla \mathcal{I}|) |\nabla \phi| + \lambda \phi_{\mathcal{M}}^2(\mathcal{A}(\omega)) \right) d\omega \quad (20)$$

where $\mathcal{A} = (s, \theta, (\mathcal{T}_x, \mathcal{T}_y))$ is a similarity transformation that consists of a scale factor $[s]$, a rotation component $[\theta]$ and a translation vector $(\mathcal{T}_x, \mathcal{T}_y)$. $\phi_{\mathcal{M}}$ is an implicit representation of the mean model according to a distance function and λ is a constant that determines the importance of the prior term. Such an objective

function aims at finding a minimal length geodesic curve that is attracted to the object boundaries and is not far from being a similarity transformation of the prior model:

$$\phi_{\mathcal{M}}(\mathcal{A}(\Gamma_{\mathcal{M}})) \rightarrow 0$$

Such an approach can be very efficient when modelling shapes of limited variation. On the other hand, one can claim that for shapes with important deviation from the mean model the method could fail. Furthermore, given the small number of constraints when determining the transformation between the image and the model space the estimation $[\mathcal{A}]$ could become a quite unstable task.

Towards a more stable approach to determine the optimal transformation between the evolving contour and the average model, in [32] a direct comparison between the contour implicit function and the model distance transform was used to enforce prior knowledge:

$$\phi(\omega) = \phi_{\mathcal{M}}(\mathcal{A}(\omega))$$

Despite the fact that distance transforms are robust to local deformations, invariant to translation and rotation, they are not invariant to scale variations. Slight modification of the above condition [30] could also lead to scale invariant term:

$$s\phi(\omega) = \phi_{\mathcal{M}}(\mathcal{A}(\omega))$$

The minimisation of the SSD between the implicit representations of the evolving contour and the distance transform of the average prior model can be considered to impose prior knowledge, or

$$E(\phi, \mathcal{A}) = \iint_{\Omega} \delta_{\alpha}(\phi) (s\phi(\omega) - \phi_{\mathcal{M}}(\mathcal{A}(\omega)))^2 d\omega \quad (21)$$

a term that is evaluated within the vicinity of the zero level-set contour (modulo the selection of α). The calculus of variations within a gradient descent method can provide the lowest potential of the cost function. Two unknown variables are to be recovered, the object position (form of function ϕ),

$$\frac{d}{d\tau}\phi = - \underbrace{\left[\frac{\partial}{\partial\phi} \delta_{\alpha}(\phi) \right]}_{\text{area force}} \left((s\phi - \phi_{\mathcal{M}}(\mathcal{A}))^2 - 2\delta_{\alpha}(\phi) s(s\phi - \phi_{\mathcal{M}}(\mathcal{A})) \right) \quad (22)$$

shape consistency force

This flow consists of two terms: (i) a shape consistency force that updates the interface towards a better local match with the prior and (ii) a force that aims at updating the level set values such that the region on which the objective functions is evaluated $(-\alpha, \alpha)$ becomes smaller and smaller in the image plane. In order to better

understand the influence of this force, one can consider a negative ϕ value, within the range of $(-\alpha, \alpha)$; Such a term does not change the position of the interface and therefore it could be omitted:

$$\frac{d}{d\tau}\phi = -2\delta_\alpha(\phi) s (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) \quad (23)$$

Towards recovering the transformation parameters $[\mathcal{A}]$ between the evolving contour and the average model, a gradient descent approach could be considered in parallel: \mathcal{A}

$$\begin{cases} \frac{d}{dt}\theta = 2\int_{\Omega}\delta_\epsilon(\phi) (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) (\nabla\phi_{\mathcal{M}}(\mathcal{A})) \cdot \frac{\partial}{\partial\theta}\mathcal{A} d\Omega \\ \frac{d}{dt}\mathcal{T}_x = 2\int_{\Omega}\delta_\epsilon(\phi) (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) (\nabla\phi_{\mathcal{M}}(\mathcal{A})) \cdot \frac{\partial}{\partial\mathcal{T}_x}\mathcal{A} d\Omega \\ \frac{d}{dt}\mathcal{T}_y = 2\int_{\Omega}\delta_\epsilon(\phi) (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) (\nabla\phi_{\mathcal{M}}(\mathcal{A})) \cdot \frac{\partial}{\partial\mathcal{T}_y}\mathcal{A} d\Omega \\ \frac{d}{dt}s = 2\int_{\Omega}\delta_\epsilon(\phi) (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) (-\phi + \nabla\phi_{\mathcal{M}}(\mathcal{A})) \cdot \frac{\partial}{\partial s}\mathcal{A} d\Omega \end{cases} \quad (24)$$

One can refer to very promising results - as shown in [Fig. (4)] - on objects that refer to limited shape variability using such a method [32]. However, often the object under consideration presents important shape variations that cannot be accounted for with simple average models. Decomposition and representation of the training set through linear shape spaces is the most common method to address such a limitation.

4.2 Prior Knowledge through Linear Shape Spaces

In [19] a principal component analysis on the registered set of the space of distance functions (training examples) was considered to recover a model that can account for important shape variations. Similar approach was consider in [2, 33, 39]. Principal component analysis refers to a linear transformation of variables that retains - for a given number n of operators - the largest amount of variation within the training data.

Let $\phi_{i=1\dots n}$ be a column vector representation of the training set of n implicit function elements registered to the same pose. We assume that the dimensionality of this vector is d . Using the technique introduced in [32] one can estimate a mean vector $\phi_{\mathcal{M}}$ that is part of the space of implicit functions and subtract it from the input to obtain zero mean vectors $\{\tilde{\phi}_i = \phi_i - \phi_{\mathcal{M}}\}$.

Given the set of training examples and the mean vector, one can define the $d \times d$ covariance matrix:

$$\sum_{\tilde{\phi}} = E \{\tilde{\phi}_i \tilde{\phi}_i^T\} \quad (25)$$

It is well known that the principal orthogonal directions of maximum variation are the eigenvectors of $\sum_{\tilde{\phi}}$.

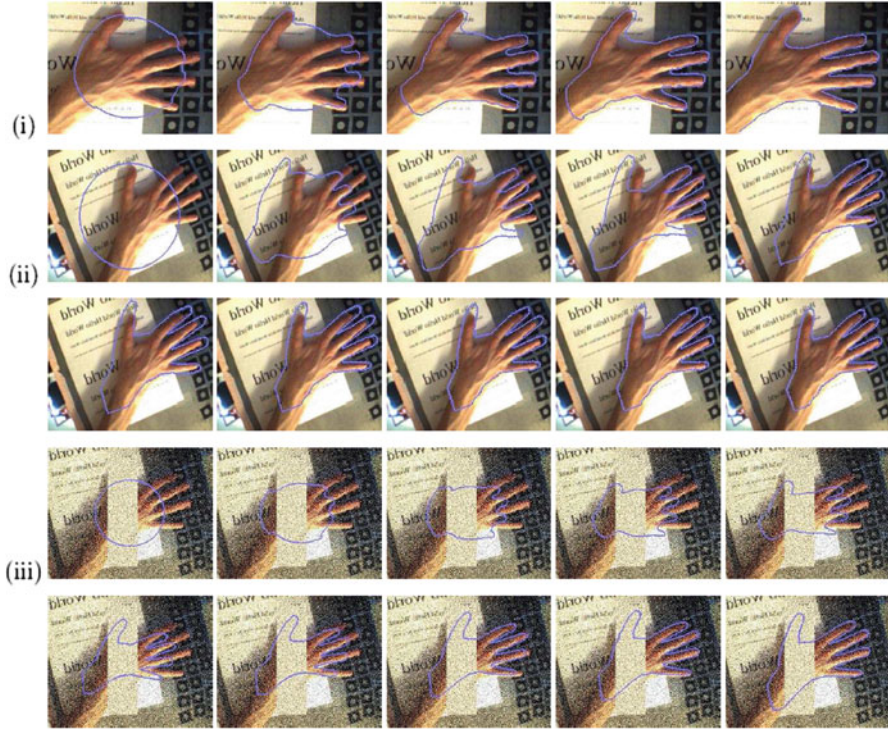


Fig. 4 Level set methods, prior knowledge, average models and similarity invariant object extraction [32] in various pose conditions (i,ii, iii)

One can approximate $\sum_{\tilde{\phi}}$ with the sample covariance matrix that is given by $[\tilde{\phi}_N \tilde{\phi}_N^T]$ where $\tilde{\phi}_N$ is the matrix formed by concatenating the set of implicit functions $\{\tilde{\phi}_i\}_{i=1\dots n}$. Then, the eigenvectors of $\sum_{\tilde{\phi}}$ can be computed through the singular value decomposition (SVD) of $\tilde{\phi}_N$:

$$\tilde{\phi}_N = \mathbf{U} \mathbf{D} \mathbf{U}^T \quad (26)$$

The eigenvectors of the covariance matrix $\sum_{\tilde{\phi}}$ are the columns of the matrix \mathbf{U} (referred to as the basis vectors henceforth) while the elements of the diagonal matrix \mathbf{D} are the square root of the corresponding eigenvalues and refer to the variance of the data in the direction of the basis vectors. Such information can be used to determine the number of basis vectors (m) required to retain a certain percentage of the variance in the data.

Then, one can consider a linear shape space that consists of the (m) basis vectors required to retain a certain percentage of the training set:

$$\phi = \phi_{\mathcal{M}} + \sum_{j=1}^m \lambda_j \mathbf{U}_j \quad (27)$$

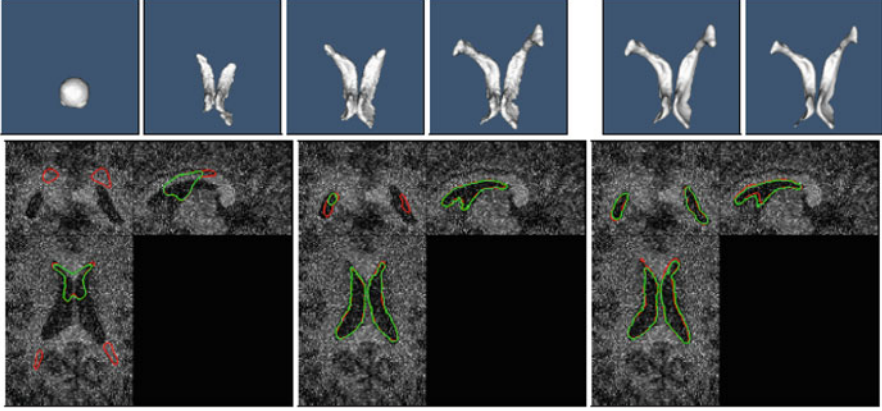


Fig. 5 Level set methods, prior knowledge, linear shape spaces and Object Extraction [33]; segmentation of lateral brain ventricles (Top Left) surface evolution, (Top Right) projected surface in the learning space and ground-truth surface (from the training set), (Bottom) surface cut and its projection in the learning space during surface evolution

Such linear space can now be used as prior model that refers to a global transformation \mathcal{A} of the average model $\phi_{\mathcal{M}}$ and its local deformation $\lambda = (\lambda_1, \dots, \lambda_m)$ through a linear combination of the the basis vectors \mathbf{U}_j . Then, object extraction is equivalent with finding a shape for which there exists such a transformation that will map each value of current representation to the “best” level set representation belonging to the class of the training shapes:

$$E(\phi, \mathcal{A}, \lambda) = \int_{\Omega} \delta_{\epsilon}(\phi) \left(s\phi - \left(\phi_{\mathcal{M}}(\mathcal{A}) + \sum_{j=1}^m \lambda_j \mathbf{U}_j(\mathcal{A}) \right) \right)^2 d\Omega \quad (28)$$

where the rotation factor $\mathbf{U}_j(\mathcal{A})$ has to be accounted for when applying the principal modes of variations to deform the average shape.

In order to minimise the above functional with respect to the evolving level set representation, the global linear transformation \mathcal{A} and the modes weights λ_j , we use the calculus of variations. The deformation of ϕ is guided by a flow similar to (1.22) that is also the case with respect to the pose parameters \mathcal{A} as shown in (). Last, but not least the differentiation with respect to the coefficients $\lambda = (\lambda_1, \dots, \lambda_m)$ leads to a linear system that has a closed form solution $\bar{\mathbf{V}}\lambda = b$ with:

$$\begin{cases} \bar{\mathbf{V}}(i, j) = \int_{\Omega} \delta_{\epsilon}(\phi) \mathbf{U}_i(\mathcal{A}) \mathbf{U}_j(\mathcal{A}) \\ b(i) = \int_{\Omega} \delta_{\epsilon}(\phi) (s\phi - \phi_{\mathcal{M}}(\mathcal{A})) \mathbf{U}_i(\mathcal{A}) \end{cases} \quad (29)$$

where $\bar{\mathbf{V}}$ is a $m \times m$ positive definite matrix. Such an approach as shown in [Fig. (5)] - can cope with important shape variations under the assumption that the distribution of the training set is Gaussian and therefore its PCA is valid.

5 Discussion

In this chapter, we have presented an approach to object extraction through the level set method that is implicit, intrinsic, parameter free and can account for topological changes. First, we have introduced a connection between the active contours, propagation of curves and their level set implementation. Then, we have considered the notion of implicit functions to represent shapes and define objective functions in such spaces to perform object extraction and segmentation. Edge-driven as well as global statistical-based region-defined segmentation criteria were presented. In the last part of the chapter we have presented prominent techniques to account for prior knowledge on the object to be recovered. To this end, we have introduced constraints of increasing complexity proportional to the spectrum of expected shape deformations that constrains the evolving interface according to the prior knowledge. Therefore one can conclude that the level set method is an efficient technique to address object extraction, is able to deal with important shape deformations, topological changes, can integrate visual cues of different nature and can account for corrupted, incomplete and occluded data.

References

1. O. Amadiou, E. Debreuve, M. Barlaud, and G. Aubert. Inward and Outward Curve Evolution Using Level Set Method. In *IEEE International Conference on Image Processing*, volume III, pages 188–192, 1999.
2. X. Bresson, P. Vanderghelynst, and J. Thiran. A Priori Information in Image Segmentation: Energy Functional based on Shape Statistical Model and Image Information. In *IEEE International Conference on Image Processing*, volume 3, pages 428–428, Barcelona, Spain, 2003.
3. V. Caselles, F. Catté, B. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.
4. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. In *IEEE International Conference in Computer Vision*, pages 694–699, 1995.
5. T. Chan and L. Vese. An Active Contour Model without Edges. In *International Conference on Scale-Space Theories in Computer Vision*, pages 141–151, 1999.
6. Y. Chen, H. Thiruvankadam, H. Tagare, F. Huang, and D. Wilson. On the Incorporation of Shape Priors into Geometric Active Contours. In *IEEE Workshop in Variational and Level Set Methods*, pages 145–152, 2001.
7. D. Chopp. Computing Minimal Surfaces via Level Set Curvature Flow. *Journal of Computational Physics*, 106:77–91, 1993.
8. L. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53:211–218, 1991.
9. D. Cremers, N. Sochen, and C. Schnorr. Multiphase Dynamic Labeling for Variational Recognition-driven Image Segmentation. In *European Conference on Computer Vision*, pages 74–86, Prague, Czech Republic, 2004.
10. A. Dervieux and F. Thomasset. A finite element method for the simulation of rayleigh-taylor instability. *Lecture Notes in Mathematics*, 771:145–159, 1979.
11. S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

12. R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast Geodesic Active Contours. *IEEE Transactions on Image Processing*, 10:1467–1475, 2001.
13. J. Gomes and O. Faugeras. Reconciling distance functions and level sets. *Journal of Visual Communication and Image Representation*, 11:209–223, 2000.
14. R. Haralick. Digital step edges from zero crossing of second directional derivatives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:58–68, 1984.
15. X. Huang, N. Paragios, and D. Metaxas. Registration of Structures in Arbitrary Dimensions: Implicit Representations, Mutual Information & Free-Form Deformations. Technical Report DCS-TR-0520, Division of Computer & Information Science, Rutgers University, 2003.
16. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. In *IEEE International Conference in Computer Vision*, pages 261–268, 1987.
17. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *IEEE International Conference in Computer Vision*, pages 810–815, 1995.
18. J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky. Non-Parametric Methods for Image Segmentation using Information Theory and Curve Evolution. In *IEEE International Conference on Image Processing*, 2002.
19. M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1:316–322, 2000.
20. W. Lorensen and H. Cline. Marching cubes: a high resolution 3D surface construction algorithm. In *ACM SIGGRAPH*, volume 21, pages 163–170, 1987.
21. R. Malladi, J. Sethian, and B. Vemuri. Evolutionary fronts for topology independent shape modeling and recovery. In *European Conference on Computer Vision*, pages 1–13, 1994.
22. D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 22–26, 1985.
23. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
24. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed : Algorithms based on the Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.
25. N. Paragios. *Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision*. PhD thesis, I.N.R.I.A./ University of Nice-Sophia Antipolis, 2000. <http://www.inria.fr/RRRT/TU-0636.html>.
26. N. Paragios and R. Deriche. A PDE-based Level Set approach for Detection and Tracking of moving objects. In *IEEE International Conference in Computer Vision*, pages 1139–1145, 1998.
27. N. Paragios and R. Deriche. Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:266–280, 2000.
28. N. Paragios and R. Deriche. Geodesic Active Regions: A New Framework to Deal with Frame Partition Problems in Computer Vision. *Journal of Visual Communication and Image Representation*, 13:249–268, 2002.
29. N. Paragios, O. Mellina-Gottardo, and V. Ramesh. Gradient Vector Flow Fast Geodesic Active Contours. In *IEEE International Conference in Computer Vision*, pages I:67–73, 2001.
30. N. Paragios, M. Rousson, and V. Ramesh. Non-Rigid Registration Using Distance Functions. *Computer Vision and Image Understanding*, 2003. to appear.
31. M. Rousson and R. Deriche. A Variational Framework for Active and Adaptive Segmentation of Vector Valued Images. Technical Report 4515, INRIA, France, 2002.
32. M. Rousson and N. Paragios. Shape Priors for Level Set Representations. In *European Conference on Computer Vision*, pages II:78–93, Copenhagen, Denmark, 2002.
33. M. Rousson, N. Paragios, and R. Deriche. Implicit Active Shape Models for 3D Segmentation in MR Imaging. In *Medical Imaging Computing and Computer-Assisted Intervention*, 2004.
34. C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia. A Level Set Model for Image Classification. *International Journal of Computer Vision*, 40:187–197, 2000.

35. J. Sethian. A Review of the Theory, Algorithms, and Applications of Level Set Methods for Propagating Interfaces. *Cambridge University Press*, pages 487–499, 1995.
36. J. Sethian. *Level Set Methods*. Cambridge University Press, 1996.
37. L. Staib and S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:1061–1075, 1992.
38. M. Sussman, P. Smereka, and S. Osher. A Level Set Method for Computing Solutions to Incompressible Two-Phase Flow. *Journal of Computational Physics*, 114:146–159, 1994.
39. A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, A. Grimson, and A. Willsky. Model-based Curve Evolution Technique for Image Segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume I, pages 463–468, 2001.
40. J. Tsitsiklis. Efficient Algorithms for Globally Optimal Trajectories. In *33rd Conference on Decision and Control*, pages 1368–1373, 1994.
41. L. Vese and T. Chan. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *International Journal of Computer Vision*, 50:271–293, 2002.
42. J. Weickert and G. Kuhne. Fast Methods for Implicit Active Contours. In S. Osher and n. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision and Graphics*, pages 43–58. Springer, 2003.
43. A. Yezzi, A. Tsai, and A. Willsky. A Statistical Approach to Snakes for Bimodal and Trimodal Imagery. In *IEEE International Conference in Computer Vision*, pages 898–903, 1999.
44. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational Level Set Approach to Multiphase Motion. *Journal of Computational Physics*, 127:179–195, 1996.
45. S. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:884–900, 1996.

Kernel Methods in Medical Imaging

G. Charpiat, M. Hofmann, and B. Schölkopf

Abstract We introduce machine learning techniques, more specifically kernel methods, and show how they can be used for medical imaging. After a tutorial presentation of machine learning concepts and tools, including Support Vector Machine (SVM), kernel ridge regression and kernel PCA, we present an application of these tools to the prediction of Computed Tomography (CT) images based on Magnetic Resonance (MR) images.

1 Introduction

Machine learning has shown dramatic progress over the last decades, creating tools like the well-known Support Vector Machine (SVM), which have been intensively applied to many different fields and have proved their efficiency. Learning tools have often changed the whole perspective of the issues they have been applied to. For example, in computer vision, the detection of objects in images, and the automatic classification of images into categories (landscape, car, etc.) rely now most often on intensive patch-based learning, whereas it was previously commonly thought that a complete image segmentation would be required. The results of this machine learning approach are often surprisingly good, showing that under certain conditions, many tasks are much easier to solve by incorporating prior knowledge retrieved from a set of examples.

G. Charpiat (✉) • M. Hofmann • B. Schölkopf
Max Planck Institute for Biological Cybernetics, Spemannstr. 38, Tuebingen 72076, Germany
e-mail: Guillaume.Charpiat@inria.fr; hofmann@petmr.com;
bernhard.schoelkopf@tuebingen.mpg.de

In medical imaging, approaches are often *example*-based, in the sense that the aim often consists in the automatization of a task already performed by hand by medical people on a few examples, such as segmentation, registration, detection (of tumors, of organs) or classification. As medical imaging deals with images, there is also much inspiration to get from what has already been achieved in computer vision, in object detection [9] as well in shape priors [5].

We start here with a tutorial on machine learning techniques. We present basic concepts, and then focus on kernel methods. We introduce standard tools like kernel ridge regression, SVM and kernel PCA. Then we apply some of these tools to the case of medical image prediction, when the Magnetic Resonance scan of a patient is known and we would like to guess what the corresponding Computed Tomography scan would look like.

2 Machine learning with kernels

This section describes the central ideas of kernel methods in a nutshell by providing an overview of the basic concepts. We first state mathematically the problems of classification and regression. Then we introduce the concept of *kernel* and explain the *kernel trick* which leads to kernel PCA as well as kernel ridge regression. The last concept introduced is the *support vector* (SV), which is the basis of the SVM. We have tried to keep this tutorial as basic as possible and refer to [11] for further details.

2.1 Basics

Classification and Regression

Suppose we are given a set of m objects $(x_i)_{1 \leq i \leq m} \in \mathcal{X}^m$ with labels $(y_i)_{1 \leq i \leq m} \in \mathcal{Y}^m$. If the number of possible labels is finite and small, then we can be interested in *classification*, i.e. in finding the label to assign to a new object based on the given examples. Otherwise, if the labels are values in a vector space, we can be interested in *regression*, i.e. in extrapolating previously observed values to any new object. Thus classification and regression tasks can be embedded in a similar framework, one aiming to predict discrete labels and the other one to continuous values.

The objects $(x_i)_{1 \leq i \leq m}$ are often named *patterns*, or *cases*, *inputs*, *instances*, or *observations*. The $(y_i)_{1 \leq i \leq m}$ are called *labels*, or *targets*, *outputs* or sometimes also *observations*. The set of all correspondences $(x_i, y_i)_{1 \leq i \leq m}$ given as examples is called the *training set*, whereas we name *test set* the set of new objects for which we would like to guess the label by extracting knowledge from the examples in the training set.

In both cases, classification or regression, we aim to generalize the correspondences (x_i, y_i) to a function f defined on the set \mathcal{X} of all possible objects and with

values in the set \mathcal{Y} of all possible labels. The label predicted for a new test object x would then be $f(x)$. Here we have no particular assumption on the spaces \mathcal{X} and \mathcal{Y} except that \mathcal{Y} should be a vector space if we are interested in regression (in order to extrapolate continuously between any two values). But we have a strong intuitive assumption on f : it should generalize as well as possible the given examples, i.e. if x is *close* to an already observed input x_i , its output $f(x)$ should be *close* to the already observed output y_i . The whole difficulty consists in defining precisely what we mean by “close” in the spaces \mathcal{X} and \mathcal{Y} . More precisely, we need to quantify the similarity of inputs in \mathcal{X} and the cost of assigning wrong outputs in \mathcal{Y} .

Loss function

Generally, expressing a distance or similarity measure in \mathcal{Y} is easy. In the case of regression, the Euclidean distance in \mathcal{Y} is often a simple, convenient choice. However we can consider other functions than distances, provided they express the cost of assigning a wrong label. We call the *loss function* the sum of the costs (or losses) of all mistakes made when we consider a particular possible solution f and apply it to all known examples. For instance we can choose:

$$L(f, (x_i, y_i)_{1 \leq i \leq m}) = \sum_{i=1}^m \|f(x_i) - y_i\|_{\mathcal{Y}}$$

Duality between features and similarity measures

On the other hand, expressing a similarity measure in \mathcal{X} is much more difficult and lies at the core of machine learning. Either the space \mathcal{X} has been carefully chosen so that the representation of the observed objects x_i are meaningful, in the sense that their “natural” distance in \mathcal{X} (say the Euclidean distance if \mathcal{X} is a vector space) is meaningful, in which case learning will be easy; either \mathcal{X} is non-trivial and we need to choose a set of N sensible *features* (seen as a function Φ from \mathcal{X} to $\mathcal{H} = \mathbb{R}^N$), so that if we compute these *features* $\Phi(x_i)$ for each x_i , we can consider a more natural distance in the *feature space* \mathcal{H} . From a certain point of view, choosing a sensible *feature map* Φ or choosing a sensible distance in \mathcal{X} (or in the feature space \mathcal{H}) are equivalent problems, and hence equivalently hard in the general case.

Optimization problem over functions

The problem of classification or regression can be written as an optimization problem over all possible functions f : find the best function f from \mathcal{X} to \mathcal{Y} such that it minimizes

$$L(f, (x_i, y_i)_{1 \leq i \leq m}) + R(f)$$

where $R(f)$ is a regularizer constraining f to be *smooth* in some way with respect to the similarity measure chosen in \mathcal{X} . Note that we could also have restricted f to be a member of a small function space \mathcal{F} . There are very nice theoretical results concerning the function space in the kernel case (see for example Sect. 2.3 about ridge regression).

2.2 Kernels

This section aims to define kernels and to explain all facets of the concept. It is a preliminary step to the following sections dedicated to kernel algorithms themselves.

A *kernel* is any symmetric similarity measure on \mathcal{X}

$$\begin{aligned} k : \mathcal{X} \times \mathcal{X} &\rightarrow \mathbb{R} \\ (x, x') &\mapsto k(x, x'), \end{aligned}$$

that is, a symmetric function that, given two inputs x and x' , returns a real number characterizing their similarity (cf. [1, 3, 4, 7, 10]).

Kernels as inner products in the feature space

In the general case, either \mathcal{X} is not a vector space, or the natural Euclidean inner product in \mathcal{X} is not particularly relevant as a similarity measure. Most often, a set of possibly-meaningful *features* is available, and we can consequently use the *feature map*

$$\begin{aligned} \Phi : \mathcal{X} &\rightarrow \mathcal{H} \\ x &\mapsto \mathbf{x} := \Phi(x). \end{aligned}$$

Φ will typically be a nonlinear map with values in a vector space. It could for example compute products of components of the input x . We have used a bold face \mathbf{x} to denote the vectorial representation of x in the feature space \mathcal{H} . We will follow this convention throughout the chapter.

We can use the non-linear embedding of the data into the linear space \mathcal{H} via Φ to define a similarity measure from the dot product in \mathcal{H} ,

$$k(x, x') := \langle \mathbf{x}, \mathbf{x}' \rangle_{\mathcal{H}} = \langle \Phi(x), \Phi(x') \rangle_{\mathcal{H}}. \quad (1)$$

The freedom to choose the mapping Φ will enable us to design a large variety of similarity measures and learning algorithms. The transformation of x_i into $\Phi(x_i) = \mathbf{x}_i$ can be seen as a change of the inputs, i.e. as a new model of the

initial problem. However, we will see later that, in some cases, we won't need to do this transformation explicitly, which is very convenient if the number of features considered (or the dimension of \mathcal{H}) is high.

Geometrical interpretation and kernel trick

Through the definition of k , we can provide geometric interpretation of the input data:

$$\|\mathbf{x}\|_{\mathcal{H}} = \|\Phi(x)\|_{\mathcal{H}} = \sqrt{\langle \Phi(x), \Phi(x) \rangle_{\mathcal{H}}} = \sqrt{k(x, x)}$$

is the *length* (or *norm*) of \mathbf{x} in the feature space. Similarly, $k(x, x')$ computes the cosine of the angle between the vectors \mathbf{x} and \mathbf{x}' , provided they are normalized to length 1. Likewise, the distance between two vectors is computed as the length of the difference vector:

$$\|\mathbf{x} - \mathbf{x}'\|_{\mathcal{H}}^2 = \|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2\langle \Phi(x), \Phi(x') \rangle = k(x, x) + k(x', x') - 2k(x, x').$$

The interesting point is that we could consider any such similarity measure k and forget about the associated Φ : we would still be able to compute lengths, distances and angles with the only knowledge of k thanks to these formulas. This framework allows us to deal with the patterns geometrically through a understated non-linear embedding, and thus lets us study learning algorithms using linear algebra and analytic geometry. This is known as the *kernel trick*: any algorithm dedicated to Euclidean geometry involving only distances, lengths and angles can be *kernelized* by replacing all occurrences of these geometric quantities by their expressions as a function of k . Next section is dedicated to such *kernelizations*.

Examples of kernels

Let us introduce the most-commonly used kernels. They are namely: the polynomial kernel

$$k(x, x') = \langle x, x' \rangle^d,$$

and the Gaussian

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$$

for suitable choices of d and σ . Let us focus on the Gaussian case: the similarity measure $k(x, x')$ between x and x' is always positive, and is maximal when $x = x'$. All points \mathbf{x} have the same unit norm (since $k(x, x) = 1 \forall x$) and consequently the images of all points x in the associated feature space \mathcal{H} lie on the unit sphere.

Reproducing kernels as feature maps

One could wonder what is the feature map Φ which was used to build the Gaussian kernel. In fact kernel theory goes far beyond the way we introduced kernels. Let us consider any symmetric function k , not necessarily related to a feature map. Let us suppose also that k , seen as an operator, is positive definite, that is to say that for any non-zero L^2 function $\alpha : \mathcal{X} \rightarrow \mathbb{R}$:

$$\int_{\mathcal{X} \times \mathcal{X}} \alpha(x)k(x, x')\alpha(x') dx dx' > 0.$$

Note that to be able to integrate over $x \in \mathcal{X}$, we need a measure on \mathcal{X} . This measure is often thought of as a *probability measure* over \mathcal{X} , giving more weight to objects that are more likely to appear.

Then we can define from this kernel k an associated feature map by:

$$\begin{aligned} \Phi : \mathcal{X} &\rightarrow \mathcal{F}(\mathcal{X}) \\ x &\mapsto \mathbf{x} := k(x, \cdot). \end{aligned} \quad (2)$$

This image of any input x by Φ is the function

$$\begin{aligned} k(x, \cdot) : \mathcal{X} &\rightarrow \mathbb{R} \\ x' &\mapsto k(x, x'). \end{aligned} \quad (3)$$

Φ has now values in the space $\mathcal{F}(\mathcal{X})$ of functions over \mathcal{X} instead of having values in just a finite dimensioned vector space like \mathbb{R}^N .

The magic comes from Moore-Aronszajn theorem [2] which states that it is *always* possible, for any symmetric positive definite function k , to build a *reproducing kernel Hilbert space* (RKHS) $\mathcal{H} \subset \mathcal{F}(\mathcal{X})$ so that

$$\forall x, x' \in \mathcal{X}, \quad k(x, x') = \langle k(x, \cdot), k(x', \cdot) \rangle_{\mathcal{H}} = \langle \Phi(x), \Phi(x') \rangle_{\mathcal{H}}. \quad (4)$$

Because of such a property, symmetric positive definite kernels are also called *reproducing kernels*. This theorem highlights the duality between *reproducing kernels* k and *feature maps* Φ : choosing the feature space or choosing the kernel is equivalent, since one determines the other.

The Gaussian case (details)

We can make explicit the inner product on \mathcal{H} in the Gaussian case. The associated norm is

$$\|f\|_{\mathcal{H}}^2 = \int_{\mathcal{X}} \sum_{n=0}^{\infty} \frac{\sigma^{2n}}{n! 2^n} \left(\frac{d^n}{dx^n} f \right)^2(x) dx = \sum_{n=0}^{\infty} \frac{\sigma^{2n}}{n! 2^n} \left\| \frac{d^n}{dx^n} f \right\|_{L^2(\mathcal{X})}^2$$

which penalizes all fast variations of f at all derivative orders. We refer to [6] for a more general mathematical study of radial basis functions. Intuitively, consider the operator $P = e^{-\frac{\sigma^2}{2} \frac{d^2}{dx^2}} := \sum_n \frac{(-\sigma^2/2)^n}{n!} \frac{d^{2n}}{dx^{2n}}$. In the Fourier domain, it writes $e^{-\sigma^2 w^2/2}$, whereas $k(x, \cdot)$ becomes $\sigma e^{-\sigma^2 w^2/2} e^{-iwx}$. Thus $\frac{1}{\sigma} P(k(x, \cdot)) = \delta_x(\cdot)$ is a Dirac peak in the space $\mathcal{D}(\mathcal{X})$ of distributions over \mathcal{X} . The inner product $\langle f, g \rangle_{\mathcal{H}} := \langle \frac{1}{\sigma} P(f), g \rangle_{\mathcal{D}(\mathcal{X})}$ on \mathcal{H} will therefore satisfy:

$$\langle k(x, \cdot), f \rangle_{\mathcal{H}} := \left\langle \frac{1}{\sigma} P(k(x, \cdot)), f \right\rangle_{\mathcal{D}(\mathcal{X})} = \langle \delta_x, f \rangle_{\mathcal{D}(\mathcal{X})} = f(x)$$

hence, for the particular case $f = k(x', \cdot)$,

$$\langle k(x, \cdot), k(x', \cdot) \rangle_{\mathcal{H}} = k(x, x').$$

The overfitting problem

The kernel k should be chosen carefully, since it is the core of the generalization process: if the neighborhood induced by k is too small (for instance if k is a Gaussian with a tiny standard deviation σ), then we will overfit the given examples without being able to generalize to new points (which would be found very dissimilar to all examples). On the contrary, if the neighborhood is too large (for instance if k is a Gaussian with a standard deviation so huge that all examples are considered as very similar), then it is not possible to distinguish any clusters or classes.

Kernels as regularizers

We introduced initially kernels as similarity measures on the space \mathcal{X} of inputs. But with the reproducing kernel framework, the choice of a kernel implies a structure on the space of functions from \mathcal{X} to \mathbb{R} , in particular it defines a norm on this space. Consequently choosing a kernel is the same as choosing a *regularizer* on the function space.

Let us go back to the initial problem, and, for the sake of simplicity, let us consider the case where the output space \mathcal{Y} is included in \mathbb{R} . We expressed the classification or regression problem as the search for the best function f from \mathcal{X} to \mathcal{Y} that minimizes a loss plus a regularizer on f . We have found here a natural way to choose the regularizer according to the kernel, i.e. $R(f) = \|f\|_{\mathcal{H}}^2$.

A whole class of problems involving this norm can be shown [11] to have solutions in the span of functions $k(x_i, \cdot)$, i.e. solutions f have the form $f(x) = \sum_i \alpha_i k(x_i, x)$. Interestingly, this allows the reduction of the search space (the function space) to a linear, finite-dimensional subspace, while involving non-linear quantities (the kernels $k(x, x')$ or the features $\Phi(x)$).

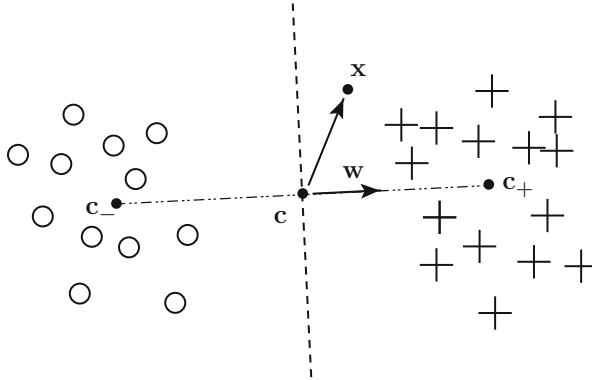


Fig. 1 A very simple classifier in the feature space: associate to any new point \mathbf{x} the class whose mean \mathbf{c}_i is the closest. The decision boundary is an hyperplane

2.3 Kernelization of existing linear algorithms

We now have all the concepts required to transform existing algorithms dealing linearly with data into kernel methods. We consider standard, simple algorithms such as PCA and linear regression and build out of them more efficient tools which take advantage of the prior knowledge provided by the definition of a kernel and of their ability to deal linearly with non-linear quantities.

A very simple hyperplanar classifier

To show the spirit of *kernelization*, let us first describe a very simple learning algorithm for binary classification. The label space \mathcal{Y} contains only two elements, $+1$ and -1 , and the training set consists of labeled examples of the two classes. The basic idea is to assign any previously unseen pattern x to the class with closest mean. Let us work directly in the feature space \mathcal{H} and deal with $\mathbf{x} = \Phi(x)$ instead of x since the metric which makes sense is the one in the feature space. In \mathcal{H} , the means of the two classes are:

$$\mathbf{c}_+ = \frac{1}{m_+} \sum_{\{i|y_i=+1\}} \mathbf{x}_i \quad \text{and} \quad \mathbf{c}_- = \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \mathbf{x}_i, \quad (5)$$

where m_+ and m_- are the number of examples with positive and negative labels, respectively. Half way between \mathbf{c}_+ and \mathbf{c}_- lies the point $\mathbf{c} := (\mathbf{c}_+ + \mathbf{c}_-)/2$. We compute the class of \mathbf{x} , based on the angle between the vector $\mathbf{x} - \mathbf{c}$ and the vector $\mathbf{w} := \mathbf{c}_+ - \mathbf{c}_-$ (see Fig. 1):

$$\begin{aligned} y &= \text{sgn} \langle (\mathbf{x} - \mathbf{c}), \mathbf{w} \rangle_{\mathcal{H}} = \text{sgn} \langle (\mathbf{x} - (\mathbf{c}_+ + \mathbf{c}_-)/2), (\mathbf{c}_+ - \mathbf{c}_-) \rangle_{\mathcal{H}} \\ &= \text{sgn} (\langle \mathbf{x}, \mathbf{c}_+ \rangle_{\mathcal{H}} - \langle \mathbf{x}, \mathbf{c}_- \rangle_{\mathcal{H}} + b) \end{aligned} \quad (6)$$

$$\text{where we have defined the offset } b := \frac{1}{2}(\|\mathbf{c}_-\|_{\mathcal{H}}^2 - \|\mathbf{c}_+\|_{\mathcal{H}}^2). \quad (7)$$

Note that (6) induces a decision boundary which has the form of a hyperplane in the feature space. We can now call the kernel trick in order to express all quantities as a function of the kernel, which is the only thing we can easily compute (unless Φ is explicit and simple). But this trick deals only with norms, distances and angles of features points of the form $\mathbf{x} = \Phi(x)$, for which we already know x . Therefore we need to express the vectors \mathbf{c}_i and \mathbf{w} in terms of $\mathbf{x}_1, \dots, \mathbf{x}_m$.

To this end, substitute (5) into (6) to get the *decision function*

$$\begin{aligned} y &= \text{sgn} \left(\frac{1}{m_+} \sum_{\{i|y_i=+1\}} \langle \mathbf{x}, \mathbf{x}_i \rangle_{\mathcal{H}} - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \langle \mathbf{x}, \mathbf{x}_i \rangle_{\mathcal{H}} + b \right) \\ &= \text{sgn} \left(\frac{1}{m_+} \sum_{\{i|y_i=+1\}} k(x, x_i) - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} k(x, x_i) + b \right). \end{aligned} \quad (8)$$

Similarly, the offset becomes

$$b := \frac{1}{2} \left(\frac{1}{m_-} \sum_{\{(i,j)|y_i=y_j=-1\}} k(x_i, x_j) - \frac{1}{m_+} \sum_{\{(i,j)|y_i=y_j=+1\}} k(x_i, x_j) \right). \quad (9)$$

Surprisingly, it turns out that this rather simple-minded approach contains a well-known statistical classification method as a special case. Assume that the class means have the same distance to the origin (hence $b = 0$, cf. (7)), and that k can be viewed as a probability density when one of its arguments is fixed. By this we mean that it is positive and that $\forall x' \in \mathcal{X}, \int_{\mathcal{X}} k(x, x') dx = 1$. In this case, (8) takes the form of the so-called Bayes classifier separating the two classes, subject to the assumption that the two classes of patterns were generated by sampling from two probability distributions that are correctly estimated by the *Parzen windows* estimators of the two class densities,

$$p_+(x) := \frac{1}{m_+} \sum_{\{i|y_i=+1\}} k(x, x_i) \text{ and } p_-(x) := \frac{1}{m_-} \sum_{\{i|y_i=-1\}} k(x, x_i). \quad (10)$$

Given some point x , the label is then simply computed by checking which of the two values $p_+(x)$ or $p_-(x)$ is larger, which leads directly to (8). Note that this decision is the best we can do if we have no prior information about the probabilities of the two classes.

The classifier (8) is a particular case of a more general family of classifiers, which are of the form of an affine combination of kernels on the input domain, $y = \text{sgn} \left(\sum_{i=1}^m \alpha_i k(x, x_i) + b \right)$. The affine combination corresponds to a separating hyperplane in the feature space. In this sense, the α_i can be considered a *dual representation* of the hyperplane's normal vector [7]. These classifiers are example-

based in the sense that the kernels are centered on the training patterns; that is, one of the two arguments of the kernel is always a training pattern. A test point is classified by comparing it to all the training points with a nonzero weight α_i . One of the great benefits that SVM brings in the next section is the assignment of a zero weight to most training points and the sensible selection of the ones kept for classification.

Principal component analysis

Suppose we are given a set of unlabeled points, or a set of points of the same class. In the case of a vector space, we could perform a principal component analysis (PCA) to extract the main axes of the cloud of points. These main axes can then be used as a low-dimensional coordinate system expressing most of the information contained in the initial vector coordinates.

PCA in feature space leads to an algorithm called *kernel PCA* [12]. By solving an eigenvalue problem, the algorithm computes nonlinear feature extraction functions

$$f_n(x) = \sum_{i=1}^m \alpha_i^n k(x_i, x), \quad (11)$$

where, up to a normalizing constant, the α_i^n are the components of the n th eigenvector of the kernel matrix $K_{ij} := (k(x_i, x_j))$.

In a nutshell, this can be understood as follows. To perform PCA in \mathcal{H} , we need to find eigenvectors \mathbf{v} and eigenvalues λ of the so-called *covariance matrix* \mathbf{C} in the feature space, where

$$\mathbf{C} := \frac{1}{m} \sum_{i=1}^m \Phi(x_i) \Phi(x_i)^\top. \quad (12)$$

Here, $\Phi(x_i)^\top$ denotes the transpose of $\Phi(x_i)$. When \mathcal{H} is very high dimensional, the computational costs of doing this directly are prohibitive. Fortunately, one can show that all solutions to

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} \quad (13)$$

with $\lambda \neq 0$ must lie in the span of Φ -images of the training data. Thus, we may expand the solution \mathbf{v} as

$$\mathbf{v} = \sum_{i=1}^m \alpha_i \Phi(x_i), \quad (14)$$

thereby reducing the problem to that of finding the α_i . It turns out that this leads to a dual eigenvalue problem for the expansion coefficients,

$$K\boldsymbol{\alpha} = m\lambda\boldsymbol{\alpha}, \quad (15)$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)^\top$.

To extract nonlinear features from a test point x , we compute the dot product between $\Phi(x)$ and the n th normalized eigenvector in feature space,

$$\langle \mathbf{v}^n, \Phi(x) \rangle = \sum_{i=1}^m \alpha_i^n k(x_i, x). \quad (16)$$

Usually, this will be computationally far less expensive than taking the dot product in the feature space explicitly.

Kernel ridge regression and the representer theorem

Let us now consider the case of regression: we know the values $y_i \in \mathbb{R}$ of a function at m given points $(x_i)_{1 \leq i \leq m}$ and we would like to interpolate it to any new point $x \in \mathcal{X}$. The notion of regression requires the one of regularization, so we choose a kernel k and use the associated norm $\|\cdot\|_{\mathcal{H}}$. The problem can be expressed mathematically as the search for the best function $f : \mathcal{X} \rightarrow \mathbb{R}$ which minimizes a weighted sum of the prediction errors $(f(x_i) - y_i)^2$ at known points and the regularity cost $\|f\|_{\mathcal{H}}$:

$$\inf_{f: \mathcal{X} \rightarrow \mathbb{R}} \left\{ \sum_{i=1}^m (f(x_i) - y_i)^2 + \lambda \|f\|_{\mathcal{H}}^2 \right\} \quad (17)$$

Representer Theorem *The solution f of (17) in the RKHS belongs to the span of functions $k(x_i, \cdot)$ and thus admits a representation of the form*

$$f(x) = \sum_{j=1}^m \alpha_j k(x_j, x). \quad (18)$$

More details can be found in ([11], p. 89). Using (18) and (4), the problem (17) becomes:

$$\inf_{\alpha \in \mathbb{R}^m} \left\{ \sum_{i=1}^m \left(\sum_j \alpha_j k(x_j, x_i) - y_i \right)^2 + \lambda \sum_{i,j} \alpha_i \alpha_j k(x_i, x_j) \right\}. \quad (19)$$

By computing the derivative with respect to α , denoting by K the $m \times m$ matrix $(k(x_i, x_j))_{i,j}$, and by Y the vector $(y_i)_{1 \leq i \leq m}$ we obtain:

$$2K(K\alpha - Y) + 2\lambda K\alpha = 0$$

which leads, since K is positive definite, to the linear system:

$$(K + \lambda \text{Id})\alpha = Y. \quad (20)$$

where Id is the identity matrix.

2.4 Support vectors

The kernelized examples in the previous section are able to deal linearly with the non-linear priors on the data (i.e., the kernel, which induces a feature space and a metric therein) and are consequently able to deal with far more general tasks than usual linear classification or regression. However the computation of the label to assign to a new test point involves its distances to all training points, and consequently these algorithms are naturally slow if the training set is big. Instead of using tricks to reduce the training set size or to avoid the computation of all distances for each new point, one can wonder whether there would exist another, similar approach, which would naturally and directly lead to a huge compression of the training data, keeping only a few meaningful training points to predict the labels of new test points. Such an approach does exist. We present here the fundamentals of support vector classification.

Hyperplanar classifier in feature space and margin

We are given a set of points x_i with a binary label $y_i \in \{-1, 1\}$ and we would like to attribute to any new point $x \in \mathcal{X}$ a class label $f(x)$. We consider a kernel k and search for the best hyperplane in the feature space \mathcal{H} which separates the training points $\mathbf{x}_i = \Phi(x_i)$ into two classes, so that f has the form:

$$f(x) = \text{sgn} \left(\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{H}} + b \right) \quad (21)$$

where $\mathbf{w} \in \mathcal{H}$ is a vector normal to the hyperplane and $b \in \mathbb{R}$ is the shift of the hyperplane. Let us rescale the problem by adding the constraint that the closest data point \mathbf{x}_i to the hyperplane satisfies

$$|\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{H}} + b| = 1. \quad (22)$$

Note that the margin, i.e. the distance between the hyperplane and the closest point, is then $1/\|\mathbf{w}\|_{\mathcal{H}}$. We would like the margin to be as large as possible in order to ensure the quality and the robustness of the classification (see Fig. 2). Therefore we would like to minimize $\|\mathbf{w}\|_{\mathcal{H}}$.

We would like also the predictions $f(x_i)$ on training points to be as good as possible. Since the labels are binary, i.e. $y_i \in \{-1, 1\}$, a correct labelling $f(x_i)$ of the point x_i means $y_i f(x_i) > 0$. Because of constraint (22), this is equivalent to:

$$\forall i, \quad y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{H}} + b) \geq 1. \quad (23)$$

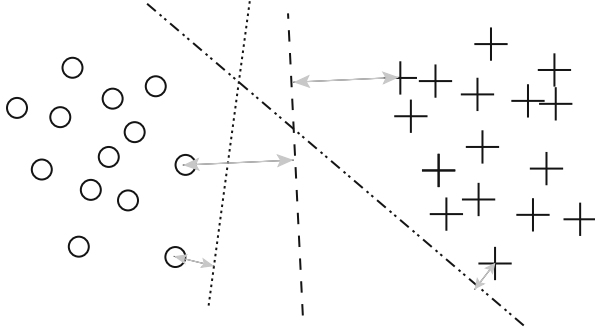


Fig. 2 Example of a good and two bad hyperplane classifiers for a same training set. The larger the margin is, the better the classifier is likely to perform

Soft margin

However, in practice, it may happen that the two classes overlap in the feature space and consequently cannot be separated by an hyperplane satisfying (23) for all examples i . Outliers may also be present in the training set and it may be better to relax the constraints (23) than to overfit the data. Let us denote by ξ_i non-negative slack variables, and relax (23) to:

$$\forall i, \quad y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{H}} + b) \geq 1 - \xi_i. \tag{24}$$

We would prefer the sum of the slacks $\sum_i \xi_i$ to be as small as possible, so we build a *soft margin classifier* by solving

$$\underset{\mathbf{w} \in \mathcal{H}, b \in \mathbb{R}, \xi \in \mathbb{R}_+^m}{\text{minimize}} \quad \tau(\mathbf{w}, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \tag{25}$$

$$\text{subject to} \quad \forall i, \quad y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) - 1 + \xi_i \geq 0. \tag{26}$$

where the constant $C > 0$ determines the trade-off between margin maximization and training error minimization.

Lagrangian approach and dual problem

The *constrained optimization problem* (25,26) can be solved by introducing *Lagrangian multipliers* $\alpha_i \geq 0$ and a *Lagrangian*

$$L(\mathbf{w}, b, \xi, \alpha) = \tau(\mathbf{w}, \xi) - \sum_{i=1}^m \alpha_i (y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle + b) - 1 + \xi_i). \tag{27}$$

and by minimizing it with respect to \mathbf{w} , b and ξ while maximizing it with respect to α . This additional maximization is a practical way to enforce the constraints (26). Indeed, for given \mathbf{w} , b and ξ , if one constraint i was violated in (27), then the corresponding $y_i(\langle \mathbf{x}_i, \mathbf{w} \rangle + b) - 1 + \xi_i$ would be negative, and thus maximizing L w.r.t. α_i would lead to infinity. Similarly, for given \mathbf{w} , b and ξ , the α_i that maximize (27) are zero if the corresponding constraints are strictly satisfied (i.e. $y_i(\langle \mathbf{x}_i, \mathbf{w} \rangle + b) - 1 + \xi_i > 0$). This is essentially the Karush-Kuhn-Tucker (KKT) complementarity conditions of optimization theory. Consequently only a few α_i will be non-zero, leading to a sparse representation of the training data.

Maximizing L w.r.t. the primal variables \mathbf{w} , b and ξ leads to:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \text{ and } \frac{\partial L}{\partial b} = 0 \text{ and } \forall i, \frac{\partial L}{\partial \xi_i} = 0 \text{ or } \left\{ \xi_i = 0 \text{ and } \frac{\partial L}{\partial \xi_i} \geq 0 \right\} \quad (28)$$

$$\text{which are respectively equivalent to } \mathbf{w} = \sum_{i=1}^m \alpha_i y_i \mathbf{x}_i \quad (29)$$

$$\text{and } \sum_{i=1}^m \alpha_i y_i = 0 \text{ and } \forall i, \alpha_i = C \text{ or } \{ \xi_i = 0 \text{ and } \alpha_i \leq C \}. \quad (30)$$

Incorporating (29, 30) into (27) makes \mathbf{w} , b and ξ vanish, and together with the kernel trick (4) we obtain

$$\text{maximize}_{\alpha \in \mathbb{R}^m} \quad W(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (31)$$

$$\text{subject to } \sum_{i=1}^m \alpha_i y_i = 0 \text{ and } \forall i, 0 \leq \alpha_i \leq C. \quad (32)$$

Note that there is an alternative parametrization of SV classifiers where a free parameter ν is used instead of C , with ν asymptotically characterizing the fraction of points falling into the margin, and the fraction of support vectors.

Support vector machine

Once the quadratic energy (31) in α with linear constraints (32) has been maximized, equation (29) gives us an algorithm of the form (21) we were searching for:

$$f(x) = \text{sgn} \left(\sum_i \alpha_i y_i k(x_i, x) + b \right) \quad (33)$$

with $\alpha_i = 0$ for most i . The few data points x_i which have a non-zero coefficient α_i are called *support vectors*. To compute the value of the threshold b , one uses equation (30) which states that for any support vector x_i with $\alpha_i < C$, the slack ξ_i

is zero and consequently the constraint (24) becomes:

$$b = y_i - \sum_j \alpha_j k(x_j, x_i). \quad (34)$$

3 Application to Intermodality Image Prediction

As a medical application of the above methods, we look at intermodality image prediction, i.e. the task of predicting an image of a subject (for instance a Computed Tomography (CT) scan), from an image of the same subject in a different modality (here, a Magnetic Resonance (MR) scan), given a training set of corresponding MR-CT pairs from different patients.

3.1 *The MR-CT issue*

MR-based CT prediction is needed for example for attenuation correction in Positron Emission Tomography (PET). The 511 keV radiation in PET gets attenuated while passing through tissue, and correcting for this effect requires knowledge of the so-called attenuation map, which can be derived from a CT image. In modern scanners, CT images are therefore often acquired alongside the PET image, typically in combined PET/CT scanners. However, the CT scan involves important additional radiation exposure for the patient. Moreover, novel PET/MR scanners are not equipped with a CT scanner, and thus it is desirable to perform the attenuation correction based on the MR, by estimating a “pseudo” CT image from the MR.

MR and CT scanners detect different properties of the matter, and consequently there is no one-to-one correspondence between the intensities in the MR image and the CT intensities. In particular, bone and air both yield no signal in all standard MR sequences, whereas their intensities in CT images are on opposite ends of the scale. For this application, it is therefore crucial to distinguish bone from air, and the MR intensity alone contains no helpful information for this problem.

3.2 *Atlas registration vs. patches*

Atlas registration is the process of aligning a new image with a template image already segmented into bone, air and tissue regions. This yields a segmentation for the new image. The implicit assumption is that there exists a continuous one-to-one transformation between the new patient and the template, and that this transformation can be easily computed. In the case of medical scans, it turns out that these assumptions are not always satisfied, for instance pockets of gas in the abdominal region are unlikely to occur in the same number and shape for

different patients. Even if the assumptions were satisfied, one may be trying to solve a problem more difficult than necessary by searching for a topology-preserving transformation.

Even a rough registration, which does not assume one-to-one correspondence between images, brings useful information since the location of a point in a scan is clearly correlated with the type of tissue which can be found at that point. This correlation is not always decisive enough to determine fully the tissue class, for instance when several scenarios can be thought of at a same place (abdomen or random pocket of gas), or when the registration lacks accuracy.

On the other hand, a patch-based approach would consist in extracting local information from a patch in the MR image centered on the pixel considered, and in classifying this pixel according to similar patches previously observed. This would not require any prior registration, would not assume a one-to-one correspondence between all MR scans and consequently would be able to deal with several possibilities of scenarios for the same location. It would, in some way, build a prediction by picking parts from different examples. This approach is much more flexible than template registration. However it ignores the important information given by the location.

We proposed in [8] to make simultaneous use of both the local and global information given by patches and registration, respectively. We first estimate a rough registration of the test image to a template image, and call *normalized coordinates* the resulting new positions of pixels.

3.3 Image prediction using kernel methods

The key in working with kernel methods is in designing a kernel, or features, which are adapted to the application. In our case an input will be a pair $x_i = (p_i, c_i)$ of the local patch p_i and its normalized coordinates c_i ; we define a similarity measure between inputs by

$$k(x_i, x_j) = \exp\left(\frac{-\|p_i - p_j\|^2}{2\sigma_{\text{patch}}^2}\right) \exp\left(\frac{-\|c_i - c_j\|^2}{2\sigma_{\text{pos}}^2}\right). \quad (35)$$

The parameters σ_{patch} and σ_{pos} involved express the weighting between the different information sources. Their optimal values can be determined by the standard technique of *cross-validation*: to estimate the relevance of any particular choice of $(\sigma_{\text{patch}}, \sigma_{\text{pos}})$, the training set is partitioned into n subsets, and each subset is used for testing the algorithm trained with these parameters on the remaining $n - 1$ subsets. The sum of the losses of all subsets is the energy to minimize with respect to the parameters.

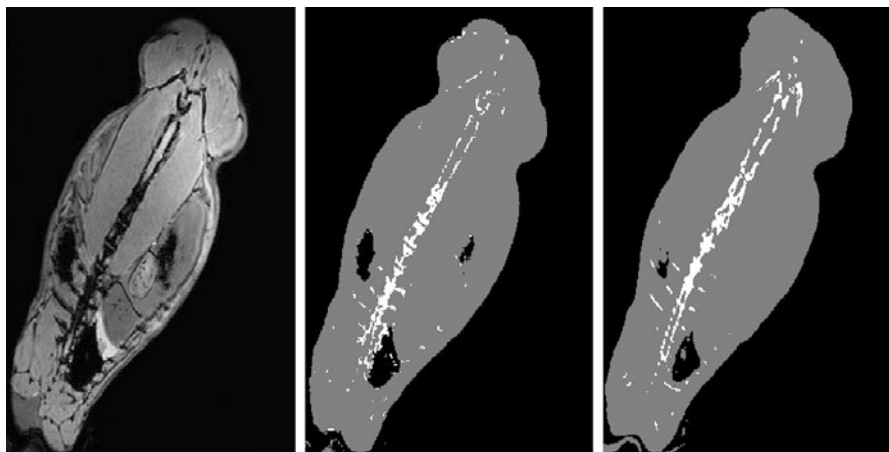


Fig. 3 Left: MR (T2 MEDIC) of a rabbit. Middle: Three class labels as predicted using SVM. Right: Three class labels obtained by thresholding a CT image of the rabbit. Many differences between b) and c) are due not to false classifications, but to some slight movement between MR and CT scans, which explains the misalignment between the test image a) and the ground truth c)

For our application, cross-validation typically yields optimal values for σ_{pos} that are far bigger than 1. This implies that registration errors of a few pixels will not affect the accuracy of our algorithm.

In the CT prediction problem, we may be interested in the classification of MR pixels into three classes, namely bone, air and tissue, because in first approximation there is a one-to-one correspondence between these classes and the CT values. We build three binary classifiers with SVM, one for each class against the two others, or more exactly we compute the quantity whose sign is checked in (33), and then return the class which achieves the greatest score. We show examples of results in Fig. 3.

3.4 Local learning

If the position is very informative, we can learn locally, i.e. cut the template image into regions and train independently a classifier/regression for each region. For brain images for example, intersubject variability is much smaller than for whole body images. Thus non-rigid registration between subjects is possible with only minor misalignments, and it is reasonable to compare patches only within a localized neighborhood. We use kernel ridge regression in order to take into account the variability of CT values. More precisely, from pairs of patches and normalized coordinates, we do not predict the CT value itself but the variation between the CT value and the one in the template at that position. Results are shown in Fig. 4.

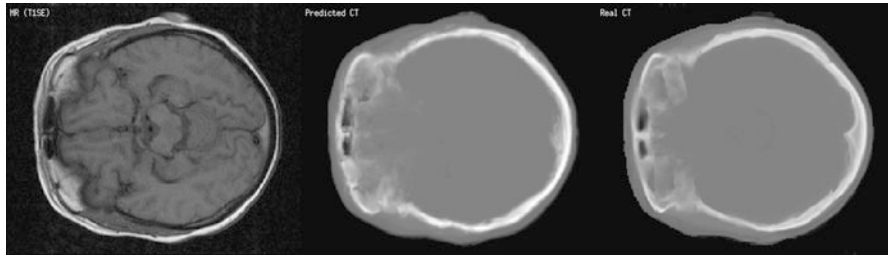


Fig. 4 New MR image; pseudo-CT predicted by our method; ground truth

4 Discussion

After a tutorial on kernel methods, we have presented a way to use these machine learning tools to extract information from a set of medical images for MR-based CT prediction, in a framework which makes use of both local and global information. This presents the advantage of requiring neither a precise registration between template and test image, nor a one-to-one correspondence between them. We hope we have woken up the reader’s enthusiasm for machine learning in medical imaging, there are still plenty of other ways to use machine learning tools in this field!

References

1. M. A. Aizerman, É. M. Braverman, and L. I. Rozonoér. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25:821–837, 1964.
2. N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68:337–404, 1950.
3. C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups*. Springer-Verlag, New York, 1984.
4. B. E. Boser, I. M. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In D. Haussler, editor, *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152, Pittsburgh, PA, July 1992. ACM Press.
5. D. Cremers, T. Kohlberger, and C. Schnörr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, 2003.
6. J. Duchon. Spline minimizing rotation-invariant semi-norms in Sobolev spaces. In W. Schempp and K. Zeller, editors, *Constructive theory of functions os several variables, Lecture Notes in Mathematics*, 571. Springer-Verlag, Berlin, 1977.
7. I. Guyon, B. Boser, and V. Vapnik. Automatic capacity tuning of very large VC-dimension classifiers. In S. J. Hanson, J. D. Cowan, and C. L. Giles, editors, *Advances in Neural Information Processing Systems*, volume 5, pages 147–155. Morgan Kaufmann, San Mateo, CA, 1993.
8. M. Hofmann, F. Steinke, V. Scheel, G. Charpiat, M. Brady, B. Schölkopf, and B. Pichler. MR-based PET attenuation correction – method and validation. In *IEEE Medical Imaging Conference*, 2007.

9. F. Jurie and C. Schmid. Scale-invariant shape features for recognition of object categories. In *International Conference on Computer Vision & Pattern Recognition*, volume II, pages 90–96, 2004.
10. J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society, London*, A 209:415–446, 1909.
11. B. Schölkopf and A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press, December 2001.
12. B. Schölkopf, A. J. Smola, and K.-R. Müller. Kernel principal component analysis. In B. Schölkopf, C. J. C. Burges, and A. J. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*, pages 327–352. MIT Press, Cambridge, MA, 1999. Short version appeared in *Neural Computation* 10:1299–1319, 1998.

Geometric Deformable Models

Y. Bai, X. Han, and J.L. Prince

Abstract Geometric deformable models are deformable models that are implemented using the level set method. They have been extensively studied and widely used in a variety of applications in biomedical image analysis. In this chapter, the general geometric deformable model framework is first presented and then recent developments on topology, prior shape, intensity and motion, resolution, efficiency, robust optimization, and multiple objects are reviewed. Key equations and motivating and demonstrative examples are provided for many methods and guidelines for appropriate use are noted.

1 Introduction

Deformable models (also called “snakes” and “active contours”) have been extensively studied and widely used in biomedical image analysis applications such as image segmentation, geometrical modeling, surgery simulation, etc. Deformable models are curves or surfaces that deform within two-dimensional (2D) or three-dimensional (3D) digital images under the influence of both internal and external forces and/or user defined constraints. Traditional deformable models [27, 47, 70, 73, 99, 120] are represented using explicit parametric forms during deformation,

Y. Bai (✉)
HeartFlow, Inc., 1400B Seaport Blvd, Redwood City, CA 94063, USA
e-mail: yingbai@gmail.com

X. Han
Elekta Inc., 13723 Riverport Dr., Suite 100, St. Louis, MO, 63043, USA
e-mail: xiao.han@elekta.com

J.L. Prince
Department of Electrical and Computer Engineering, Johns Hopkins University,
201B Clark Hall, 3400 N Charles St, Baltimore, MD 21218, USA
e-mail: prince@jhu.edu

and are implemented in a Lagrangian framework. Parametric models have been very successful in a broad cross-section of applications; however, topological inflexibility, self-intersections, and reparametrization are often stated as difficulties with these models.

Geometric deformable models (GDMs) [17,20,68,127] are represented implicitly without the need for parameterization during contour evolution. GDMs deform according to an Eulerian formulation known as the level set method [76,95]. They provide an elegant solution to the primary limitations of parametric deformable models: they are independent of parametrization, they are numerically stable yielding no self-intersections, and they permit automatic change of topology. A comprehensive accounting of the connections between PDMs and GDMs is provided in [119,121].

Most earlier GDMs are driven by forces derived from local edge information or global regional statistics. There has been considerable recent research devoted to the integration of high-level prior knowledge into GDM segmentation. Of particular interest is the knowledge of object topology, object shape, and image appearance. Many objects in biomedical images have known topologies, and conforming to the known object topology is often necessary both for meaningful segmentations and for correct shape modeling and analysis. Knowledge of the typical shape and image appearance of an object is another important cue that can be integrated into a GDM to improve its performance. Given a set of training data, one can learn the intensity distribution of each object region, the geometric relationships between neighboring objects, the overall shape of each object, and/or the pattern of motion of an object in an image sequence. Use of these kinds of prior information can help constrain the computation to within a reduced space of admissible solutions and thereby improve an algorithm's robustness with respect to image noise, weak or missing object boundaries, and other artifacts often encountered in real medical image data.

There are several other considerations in developing GDMs. First, there is a tradeoff between the model resolution and the computational efficiency. The resolution of a GDM is inherently limited by the resolution of the sampling grid used to implement the model. A highly refined grid improves the achievable resolution at the cost of increasing the computation time. Second, it is difficult to define a GDM for the simultaneous segmentation of multiple objects. The classical GDM is a so-called "two-phase" representation that can only segment a single object relative to a single background. Third, there is a question of robustness. Ideally, when the segmentation problem is formulated as a variational problem, the solution—i.e., the GDM algorithm—should be able to start from an arbitrary initialization and converge to a global minimizer of the energy function. In reality, the classical gradient descent approach often converges to a local minimum, which is not the desired object.

In this chapter, we briefly review the basics of GDMs and then discuss recent developments that aim to address the challenges mentioned above. The chapter is organized as follows. In Sect. 2, we first revisit the classical GDM framework, including the topology-preserving method of Han et al. [42] and the shape-based

priors of Leventon et al. [57]. We then survey the recent research developments of GDMs to achieve more flexible and general topology control, integration of statistical priors of shape, intensity and motion, higher resolution and better efficiency, robust optimization, and extensions to the segmentation of multiple objects. For other excellent tutorials and reviews of earlier development of deformable models, we refer interested readers to [71, 75, 119, 121].

2 Overview

2.1 Basic Framework

GDMs are based on the theory of front evolution [2, 52, 90] and are implemented using the level set numerical method [76, 95]. The model contour(s) (curves in 2D or surfaces in 3D) are embedded as the zero level set of a higher dimensional level set function $\Phi(\mathbf{x}, t)$ and propagates implicitly through the temporal evolution of $\Phi(\mathbf{x}, t)$. Due to numerical stability and computational convenience, $\Phi(\mathbf{x}, t)$ is often chosen to be a signed distance function of the embedded contour(s). Such a signed distance function can be computed very efficiently using the *fast marching* method [94, 109] or the *fast sweeping* method [130].

For image segmentation problems, a GDM is either formulated as an energy minimization problem where the solution is sought through gradient descent optimization, or by directly designing the various forces that drive the model contour(s) towards desired object boundaries. In either case, the final evolution equation regarding the level set function $\Phi(\mathbf{x}, t)$ can be summarized in the following general form:

$$\frac{\partial \Phi(\mathbf{x}, t)}{\partial t} = [F_{\text{prop}}(\mathbf{x}, t) + F_{\text{curv}}(\mathbf{x}, t)] |\nabla \Phi(\mathbf{x}, t)| + \mathbf{F}_{\text{adv}}(\mathbf{x}, t) \cdot \nabla \Phi(\mathbf{x}, t),$$

where F_{prop} , F_{curv} , and F_{adv} are spatially-varying force terms that drive the front evolution. In particular, F_{prop} is an expansion or contraction force, F_{curv} is the part of the force that depends on the intrinsic geometry, especially the curvature of the implicit contour and/or its derivatives, and F_{adv} is an advection force that passively transports the contour.

Numerical schemes to solve the above level set PDE must be carefully designed. The time derivative can be approximated by a forward difference scheme. The spatial derivatives are computed using upwind scheme for terms F_{prop} and F_{adv} , and using central difference scheme for term F_{curv} [76, 95]. No parameterization of the deforming contour is needed during the evolution. The parametric representation is computed (if necessary) using an isocontour algorithm (e.g., [65]) only after the evolution is complete.

Classical GDMs use edge information to design driving forces [16,49,50,68,98]. For example, the well-known *geodesic active contour* (GAC) model proposed by Caselles et al. [16, 17, 18] and Kichenassamy et al. [49, 50] is designed to find the best minimal-length smooth contour that takes into account the image gradient information. In a GAC model, $F_{\text{prop}}(\mathbf{x}, t) = cg(\mathbf{x})$, $F_{\text{curv}}(\mathbf{x}, t) = \kappa(\mathbf{x}, t)g(\mathbf{x})$, and $F_{\text{adv}}(\mathbf{x}, t) = \nabla g(\mathbf{x})$, where c is a constant and $g(\cdot)$ is an image-derived metric, which is often a monotonically decreasing function of the gradient magnitude of the image I . Also, $\kappa(\mathbf{x}, t)$ is the (mean) curvature of the level set of $\Phi(\cdot, t)$ that passes through \mathbf{x} . As edge-based models rely primarily on local gradient information, they are extremely sensitive to noise and spurious edges. Initialization close to the final boundary is needed to avoid becoming stuck in local minima. Leakage through weak edges is another typical problem with these models.

To increase the robustness of the edge-based models, a number of region-based models have been proposed [20, 26, 78, 80, 84, 116, 128, 132]. Most region-based methods assume that an image consists of a finite number of regions, parametrized by a predetermined set of features (e.g., means, variances, textures) which may be inferred or estimated from the image data. Contours are then driven by regional forces such that a maximal separation of the regional features is achieved. Most often the region features are re-evaluated each time the contour evolves. More recently, models that are based on non-parametric region statistics [51] and shape optimization principles [5,44] have also been proposed.

2.2 Topology Preservation

Incorporation of topology control in the GDM framework was introduced by Han et al. [42] in the topology preserving geometric deformable model (TGDM). Assuming that the topology of the zero level set (i.e., the embedded contour) is homeomorphic to the topology of the boundary of the digital object it encircles, TGDM preserves the topology of the implicit contour by preserving the topology of the corresponding digital object. This is achieved by adopting the concept of *simple point* from the digital topology theory [55] to monitor the sign changes of the level set function during its evolution, and preventing sign changes at non-simple points. The basic principle is illustrated in Fig. 1. In Fig. 1(a), the point x is a simple point. A sign change at x yields a contour that has the same topology as the original contour, as shown in Fig. 1(b). In contrast, the point y in Fig. 1(b) is not a simple point. A sign change at y will split the original contour into two, as shown in Fig. 1(c). A simple point criterion check can thus predict and prevent possible topology changes. In addition, a connectivity-consistent isocontour algorithm is necessary to extract a topologically-consistent explicit contour representation from the embedded level set function at convergence. More details of the algorithm and its implementation can be found in [42].

TGDMs are useful when the object topology is known *a priori*. As mentioned earlier, in cortical segmentation it is known that a cortical surface (closed at the

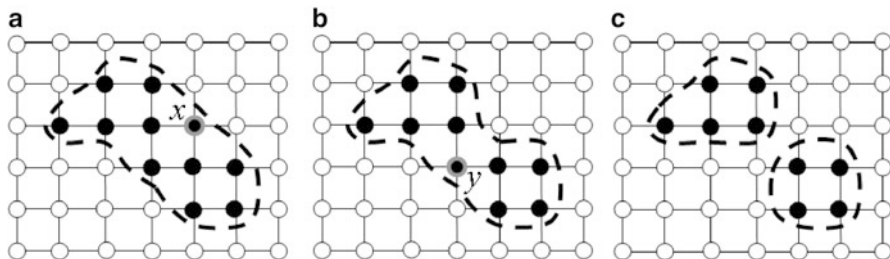


Fig. 1 Topology equivalence of implicit contour topology and the digital object boundary topology: 4-connectivity for dark points and 8-connectivity for others. (a) Original contour. (b) The contour passes over a simple point. (c) The contour splits at a nonsimple point

bottom) should be topologically equivalent to a sphere with no handles. Due to image noise and artifacts, standard GDMs typically yield a surface segmentation with many handles, as shown in Fig. 2(b). A TGDM can guarantee the correct topology with the same input image and image force design, as shown in Fig. 2(c). Another example is the segmentation of the human pelvis, whose boundary surface is typically assumed to be of genus 3 (with three handles). As shown in bottom row of Fig. 2(b) and Fig. 2(c), a standard GDM yields a segmentation with many undesirable handles, while the TGDM result does not have these extra handles.

2.3 Shape Priors

Segmentation algorithms that only make use of low-level image information (such as intensity gradients) often fail to produce satisfactory results in medical imaging applications due to image noise, limited image resolution and contrast, and other imaging artifacts that often present in typical medical image data. In these cases, simple geometric regularization no longer suffices, and higher-level prior knowledge about the shape of desired object(s) can help.

The first published method that applies shape priors in GDM segmentation was proposed by Leventon et al. [57]. The method consists of two major stages. In the first stage, a statistical object shape model is computed from a set of training samples $\Gamma = \{\Phi_1, \Phi_2, \dots, \Phi_n\}$, where each training shape Φ_i , $1 \leq i \leq n$ is embedded as the zero level set of a signed distance function. Using principal component analysis (PCA), the covariance matrix of the training set is decomposed as $U\Sigma U^T$, where U is a matrix whose column vectors represent the set of orthogonal modes of shape variation and Σ is a diagonal matrix of corresponding singular values. The first k columns of U form the eigenspace of the (typical) shapes. One can project a given shape Φ onto this shape space using

$$\alpha = U_k^T (\Phi - \bar{\Phi})$$

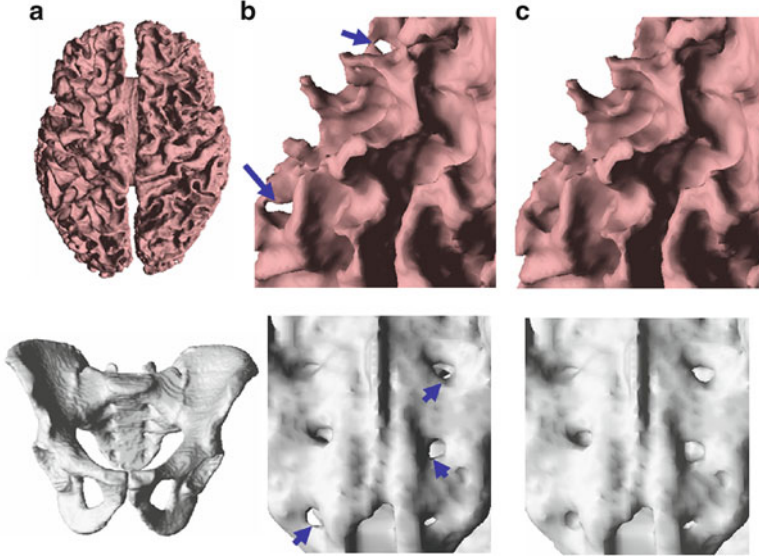


Fig. 2 Topologically-constrained segmentations. (a) shows the reconstructed surface; (b) shows close-up views of the standard GDM results; and (c) shows close-up views of the TGDM results

where $\bar{\Phi}$ is the mean shape, and α is the k -dimensional vector of coefficients that represent Φ in the space spanned by U_k . The shape model construction is illustrated in the top panel of Fig. 3. A training set of corpus callosa segmentation is analyzed by PCA and the three primary modes of variations of the shape distribution are shown in the figure. Assuming a Gaussian distribution for α , the probability of a new shape can be computed as:

$$P(\alpha) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_k|}} \exp\left(-\frac{1}{2} \alpha^T \Sigma_k^{-1} \alpha\right)$$

where Σ_k contains the first k rows and columns of Σ .

In the second stage of the method, an active contour model is evolved both locally, based on image gradient and curvature information, and globally towards a maximum *a posteriori* (MAP) estimate of the shape and pose. At each iteration, the MAP estimates of the shape coefficients α^* and the pose parameter p^* are first computed, and the MAP estimate of the shape Φ^* is then determined. Incorporating the shape estimation into a traditional GAC model leads to a new level set evolution equation in the following form:

$$\frac{\partial \Phi}{\partial t} = (c + \kappa)g|\nabla \Phi(t)| + \nabla g \cdot \nabla \Phi(t) + \beta(\Phi^*(t) - \Phi(t)).$$

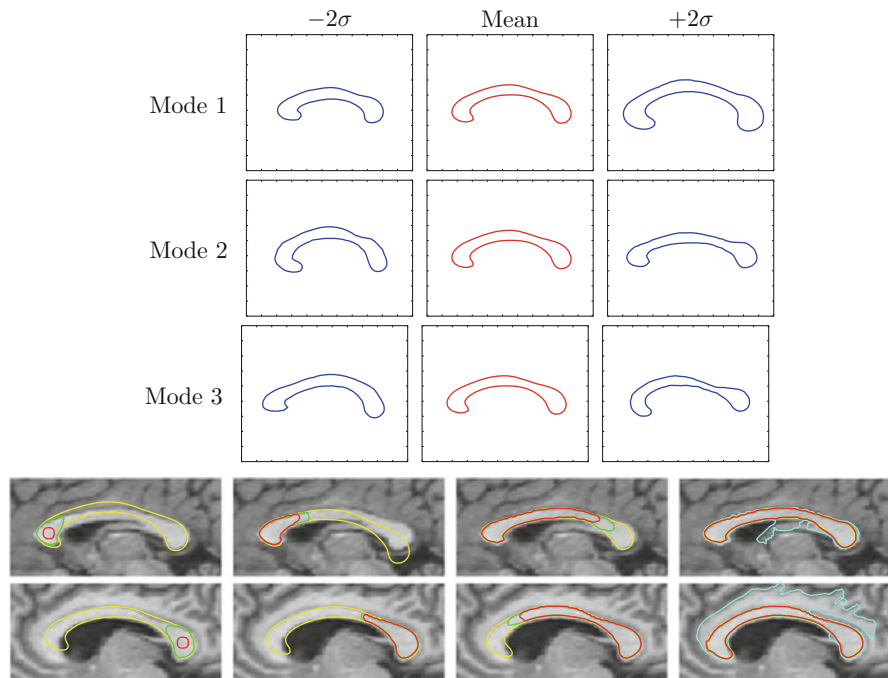


Fig. 3 Top: The three primary modes of variations of the corpus callosum training data set. Bottom: Four steps in the segmentation of two different corpora callosa. The last image in each case shows the final segmentation in red. The cyan contour is the standard evolution without the shape influence. Image courtesy of the authors of [57]

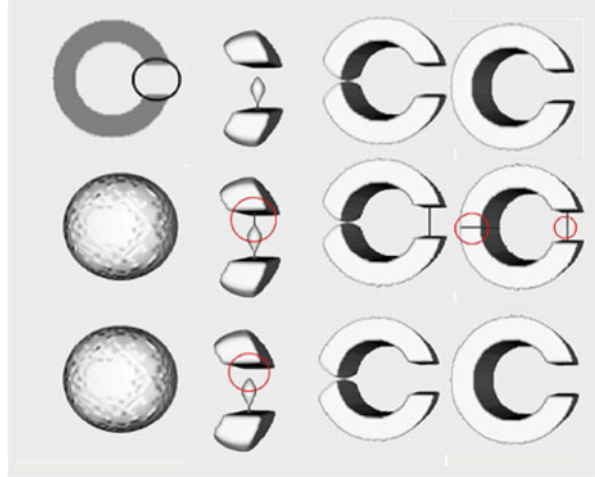
The first two terms are the typical GAC model (as mentioned earlier) and the last term is the shape prior term. The parameter β is used to balance the influence of the shape model and the GAC model. The bottom panel of Fig. 3 compares the performance of the GAC model with and without using the shape prior. Due to weak edges, the GAC without the shape prior leaks out and fails to capture the desired shape; whereas the GAC with the proposed shape prior is well constrained and successfully converges to the boundary of the corpus callosum.

3 GDMs with Topology Control

Since the introduction of the original TGDM, several alternative approaches or extensions have been proposed in the literature. We briefly summarize these approaches in this section.

In the original TGDM, strict topology preservation is enforced, i.e., the number of connected components, cavities and handles are all constrained to remain unchanged [42]. Segonne et al. [92, 93] made the observation that TGDM can

Fig. 4 Segmentation of a “C” shape using a spherical initialization. First row: GDM without topology control; second row: TGDM result; third row: GGDM result. Image courtesy of the authors of [93]



sometimes generate “topological barriers” that lead to geometric inconsistencies (cf. Fig. 4). They proposed to loosen the topological constraint so that only genus (i.e., the number of handles) is preserved. Their *genus-preserving* GDM (GGDM) allows connected components to merge, split, or vanish but no topological defects (i.e., handles) can be generated. To this end, the concept of “simple point” is extended to a “multisimple point” criterion, which is checked whenever the level set function is about to change sign at a computational point. Fig. 4 illustrates the benefit of the GGDM method.

In both TGDM and GGDM, the topological constraint is enforced in the discrete sense. A few efforts have been made to formulate topology preservation in the continuous domain. Shi and Karl [97] proposed to use the so-called *differentiable minimum shape distance* (DMSD) to define a topological prior. DMSD is defined as the shortest distance between any two points on two closed curves. Cast in a variational framework, a topological prior term is introduced that is inversely proportional to the DMSD. The term approaches infinity when the DMSD between two curves approaches zero, thus preventing two curves from merging (though they can still split).

Alexandrov and Santosa [1] proposed a topology-preserving level set method for shape optimization. A topological constraint is built into the optimization problem by adding the following *logarithmic barrier* energy functional:

$$H(\Phi) = - \int_{\partial D} \log[\Phi(x + d \nabla \Phi(x))] ds - \int_{\partial D} \log[-\Phi(x - l \nabla \Phi(x))] ds$$

where Φ denotes the level set function, ∂D denotes the contour boundary, ds is the arc-length measure, and $d > 0$ and $l > 0$ are real numbers. These two terms probe in the vicinity of the contour boundary in the inner and outer normal directions, respectively, and penalize cases when points away from the zero level set have small

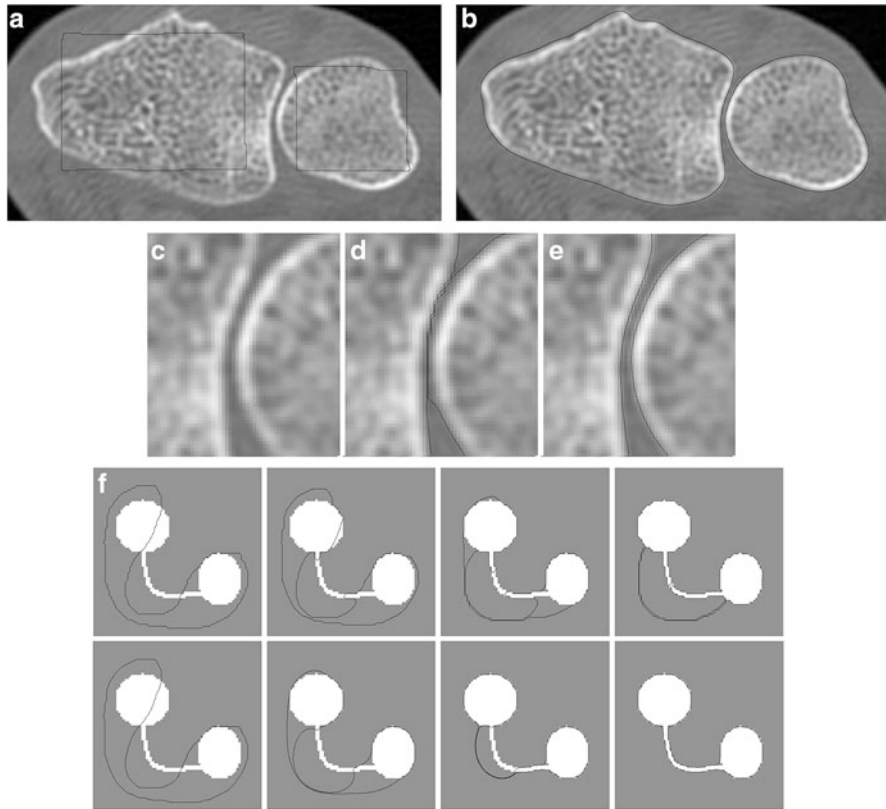


Fig. 5 (a)–(e): carpal bone segmentation. (f): a toy problem in which the “stuck” situation of TGDM is avoided by the global regularizing flow. (See text for details.) Image courtesy of the authors of [103]

absolute distance values. The energy functional has the effect similar to building “barriers” between merging or splitting contours, thus preventing such topological changes to happen.

Sundaramoorthi and Yezzi [103] proposed to use a PDE-based geometric flow to achieve topology preservation. The flow minimizes the following energy functional:

$$E_\gamma(C) = \frac{1}{2} \int \int_{C \times C} \frac{d\hat{s} ds}{\|C(\hat{s}) - C(s)\|^\gamma}$$

where C is a contour, $d\hat{s}$ and ds are the arc-length measures, $\|\cdot\|$ is the Euclidean norm, and $\gamma > 0$ is a free parameter. Minimization of this energy functional leads to a repulsive force that prevents the model contour from self-intersecting or splitting. It also imposes a global regularization of the evolving contour. Fig. 5 demonstrates the benefit of this global regularizing force. Figs. 5(a)–(e) compare the segmentation

results of TGDM and the global regularizing flow for a carpal bone image. Fig. 5(a) shows the contour initialization. Fig. 5(b) shows the final segmentation of the global regularization flow. A magnified view of the bone joint part is shown in Fig. 5(c), and the results of TGDM and the new method in this area are shown in Figs. 5(d) and (e), respectively. Clearly, the latter approach keeps the contours more separated. Fig. 5(f) is another demonstration using a toy problem. Since TGDM only uses a hard constraint that does not come into play until topology is about to change in the next step, it does not regularize the contour as the global regularizing flow does. A similar double integral energy was also used by Rochery et al. [83] for the extraction of a road network and by Guyader and Vese [40] who integrate this energy over regions and formulate it directly in the level set framework.

4 GDMs Incorporating Shape Priors

Following the work of Leventon et al., many GDM methods that incorporate prior shape information have been recently proposed [24, 25, 87, 88, 107, 108, 124]. An extensive review can be found in [32]. We briefly summarize the major contributions.

Most of the shape-constrained GDM methods assume a linear model for shape variations, which tend to have two major limitations [21, 30, 31, 86]. First, the training shapes do not always satisfy a Gaussian distribution as typically assumed. Second, the space of signed distance functions is not linear, i.e. a linear combination of signed distance functions is in general no longer a signed distance function. To address these limitations, Cremers et al. [30,31,86] proposed a statistical shape prior based on an extension of classical *kernel density estimators* (cf. [81,85]) to the level set domain. This prior statistically approximates an arbitrary distribution of training shapes (without making the restrictive assumption of a Gaussian distribution). In the limit of infinite sample size, the distribution inferred by the kernel density estimator converges towards a distribution on the manifold of signed distance functions. In addition, the cited works also embed an intrinsic alignment in the energy function so that the shape prior is invariant to certain group transformations such as translation and scaling. Fig. 6 compares the segmentation result of using a kernel prior against the results of using a uniform prior, a linear prior, and a manual segmentation. It can be seen that in this example, the result of using a kernel prior is closest to the manual segmentation.

Another development in shape prior modeling is the incorporation of object dynamics. In applications such as cardiac segmentation and tracking, it is important to take into account temporal correlations of the images. Cremers et al. [29] proposed to extend the shape modeling to learn the temporal dynamics of a deforming shape for the segmentation of an image sequence. The *dynamical statistical shape model* is constructed by approximating the shape vectors of a sequence of silhouettes by a Markov chain, which is then integrated into a Bayesian segmentation framework. Kohlberger et al. [54] proposed to treat time as an ordinary

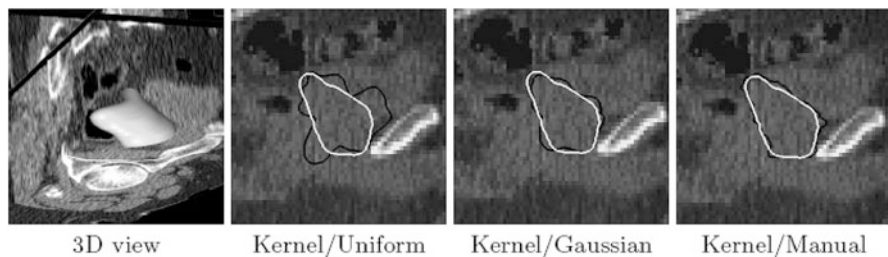


Fig. 6 Comparison of the segmentations obtained with the kernel prior (white) and with alternative approaches (black). Image courtesy of the authors of [86]

fourth dimension and applied a 4D PCA analysis on the training sequences to derive a 4D shape model. In this method, a whole volume sequence is segmented at the same time as a 4D image.

5 Fast GDMs

Since GDMs represent the model contour(s) using a higher-dimensional level set function, a heavy computational cost can be incurred with a naive implementation. To improve computational efficiency, narrowband methods in conjunction with reinitialization techniques [82, 95] are widely used to restrict the computation to the neighborhood of the evolving contour(s). However, the overall computation load can still be prohibitive when the grid size is large.

Several fast implementations of GDMs have been proposed. Goldenberg et al. [38] and Weickert et al. [113] proposed to adapt the *additive operator splitting* (AOS) scheme [114] for GDMs, which relaxes the stability constraint on the size of the time step associated with the explicit numerical schemes. The AOS scheme is very stable; but when large time steps are used, splitting artifacts may arise due to reduced rotational invariance. Kenigsberg et al. [48] and Papandreou and Maragos [77] proposed to use *multigrid* techniques to address this problem and to allow the use of even larger time steps than the AOS scheme. When sub-voxel accuracy is not of concern, Shi and Karl [96]’s method can provide a very high efficiency since it eliminates the need to solve the level set PDE. The method directly tests an optimality condition for the final curve location based on the speed functions, and uses only simple operations like insertion and deletion on two lists of boundary points to evolve the curve.

The reinitialization of the level set function can also be accelerated or even omitted. Krissian and Westin [56] proposed a fast implementation of the Chamfer distance to save computation time while maintaining the sub-voxel accuracy of the interface. Li et al. [58, 59] proposed a *distance-preserving* energy function that forces the level set function to be close to a signed distance function, which

eliminates the need for re-initialization and improves the overall efficiency. Another distance-preserving level set method was later proposed by Estellers et al. [35]; it is more efficient due to the use of a splitting strategy and advanced ℓ_1 optimization techniques.

Several new methods, constituting a new variational model for image segmentation that is closely related to the GDM, have been recently proposed [13, 19, 39]. These methods represent objects by soft membership functions rather than signed distance functions and the smoothness of the segmentation results is controlled by total variation regularization. The resulting models are convex and thus global optimal solutions can be guaranteed to be found. Also, the development of efficient convex relaxation methods [13, 39] allows such models to be computed much faster than traditional GDMs. One weakness, however, is that geometric properties of objects such as surface curvature and distances between coupled surfaces cannot be easily modeled and there is no control of the final segmentation topology. Such models have also been extended to the segmentation of multiple objects (cf. [6, 60]).

6 Adaptive Grid GDMs

The development of adaptive grid GDMs was motivated by the observation that the resolution of a GDM is directly limited by the resolution of the computational grid. One can improve the model resolution by using highly refined grids at the cost of losing efficiency and increasing the size—i.e., number of vertices—of the resulting contour(s). A more elegant approach to address the resolution and efficiency tradeoff is to use *adaptive grid techniques* [53], which locally refine or deform a coarse grid to concentrate computational efforts where more accuracy is needed. Incorporation of topological constraints also becomes feasible. Two types of adaptive grids have been used: the *moving grid* and the *quadtree/octree grid*. We briefly summarize these two types of approaches in the following.

A 2D *moving grid* GDM method was introduced in [41]; it maintains a fixed reference grid, but moves the actual physical grid points according to the desired image features [15, 63]. The adaptively deformed grid is obtained by first solving a Poisson equation using a DCT solver, and then solving an ordinary differential equation. After that, the level set PDE is solved on the deformed grid with narrow-banding. Since a uniform reference grid is always kept, the topology preserving principle on uniform grids (cf. Sect. 2.2) can be directly applied, by performing the simple point check directly on the reference grid.

An *octree grid* in 3D (or similarly a quadtree grid in 2D) is a hierarchical cartesian grid that is locally refined in regions of interest. These adaptive grids are widely used to improve accuracy in the solution of PDE's [74, 104] and in medical image segmentation [8, 34, 122]. The cost for generating an adaptive octree grid is much smaller compared to that for a moving grid, since no partial differential equations

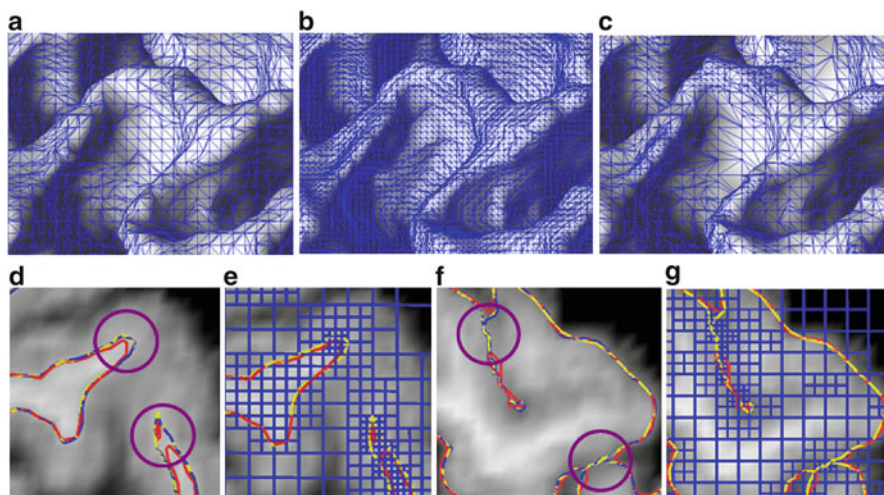


Fig. 7 Extraction of inner and outer surfaces using three computational grids. (a)–(c) triangle meshes: (a) coarse uniform grid TGDM result; (b) fine uniform grid TGDM result; (c) OTGDM result. (d) and (f): close-up views of inner and outer surfaces reconstructed by three types of grid: red–coarse uniform grid TGDM result, blue–fine uniform grid TGDM result, yellow–OTGDM result; (e) and (g): close-up views of octree grids used by OTGDM (shown in blue)

need to be solved. Implementation of topological constraints as introduced in TGDM is challenging, however, since the theories of common digital topology on uniform grids cannot be directly applied to an octree grid. Bai et al. [7, 8] proposed a new digital topology framework on octree grids that include new definitions of neighborhood, connectivities, and simple points. Based on this extension, a new octree-grid TGDM (OTGDM) method was introduced in [8, 9]. Adaptive octree grids are generated according to boundary curvature estimates, and the OTGDMs preserve the surface topology while evolving on octree grids.

A comparison is shown in Fig. 7 for the use of TGDM to extract inner and outer cortical surfaces from a brain image on three types of grids: a coarse uniform computational grid (of the original image size); a fine uniform computational grid (double size of the original image); and an octree grid whose finest grid resolution is same as the fine uniform grid. Figs. 7(a)–(c) show close-up views of the three inner surface results. It can be seen that the OTGDM result provides an adaptive multi-resolution representation of the surface mesh, which has large triangles in regions that are relatively flat and small triangles in regions with high curvatures. Figs. 7(d)–(g) show close-up cross-sectional views of the inner and outer surface results and the octree grids generated. It can be observed that the fine grid TGDM result and the OTGDM result capture anatomical details (such as the deeply folded sulci and gyri indicated by the circles) better than the coarse grid TGDM result.

7 Miscellaneous

There are many other interesting developments in GDM in addition to those categorized above. Due to the lack of space, we briefly summarize some of them in the following. For others such as GDM methods for the segmentation of tubular structures or tensor-valued images, we refer interested readers to [36, 37, 43, 45, 46, 66, 67, 72, 110, 112, 123].

7.1 Multiple Objects

The conventional level set implementation of GDMs can only deal with a two-phase image, i.e. a single object on a single background. Extensions to multiple objects in which n level set functions are used to model n objects with different characteristics have been proposed [79, 89, 129, 131]. In [106, 125], a joint shape modeling of n neighboring structures is also implemented using n level set functions. Such a strategy incurs great computational cost when the number of objects is large [3, 14, 62, 64, 69, 79, 89, 105, 111, 129, 131].

The *multi-phase level set* (MPLS) method [111] is an elegant framework proposed to address the above issue. MPLS generalizes the Chan and Vese model [20] to segment images with more than two regions. Using heaviside functions, the MPLS method needs only $\log n$ level set functions for n phases in the piecewise constant case, and can represent boundaries with complex topologies including triple junctions. In the piecewise smooth case, only two level set functions formally suffice to represent any partition, based on the *four color theorem* [4]. Fig. 8 shows how the model works on a color image, where three level set functions are used to represent up to eight phases (or colors). In this example, the MPLS method detects six regions and their junctions, which would require at least six level set functions using the conventional approach. The MPLS framework has been adopted by many others to extend their work to deal with multiple regions, such as Bertelli et al. [10] who extended their graph-partitioning method in [100], Kim et al. [51] who extended their mutual information based approach, and Cremers et al. [33] who integrated multiple competing shape priors into shape-based GDMs.

The multi-object geometric deformable model (MGDM) was recently developed to further improve memory requirements, flexibility in speed specification, and topology control [12]. MGDM represents multiple objects with label and distance functions rather than separate level set functions. To good approximation only four functions in 2-D and six functions in 3-D are required to represent any number of objects and to carry out shape evolution without reconstructing independent level set functions. Boundary-specific speed functions and topology control of individual objects and groups of objects can be specified. MGDM was used to parcellate the cerebellum into individual lobules from magnetic resonance brain images in [11].

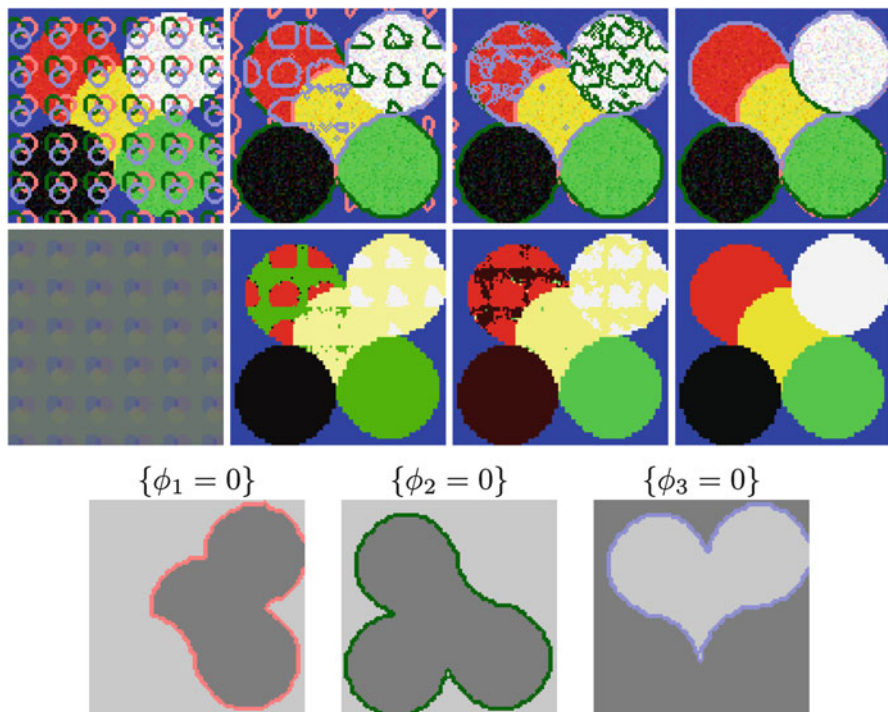


Fig. 8 Color noisy picture with junctions. Three level set functions are used to represent up to eight constant regions. Six segments are detected. Bottom row shows the final zero-level sets of ϕ_1, ϕ_2, ϕ_3 . Image courtesy of the authors of [111]

7.2 Long-range Interactions

Most existing GDMs allow only local interactions or competitions between different parts of an implicit contour or between multiple contours. Recently, it has been shown that enabling long-range interactions can improve the robustness of GDMs and enable modeling of shapes with complex geometries.

Sebastian et al. [91] proposed a *skeletally coupled deformable model* which combines the advantages of curve evolution deformable models, seeded region growing and region competition. The method uses a curve evolution implementation of region growing from initialized seeds, where growth is modulated by a skeletally-mediated competition between neighboring regions. The inter-seed skeleton, which is interpreted as the predicted boundary of collision between two regions, is used to couple the growth of seeds and to mediate long-range competition between them. The long range predicted competition made possible by the inter-seed skeletons helps achieve a more global minimizer.

Rochery et al. [83] extended the GDM formulation to *higher order energy functionals*. Instead of a single integral over the contour, the new functionals consist

of arbitrary order polynomials that include multiple integrals over the contour so that arbitrary long-range interactions between subsets of the contour can be modeled. We note that the authors of [103] and [40] also proposed to use quadratic energy functionals, although the methods were specifically designed for topology preservation purposes.

Xiang et al. [115] proposed a physics-based active contour model using a long-ranged interaction between image boundaries and the moving contours, which was inspired by the *elastic interaction* effects between line defects in solids. Another interesting physics-based model called *magnetostatic active contour model* was recently proposed by Xie and Mirmehdi [117, 118]. Their model simulates magnetic interactions between the evolving contour and target object boundaries to improve the method robustness against arbitrary model initialization and weak edges. The model was further generalized and extended to 3D in [126].

7.3 Robust Optimization

In most GDM formulations, the final solution is sought through gradient descent types of approaches that can easily get trapped in undesirable local minima. A multi-resolution implementation can partially address this problem. Recently, more efforts are made towards the design of novel optimization methods that are robust to image noise and insensitive to model initialization.

Li and Yezzi [61] proposed a *dual front* implementation of GDMs that was motivated by minimal path techniques [28]. The method seeks a global optimum inside an active region surrounding the current model position. By tuning the size of the active region, it achieves an optimal solution with variable degrees of localness and globalness.

Sundaramoorthi and Yezzi [101, 102] and Charpiat et al. [22, 23] observed that using the canonical L^2 norm as the Riemannian metric in gradient flow optimization often leads to undesirable local minimum and irregular flows. They both proposed to optimize the solution in other functional spaces such as the Sobolev space. The resulting *Sobolev gradient flows* are more global in the sense that the deformation of each point is affected by all the other points on the contour. These flows also favor global motions (such as translations) over local deformations, which helps avoid getting trapped at undesired local optima.

8 Conclusion

In this chapter, we reviewed the basics and recent developments of geometric deformable models based on the level set method. The surveyed work clearly demonstrates the importance of GDMs as a valuable framework for image segmentation, allowing easy integration of both image boundary and region information, as

well as prior knowledge about object topology and shape. Undoubtedly, they will continue to play an important role in various image processing and computer vision applications for a long time.

References

1. O. Alexandrov and F. Santosa. A topology-preserving level set method for shape optimization. *J. Comput. Phys.*, 204(1):121–130, 2005.
2. L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axioms and fundamental equations of image processing. *Archive for Rational Mechanics and Analysis*, 123:199–257, 1993.
3. E. D. Angelini, T. Song, B. D. Mensh, and A. Laine. Multi-phase three-dimensional level set segmentation of brain MRI. In *Medical Image Computing and Computer-Assisted Intervention*, volume 3216, pages 318–326, 2004.
4. K. Appel and W. Haken. Every planar map is four colorable. *Illinois Journal of Mathematics*, 21:429–567, 1977.
5. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image segmentation using active contours: Calculus of variations or shape gradients? *SIAM Journal of Applied Mathematics*, 63:2128–2154, 2003.
6. E. Bae, J. Yuan, and X. C. Tai. Global minimization for continuous multiphase partitioning problems using a dual approach. *Int. J. Comput. Vis.*, 92:112–129, 2011.
7. Y. Bai, X. Han, and J. L. Prince. Octree-based topology-preserving isosurface simplification. In *Computer Vision and Pattern Recognition Workshop*, page 81, New York, June 2006.
8. Y. Bai, X. Han, and J. L. Prince. Octree grid topology preserving geometric deformable models for 3D medical image segmentation. In *Inf Process Med Imaging*, volume 20, pages 556–568, 2007.
9. Y. Bai, X. Han, and J. L. Prince. *Advances in Imaging and Electron Physics*, volume 181, chapter Octree-grid Topology-preserving Geometric Deformable Model, pages 1–34. 2014.
10. L. Bertelli, B. Sumengen, B. S. Manjunath, and F. Gibou. A variational framework for multi-region pairwise similarity-based image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, pages 1400 – 1414, 2008.
11. J. Bogovic, P. -L. Bazin, S. Ying, and J. Prince. Automated segmentation of the cerebellar lobules using boundary specific classification and evolution. In *Information Processing in Medical Imaging*, pages 62–73, 2013.
12. J. Bogovic, J. Prince, and P. -L. Bazin. A multiple object geometric deformable model for image segmentation. *Comput. Vis. Image Underst.*, 117:145–157, 2013.
13. X. Bresson, S. Esedoglu, P. Vandergheynst, J. -P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *J. Math. Imaging Vis.*, 28:151–167, 2007.
14. T. Brox and J. Weickert. Level set segmentation with multiple regions. *IEEE T. Image Process.*, 10:3213– 3218, 2006.
15. W. Cao, W. Huang, and R. D. Russell. A moving mesh method based on the geometric conservation law. *SIAM J. Sci. Comput.*, 24:118–142, 2002.
16. V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1–31, 1993.
17. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Int. J. Comput. Vision*, 22:61–79, 1997.
18. V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. Minimal surfaces based object segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, 19:394–398, 1997.
19. T. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J. Appl. Math.*, 66:1632–1648, 2006.

20. T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. Image Proc.*, 10(2):266–277, 2001.
21. G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Found. Comput. Math.*, 5:1–58, 2005.
22. G. Charpiat, R. Keriven, J. -P. Pons, and O. Faugeras. Designing spatially coherent minimizing flows for variational problems based on active contours. In *IEEE International Conference on Computer Vision*, volume 2, pages 1403–1408, 2005.
23. G. Charpiat, P. Maurel, J. -P. Pons, R. Keriven, and O. Faugeras. Generalized gradients: Priors on minimization flows. *Int. J. Comput. Vision*, 73(3):325 – 344, 2007.
24. Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggs, and E.A. Geiser. Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vision*, 50:315–328, 2002.
25. Y. Chen, S. Thiruvenkadam, F. Huang, D. Wilson, E. A. G. Md, and H. D. Tagare. On the incorporation of shape priors into geometric active contours. In *Variational and Level Set Methods in Computer Vision*, pages 145–152, 2001.
26. L. Cohen, E. Bardinet, and N. Ayache. Surface reconstruction using active contour models. In *SPIE on Geometric Methods in Computer Vision*, 1993.
27. L. D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. Pattern Anal. Machine Intell.*, 15:1131–1147, 1993.
28. L. D. Cohen and R. Kimmel. Global minimum for active contour models: A minimal path approach. *Int. J. Comput. Vision*, 24:57–78, 1997.
29. D. Cremers and G. Funka-Lea. Dynamical statistical shape priors for level set based sequence segmentation. In *Variational, Geometric, and Level Set Methods in Computer Vision*, volume 3752, pages 210–221, 2005.
30. D. Cremers, S. J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. In *Pattern Recognition (Proc. DAGM)*, volume 3175, pages 36–44, 2004.
31. D. Cremers, S. J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *Int. J. Comput. Vision*, 69:335 – 351, 2006.
32. D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *Int. J. Comput. Vision*, 72: 195–215, 2007.
33. D. Cremers, N. Sochen, and C. Schnörr. A multiphase dynamic labeling model for variational recognition-driven image segmentation. *Int. J. Comput. Vision*, 66:67–81, 2006.
34. M. Droske, B. Meyer, C. Schaller, and M. Rumpf. An adaptive level set method for medical image segmentation. In *Information Processing in Medical Imaging*, volume 2082, pages 416–422, 2001.
35. V. Estellers, D. Zosso, R. Lai, J. -P. Thiran, S. Osher, and X. Bresson. An efficient algorithm for level set method preserving distance function. *IEEE T. Image Process.*, 21:4722–34, 2012.
36. C. Feddern, J. Weickert, and B. Burgeth. Level-set methods for tensor-valued images. In *Proc. 2nd IEEE Workshop Variational, Geometric and Level Set Methods in Computer Vision*, pages 65–72, 2003.
37. C. Feddern, J. Weickert, B. Burgeth, and M. Welk. Curvature-driven PDE methods for matrix-valued images. *Int. J. Comput. Vision*, 69:93–107, 2006.
38. R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast geodesic active contours. *IEEE T. Image. Process.*, 10(10):1467 – 1475, 2001.
39. T. Goldstein, X. Bresson, and S. Osher. Geometric applications of the split Bregman method: Segmentation and surface reconstruction. *J. Sci. Comput.*, 45:272–293, 2010.
40. C. L. Guyader and L. Vese. Self-repelling snakes for topology-preserving segmentation models. Technical Report 07-20, UCLA, 2007.
41. X. Han, C. Xu, and J. L. Prince. A 2D moving grid geometric deformable model. In *Computer Vision and Pattern Recognition*, pages I:153–160, Madison, Wisconsin, June 2003.
42. X. Han, C. Xu, and J. L. Prince. A topology preserving level set method for geometric deformable models. *IEEE Trans. Pattern Anal. Machine Intell.*, 25:755–768, 2003.

43. M. Hernandez and A. F. Frangi. Non-parametric geodesic active regions: Method and evaluation for cerebral aneurysms segmentation in 3DRA and CTA. *Med. Image Anal.*, 11:224–241, 2007.
44. S. Jehan-Besson, M. Barlaud, and G. Aubert. Dream2s: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation. *Int. J. Comput. Vision*, 53:45–70, 2003.
45. L. Jonassona, X. Bressona, P. Hagmanna, O. Cuisenaire, R. Meulib, and J. -P. Thiran. White matter fiber tract segmentation in DT-MRI using geometric flows. *Med. Image Anal.*, 9: 223–236, 2005.
46. L. Jonassona, P. Hagmanna, C. Polloa, X. Bressona, C. R. Wilsona, R. Meulib, and J. -P. Thiran. A level set method for segmentation of the thalamus and its nuclei in DT-MRI. *Signal Process.*, 87:309–321, 2007.
47. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Intl. J. Comp. Vision*, 1(4):321–331, 1988.
48. A. Kenigsberg, R. Kimmel, and I. Yavneh. A multigrid approach for fast geodesic active contours. Technical report, Technion - I.I.T, Haifa 32000, Israel, 2004.
49. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contours. In *International Conference on Computer Vision*, pages 810–815, Boston, USA, 1995.
50. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Conformal curvature flows: From phase transitions to active vision. *Arch. Ration. Mech. Anal.*, 134:275–301, 1996.
51. J. Kim, J. W. Fisher, A. Yezzi, Mujdatetin, and A. S. Willsky. A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE T. Image Process.*, 14:1486–1502, 2005.
52. B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations I: the components of two-dimensional shape and the reaction-diffusion space. *Int. J. Comput. Vision*, 15:189–224, 1995.
53. P. Knupp and S. Steinberg. *Fundamentals of Grid Generation*. CRC Press, Boca Raton, FL, 1994.
54. T. Kohlberger, D. Cremers, M. Rousson, R. Ramaraj, and G. Funka-Lea. 4D shape priors for a level set segmentation of the left myocardium in SPECT sequences. In *Med Image Comput Comput Assist Interv.*, volume 9, pages 92–100, 2006.
55. T. Y. Kong and A. Rosenfeld. Digital topology: Introduction and survey. *CVGIP: Image Understanding*, 48:357–393, 1989.
56. K. Krissian and C. -F. Westin. Fast sub-voxel re-initialization of the distance map for level set methods. *Pattern Recogn. Lett.*, 26:1532–1542, 2005.
57. M. Leventon, E. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition*, volume 1, pages 316–322, 2000.
58. C. Li, C. Xu, C. Cui, and M. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE T. Image Process.*, 19:3243–3254, 2010.
59. C. Li, C. Xu, C. Gui, and M. Fox. Level set evolution without re-initialization: a new variational formulation. In *Computer Vision and Pattern Recognition*, volume 1, pages 430–436, 2005.
60. F. Li, C. Shen, and C. Li. Multiphase soft segmentation with total variation and H^1 regularization. *J. Math. Imaging Vis.*, 37:98–111, 2010.
61. H. Li and A. J. Yezzi. Local or global minima: Flexible dual-front active contours. *IEEE Trans. Pattern Anal. Machine Intell.*, 29(1):1–14, 2007.
62. S. Li, T. Fevens, A. Krzyzak, C. Jin, and S. Li. Fast and robust clinical triple-region image segmentation using one level set function. In *Med Image Comput Comput Assist Interv.*, volume 9, pages 766–773, 2006.
63. G. Liao, F. Liu, G. de la Pena, D. Peng, and S. Osher. Level-set-based deformation methods for adaptive grids. *J. Comput. Phys.*, 159:103–122, 2000.
64. A. Litvin and W. C. Karl. Coupled shape distribution-based segmentation of multiple objects. In *Information Processing in Medical Imaging*, volume 3565, pages 345–356, 2005.

65. W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. In *ACM SIGGRAPH Computer Graphics*, volume 21, pages 163–169, 1987.
66. L. M. Lorigo, O. Faugeras, and W. Grimson. Co-dimension 2 geodesic active contours for MRA segmentation. In *Information Processing in Medical Imaging*, volume 1613, pages 126–139, 1999.
67. L. M. Lorigo, O. D. Faugeras, W. E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C. -F. Westin. Curves: Curve evolution for vessel segmentation. *Med. Image Anal.*, 5: 195–206, 2001.
68. R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. Pattern Anal. Machine Intell.*, 17:158–175, 1995.
69. A. -R. Mansouri, A. Mitiche, and C. Vázquez. Multiregion competition: a level set extension of region competition to multiple region image partitioning. *Comput. Vis. Image Underst.*, 101:137–150, 2006.
70. T. McInerney and D. Terzopoulos. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *Comput. Med. Imag. Grap.*, 19:69–83, 1995.
71. T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Med. Image Anal.*, 1:91–108, 1996.
72. J. Melonakos, E. Pichon, S. Angenent, and A. Tannenbaum. Finsler active contours. *IEEE Trans. Pattern Anal. Machine Intell.*, 30:412–423, 2008.
73. D. Metaxas. *Physics-Based Deformable Models: Applications to Computer Vision, Graphics and Medical Imaging*. Kluwer Academic Publishers, 1996.
74. R. B. Milne. *Adaptive Level Sets Methods Interfaces*. PhD thesis, Dept. Math., UC Berkely, 1995.
75. J. Montagnat, H. Delingette, and N. Ayache. A review of deformable surfaces: Topology, geometry and deformation. *Image Vision Comput.*, 19:1023–1040, 2001.
76. S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.
77. G. Papandreou and P. Maragos. Multigrid geometric active contour models. *IEEE T. Image. Process.*, 16:229–240, 2007.
78. N. Paragios and R. Deriche. Unifying boundary and region-based information for geodesic active tracking. In *Computer Vision and Pattern Recognition*, volume 2, pages 300–305, 1999.
79. N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: A level set approach. In *European Conference in Computer Vision*, volume 1843, pages 224–240, 2000.
80. N. Paragios and R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans. Pattern Anal. Machine Intell.*, 22(3):1–15, 2000.
81. E. Parzen. On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics*, 33:1065C1076, 1962.
82. D. Peng, B. Merriman, S. Osher, H. Zhao, and M. Kang. A PDE-based fast local level set method. *J. Comput. Phys.*, 155:410–438, 1999.
83. M. Rochery, I. H. J. C. Information, and J. Zerubia. Higher order active contours. *Int. J. Comput. Vision*, 69:27–42, 2006.
84. R. Ronfard. Region-based strategies for active contour models. *Int. J. Comput. Vision*, 13: 229–251, 1994.
85. M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 27:832–837, 1956.
86. M. Rousson and D. Cremers. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In *Med Image Comput Comput Assist Interv.*, volume 8, pages 757–764, 2005.
87. M. Rousson and N. Paragios. Shape priors for level set representations. In *European Conference on Computer Vision*, volume 2351, pages 78–92, 2002.
88. M. Rousson, N. Paragios, and R. Deriche. Implicit active shape models for 3D segmentation in MR imaging. In *Medical Image Computing and Computer-Assisted Intervention*, volume 3216, pages 209–216, 2004.

89. C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia. A level set model for image classification. *Int. J. Comput. Vision*, 40:187–197, 2000.
90. G. Sapiro and A. Tannenbaum. Affine invariant scale-space. *Int. J. Comput. Vision*, 11:25–44, 1993.
91. T. B. Sebastian, H. Tek, J. J. Crisco, S. W. Wolfe, and B. B. Kimia. Segmentation of carpal bones from 3D CT images using skeletally coupled deformable models. *Med. Image Anal.*, 7:21–45, 2003.
92. F. Ségonne. Active contours under topology control genus preserving level sets. *Int. J. Comput. Vision*, 79:107–117, 2008.
93. F. Ségonne, J. -P. Pons, E. Grimson, and B. Fischl. Active contours under topology control genus preserving level sets. In *Computer Vision for Biomedical Image Applications*, volume 3765, pages 135–145, 2005.
94. J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci.*, 93:1591–1595, 1996.
95. J. A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge, UK, 2nd edition, 1999.
96. Y. Shi and W. Karl. A fast level set method without solving PDEs. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 97–100, 2005.
97. Y. Shi and W. C. Karl. Differentiable minimin shape distance for incorporating topological priors in biomedical imaging. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, volume 2, pages 1247–1250, 2004.
98. K. Siddiqi, Y. B. Lauziere, A. Tannenbaum, and S. W. Zucker. Area and length minimizing flow for shape segmentation. *IEEE T. Image. Process.*, 7:433–443, 1998.
99. L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. Pattern Anal. Machine Intell.*, 15:1061–1075, 1992.
100. B. Sumengen and B. Manjunath. Graph partitioning active contours (GPAC) for image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, 28:509– 521, 2006.
101. G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Sobolev active contours. In *Variational, Geometric, and Level Set Methods in Computer Vision*, volume 3752, pages 109–120, 2005.
102. G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Sobolev active contours. *Int. J. Comput. Vision*, 73(3):345–366, 2006.
103. G. Sundaramoorthi and A. J. Yezzi. Global regularizing flow with topology preservation for active contours and polygons. *IEEE T. Image. Process.*, 16(3):803–812, 2007.
104. M. Sussman, A. S. Almgren, J. B. Bell, P. Colella, L. H. Howell, and M. L. Welcome. An adaptive level set approach for incompressible two-phase flow. *J. Comput. Phys.*, 148: 81–124, 1999.
105. L. Tan and N. Zabarar. Modeling the growth and interaction of multiple dendrites in solidification using a level set method. *J. Comput. Phys.*, 226:131–155, 2007.
106. A. Tsai, W. Wells, C. Tempny, E. Grimson, and A. Willsky. Mutual information in coupled multi-shape model for medical image segmentation. *Med. Image Anal.*, 4:429–445, 2004.
107. A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, W. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE T. Med. Imaging.*, 22:137–154, 2003.
108. A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, E. Grimson, and A. Willsky. Model-based curve evolution technique for image segmentation. In *Computer Vision and Pattern Recognition*, volume 1, pages 463–468, 2001.
109. J. N. Tsitsiklis. Efficient algorithm for globally optimal trajectories. *IEEE T. Automat. Contr.*, 40(9):1528–1538, 1995.
110. A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *IEEE Trans. Pattern Anal. Machine Intell.*, 24:1565– 1578, 2002.
111. L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *Int. J. Comput. Vision*, 50:271–293, 2002.
112. Z. Wang and B. C. Vemuri. Tensor field segmentation using region based active contour model. In *European Conference on Computer Vision*, volume 3024, pages 304–315, 2004.

113. J. Weickert and G. Kuhne. Fast methods for implicit active contour models. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer, 2003.
114. J. Weickert, B. Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE T. Image. Process.*, 7:398–410, 1998.
115. Y. Xiang, A. C. Chung, and J. Ye. A new active contour method based on elastic interaction. In *Computer Vision and Pattern Recognition*, volume 1, pages 452–457, 2005.
116. X. Xie and M. Mirmehdi. RAGS: Region-aided geometric snake. *IEEE T. Image Process.*, 13:640–652, 2004.
117. X. Xie and M. Mirmehdi. Magnetostatic field for the active contour model: A study in convergence. In *British Machine Vision Conference*, pages 127–136, 2006.
118. X. Xie and M. Mirmehdi. MAC: Magnetostatic active contour model. *IEEE Trans. Pattern Anal. Machine Intell.*, 30:632–646, 2008.
119. C. Xu, D. L. Pham, and J. L. Prince. *Handbook of Medical Imaging – Volume 2: Medical Image Processing and Analysis*, chapter Image Segmentation Using Deformable Models, pages 129–174. SPIE Press, 2000.
120. C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. Imag. Proc.*, 7(3):359–369, 1998.
121. C. Xu, A. Yezzi, and J. L. Prince. A summary of geometric level-set analogues for a general class of parametric active contour and surface models. In *Variational and Level Set Methods in Computer Vision*, pages 104–111, 2001.
122. M. Xu, P. M. Thompson, and A. W. Toga. An adaptive level set segmentation on a triangulated mesh. *IEEE T. Med. Imaging.*, 23(2):191–201, 2004.
123. P. Yan and A. A. Kassim. Segmentation of volumetric MRA images by using capillary active contour. *Med. Image Anal.*, 10:317–329, 2006.
124. J. Yang and J. S. Duncan. 3D image segmentation of deformable objects with shape-appearance joint prior models. In *Medical Image Computing and Computer-Assisted Intervention*, volume 2878, pages 573–580, 2003.
125. J. Yang, L. H. Staib, and J. S. Duncan. Neighbor-constrained segmentation with 3D deformable models. *IEEE T. Med. Imaging.*, 23:940–948, 2004.
126. S. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu. Geometrically induced force interaction for three dimensional deformable models. *IEEE T. Image Process.*, 20:1373–1387, 2011.
127. A. Yezzi, S. Kichenamy, A. Kumar, P. Olver, and A. Tannebaum. A geometric snake model for segmentation of medical imagery. *IEEE T. Med. Imaging.*, 16:199–209, 1997.
128. A. Yezzi, A. Tsai, and A. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *International Conference on Computer Vision*, volume 2, pages 898–903, Corfu, Greece, 1999.
129. A. Yezzi, A. Tsai, and A. Willsky. Medical image segmentation via coupled curve evolution equations with global constraints. In *Mathematical Methods in Biomedical Image Analysis*, pages 12–19, 2000.
130. H. Zhao. Fast sweeping method for Eikonal equations. *Math. Computation*, 74:603–627, 2004.
131. H. K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *J. Comput. Phys.*, 127:179–195, 1996.
132. S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, 18:884–900, 1996.

Active Shape and Appearance Models

T.F. Cootes, M.G. Roberts, K.O. Babalola, and C.J. Taylor

Abstract Statistical models of shape and appearance are powerful tools for medical image analysis. The shape models can capture the mean and variation in shape of a structure or set of structures across a population. They can be used to help interpret new images by finding the parameters which best match an instance of the model to the image. Two widely used methods for matching are the Active Shape Model and the Active Appearance Model. We describe the models and the matching algorithms, and give examples of their use.

1 Introduction

Although organs and structures in the body can exhibit a huge range of variation across a population, the shape of many can be characterised as being a transformed version of some template or reference shape. For instance, almost everyone’s face can be thought of as a variant of an ‘average’ face, with two eyes, a nose and a mouth, though in different positions on each individual. Similarly, almost every human femur has a shape which is a transformed version of the average. We can

T.F. Cootes (✉)
Centre for Imaging Sciences, University of Manchester Stopford Building, Oxford Road,
Manchester M13 9PT, England
e-mail: t.cootes@manchester.ac.uk

M.G. Roberts
Institute of Population Health Sciences, University of Manchester, UK

K.O. Babalola
Research Associate, Manchester Metropolitan University, UK

C.J. Taylor
Associate Vice President Research, University of Manchester, UK

thus capture the range of shapes of such objects by recording the average shape, and the ways in which they may vary across a population. In this chapter we describe a simple method of achieving this.

We represent a shape using a set of points (sometimes called ‘landmarks’), which define equivalent positions on each example. It should be noted that this can only be applied to objects whose shape can be consistently described in this manner. It cannot be applied to structures where we cannot define a simple correspondence across shapes. For instance, since two trees may have different numbers of branches, there is no simple way of placing meaningful landmarks across a set of trees.

For anatomical structures where this assumption holds, we can construct statistical shape models to summarise the variation across a population. Such models can be used for image interpretation. The parameters of a model define a specific shape. Given a new image, our goal is to find the model parameters which generate a shape as close as possible to that of the object in the image. This requires additional information in the form of a description of how any shape appears - what patterns of intensity information are associated with a particular shape. This too can be learned from a training set of images. However, there are many ways in which we can represent the intensity information. For instance, we could model the intensities across the whole of the object, or we can focus on areas around the boundaries of interest. Which approach is most useful depends on the application.

Given a model of shape, and some representation of intensities associated with the shape, matching the model to the image becomes an optimisation problem. Since we typically have many model parameters, it is a potentially difficult one. We describe two approaches, the Active Shape Model and the Active Appearance Model, both of which have been found to be effective. Both are iterative, local search algorithms. Thus they both require a sensible initialisation if they are to avoid falling into local minima.

2 Statistical Shape Models

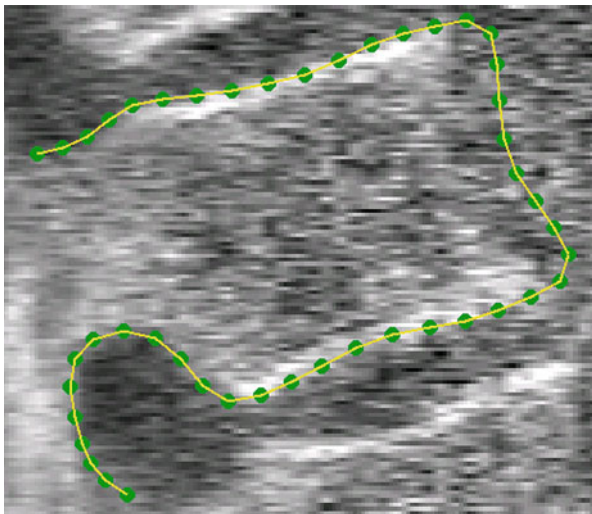
Shape is usually defined as that quality of a configuration of points which is invariant under some transformation. In two or three dimensions we usually consider either the similarity transformation (translation, rotation and scaling) or the affine transformation.

Let $T_{\mathbf{t}}(\mathbf{x})$ apply a transformation defined by parameters \mathbf{t} . In 2D, similarity transformations have 4 parameters, affine transformations are defined by 6 parameters. The configurations of points defined by \mathbf{x} and $T_{\mathbf{t}}(\mathbf{x})$ are considered to have the same *shape*. Shape differences are those changes that cannot be explained by application of such a global transformation. If we use n 2D points, $\{(x_j, y_j)\}$, to describe the shape, then we can represent the shape as the $2n$ element vector, \mathbf{x} , where

$$\mathbf{x} = (x_1, \dots, x_n, y_1, \dots, y_n)^T \quad (1)$$

For instance, Fig. 1 gives an example of a set of 46 points used to define the shape of the outline of a vertebra in a DXA image

Fig. 1 Set of points used to define the shape of a vertebra in a DXA image (image enhanced to aid visibility of structures)



Given s training examples, we generate s such vectors \mathbf{x}_i ($i = 1 \dots s$). Before we can perform statistical analysis on these vectors it is important that the shapes represented are in the same co-ordinate frame. This can be achieved by using Procrustes Analysis [15]. This transforms each shape in a set, \mathbf{x}_i , so that the sum of squared distances of the shape to the mean ($D = \sum |\mathcal{S}(\mathbf{x}_i) - \bar{\mathbf{x}}|^2$) is minimised.

Let the vector \mathbf{x}_i contain the n coordinates of the i^{th} shape. These vectors form a distribution in $2n$ dimensional space. If we can model this distribution, we can generate new examples, similar to those in the original training set, and we can examine new shapes to decide whether they are plausible examples.

To simplify the problem, we first wish to reduce the dimensionality of the data from nd to something more manageable. An effective approach is to apply Principal Component Analysis (PCA) to the data [10]. The data form a cloud of points in the $2n$ -D space. PCA computes the main axes of this cloud, allowing one to approximate any of the original points using a model with fewer than $2n$ parameters. The result is a linear model of the form

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \quad (2)$$

where $\bar{\mathbf{x}}$ is the mean of the data, $\mathbf{P} = (\phi_1 | \phi_2 | \dots | \phi_t)$ contains the t eigenvectors of the covariance matrix of the training set, corresponding to the largest eigenvalues, \mathbf{b} is a t dimensional parameter vector.

Equation (2) generates a shape in the model frame, where the effects of the global transformation have been removed by the Procrustes Analysis. A shape in the image frame, \mathbf{X} , can be generated by applying a suitable transformation to the points, \mathbf{x} : $\mathbf{X} = T_t(\mathbf{x})$. Typically T_t will be a similarity transformation described by a scaling, s , an in-plane rotation, θ , and a translation (t_x, t_y) .

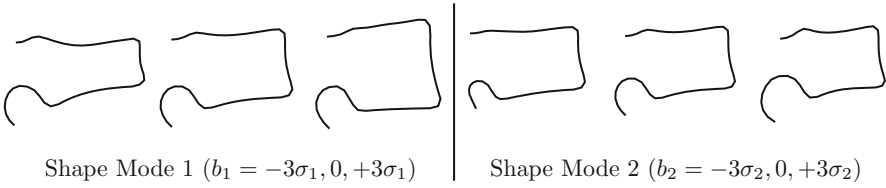


Fig. 2 First two modes of a shape model of a vertebra (Parameters varied by ± 3 s.d. from the mean)

The best choice of parameters for a given shape \mathbf{x} is given by

$$\mathbf{b} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}}) \quad (3)$$

The vector \mathbf{b} defines a set of parameters of a deformable model. By varying the elements of \mathbf{b} we can vary the shape, \mathbf{x} , using Eq. (2). The variance of the i^{th} parameter, b_i , across the training set is given by λ_i . By applying suitable limits to the model parameters we ensure that the shape generated is similar to those in the original training set.

If the original data, $\{\mathbf{x}_i\}$, is distributed as a multivariate Gaussian, then the parameters \mathbf{b} are distributed as an axis-aligned Gaussian, $p(\mathbf{b}) = N(\mathbf{0}, \Lambda)$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_l)$. In practise it has been found that the assumption of normality works well for many anatomical structures.

2.1 Examples of shape models

Figure 2 shows the effect of changing the first two shape parameters on a model of the outline of a vertebra, trained from 350 examples such as that shown in Fig. 1.

Figure 3 shows the effect of varying the parameter controlling the first mode of shape variation of a 3D model of a set of deep brain structures, constructed from shapes extracted from MR images of 69 different subjects¹. Note that since the method only represents points, it easily model multiple structures.

3 Active Shape Models

Given a new image containing the structure of interest, we wish to find the pose parameters, \mathbf{t} , and the shape parameters, \mathbf{b} , which best approximate the shape of the object. To do this we require a model of how well a given shape would match to the image.

¹Provided by David Kennedy at the Centre for Morphometric Analysis

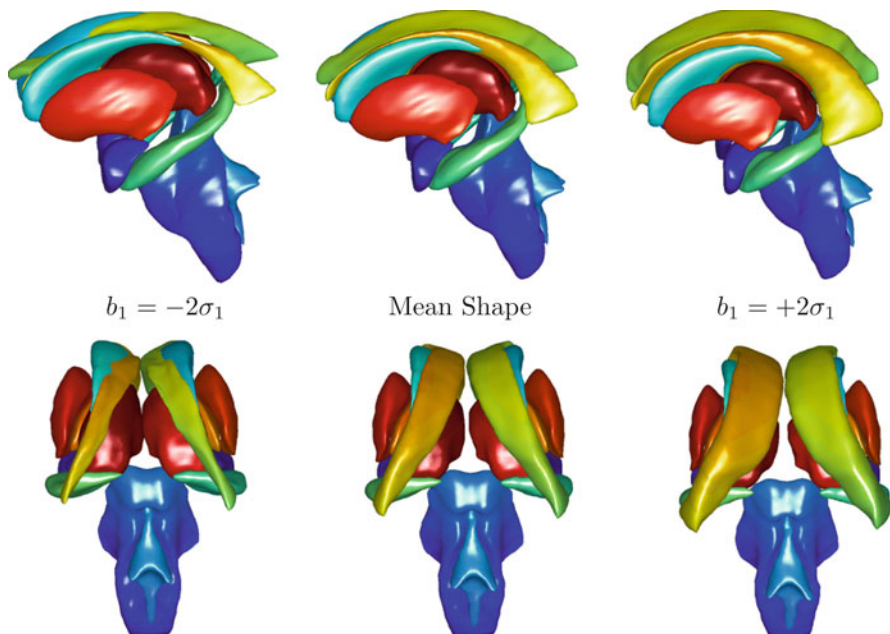


Fig. 3 Two views of the first mode of a shape model of structures in the brain

Suppose we have a set of local models, each of which can estimate how well the image information around a point matches that expected from the training set, returning a value $q_i(X_i, Y_i)$. We then seek the shape and pose parameters which generate points $\mathbf{X} = \{X_i, Y_i\}$ so as to maximise

$$Q_{asm}(\mathbf{t}, \mathbf{b}) = \sum_{i=1}^n q_i(X_i, Y_i) \quad (4)$$

where $\mathbf{X} = T_t(\bar{\mathbf{x}} + \mathbf{Pb})$.

The basic Active Shape Model algorithm is a simple iterative approach to locating this optima. Each iteration involves two steps

1. Search around each point (X_i, Y_i) for a better position, (X'_i, Y'_i)
2. Update the pose and shape parameters to best fit the model to the new positions \mathbf{X}'

By decoupling the search and update steps the method can be fast. Since each point is searched for independently, it is simple to implement and can make use of parallel processing where available.

Because of the local nature of the search, there is a danger of falling into local minima. To reduce the chance of this it is common to use a multi-resolution framework, in which we first search at coarse resolutions, with low resolution models, then refine at finer resolutions as the search progresses. This has been shown to significantly improve the speed, robustness and final accuracy of the method.

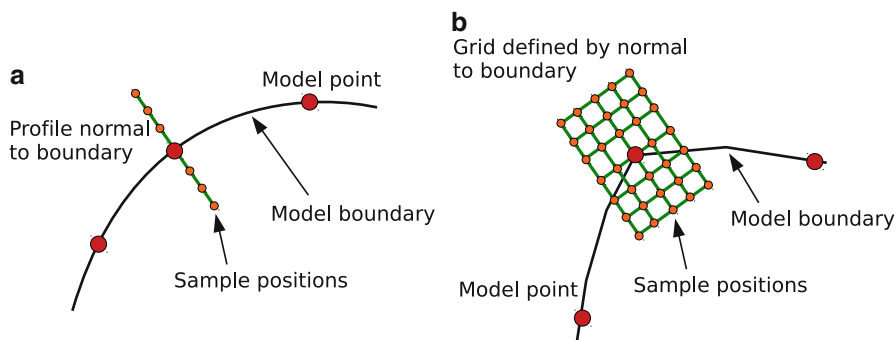


Fig. 4 ASM search requires models of the image around each point

The approach has been found to be very effective, and numerous variants explored. In the following we will summarise some of the more significant approaches to each step.

3.1 Searching for Model Points

The simplest approach is to have $q_i(X, Y)$ return the edge strength at a point. However we will usually have a direction and scale associated with each point, allowing the local models to be more specific. For instance, since most points are defined along boundaries (or surfaces in 3D), the normal to the boundary/surface defines a direction, and a scale can be estimated from the global pose transformation.

Since we have a training set, a natural approach is to learn statistical models of the local structure from that around each point in the known examples [10]. For each model point in each image, we can sample at a set of nearby positions. For instance, at points along smooth curves we can sample profiles (Fig. 4a), whereas at corner points we can sample on a grid around the point (Fig. 4b).

The samples around model point i in training image j can be concatenated into a vector, \mathbf{g}_{ij} . Typically each such vector is normalised to introduce invariance to overall intensity effects (for instance, by subtracting the mean intensity and scaling so the vector is of unit length). We then can estimate the probability density distribution $p_i(\mathbf{g})$, which may be a Gaussian, or something more sophisticated if sufficient samples are available.

Given such a PDF for each point, we can evaluate a new proposed position by sampling the image around the point into a new vector \mathbf{g} , and setting $q_i(\mathbf{X}_i) = \log p_i(\mathbf{g})$.

To search for a better position, we evaluate the PDF at a number of displaced positions, and choose the best. For profile models, we typically search at positions along a profile normal to the current boundary. For grid based models, we would test points on a grid. There are natural extensions to 3D volume images.

Rather than simply sample image intensities, it is possible to sample other features. For instance van Ginneken et al. [36] sampled Gabor features at multiple scales.

This approach has been shown to be effective in many applications [10]. However, van Ginneken et al. [36] pointed out that improved estimates of the model point positions can be obtained by explicitly training classifiers for each point, to discriminate the correct position from incorrect nearby positions. The approach is to treat samples at the model points in the training set as positive examples, then create a set of negative examples by sampling nearby. These examples are used to train a classifier. During search samples around a set of candidate positions are evaluated, and the one most strongly classified as positive is used. It is assumed that the classifier gives a continuous response which is larger for positive examples, and can thus be used to rank the candidates.

In the case of [36] a nearest neighbour classifier was used, but encouraging results have been obtained with a range of methods. For instance, Li and Ito [37] used Adaboost to train a classifier for facial feature point detection.

A related approach is to attempt to learn an objective function which has its minimum at the correct position, and has a smooth form to allow efficient optimisation [28].

3.2 Updating the Model Parameters

After locating the best local match for each model point, \mathbf{X}' , we update the model pose and shape parameters so as to best match these new points. This can be considered as a regularisation step, forcing points to form a ‘reasonable’ shape - one defined by the original model.

The simplest approach is to find \mathbf{t} and \mathbf{b} so as to minimise

$$|\mathbf{X}' - T_{\mathbf{t}}(\bar{\mathbf{x}} + \mathbf{P}\mathbf{b})|^2 \quad (5)$$

For anything other than when $T()$ is a translation, this is a non-linear sum of squares problem, so can be efficiently solved with Levenberg-Marquardt [29], or an iterative algorithm in which we alternate between solving for \mathbf{b} with fixed \mathbf{t} , and solving for \mathbf{t} with fixed \mathbf{b} , each of which has a closed form. To prevent the generation of extreme shapes, limits can be placed on the parameters. If we assume a normal distribution for the original shapes, then the parameters b_i are linearly independent with a variance λ_i . A natural constraint is then to force

$$M = \sum_{i=1}^t \frac{b_i^2}{\lambda_i} < M_{thresh} \quad (6)$$

Since M will be distributed as χ^2 with t degrees of freedom, this distribution can be used to select a suitable value of M_{thresh} (for instance, one that includes 98% of the normal samples).

This approach effectively treats every choice of shape parameters within the threshold as equally likely. An alternative is to include a prior, rather than a hard threshold, for instance

$$\log p(\mathbf{t}, \mathbf{b} | \mathbf{X}') = const - |T_{\mathbf{t}}^{-1}(\mathbf{X}') - (\bar{\mathbf{x}} + \mathbf{P}\mathbf{b})|^2 / \sigma_r^2 - \sum_{i=1}^t \frac{b_i^2}{\lambda_i} \quad (7)$$

where σ_r^2 is the variance of the residuals (those not explained by the model) in the model frame and we assume all pose parameters are equally likely. We can thus find the most likely values for \mathbf{t}, \mathbf{b} by optimising (7). Again, this is a non-linear sum of squares problem, for which fast iterative algorithms exist (for instance, see [38]).

Information about the uncertainty of the estimates of the best matching point positions can also be included into the optimisation [17].

Since the models for each point may return false matches (for instance by latching on to the wrong nearby structure), it can be effective to use a more robust method in which outliers are detected and discarded [27].

4 Active Appearance Models

The Active Shape Model treats the search for each model point as independent. However, there will usually be correlations between the local appearance around nearby points. This can be modelled, but leads to a more difficult optimisation problem. The Active Appearance Model algorithm is a local optimisation technique which can match models of appearance to new images efficiently.

4.1 Appearance Models

We assume that each example of an object in a class can be thought of as being generated by creating a version with an ‘average’ shape in a reference frame, then deforming it using the shape model.

To train a statistical model we can warp each training example back to the reference frame, and analyse the variation of the ‘shape-corrected’ intensities.

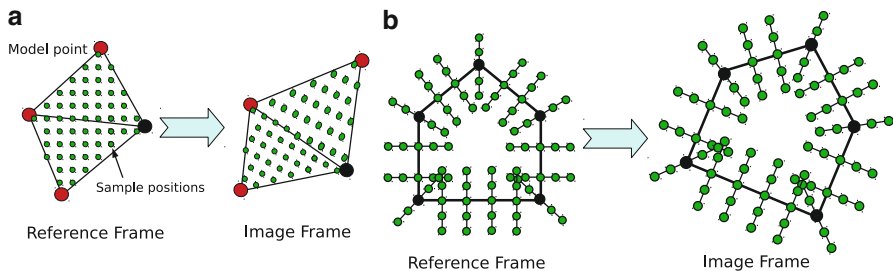


Fig. 5 Sample points can be on a dense grid, or along profiles around boundaries

More formally, let $W(\mathbf{y} : \mathbf{b}, \mathbf{t})$ apply a deformation of space such that the mean model points $\bar{\mathbf{x}}$ are mapped to the points defined by the shape model

$$W(\bar{\mathbf{x}} : \mathbf{b}, \mathbf{t}) = T_{\mathbf{t}}(\bar{\mathbf{x}} + \mathbf{Pb}) \quad (8)$$

and that other points are suitably interpolated.²

Let $\mathbf{f} = F(I, \mathbf{Y})$ be a set of features sampled at points \mathbf{Y} in image I . These can be interpolated intensities, gradient information or something more sophisticated. Note that the feature sampling will usually include some normalisation, for instance, arranging that the mean of the elements in \mathbf{f} is zero and their variance is unity, to introduce some invariance to intensity changes.

Let \mathbf{y} be a set of points in the reference frame (that of the mean of a model) indicating the normalised positions of the points in the image at which we wish to sample. This may be a dense grid of points across the region, or a subset of points around the model boundaries (see Fig. 5).

Given a set of training images, I_i , together with a set of corresponding model points \mathbf{X}_i for each image, we can construct a statistical model as follows.

1. Construct a statistical shape model from the corresponding points as described above.
2. For each image, I_i ,
 - (a) Estimate the optimal shape and pose parameters, $\mathbf{b}_i, \mathbf{t}_i$
 - (b) Sample the features³ $\mathbf{f}_i = F(I_i, W(\mathbf{y} : \mathbf{b}_i, \mathbf{t}_i))$
3. Apply PCA to the set of feature vectors $\{\mathbf{f}_i\}$ to generate a linear model of the form $\mathbf{f} = \mathbf{f} + \mathbf{F}\mathbf{a}$, where \mathbf{F} contains the most significant eigenvectors of the covariance matrix of the feature vectors.

Thus for each training example we have a shape, summarised by \mathbf{t} and \mathbf{b} , and a set of feature samples, summarised by $\mathbf{a} = \mathbf{F}^T(\mathbf{f} - \mathbf{f})$.

²We use a variant of the notation introduced by Matthews and Baker [22]

³Note that we sample the features at points based on the best shape model approximation to the shape \mathbf{X} , not on \mathbf{X} itself. Thus any approximation errors in the shape are absorbed into uncertainty in the feature samples, ensuring we are better able to reconstruct the original training set if required.

It is often the case that there is significant correlation between the shape and the feature samples. We can model this explicitly by generating joint vectors $\mathbf{j} = (\mathbf{b}^T | \mathbf{W}\mathbf{a}^T)^T$, where \mathbf{W} is a diagonal weighting matrix, chosen to account for the difference in units between the shape and the features. A useful choice of $\mathbf{W} = \alpha\mathbf{I}$, where α is chosen so that the shape and scaled feature components have similar variance. If we apply a third PCA to these joint vectors, we can construct a combined linear model of the form

$$\begin{pmatrix} \mathbf{b} \\ \mathbf{W}\mathbf{a} \end{pmatrix} = \mathbf{j} = \mathbf{Q}\mathbf{c} \quad (9)$$

which can be decomposed into separate shape and feature terms,

$$\begin{aligned} \mathbf{x} &= \bar{\mathbf{x}} + \mathbf{Q}_s\mathbf{c} \\ \mathbf{f} &= \bar{\mathbf{f}} + \mathbf{Q}_f\mathbf{c} \end{aligned} \quad (10)$$

where \mathbf{c} is a set of parameters which control both the shape and the feature model (see [7] for more details).

4.2 Image Search

Given a new image we wish to find the model parameters which generate a shape and features most similar to those in the image. This is an *interpretation through synthesis* approach.

For a given choice of parameters, $\mathbf{p} = (\mathbf{t}^T, \mathbf{c}^T)^T$, we can generate shape parameters \mathbf{b} using Eq. (9) and corresponding features \mathbf{f} using Eq. (10).

Let $\mathbf{f}' = F(I_i, W(\mathbf{y} : \mathbf{b}, \mathbf{t}))$ be the features sampled from the image given the current shape. The residual difference between those features given by the model and those sampled from the image is

$$\mathbf{r}(\mathbf{p}) = \mathbf{f}' - \mathbf{f} \quad (11)$$

The residual is a function of the model parameters. The way in which the residual varies as we vary the parameters is approximated by Jacobian, \mathbf{J} ,

$$J_{ij} = \frac{\partial r_i}{\partial p_j} \quad (12)$$

To match the model to the image we seek the parameters which minimise the sum of squares of the residual,

$$E(\mathbf{p}) = |\mathbf{r}(\mathbf{p})|^2 = |\mathbf{f}' - \mathbf{f}|^2 \quad (13)$$

This problem can be solved efficiently with a fast iterative gradient descent method. If our current parameters are \mathbf{p} , we seek an update step, $\delta\mathbf{p}$ which improves our match. It can be shown [7, 22] that a good estimate of the step is given by

$$\delta\mathbf{p} = -\mathbf{R}\mathbf{r} \quad (14)$$

where \mathbf{R} is the pseudo-inverse of the Jacobian of $\mathbf{r}(\mathbf{p})$,

$$\mathbf{R} = (\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}^T \quad (15)$$

This requires an estimate of the Jacobian at the current parameters, which can be relatively expensive to compute. However, since the residual is measured in the normalised reference frame, for many problems it is found that the Jacobian is approximately constant over a reasonable range of parameters. Thus we can precompute it, by numeric differentiation on the training set [7, 22]. Alternatively good results can be obtained by treating (14) as a regression problem and learning \mathbf{R} from large numbers of randomly sampled displacements across the training set [5].

A simple algorithm to search for the best match given this relationship is then

1. Evaluate the residual $\mathbf{r}(\mathbf{p}) = \mathbf{f}' - \mathbf{f}$
2. Evaluate the current error $E_0 = |\mathbf{r}|^2$
3. Predict the displacement, $\delta\mathbf{p} = -\mathbf{R}\mathbf{r}$
4. If $|\mathbf{r}(\mathbf{p} + \delta\mathbf{p})|^2 < E_0$ then accept the new estimate, $\mathbf{p} := \mathbf{p} + \delta\mathbf{p}$, otherwise perform line search along $\mathbf{p} + \alpha\delta\mathbf{p}$
5. Repeat until convergence

Typically only a small number of iterations are required. The algorithm is usually used in a multi-resolution framework, in which models trained at a coarse resolution are used to get an approximate estimate, which is then refined using higher resolution models.

4.3 Updating the parameters

In the above we use a simplification that the parameters should be updated in an additive fashion ($\mathbf{p} := \mathbf{p} + \delta\mathbf{p}$). However, in practise a slightly more complicated compositional approach should be used [7, 22].

Consider, for instance, the parameters defining the pose transformation, $T_{\mathbf{t}}(\mathbf{x})$. The feature sampling effectively warps the image into the reference frame, and the residual is computed in that frame. Suppose that the predicted update is $\delta\mathbf{t}$. The resulting transform should then be $T_{\mathbf{t}'}(\mathbf{x}) = T_{\mathbf{t}}(T_{\delta\mathbf{t}}(\mathbf{x}))$ (we first apply the update in the reference frame, then apply the original transformation). Thus the correct approach is to use a compositional updating scheme. Matthews and Baker [22] point out that the shape update should also be treated as a transformation, and a compositional scheme used for that, rather than simply updating the shape

parameters linearly. Referring to Eq. (8), if $\delta\mathbf{t}$ and $\delta\mathbf{b}$ are the pose and shape updates, the new parameter values, \mathbf{t}' and \mathbf{b}' should be chosen so that

$$W(\bar{\mathbf{x}} : \mathbf{b}', \mathbf{t}') \approx W(W(\bar{\mathbf{x}} : \delta\mathbf{b}, \delta\mathbf{t}) : \mathbf{b}, \mathbf{t}) \quad (16)$$

However, this is more complex, and in practise it seems that a linear scheme is generally sufficient (though slightly less efficient), as long as the pose is dealt with correctly.

4.4 Choice of Features

The simplest approach is to sample intensity values at each point. The resulting feature vector of intensities can be normalised to allow for some variations in imaging conditions. However, methods based just on intensities can be sensitive to variation in brightness across the image and other effects. More robust results can be obtained by modeling some filtered version of the original image.

Edge-based representations tend to be less sensitive to imaging parameters than raw intensity measures. Thus an obvious alternative to modeling the intensity values directly is to record the local image gradient in each direction at each pixel. Although this yields more information at each pixel, and at first glance might seem to favor edge regions over flatter regions, it is only a linear transformation of the original intensity data. Where model building involves applying a linear PCA to the samples, the resulting model is almost identical to one built from raw intensities, apart from some effects around the border, where computing the gradients includes some background information into the model.

However, nonlinear normalization of the gradient at each pixel in the region to be modeled has been found to be a useful representation. If the local gradients at a pixel are g_x, g_y , we compute normalized features $(g'_x, g'_y) = (g_x, g_y)/(g + g_0)$ where g is the magnitude of the gradient, and g_0 is the mean gradient magnitude over a region. Building texture models of this normalized value has been shown to give more robust matching than matching intensity models [34].

Stegmann and Larsen [24] demonstrated that combining multiple feature bands (e.g., intensity, hue, and edge information) improved face feature location accuracy. Scott et al. [18] have shown that including features derived from "cornerness" measures can further improve performance.

Liu [21] uses features learnt from the training set (see below).

4.5 Shape AAMs

The original AAM manipulates the combined shape and feature parameter vector, \mathbf{c} . However, a very similar formulation can be used to instead drive the shape and pose parameters, treating them as independent of the feature parameters [6, 22].

In this formulation the explicit parameters are \mathbf{t} and \mathbf{b} . The feature parameters, \mathbf{a} , are set to the best match to the normalised features sampled from the image, \mathbf{f}' ,

$$\mathbf{a} = \mathbf{F}^T (\mathbf{f}' - \bar{\mathbf{f}}) \quad (17)$$

The residual is then a function of \mathbf{t} and \mathbf{b} alone,

$$\mathbf{r}_s(\mathbf{t}, \mathbf{b}) = \mathbf{f}' - (\bar{\mathbf{f}} + \mathbf{F}\mathbf{a}) \quad (18)$$

Again the updates, $\delta\mathbf{t}$ and $\delta\mathbf{b}$ can be computed using a matrix multiplication with this residual,

$$\begin{aligned} \delta\mathbf{t} &= \mathbf{R}_{s1} \mathbf{r}_s(\mathbf{t}, \mathbf{b}) \\ \delta\mathbf{b} &= \mathbf{R}_{s2} \mathbf{r}_s(\mathbf{t}, \mathbf{b}) \end{aligned} \quad (19)$$

where the update matrices \mathbf{R}_{s1} and \mathbf{R}_{s2} are derived from numerical estimates of the Jacobian of \mathbf{r}_s , or learnt from training data.

The advantage of this approach is that there are fewer shape parameters than combined parameters, thus the method has the potential to be more efficient and more robust. However, whether it does indeed give better results than the formulation manipulating the combined parameters seems to be somewhat dependent on the nature of the application being addressed, so it is advisable to try both methods. In particular, where there is significant correlation between shape and texture parameters, the combined parameter update may be more robust.

4.6 Other Variants

Liu [21] describes a variant in which the image features are selected from a large basis set using an Adaboost approach, chosen so as to best discriminate between the correct position and nearby incorrect positions. Given this feature model, it is possible to evaluate the quality of fit of a particular choice of shape parameters. Liu presents an efficient gradient-descent search algorithm for optimising the shape and pose parameters.

Batur and Hayes demonstrated that modifying the update matrix \mathbf{R} , depending on the current estimates of the model parameters, can lead to more accurate and reliable matching [2].

Cootes and Taylor demonstrated that the estimate of the Jacobian can be refined during the search itself, leading to a better overall result [3].

Various authors have extended the AAM to use robust estimates for the model building and parameter updates [13, 16, 25].

Stegmann et al. have done extensive work with Active Appearance Models, and have made their software available [23].

The approach extends naturally to 3D volume images, either by using volumetric models [26] or profiles normal to surfaces [1].

Non-linear regressors have been used in place of the linear predictor in the update step (Eq. 14), leading to improved performance if carefully trained. For instance boosted regressors [33, 35] and Random Forest regressors [30].

5 Examples of application

5.1 *Detecting Vertebral Fracture*

Osteoporosis is a progressive skeletal disease characterized by a reduction in bone density, leading to an increased risk of fractures, particularly of the hip, vertebrae and wrist. Vertebral fractures are the most common, and appear in younger patients. The detection of such fractures is often used to diagnose and monitor osteoporosis. Typically vertebral fractures are located by expert radiologists analysing either radiographs or DXA images of the spine. A fractured vertebra typically exhibits changes in shape and more subtle changes in the appearance of the ‘endplate’.

Automating the relatively time-consuming process of detecting such fractures is challenging because of the noisy and cluttered nature of the images (see Fig. 1 for an example of the appearance of a vertebra in a DXA image). However, by using statistical models of shape and appearance it is possible to accurately and reliably segment such structures [25, 31, 32].

A training set of 315 DXA images was annotated by expert radiologists. On each image the outlines of 10 vertebra (T7-L4) were marked using a total of 405 points (Fig. 6).

A global shape model was constructed from all 10 vertebrae. Figure 6 shows the first two shape modes of the model. In addition, models were constructed for each triplet of 3 consecutive vertebrae.

Given a new image, the model is initialised by the radiologist clicking the approximate centre of each vertebra. The global shape model is then matched to the input points. The result of this is then used to initialise the triplet models. Each triplet model is matched using a constrained AAM search and the quality of its final match evaluated. The best fitting triplet is selected, and the points of its central vertebra fixed. The global shape model is updated to refine the initialisation of remaining triplets, and other triplets are tried. This leads to a dynamic ordering in which the best fitting models are used to constrain the search for the models where poorer image information is available (for full details, see [31]). It can be demonstrated that decomposing a global model into a set of overlapping smaller models in this way can lead to more accurate and robust results than searching with the global model alone. This is in part due to the global model not being flexible enough due to undertraining. The system has been shown to achieve a median accuracy of 0.55 mm on normal vertebra, and 0.88 mm on the most extreme grade 3 fractures [25].

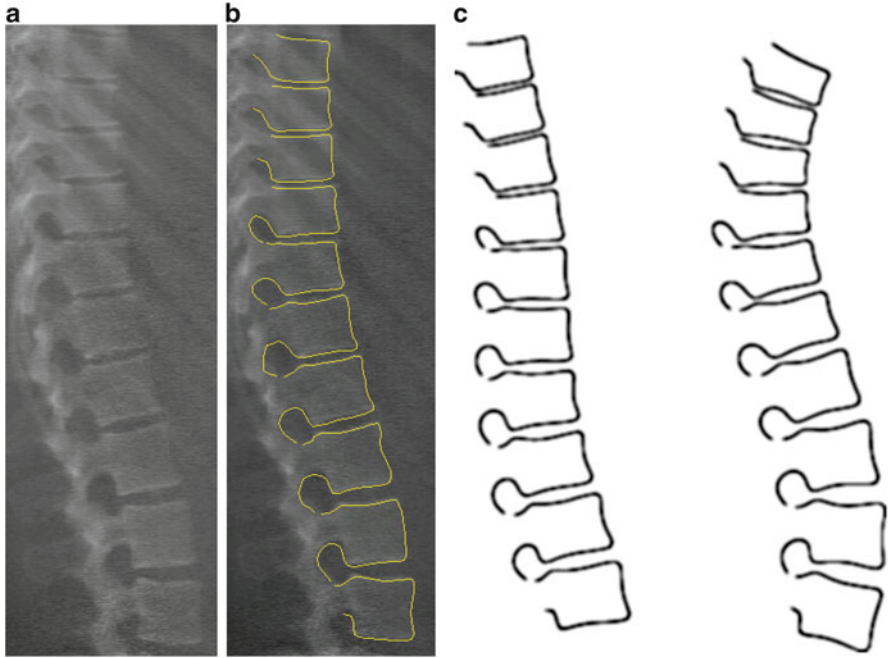


Fig. 6 Typical training image and second mode of a shape model of a spine

6 Discussion

The statistical models of shape and appearance described above are able to capture the variation in structure across a population. Such models are very useful for searching new images and for giving concise descriptions of new examples of the modelled structure.

Though they can only be used for objects which can be modelled as a deformed version of some template, they still have wide application, as many structures and organs of interest satisfy this constraint.

The choice of which approach (ASM or AAM) is most suitable is somewhat application specific. ASMs, relying on multiple local searches, can be made very efficient and are simpler to implement. AAMs are more complex, as they take account of correlations in the texture across the whole of the modelled region. However, both methods have been shown to give good results in different domains.

A limitation of using a single global shape model is that it can overconstrain the final solution, particularly when there is a lot of local variation in shape which may not be adequately captured in a small training set. Approaches to overcoming this include artificially introducing extra shape variation into the model [8, 11], applying local search to locally deform the fit of the best global result [9] and using a set of linked local models to optimise the result of a global model [31].

One of the most effective new techniques for matching shape models to new images is a variant of the Active Shape Model which uses Random Forest regressors at each point to vote for the best position of the points. Votes are accumulated from a region around the current position, and the shape model is fitted so as to maximise the total votes under each model point - see [20]. The ability of the Random Forest to capture non-linear relationships between image structure and point positions is particularly useful for achieving robust results.

Perhaps the most challenging problem associated with the use of such models is obtaining the point correspondences required for training. Though in 2D they can often be supplied manually, for large 3D datasets this is not practical. Considerable research has been done into the subject of automatically finding correspondences across large groups of shapes and images. Though it is still an active area, practical methods are now available, including [4, 12, 14, 19].

Acknowledgements We would like to thank all our colleagues in the Centre for Imaging Sciences for their help. The work was funded by the EPSRC, the MRC and the Arthritis Research Campaign.

References

1. K. O. Babalola, V. Petrovic, T. F. Cootes, C. J. Taylor, C. J. Twining, T. G. Williams, and A. Mills. Automated segmentation of the caudate nuclei using active appearance models. In *3D Segmentation in the clinic: A grand challenge*, MICCAI Workshop Proceedings, pages 57–64, 2007.
2. A. Batur and M. Hayes. Adaptive active appearance models. *IEEE Trans. Imaging Processing*, 14:1707–21, 2005.
3. T. Cootes and C. Taylor. An algorithm for tuning an active appearance model to new data. In *Proc. British Machine Vision Conference*, volume 3, pages 919–928, 2006.
4. T. F. Cootes, C. J. Twining, V. S. Petrović, K. O. Babalola, and C. J. Taylor. Computing accurate correspondences across groups of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(11):1994–2005, 2010.
5. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *ECCV*, volume 2, pages 484–498, 1998.
6. T. F. Cootes, G. J. Edwards, and C. J. Taylor. A comparative evaluation of active appearance model algorithms. In *British Machine Vision Conference*, volume 2, pages 680–689. BMVA Press, Sept. 1998.
7. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.
8. T. F. Cootes and C. J. Taylor. Combining point distribution models with shape models based on finite-element analysis. *Image and Vision Computing*, 13(5):403–409, 1995.
9. T. F. Cootes and C. J. Taylor. Combining elastic and statistical models of appearance variation. In *European Conference on Computer Vision*, volume 1, pages 149–163. Springer, 2000.
10. T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, Jan. 1995.
11. C. Davatzikos, X. Tao, and D. Shen. Hierarchical active shape models using the wavelet transform. *IEEE Trans. on Medical Imaging*, 22(3):414–423, 2003.
12. R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. 3D statistical shape models using direct optimisation of description length. In *European Conference on Computer Vision*, volume 3, pages 3–20. Springer, 2002.

13. G. Edwards, T. F. Cootes, and C. J. Taylor. Advances in active appearance models. In *Int. Conf. on Computer Vision*, pages 137–142, 1999.
14. A. Frangi, D. Rueckert, J. Schnabel, and W. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. *IEEE-TMI*, 21:1151–66, 2002.
15. C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society B*, 53(2):285–339, 1991.
16. R. Gross, I. Matthews, and S. Baker. Constructing and fitting active appearance models with occlusion. In *Proceedings of the IEEE Workshop on Face Processing in Video*, June 2004.
17. A. Hill, T. Cootes, and C. Taylor. Active shape models and the shape approximation problem. In *British Machine Vision Conference*, pages 157–166. BMVA Press, 1995.
18. I. M. Scott, T. F. Cootes, and C. J. Taylor. Improving appearance model matching using local image structure. In *Information Processing in Medical Imaging*, pages 258–269. Springer-Verlag, 2003.
19. J. Klemencic, J. Pluim, and M. Viergever. Non-rigid registration based active appearance models for 3d medical image segmentation. *Journal of Imaging Science and Technology*, 48(2):166–171, 2004.
20. C. Lindner, S. Thiagarajah, J. M. Wilkinson, arcOGEN, G. Wallis, and T. F. Cootes. Fully automatic segmentation of the proximal femur using random forest regression voting. *IEEE Trans. Medical Imaging*, 32(8):1462–1472, 2013.
21. X. Liu. Generic face alignment using boosted appearance model. In *Computer Vision and Pattern Recognition*, pages 1–8, 2007.
22. I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135 – 164, November 2004.
23. M. B. Stegmann, B. K. Ersbøll, and R. Larsen. FAME - a flexible appearance modelling environment. *IEEE Trans. on Medical Imaging*, 22(10):1319–1331, 2003.
24. M. B. Stegmann and R. Larsen. Multi-band modelling of appearance. *Image and Vision Computing*, 21(1):66–67, 2003.
25. M. G. Roberts, T. F. Cootes, and J. E. Adams. Robust active appearance models with iteratively rescaled kernels. In *Proc. British Machine Vision Conference*, volume 1, pages 302–311, 2007.
26. S. Mitchell, H. Bosch, B. F. Lelieveldt, R. van der Geest, J. Reiber, and M. Sonka. 3-D active appearance models: Segmentation of cardiac mr and ultrasound images. *Trans. Medical Imaging*, 21(9):1167–78, 2002.
27. M. Rogers and J. Graham. Robust active shape model search. In *European Conference on Computer Vision*, volume 4, pages 517–530. Springer, 2002.
28. M. Wimmer, F. Stulp, S. J. Tschechne, and B. Radig. Learning robust objective functions for model fitting in image understanding applications. In *Proc. British Machine Vision Conference*, volume 3, pages 1159–168, 2006.
29. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C (2nd Edition)*. Cambridge University Press, 1992.
30. P. Sauer, T. Cootes, and C. Taylor. Accurate regression procedures for active appearance models. In *BMVC*, 2011.
31. M. Roberts, T. Cootes, and J. Adams. Linking sequences of active appearance sub-models via constraints: an application in automated vertebral morphometry. In *14th British Machine Vision Conference*, volume 1, pages 349–358, 2003.
32. M. Roberts, T. Cootes, and J. Adams. Vertebral morphometry: Semi-automatic determination of detailed vertebral shape from dxa images using active appearance models. *Investigative Radiology*, 41(12):849–859, 2006.
33. J. Saragih and R. Goecke. A non-linear discriminative approach to AAM fitting. In *Proc. ICCV*, 2007.
34. T. F. Cootes and C. J. Taylor. On representing edge structure for model matching. In *Computer Vision and Pattern Recognition*, volume 1, pages 1114–1119, 2001.
35. P. Tresadern, P. Sauer, and T. Cootes. Additive update predictors in active appearance models. In *British Machine Vision Conference*. BMVA Press, 2010.

36. B. van Ginneken, A. F. Frangi, J. J. Stall, and B. ter Haar Romeny. Active shape model segmentation with optimal features. *IEEE-TMI*, 21:924–933, 2002.
37. Y. Li and W. Ito. Shape parameter optimization for adaboost active shape model. In *International Conference on Computer Vision*, volume 1, pages 251–258. IEEE Computer Society Press, 2005.
38. Y. Zhou, L. Gu, and H. -J. Zhang. Bayesian tangent shape model: Estimating shape and pose parameters via bayesian inference. In *Computer Vision and Pattern Recognition*, volume 1, pages 109–118, 2003.

Part II
Statistical & Physiological Models

Statistical Atlases

C. Davatzikos, R. Verma, and D. Shen

Abstract This chapter discusses the general concept of statistical atlases built from medical images. A statistical atlas is a quantitative reflection of normal variability in anatomy, function, pathology, or other imaging measurements, and it allows us to establish a baseline against which abnormal images are to be compared for diagnostic or treatment planning purposes. Constructing a statistical atlas relies on a fundamental building block, namely deformable registration, which maps imaging data from many individuals to a common coordinate system, so that statistics of normal variability, as well as abnormal deviations from it, can be performed. 3D and 4D registration methods are discussed. This chapter also discusses the statistical analyses applied to co-registered normative images, and finally briefly touches upon use of machine learning for detection of imaging patterns that distinctly deviate from the normative range to allow for individualized classification.

1 Introduction

Medical images are now used routinely in a large number of diagnostic and prognostic evaluations. Their widespread use has opened up tremendous opportunities for studying the structure and physiology of the human body, as well as the ways in which structure and function are affected by a variety of diseases and disorders.

C. Davatzikos (✉)

Department of Radiology, University of Pennsylvania, 3600 Market Street,
Suite 380, Philadelphia, PA 19104, USA
e-mail: Christos.Davatzikos@uphs.upenn.edu

R. Verma

Department of Radiology, University of Pennsylvania, PA, USA

D. Shen

Department of Radiology, University of North Carolina, NC, USA

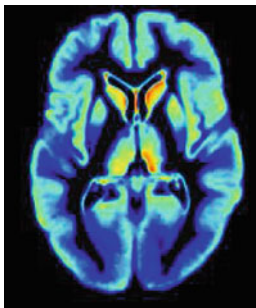


Fig. 1 A statistical atlas of the spatial distribution of gray matter (GM) in a population of elderly healthy individuals. Hot colors indicate brain regions with the highest frequency/volume of GM in the population. A new individual's spatial distribution of GM can be contrasted against this atlas, to identify regions of potentially abnormal brain atrophy

Although earlier studies typically involved a few dozens of images each, many current clinical research studies involve hundreds, and thousands of participants, often with multiple scans each. Large databases are therefore constructed rapidly, incorporating rich information of structure and function in normal and diseased states. Analysis of such a wealth of information is becoming increasingly difficult, without the availability of advanced statistical image analysis methods.

In order to be able to integrate images from different individuals, modalities, time-points, and conditions, the concept of a statistical atlas has been introduced and used extensively in the medical image analysis literature, especially in the fields of computational anatomy and statistical parametric mapping of brain functional images [8, 14, 23, 28, 30, 82, 84]. Images over certain populations. For example, a statistical atlas of the typical regional distribution of gray and white matter (GM, WM) in the brain can be constructed by spatially normalizing a number of brain images of healthy individuals into the stereotaxic space, and measuring the average and standard deviation of the amount of GM and WM in each brain region (Fig. 1). This atlas can also become more specific, for example to the age, sex, and other characteristics of the underlying population. Similarly, a statistical atlas of cardiac structure and function can provide the average myocardial wall thickness at different locations, its change over time within the cardiac cycle, and its statistical variation over a number of healthy individuals or of patients with a specific cardiac pathology. Another representative example could be an atlas of the spatial distribution of prostate cancer [98], which can be constructed from a number of patients undergoing prostatectomy, in order to guide biopsy procedures aiming to sample prostate regions that tend to present higher incidence of prostate cancer (e.g. Fig. 2).

Statistical atlases are quantitative analogs to the knowledge acquired by a radiologist during clinical training, in that they learn patterns from large number of scans, and represent anatomical or functional variability in a group of individuals. The two most common ways in which atlases are used are the following:

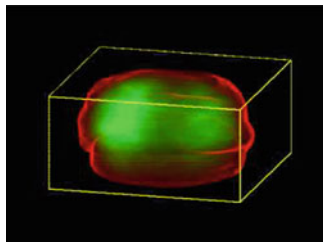


Fig. 2 A statistical atlas of the spatial distribution of prostate cancer obtained from co-registered prostatectomy specimens. Brighter green indicates areas of higher cancer incidence. This atlas can potentially be used to guide biopsy procedures aiming to obtain tissue from regions more likely to host cancer

- 1) To compare two groups of individuals, for example a group of patients with a specific disease and a group of healthy controls. Conventional approaches have used standard linear statistics [40]. However, the more and more common use of multi-parametric or multi-modality images calls for more sophisticated statistical image analysis methods that determine often nonlinear relationships among different scans, and therefore lead to the identification of subtle imaging phenotypes. Section 3 discusses some of the approaches and challenges.
- 2) To compare an individual with a group of individuals. For example, one might want to evaluate whether a patient with cognitive impairment presents the spatial pattern of brain tissue atrophy that is common to Alzheimer’s disease patients. The individual’s scans are then to be compared with the statistical atlas of AD patients. High-dimensional pattern classification methods are finding their way into the imaging literature more and more frequently, providing a new class of imaging-based diagnostic tools. Some of the work in this direction is discussed in Sect. 4.

A fundamental process in constructing a statistical atlas is that of *registration*, i.e. the process of bringing scans from different session and different individuals into a common spatial coordinate system, where they can directly be compared with each other and integrated into a statistical atlas. This process is described in Sect. 2.

2 3D and 4D Deformable registration

Image analysis methods have been studied in the literature during the past 15 years [3, 6, 12, 22, 26, 38, 46, 47, 50, 60, 65, 67, 69, 76, 77, 81]. One very promising approach for morphometric analysis is based on *shape transformations* that maps one template of anatomy (e.g. a typical brain, spinal, cardiac, or prostate image) to an image of interest. The resulting transformation measures the detailed differences between the two anatomies under consideration. So far, many methods have been proposed in the literature for obtaining the shape transformations, typically based on a method called deformable image registration.

The goal of deformable image registration is to find a transformation that maps every point in one image to its matching point in another image. Matching points should correspond to the same anatomical feature or structure. Also, the resulting shape transformation is usually required to be a diffeomorphism, i.e., a differentiable and invertible mapping between the domains of the two images. So far, various methods have been proposed for deformable image registration [1, 2, 6, 11, 18, 19, 20, 24, 31, 40, 42, 44, 49, 62, 68, 79, 81, 89, 90], which fall into *three categories*, i.e., *landmark-based* registration, *intensity-based* registration, and *feature-based* registration. Each of these methods has its own advantages and disadvantages, as described below.

The *landmark-based* methods utilize the prior knowledge of anatomical structure and thus are computationally fast since they do not evaluate a matching criterion on every single voxel in an image. Generally, a human rater is required to define a number of identifiable landmark points on a template and to manually locate their corresponding points on each of the images that are subject to the analysis. The beginning of the modern phase of this approach in medical imaging could perhaps be dated back to 1989 with Bookstein's work on landmark-based morphometrics [6]. Shape transformation is obtained by interpolating the mapping on the landmarks everywhere else in the image, using a thin plate spline model. This approach has been mainly used in 2D neuroimaging studies, often restricted to the corpus callosum, since not many anatomical structures lend themselves to 2D analysis. Moreover, defining landmarks in 3D with high accuracy and reproducibility is often a very difficult and impractical task, especially in large studies. Also, it is time-consuming to manually place a sufficient number of landmarks for accurate registration.

The *intensity-based* methods [11, 12, 17, 21, 36, 39, 60, 68, 79] aim to maximize the intensity similarity of two images, and can be fully automated since they do not require the construction of a specific anatomical model each time they are applied to a new problem. These methods assume that if a shape transformation renders two images similar, it implies anatomical correspondence between the underlying anatomies. This is a reasonable assumption, but it can easily be violated in practice, since two images can be made similar via shape transformations that do not respect the underlying anatomical correspondences.

Somewhat related to image intensity matching are the methods optimizing information theoretic criteria, in order to find appropriate shape transformations. The main advantage of these methods over intensity similarity methods is that they can be potentially used for multi-modality image registration, particularly when the tissue intensities are different in the different modality images. The most popular criterion used for multimodality image registration has been mutual information [68, 88, 91], which is maximized when the "predictability" of the warped image based on the template is maximized, and which tends to occur when the same tissue types in two images are well registered.

The *feature-based* registration methods formulate the image registration as a feature matching and optimization problem [16, 41, 85, 86]. A number of features, such as edges or curves or surfaces [15, 16, 29, 49, 65, 80, 81, 85], are typically

extracted from the images via an image analysis algorithm, or simply drawn manually, and are then used to drive a 3D deformable registration method, which effectively interpolates feature correspondence in the remainder of the image. Feature-based methods pay more attention to the biological relevance of the shape matching procedure, since they only use anatomically distinct features to determine the shape transformation, whereas image matching methods seek the transformations that maximize the similarity of images, with little warranty that the implied correspondences have anatomical meaning. However, the latter approaches take advantage of the full dataset, and not only of a relatively sparse subset of features.

Deformable registration using attribute vectors: A method that has been previously developed by our group attempts to integrate the advantages of various methods and at the same time to overcome some of their limitations, by developing an attribute vector as a morphological signature of each point, to allow the selection of the distinctive points for hierarchically guiding the image registration procedure [71, 73, 74, 95]. This is the method called Hierarchical Attribute Matching Mechanism for Elastic Registration (HAMMER). HAMMER is a hierarchical warping mechanism that has three key characteristics.

First, it places emphasis on determining anatomical correspondences, which in turn drive the 3D warping procedure. In particular, feature extraction methods have been used for determining a number of parameters from the images, to characterize at least some key anatomical features as distinctively as possible. In [73], geometric moment invariants (GMIs) were particularly used as a means for achieving this goal. GMIs are quantities that are constructed from images that are first segmented into GM, WM and CSF, or any other set of tissues of interest. They are determined from the image content around each voxel, and they quantify the anatomy in the vicinity of that voxel. GMIs of different tissues and different orders are collected into a long attribute vector for representing each voxel in an image. Ideally, attribute vectors are made as distinctive as possible for each voxel, so that anatomical matching across individual brains can be automatically determined during the image registration procedure. Fig. 3 shows a color-coded image of the degree of similarity between the GMI-based attribute vector of a point on the anterior horn of the left ventricle and the attribute vectors of every other point in the image. The GMI attribute vector of this point, as well as of many other points in the brain, is reasonably distinctive, as shown in Fig. 3. We have also explored more distinctive attribute vectors, aiming at constructing even more reliable and distinctive morphological signatures for every voxel in the image. Toward this end, wavelet coefficients [95], multiple-scale histogram features [72], local descriptor features [92], or combinations of various local features [92] were computed for hierarchical characterization of images of multi-scale neighborhoods centered on each voxel [94, 95].

Second, HAMMER addresses a fundamental problem encountered in high-dimensional image matching. In particular, the cost function being optimized typically has many local minima, which trap an iterative optimization procedure into solutions that correspond to poor matches between the template and the individual. This is partly due to the ambiguity in finding the point correspondences. For

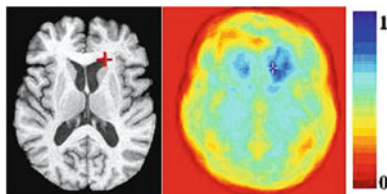


Fig. 3 The point marked by a cross has a relatively distinctive GMI-based attribute vector. The color-coded image on the right shows the degree of similarity between the attribute vector of the marked point and the attribute vector of every other point in the brain. 1 is maximum similarity and 0 is minimum similarity

example, if many candidate points in an individual image have similar attribute vectors to that of a particular template voxel, then this introduces an ambiguity and thus local minima in the corresponding energy function. In contrast, with definition of a good set of attributes, there will be a few anchor points for which correspondences can be determined rather unambiguously, perhaps because each anchor point's attribute vector is very different from all but its corresponding anchor point. In this way, correspondences on all other (non-anchor) points could be approximately determined via some sort of interpolation from the anchor points. This problem would not have local minima, since the cost function being minimized would be a lower-dimensional approximation involving only a few anchor points, compared to a cost function involving every single voxel in the two images under registration. HAMMER is exactly based on this fact, i.e., forming successive lower-dimensional cost functions, based initially only on key anchor points and gradually on more and more points. More points are considered as a better estimate of the shape transformation is obtained, thus potentially avoiding local minima.

Third, HAMMER is inverse-consistent, in terms of the driving correspondences. This means that if the individual is deformed to the template, instead of the converse, the mapping between any two driving points during this procedure would be identical. This feature is a computationally fast approximation to the problem of finding fully 3D inverse consistent shape transformations originally proposed by Christensen [48].

We have validated the HAMMER approach on brain images, and have found that it can achieve high accuracy, even in the presence of significant morphological differences between the two images being matched, such as differences in the cortical folding pattern of the brain. Fig. 4 show a representative example of the performance of HAMMER on MR images of elderly subjects from the Baltimore Longitudinal Study of Aging (BLSA) [66].

Longitudinal image registration: With a growing interest in longitudinal studies, which are important in studying development, normal, aging, early markers of Alzheimer's disease, and response to various treatments, amongst others, securing longitudinal stability of the measurements is of paramount importance. However, in the longitudinal morphometric study, people would typically measure the shape

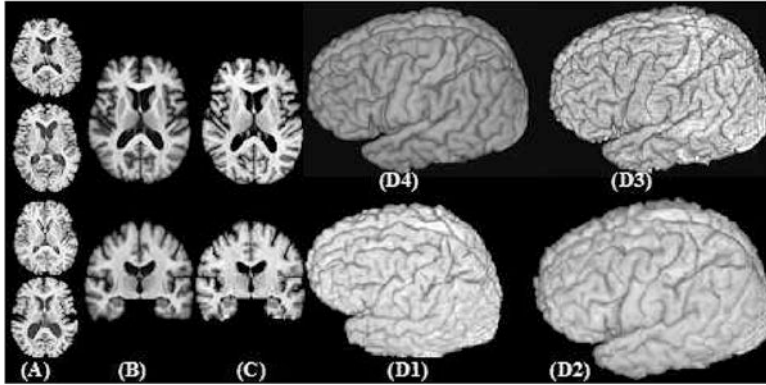


Fig. 4 Results using the HAMMER warping algorithm. (A) 4 representative sections from MR images of the BLSA database (B) Representative sections from the image formed by averaging 150 images warped by HAMMER to match the template shown in (C). (D1-D4) 3D renderings of a representative case, its warped configuration using HAMMER, the template, and the average of 150 warped images, respectively. The anatomical detail seen in (B) and (D4) is indicative of the registration accuracy

transformation during each time point, and then examine longitudinal changes in the shape transformations. This approach is valid in theory, but limited in practice. This is because the independent atlas warping for each longitudinal scan typically leads to *jittery longitudinal measurements*, particularly for small structures such as the hippocampus of the brain, due to inconsistent atlas warping across different scans in a series [75]. Although smoothed estimates of longitudinal changes can be obtained by smoothing or regressing over the measurements along the temporal dimension, *post hoc* smoothed measurements will, in general, deviate significantly from the actual image data, unless smoothing is performed concurrently with atlas warping and by taking into account the image features. It is worth noting that the issue of longitudinal measurement robustness is particularly important in measuring the progression of a normal older adult into mild cognitive impairment, which makes it possible to have the ability to detect subtle morphological changes well before severe cognitive decline appears.

In order to achieve longitudinally stable measurements, we have developed 4-dimensional image analysis techniques, such as a 4D extension of HAMMER [75]. In this approach [75], all serial scans are jointly considered as a single 4D scan, and the optimal 4D deformation is determined, thus avoid inconsistent atlas warping across different serial scans. The 4D warping approach of [75] simultaneously establishes longitudinal correspondences in the individual as well as correspondences between the template and the individual. This is different from the 3D warping methods, which aim at establishing only the inter-subject correspondences between the template and the individual in a single time-point. Specifically, 4D-HAMMER uses a fully automatic 4-dimensional atlas matching method that constrains the smoothness in both spatial and temporal domains during the

hierarchical atlas matching procedure, thereby producing smooth and accurate estimations of structural changes over time. Most importantly, morphological features and matches guiding this deformation process are determined via 4D image analysis, which significantly reduces noise and improves robustness in detecting anatomical correspondence. Put simply, image features that are consistently recognized in all time-points guide the warping procedure, whereas spurious features (such as noisy edges) that appear inconsistently at different time-points are eliminated. We have validated this approach against manually-defined brain ROI volumes by very well trained experts on serial scans [75], and we determined that it produces not only smoother and more stable measurements of longitudinal atrophy, but also significantly more accurate measurements, as demonstrated in [75].

Evaluation: The plethora of automated methods for deformable image registration has necessitated the evaluation of their relative merits. To this end, evaluation criteria and metrics using large image populations have been proposed by using richly annotated image databases, computer simulated data, and increasing the number and types of evaluation criteria [13]. However, the traditional deformable simulation methods, such as the use of analytic deformation fields or the displacement of landmarks followed by some form of interpolation [6], are often unable to construct rich (complex) and/or realistic deformations of anatomical organs. To deal with this limitation, several methods have been developed to automatically simulate realistic *inter-individual*, *intra-individual*, and *longitudinal* deformations, for validation of atlas-based segmentation, registration, and longitudinal measurement algorithms [10,96].

The *inter-individual* deformations can be simulated by a statistical approach, from the high-deformation fields of a number of examples (training samples). In [96], Wavelet-Packet Transform (WPT) of the training deformations and their Jacobians, in conjunction with a Markov Random Field (MRF) spatial regularization, were used to capture both coarse and fine characteristics of the training deformations in a statistical fashion. Simulated deformations can then be constructed by randomly sampling the resultant statistical distribution in an unconstrained or a landmark-constrained fashion. In particular, the training sample deformations could be generated by first extensively labeling and landmarking a number of images [5], and then applying a high-dimensional warping algorithm constrained by these manual labels and landmarks. Such adequately constrained warping algorithms are likely to generate deformations that are close to a gold standard, and therefore appropriate for training.

The *intra-individual* brain deformations can be generated to reflect the structural changes of an individual brain at different time points, i.e., tissue atrophy/growth of a selected structure or within a selected region of a brain. In particular, the method proposed in [51] can be used to simulate the atrophy and growth, i.e., generating a deformation field by minimizing the difference between its Jacobian determinants and the desired ones, subject to some smoothness constraints on the deformation field. The desired Jacobian determinants describe the desired volumetric changes of different tissues or different regions. Moreover, by using the labeled mesh and the

FEM solver [10], the realistic *longitudinal* atrophy in brain structures can be also generated, to mimic the patterns of change obtained from a cohort of 19 real controls and 27 probable Alzheimer's disease patients.

3 Multi-parametric statistics and group analysis

Section 2 describes the process of spatial normalization, which brings images in the same coordinate frame. These can now be incorporated into a statistical atlas for group analysis as will be discussed next.

One of the most common applications of statistical atlases is to be able to compare two groups. These groups could be groups of patients with a specific disease and a group of healthy controls, or groups of subjects divided on the basis of age, gender or some other physical characteristic. Prior to performing any group-analysis, the images of the subjects to be used are spatially normalized to a subject chosen to be the template. The method adopted for spatial normalization and group-based analysis depends completely on the type of data (scalar or high-dimensional) used for representing the subjects.

3.1 Statistical analysis of scalar maps

Traditionally, in the standard framework of voxel- and deformation-based analysis, scalar images, such as tissue density maps [2, 25] or Jacobian determinants [17, 29] are used for group analysis. In diffusion tensor images (DTI), one of the several scalar maps of fractional anisotropy or diffusivity [4] may be used to represent each of the subjects. These scalar maps of choice for representing the subjects are then smoothed by a Gaussian filter to: 1) Gaussianize the data, so that linear statistics are more applicable; 2) smooth out the noise; or 3) analyze the image context around each voxel, since structural changes are unlikely to be localized at a single voxel, but are rather more likely to encompass entire regions. Voxel-wise statistical tests, such as the general linear model, are then applied to the spatially normalized and smoothed data. The output is a voxel-wise map which identifies regions of significant difference based on some user-defined threshold. Statistical Parametric Mapping or SPM [40] has been a tool widely used for this purpose. Fig. 5 shows an example of voxel-based statistical analysis performed on a group of patients with schizophrenia versus a group of healthy controls, using MRI data. Similar analysis has been performed on anisotropy and diffusivity maps computed from DTI data [54, 55]. These regions can be further analyzed with respect to the clinical correlates, to determine the biological underpinnings of the group-changes.

In addition to the application of voxel-wise statistical tests, analysis can also be performed on a region wise basis, using regions that have been defined on a template or average atlas of subjects that has been created using the process of

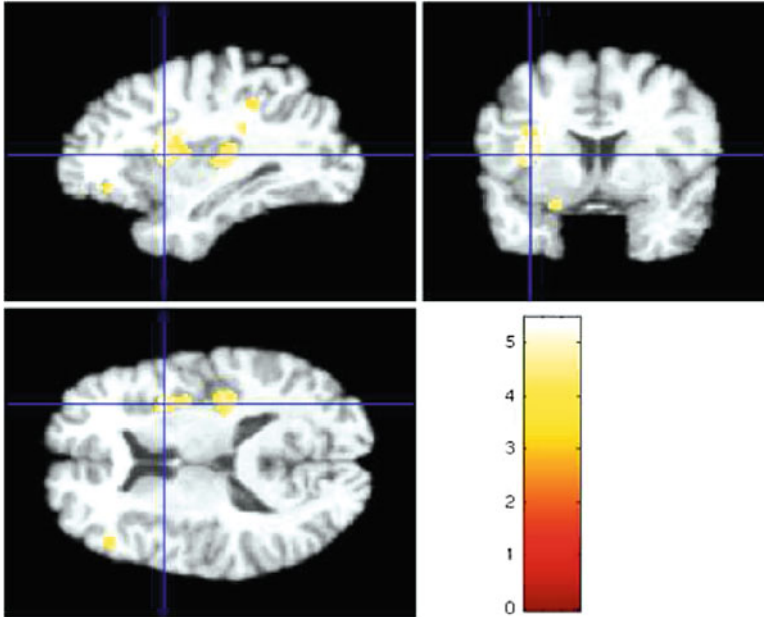


Fig. 5 Regions of difference between schizophrenia patients and healthy controls using voxel-wise application of linear statistics

spatial normalization. By the process of spatial normalization, each of these regions can be warped to each of the subjects (or vice versa) and group-based analysis may be performed using region-based values of scalar quantities that characterize the effect of pathology on these regions. For example, tissue atrophy during the course of aging or as a manifestation of disease could be evaluated percentage-wise in each of the regions that have been outlined in the atlas of ROIs.

3.2 *Statistical analysis of multi-parametric and multi-modal data*

While the general linear model (GLM) has been used effectively for the statistical analysis of the scalar map representation of the subjects, with the increasing use of multi-parametric or multi-modality images, GLM with Gaussian smoothing does not suffice for several reasons. The simple spatial filtering and subsequent linear statistics are not a valid approach for the multi-parametric data such as tensors and multi-modality data as the high-dimensional, non-linear data at each voxel are typically distributed along sub-manifolds of the embedding space. In tensors, this embedding space could be in \mathbb{R}^6 , if single voxel data is to be considered. In multi-modality data, the dimension of the embedding space will be the number of

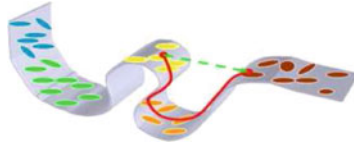


Fig. 6 Manifold structure of tensors. The gray surface represents the non-linear manifold fitted through the tensors or any high dimensional structure represented as ellipses. The green line represents the Euclidean distance between tensors treated as elements of R^6 and the red line represents the geodesic distance along the manifold that will be used for all tensor manipulations

modalities combined to represent the data. Fig. 6 demonstrates this concept, for the case of tensors, but instead of ellipses, any high-dimensional structure could be used to emulate the voxel-wise structure of multi-modality data. The filtering and the subsequent statistics need to be performed on this underlying manifold. In addition, to the underlying structure of the data, another reason for the lack of feasibility of the GLM for this data is that the shape and size of a region of interest, such as a region of growth or abnormal morphological characteristics, is not known in advance; the way in which a disease process is likely to affect the local tissue structure of the brain is highly unlikely to follow a Gaussian spatial profile of a certain pre-defined size. Thus the two main challenges that we need to address in order to form statistical atlases of this higher dimensional data and follow it with group analysis are: 1) determining the true underlying structure of the data in the form of a non-linear manifold and 2) estimating the statistical distribution of the data on that manifold.

In order to address these two issues, more sophisticated image analysis methods need to be developed that determine the often non-linear relationships among different scans and therefore lead to the identification of subtle imaging phenotypes. In relation to tensors, methods based upon Riemannian symmetric spaces [37, 58] rely upon the assumption that the tensors around a given voxel from various subjects belong to a principal geodesic (sub)-manifold and that these tensors obey a normal distribution on that sub-manifold. The basic principle of these methods is sound, namely that statistical analysis of high dimensional data must be restricted to an appropriate manifold. However, there is no guarantee that the representations of the tensors on this sub-manifold will have normal distributions, and most importantly, restricting the analysis on the manifold of positive definite symmetric tensors is of little help in hypothesis testing studies, since the tensors measured at a given voxel or neighborhood, from a particular set of brains, typically lie on a much more restricted sub-manifold of the space of symmetric positive definite matrices. For example, if all voxels in a neighborhood around the voxel under consideration belong to a particular fiber tract, then the tract geometry will itself impose an additional nonlinear structure on the sub-space of the tensors at those voxels from all subjects. Some of these issues were alleviated by the development of a manifold-based statistical analysis framework which focuses on approximating/learning the local structure of the manifold along which tensor/higher-dimensional measurements

from various individuals are distributed. The learned features belonged to a low-dimensional linear manifold parameterizing the higher-dimensional tensor manifold and were subsequently used for group-wise statistical analysis. The filtering and subsequent statistical analysis are then performed along the manifold, rather than in the embedding space.

The two frameworks that are discussed below are from the perspective of whether: 1) the manifold is explicitly determined (Sect. 3.3) or 2) the data distribution is determined by implicitly incorporating the underlying structure of the data (Sect. 3.4). While explained in the context of tensors they are applicable to high dimensional multi-parametric data.

3.3 *Statistical analysis on the estimated manifolds*

Suppose the problem is to investigate group-wise (e.g. schizophrenia patients and controls) morphological differences on a voxel-wise basis using DTI data available for these subjects. Having spatially normalized this data [93, 97], we will analyze group differences voxel-by-voxel. At a particular voxel, we form a dataset by collecting tensors from that location from all subjects. For the purposes of smoothing the data locally, tensor measurements from all voxels in a surrounding neighborhood are collected, as well as from all subjects of our study. This collection of tensors generally does not follow a Gaussian distribution in \mathbb{R}^6 , but rather lies on a sub-manifold of \mathbb{R}^6 . In such a case, manifold learning methods [9] can be used to learn the structure of the underlying manifold. Then the manifold can be linearized via “flattening” to a lower dimensional space. We use a manifold learning technique called Isomap [78] that builds a graph from these high dimensional samples, and subsequently used graph searching methods to calculate geodesics along the sub-manifold, finally applies multi dimensional scaling [9] on these geodesics, in order to flatten the sub-manifold. Standard multi-variate statistical tests such as Hotelling’s T^2 test, can then be applied on the flattened sub-manifold, determined at each voxel, to obtain a voxel-wise p -value map. Mathematical details of the method can be found in [52, 87].

Forming group averages An important aspect of statistical atlases is the ability to obtain average maps that are representative of a particular group property. This has been achieved by simple linear averaging when scalar maps are used for statistical analysis as all the maps have already been spatially normalized to a template. In the case of high-dimensional data a manifold is fitted to the data at each voxel (accumulated within a spatial neighborhood and across subjects) using manifold learning and the average is computed on the manifold [87].

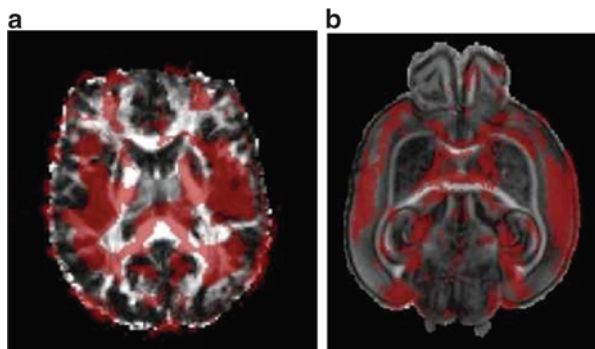
3.4 *Determining the statistical distribution on high-dimensional data*

The manifold-learning approach described above is quite effective if a sufficient number of samples are available, so that the sub-manifold is properly sampled. However in the absence of sufficient samples that do not cover the entire space, results of learning the sub-manifold can be quite unpredictable. While this can be somewhat alleviated by using neighborhood tensors from around the voxel under consideration, thereby indirectly imposing spatial smoothness, it increases the dimensionality of measurement at each voxel, thereby changing the embedding space. Also, the dataset has unknown statistical distribution on the non-linear manifold and Isomap or any other manifold learning technique flattens a sub-manifold, but does not Gaussianize or otherwise estimate the statistical distribution of tensors on the flattened manifold. Therefore, standard statistical tests may still be inappropriate for use on the flattened sub-manifold.

Hence for statistical group analysis, it is important to nonlinearly approximate the probability density from a number of samples, and to obtain a representation that will enhance group separation. This can be achieved using kernel-based methods, such as a combination of a kernel-based principal component analysis (kPCA) [70] and kernel-based Fischer discriminant analysis (kFDA) [43]. The common idea behind kernel-based techniques is to transform the samples into a higher-dimensional space, which can be used for statistical analysis and for the purposes of density estimation. These methods effectively estimate the nonlinear distribution of tensors or other higher dimensional data on the underlying manifold (without having to determine the structure of the manifold explicitly) (kPCA) and enhance their statistical separation, by finding nonlinear projections of the data which can optimally discriminate between the two groups (kFDA). Multi-variate statistical tests can be applied to these kernel-based features, in order to obtain a voxel-wise p -value map. Details can be found in [53].

As an essential final step, regions with significant differences between the two groups are identified from these parametric or non-parametric p -value maps obtained either by manifold learning or kernel-based methods by accounting for multiple comparisons [43, 64]. This can be achieved using a more complex form of the permutation test by controlling the family-wise error rate due to multiple comparisons [64] or by controlling the false discovery rate (FDR) using a suitable p -value threshold [43]. Fig. 7(a) identifies regions of difference between schizophrenia patients and controls using kernel-based methods. Fig. 7(b) shows the kernel based framework applied to mouse images (the two groups being the young and the old) at a much higher level of significance. These regions survive multi-comparison testing using FDR. Similar maps can be computed using manifold learning.

Fig. 7 Regions of significant differences between (a) schizophrenia patients and controls and (b) young and old mice obtained using statistical analysis of multi-parametric data



In the above discussion involving manifold learning and kernel based methods, although the data at each voxel were tensors, the frameworks can easily lend themselves to higher dimensional data as it derived from multi-modality data. The difference would be the manifold learned and the nature of the embedding space.

4 Individual patient analysis and high-dimensional pattern classification

Section 3 discussed statistical image analysis methods for determining group differences. Although informative from a biologic point of view, group analyses are not meant to provide diagnostic tools for individuals. This is because two groups can have highly overlapping values of a structural or functional variable, e.g. the volume of a structure, but having a sufficiently large sample size, a group analysis will identify significant group differences, even of the smallest magnitude, if they are present. For example, the joint histogram of the volume of the hippocampus and the volume of the entorhinal cortex (ERC) in a group of healthy elderly individuals and a group with mild cognitive impairment (MCI) might show that the former tend to have larger volumes than the latter in a statistical sense.

However, if we are given a new individual's volumes of the hippocampus and the ERC, we will not be able to correctly classify this individual, due to the overlap between the two joint histograms. In order to address this issue, high-dimensional pattern classification methods have been pursued with increasing frequency in the recent literature [32, 33, 34, 35, 45, 46, 56, 59, 83]. A pattern is formed by a collection of image-derived measurements, typically obtained from a number of different anatomical regions. A statistical atlas is of fundamental importance in these methods, since it represents the range of variation of the features that are used to construct the pattern. A simple example is illustrated in Fig. 8, which displays the z-score map of an Alzheimer's patient's regional distribution of gray matter tissue in the brain. Pattern recognition methods are trained to recognize such spatio-temporal patterns of structure and function.

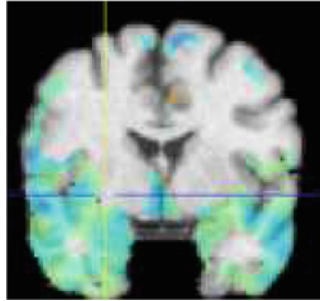


Fig. 8 An individual is compared against a statistical atlas of normal elderly individuals, in order to determine whether the pattern of brain atrophy is typical of a normal elderly or not. The figure displays a z-score map, i.e. a voxel-by-voxel evaluation of the individual's gray matter volume against the statistical atlas. A pattern of fronto-temporal atrophy (blue regions) indicates AD-like pathology

One of the motivating factors behind these developments is the complex and spatio-temporally distributed nature of the changes that many diseases cause, particularly in the brain and the heart. For example, in Alzheimer's Disease, the anatomical structures that carry most discriminative power are likely to depend on the stage of the disease, as the disease progressively spreads throughout various brain regions [7], but also on age and other demographic and genetic factors [61], since disease is to be distinguished from complex and progressively changing background normal variations in anatomy and function that may depend on demographic and/or genetic background. Moreover, Alzheimer's disease might cause changes of the image characteristics beyond those measured by volumetrics, such as for example brightening or darkening of an MR image due to demyelination, deposition of minerals, or other macro- or micro-structural changes caused by disease. Vascular disease also causes well-known MR signal changes, for example in the white matter of the brain (e.g. brightening of T₂-weighted signal). It is thus becoming clear that multiple modalities and multiple anatomical regions must be considered jointly in a (possibly nonlinear) multi-variate classification fashion, in order to achieve the desirable diagnostic power. Moreover, regions that are relatively less affected by disease should also be considered along with known to be affected regions (which, for the example of Alzheimer's Disease might include primarily temporal lobe structures, in relatively early disease stages), since differential atrophy or image intensity changes between these regions are likely to further amplify diagnostic accuracy and discrimination from a background of normal variation. Certain cardiac diseases also have subtle and spatio-temporally complex patterns of structural and functional change. For example, arrhythmogenic right ventricular disease involves spatial patterns of structural and physiological change that are not always easy to distinguish from normal inter-individual variability.

A fundamental challenge faced by high-dimensional pattern classification methods in medical imaging is the curse of dimensionality, i.e. the fact that imaging measurements have vastly larger dimensionality than the number of sample avail-

able in the typical study. Extraction, selection, and reduction of all spatio-temporal information included in a medical scan to a small number of features that optimally distinguishes between two or more groups is an open problem. Some approaches have employed global dimensionality reduction methods, such as principal or independent component analysis [32,63], before feeding the reduced features into a pattern classifier.

A more localized approach has been developed in our laboratory, and is termed COMPARE (Classification of Morphological Patterns using Adaptive Regional Elements). This approach was described in [57] and examines spatio-temporal patterns of regional brain atrophy, by hierarchically decomposing an image into images of different scales, each of which capturing structural and/or functional characteristics of interest at a different degree of spatial resolution. The most important parameters are then selected and used in conjunction with a nonlinear pattern classification technique to form a hyper-surface, the high-dimensional analog to a surface, which is constructed in a way that it optimally separates two groups of interest, for example normal controls and patients of a particular disease. Effectively, that approach defines a nonlinear combination of a large number of image-derived measurements from the entire anatomy of interest, each taken at a different scale that typically depends on the size of the respective anatomical structure and the size of the region that is most affected by the disease. This nonlinear combination of volumetric measurements is the best way, according to the respective optimality criteria, to distinguish between two groups, and therefore to perform diagnosis via classification of a new scan into patients or normal controls. In [27] excellent separation was obtained by high-dimensional nonlinear classification applied to a population of healthy controls and MCI patients that were not separable using the commonly used volumetric measurements of the hippocampus and the entorhinal cortex, further indicating that appropriate increase of dimensionality helps separate otherwise unseparable data.

In summary, the availability of large numbers of medical image datasets has necessitated the development and validation of image analysis tools that capture the range of variation of image-derived structural and functional characteristics, over populations of patients and healthy subjects. A new generation of techniques for deformable registration, statistical analysis, and pattern classification has appeared in the literature over the past decade, aiming to help identify anatomical and functional differences across different groups, but also to classify individual scans against baseline statistical atlases of normal and diseased populations. These new tools are gradually being adopted in clinical studies.

References

1. D. Alexander and J. Gee. Elastic matching of diffusion tensor images. *Computer Vision and Image Understanding*, 77:233–250, 1999.
2. J. Ashburner and K. Friston. Voxel-based morphometry: the methods. *Neuroimage*, 11(6): 805–821, 2000.

3. J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston. Identifying global anatomical differences: Deformation-based morphometry. *Human Brain Mapping*, 6(5-6):348–357, 1998.
4. P. J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance, Series B*, 111:209–219, 1996.
5. S. F. K. Boesen, J. Huang, J. Germann, J. Stern, D. L. Collins, A. C. Evans, and D. A. Rottenberg. Inter-rater reproducibility of 3d cortical and subcortical landmark points. In *11th Annual Meeting of the Organization for Human Brain Mapping*, Toronto, Canada, 2005.
6. F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
7. H. Braak, E. Braak, J. Bohl, and H. Bratzke. Evolution of alzheimer’s disease related cortical lesions. *Journal of Neural Transmission. Supplementum*, 54:97–106, 1998.
8. R. Bryan, C. Davatzikos, M. Vaillant, J. Prince, S. Letovsky, R. Raghavan, W. Nowinski, G. Salamon, N. Murayama, O. Levrier, and M. Zilbovicius. Creation of population-based anatomic atlases with a brain image database. In *First International Conference on Functional Brain Mapping*, page 72, 1995.
9. C. Burges. Geometric methods for feature extraction and dimensional reduction. In L. Rokach and O. Maimon, editors, *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*. Kluwer Academic Publishers, 2005.
10. O. Camara, R. I. Scahill, J. A. Schnabel, W. R. Crum, G. R. Ridgway, D. L. G. Hill, and N. C. Fox. Accuracy assessment of global and local atrophy measurement techniques with realistic simulated longitudinal data. In *MICCAI*, pages 785–792, 2007.
11. G. Christensen and H. Johnson. Consistent image registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582, 2001.
12. G. Christensen, R. Rabbitt, and R. Miller. 3d brain mapping using a deformable neuroanatomy. *Physics in medicine and biology*, 39:609–618, 1994.
13. G. E. Christensen, X. Geng, J. G. Kuhl, J. Bruss, T. J. Grabowski, I. A. Pirwani, M. W. Vannier, J. S. Allen, and H. Damasio. Introduction to the non-rigid image registration evaluation project (nirep). In *WBIR*, pages 128–135, 2006.
14. G. E. Christensen, H. J. Johnson, and M. W. Vannier. Synthesizing average 3d anatomical shapes. *NeuroImage*, 32(1):146–158, 2006.
15. H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.
16. H. Chui, L. Win, R. Schultz, J. Duncan, and A. Rangarajan. A unified feature registration method for brain mapping. In *Information Processing in Medical Imaging*, pages 300–314, Davis, CA, USA, 2001.
17. M. Chung, K. Worsley, T. Paus, C. Cherif, D. Collins, J. Giedd, J. Rapoport, and A. Evans. A unified statistical approach to deformation-based morphometry. *Neuroimage*, 14(3): 595–606, 2001.
18. A. Collignon. *Multi-modality medical image registration by maximization of mutual information*. Ph.d. dissertation, Catholic Univ. Leuven, 1998.
19. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In Y. Bizais, C. Barillot, and R. D. Paola, editors, *Information Processing in Medical Imaging*, pages 263–274. Kluwer, Dordrecht, The Netherlands, 1995.
20. D. Collins and A. Evans. Automatic 3d estimation of gross morphometric variability in the human brain. *Neuroimage*, 3(3):S129, 1996.
21. D. Collins, P. Neelin, T. Peters, and A. Evans. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of Computer Assisted Tomography*, 18:192–205, 1994.
22. L. Collins, T. Peters, W. Dai, and A. Evans. Model-based segmentation of individual brain structures from mri data. In *Proceedings of SPIE Conference on Visualization in Biomedical Computing*, volume 1808, pages 10–23, 1992.

23. J. Csernansky, L. Wang, S. Joshi, J. Ratnanather, and M. Miller. Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. *NeuroImage*, 23(1): 56–68, 2004.
24. C. Davatzikos. Spatial transformation and registration of brain images using elastically deformable models. *Computer Vision and Image Understanding*, 66:207–222, 1997.
25. C. Davatzikos. Voxel based morphometric analysis using shape transformations. In M. F. Glabus, editor, *International Review of Neurobiology, Neuroimaging Part A*, volume 66. Elsevier Inc., 2005.
26. C. Davatzikos and R. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. *IEEE Transactions on Medical Imaging*, 15(6):785–795, 1996.
27. C. Davatzikos, Y. Fan, X. Wu, D. Shen, and S. M. Resnick. Detection of prodromal alzheimer’s disease via pattern classification of mri. *Neurobiology of Aging*, page in press, 2007.
28. C. Davatzikos, A. Genc, D. Xu, and S. Resnick. Voxel-based morphometry using the ravens maps: Methods and validation using simulated longitudinal atrophy. *NeuroImage*, 14(6): 1361–1369, 2001.
29. C. Davatzikos, M. Vaillant, S. Resnick, J. Prince, S. Letovsky, and R. Bryan. A computerized approach for morphological analysis of the corpus callosum. *Journal of Computer Assisted Tomography*, 20:88–97, 1996.
30. C. Davatzikos and R. Verma. Constructing statistical brain atlases from diffusion tensor fields. In *ISMRM Workshop on methods for quantitative diffusion of human brain*, Lake Louise, Alberta, Canada, 2005.
31. B. Dawant, S. Hartmann, and S. Gadamsetty. Brain atlas deformation in the presence of large space-occupying tumours. In *MICCAI*, volume 1679, pages 589–596, 1999.
32. S. Duchesne, A. Caroli, C. Geroldi, G. Frisoni, and D. Collins. Predicting clinical variable from mri features: application to mmse in mci. In *MICCAI*, volume 8, pages 392–399, 2005.
33. Y. Fan, D. Shen, and C. Davatzikos. Classification of structural images via high-dimensional image warping, robust feature extraction, and svm. In J. S. Duncan and G. Gerig, editors, *MICCAI*, volume 3749 / 2005 of *Lecture Notes in Computer Science*, pages 1–8, Palm Springs, California, USA, 2005. Springer Berlin / Heidelberg.
34. Y. Fan, D. Shen, and C. Davatzikos. Decoding cognitive states from fmri images of subjects by machine learning and multivariate classification. In *IEEE Workshop on Mathematical Methods in Biomedical Image (MMBIA 2006)*, New York City, NY, USA, 2006.
35. Y. Fan, D. Shen, R. C. Gur, R. E. Gur, and C. Davatzikos. Compare: Classification of morphological patterns using adaptive regional elements. *IEEE Transactions on Medical Imaging*, 26(1):93–105, 2007.
36. M. Ferrant, S. Warfield, C. Guttman, R. Mulkern, F. Jolesz, and R. Kikinis. 3d image matching using a finite element based elastic deformation model. In *MICCAI*, pages 202–209, 1999.
37. P. T. Fletcher and S. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision Approaches to Medical Image Analysis*, volume 3117 of *LNCS: CVAMIA*, pages 87–98, 2004.
38. P. Freeborough and N. Fox. Modeling brain deformations in alzheimer’s disease by fluid registration of serial 3d mr images. *Journal of Computer Assisted Tomography*, 22:838–843, 1998.
39. K. Friston, J. Ashburner, C. Frith, J. Poline, J. Heather, and R. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.
40. K. Friston, A. Holmes, K. Worsley, J. Poline, C. Frith, and R. Frackowiak. Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, 2(4):189–210, 1995.
41. J. Gee, C. Barillot, L. Briquer, D. Haynor, and R. Bajcsy. Matching structural images of the human brain using statistical and geometrical image features. *Proc. SPIE Visualization in Biomedical Computing*, 2359:191–204, 1994.
42. J. Gee, M. Reivich, and R. Bajcsy. Elastically deforming 3d atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17:225–236, 1993.

43. C. R. Genovese, N. A. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4):870–878, 2002.
44. G. Gerig, M. Styner, and J. Lieberman. Shape versus size: Improved understanding of the morphology of brain structures. In *MICCAI*, Utrecht, the Netherlands, 2001. Kluwer.
45. P. Golland, B. Fischl, M. Spiridon, N. Kanwisher, R. Buckner, M. Shenton, R. Kikinis, A. Dale, and W. Grimson. Discriminative analysis for image-based studies. In R. K. T. Dohi, editor, *MICCAI*, volume LNCS 2488, pages 508–515, Tokyo, Japan, 2002. Springer-Verlag GmbH.
46. P. Golland, W. Grimson, and R. Kikinis. Statistical shape analysis using fixed topology skeletons: corpus callosum study. *Lecture Notes in Computer Science*, 1613:382–387, 1999.
47. P. Golland, W. E. L. Grimson, M. Shenton, and R. Kikinis. Deformation analysis for shape based classification. *Lecture Notes in Computer Science*, 2082:517–530, 2001.
48. H. Johnson and G. Christensen. Landmark and intensity-based consistent thin-plate spline image registration. In *Proceedings of the Conference on Information Processing in Medical Imaging*, LNCS, volume 2081, pages 329–343, 2001.
49. S. Joshi, M. Miller, G. Christensen, A. Banerjee, T. Coogan, and U. Grenander. Hierarchical brain mapping via a generalized dirichlet solution for mapping brain manifolds. *Proceedings of the SPIE Conference on Geom. Methods in Applied Imaging*, 2573:278–289, 1995.
50. S. Joshi, S. Pizer, P. Fletcher, A. Thall, and G. Tracton. Multi-scale 3-d deformable model segmentation based on medial description. *Lecture Notes in Computer Science*, 2082:64–77, 2001.
51. B. Karacali and C. Davatzikos. Simulation of tissue atrophy using a topology preserving transformation model. *IEEE Transactions on Medical Imaging*, 25(5):649–652, 2006.
52. P. Khurd, R. Verma, and C. Davatzikos. On characterizing and analyzing diffusion tensor images by learning their underlying manifold structure. In *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis*, Ney York, NY, 2006.
53. P. Khurd, R. Verma, and C. Davatzikos. Kernel-based manifold learning for statistical analysis of diffusion tensor images. In *Information Processing in Medical Imaging (IPMI)*, volume 4584, pages 581–593, Netherlands, 2007.
54. M. Kubicki, R. W. McCarley, C. F. Westin, H. J. Park, S. Maier, R. Kikinis, F. A. Jolesz, and M. E. Shenton. A review of diffusion tensor imaging studies in schizophrenia. *Journal of Psychiatric Research*, 41(1-2):15–30, 2007.
55. M. Kubicki, M. E. Shenton, D. F. Salisbury, Y. Hirayasu, K. Kasai, R. Kikinis, F. A. Jolesz, and R. W. McCarley. Voxel-based morphometric analysis of gray matter in first episode schizophrenia. *Neuroimage*, 17(4):1711–1719, 2002.
56. Z. Lao, D. Shen, and C. Davatzikos. Statistical shape model for automatic skull-stripping of brain images. In *2002 IEEE International Symposium on Biomedical Imaging: Macro to Nano*, pages 855–858, Washington, D.C., 2002.
57. Z. Lao, D. Shen, Z. Xue, B. Karacali, S. Resnick, and C. Davatzikos. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*, 21(1):46–57, 2004.
58. C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor mri processing. *Journal of Mathematical Imaging and Vision [Special issue Mathematics and Image Analysis]*, 2006 (to appear).
59. Y. Liu, L. Teverovskiy, O. Carmichael, R. Kikinis, M. Shenton, C. Carter, V. Stenger, S. Davis, H. Aizenstein, J. Becker, O. Lopez, and C. Meltzer. Discriminative mr image feature analysis for automatic schizophrenia and alzheimer’s disease classification. In C. Barillot, D. R. Haynor, and P. Hellier, editors, *MICCAI*, volume LNCS 3216, pages 393–401, Saint-Malo, France, 2004. Springer-Verlag GmbH.
60. M. Miller, G. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable neuroanatomies. *Proceedings of the National Academy of Sciences*, 90:11944–11948, 1993.
61. S. Moffat, C. Szekely, A. Zonderman, N. Kabani, and S. Resnick. Longitudinal change in hippocampal volume as a function of apolipoprotein e genotype. *neurology 2000; Neurology*, 55(1):134–136–, 2000.

62. A. Mohamed, D. Shen, and C. Davatzikos. Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. In J. S. Duncan and G. Gerig, editors, *MICCAI*, volume 3750 / 2005 of *Lecture Notes in Computer Science*, pages 263–270, Palm Springs, CA, 2005. Springer-Verlag GmbH.
63. J. Mourao-Miranda, A. L. Bokde, C. Born, H. Hampel, and M. Stetter. Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional mri data. *Neuroimage*, 28(4):980–995, 2005.
64. T. Nichols and A. Holmes. Non-parametric permutation tests for functional neuroimaging: A primer with examples. In *Human Brain Mapping*, volume 15, pages 1–25, 2001.
65. S. Pizer, D. S. Fritsch, P. A. Yushkevich, V. E. Johnson, and E. L. Chaney. Segmentation, registration and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18(10):851–865, 1999.
66. S. Resnick, A. Goldszal, C. Davatzikos, S. Golski, M. Kraut, E. Metter, R. Bryan, and A. Zonderman. One-year age changes in mri brain volumes in older adults. *Cerebral Cortex*, 10(5):464–472, 2000.
67. J. Rexilius, S. Warfield, C. Guttman, X. Wei, R. Benson, L. Wolfson, M. Shenton, H. Handels, and R. Kikinis. A novel nonrigid registration algorithm and applications. In *MICCAI*, pages 202–209, 1999.
68. D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Non-rigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
69. S. Sandor and R. Leahy. Surface based labelling of cortical anatomy using a deformable atlas. *IEEE Transactions on Medical Imaging*, 16(1):41–54, 1997.
70. B. Schölkopf and A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press, 1st edition (december 15, 2001) edition, 2001.
71. D. Shen. 4d image warping for measurement of longitudinal brain changes. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, volume 1, Arlington, Va., 2004.
72. D. Shen. Image registration by hierarchical matching of local spatial intensity histograms. In C. Barillot, D. R. Haynor, and P. Hellier, editors, *MICCAI*, volume 3216 / 2004 of *Lecture Notes in Computer Science*, pages 582–590, St. Malo, France, 2004. Springer-Verlag GmbH.
73. D. Shen and C. Davatzikos. Hammer: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, 21(11):1421–1439, 2002.
74. D. Shen and C. Davatzikos. Very high resolution morphometry using mass-preserving deformations and hammer elastic registration. *NeuroImage*, 18(1):28–41, 2003.
75. D. Shen and C. Davatzikos. Measuring temporal morphological changes robustly in brain mr images via 4-dimensional template warping. *NeuroImage*, 21(4):1508–1517, 2004.
76. M. Styner and G. Gerig. Medial models incorporating object variability for 3d shape analysis. *Lecture Notes in Computer Science*, 2082:502–516, 2001.
77. G. Szekely, A. Kelemen, C. Brechbuhler, and G. Gerig. Segmentation of 2-d and 3-d objects from mri volume data using constrained deformations of flexible fourier contour and surface models. *Medical Image Analysis*, 1:19–34, 1996.
78. J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
79. J. Thirion. Non-rigid matching using deamons. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1996.
80. J. Thirion, O. Monga, S. Benayoun, A. Gueziec, and N. Ayache. Automatic registration of 3-d images using surface curvature. *SPIE Proceedings, Mathematical Methods in Medical Imaging*, 1768.:206–216, 1992.
81. P. Thompson, D. MacDonald, M. Mega, C. Holmes, A. Evans, and A. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21(4):567–581, 1997.
82. P. M. Thompson, M. Mega, R. Woods, C. Zoumalan, C. Lindshield, R. Blanton, J. Moussai, C. Holmes, J. Cummings, and A. Toga. Cortical change in alzheimer’s disease detected with a disease-specific population-based brain atlas. *Cerebral Cortex*, 11(1):1–16, 2001.

83. S. Timoner, P. Golland, R. Kikinis, M. Shenton, W. Grimson, and W. M. Wells. Performance issues in shape classification. In *MICCAI*, pages 355–362, Tokyo, Japan, 2002. Springer-Verlag.
84. A. W. Toga, P. M. Thompson, M. S. Mega, K. L. Narr, and R. E. Blanton. Probabilistic approaches for atlas normal and disease-specific brain variability. *Anatomy and Embryology*, 204(4):267–282, 2001.
85. M. Vaillant and C. Davatzikos. Hierarchical matching of cortical features for deformable brain image registration. *Lecture Notes in Computer Science: Information Processing in Medical Imaging*, 1613:182–195, 1999.
86. R. Verma and C. Davatzikos. Matching of diffusion tensor images using gabor features. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 396–399, Arlington, Va., 2004.
87. R. Verma, P. Khurd, and C. Davatzikos. On analyzing diffusion tensor images by identifying manifold structure using isomaps. *IEEE Transactions on Medical Imaging*, 26(6):772–778, 2007.
88. P. Viola. *Alignment by maximization of mutual information*, Ph.D. dissertation. Ph.d. dissertation, Massachusetts Inst. Technol., 1995.
89. Y. Wang, B. S. Peterson, and L. H. Staib. 3d brain surface matching based on geodesics and local geometry. *Computer Vision and Image Understanding*, 89(2-3):252–271, 2003.
90. W. M. Wells, III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.
91. W. M. Wells, III, P. Viola, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. In *Medical Robotics and Computer Assisted Surgery*, pages 55–62. New York: Wiley, 1995.
92. G. Wu, F. Qi, and D. Shen. Learning best features for deformable registration of mr brains. In *MICCAI*, volume 3349, pages 179–187, Palm Springs, CA, 2005.
93. D. Xu, S. Mori, D. Shen, P. C. M. van Zijl, and C. Davatzikos. Spatial normalization of diffusion tensor fields. *Magnetic Resonance in Medicine*, 50(1):175–182, 2003.
94. Z. Xue, D. Shen, and C. Davatzikos. Correspondence detection using wavelet-based attribute vectors. In *MICCAI*, volume 3349, pages 762–770, Montreal, Canada, 2003. Springer-Verlag Heidelberg.
95. Z. Xue, D. Shen, and C. Davatzikos. Determining correspondence in 3d mr brain images using attribute vectors as morphological signatures of voxels. *IEEE Transactions on Medical Imaging*, 23(10):1276–1291, 2004.
96. Z. Xue, D. Shen, B. Karacali, J. Stern, D. Rottenberg, and C. Davatzikos. Simulating deformations of mr brain images for validation of atlas-based segmentation and registration algorithms. *NeuroImage*, 33(3):855–866, 2006.
97. J. Yang, D. Shen, C. Misra, X. Wu, S. Resnick, C. Davatzikos, and R. Verma. Spatial normalization of diffusion tensor images based on anisotropic segmentation. In *SPIE*, San Diego, CA, 2008.
98. Y. Zhan, D. Shen, J. Zeng, L. Sun, G. Fichtinger, J. Moul, and C. Davatzikos. Targeted prostate biopsy using statistical image analysis. *IEEE Transactions on Medical Imaging*, 26(6):779–788, 2007.

Statistical Computing on Non-Linear Spaces for Computational Anatomy

X. Pennec and P. Fillard

Abstract Computational anatomy is an emerging discipline that aims at analyzing and modeling the individual anatomy of organs and their biological variability across a population. However, understanding and modeling the shape of organs is made difficult by the absence of physical models for comparing different subjects, the complexity of shapes, and the high number of degrees of freedom implied. Moreover, the geometric nature of the anatomical features usually extracted raises the need for statistics on objects like curves, surfaces and deformations that do not belong to standard Euclidean spaces. We explain in this chapter how the Riemannian structure can provide a powerful framework to build generic statistical computing tools. We show that few computational tools derive for each Riemannian metric can be used in practice as the basic atoms to build more complex generic algorithms such as interpolation, filtering and anisotropic diffusion on fields of geometric features. This computational framework is illustrated with the analysis of the shape of the scoliotic spine and the modeling of the brain variability from sulcal lines where the results suggest new anatomical findings.

1 Computational Anatomy: Aims and Methods

Anatomy is the science that studies the structure and the relationship in space of different organs and tissues in living systems. Before the renaissance, anatomical descriptions were mainly based on animal models and the physiology was more philosophical than scientific. Modern anatomy really began with the authorized

X. Pennec (✉) • P. Fillard

Inria, 2004 Route des Lucioles, Sophia Antipolis 06600, France

e-mail: xavier.pennec@inria.fr; pierre.fillard@inria.fr

dissection of human cadavers, giving birth to the “*De humani corporis fabrica*” published by in 1543 by Vesale (1514-1564), and was strongly pushed by the progresses in surgery, as exemplified by the “*Universal anatomy of the human body*” (1561-62) of the great surgeon Ambroise Paré (1509-1590). During the following centuries, many progresses were done in anatomy thanks to new observation tools like microscopy and histology, going down to the level of cells in the 19th and 20th centuries. However, in-vivo and in-situ imaging is radically renewing the field since the 1980ies. An ever growing number of imaging modalities allows observing both the anatomy and the function at many spatial scales (from cells to the whole body) and at multiple time scales: milliseconds (e.g. beating heart), years (growth or aging), or even ages (evolution of species). Moreover, the non-invasive aspect allows repeating the observations on multiple subjects. This has a strong impact on the goals of the anatomy which are changing from the description of a *representative individual* to the description of the structure and organization of organs at the *population level*. The huge amount of information generated also raises the need for computerized methods to extract and structure information. This led in the last 10 to 20 years to the gradual evolution of *descriptive atlases* into interactive and *generative models*, allowing the simulation of new observations. Typical examples are given for the brain by the MNI 305 [25] and ICBM 152 [46] templates that are the basis of the Brain Web MRI simulation engine [20]. In the orthopedic domain, one may cite the “bone morphing” method [34,63] that allows to simulate the shape of bones.

The combination of these new observation means and of the computerized methods is at the heart of computational anatomy, an emerging discipline at the interface of geometry, statistics and image analysis which aims at developing algorithms to model and analyze the biological shape of tissues and organs. The goal is not only to estimate representative organ anatomies across diseases, populations, species or ages but also to model the organ development across time (growth or aging) and to establish their variability. Another goal is to correlate this variability information with other functional, genetic or structural information (e.g. fiber bundles extracted from diffusion tensor images). From an applicative point of view, a first objective is to understand and to model how life is functioning at the population level, for instance by classifying pathologies from structural deviations (taxonomy) and by integrating individual measures at the population level to relate anatomy and function. For instance, the goal of spatial normalization of subjects in neuroscience is to map all the anatomies into a common reference system. A second application objective is to provide better quantitative and objective measures to detect, understand and correct dysfunctions at the individual level in order to help therapy planning (before), control (during) and follow-up (after).

The method is generally to map some generic (atlas-based) knowledge to patients-specific data through atlas-patient registration. In the case of observations of the same subject, many geometrical and physically based registration methods were proposed to faithfully model and recover the deformations. However, in the case of different subjects, the absence of physical models relating the anatomies leads to a reliance on statistics to learn the geometrical relationship from

many observations. This is usually done by identifying anatomically representative geometric features (points, tensors, curves, surfaces, volume transformations), and then modeling their statistical distribution across the population, for instance via a mean shape and covariance structure analysis after a group-wise matching. In the case of the brain, one can rely on a hierarchy of structural models ranging from anatomical or functional landmarks like the AC and PC points [15, 66], curves like crest lines [65] or sulcal lines [29, 41, 44], surfaces like the cortical surface or sulcal ribbons [1, 68, 72], images seen as 3D functions, which lead to voxel-based morphometry (VBM) [6], diffusion imaging or rigid, multi-affine or diffeomorphic transformations [2, 49, 70], leading to Tensor-based morphometry (TBM). However, one crucial point is that these features usually belong to curved manifolds rather than to Euclidean spaces, which precludes the use of classical linear statistics. For instance, the average of points on a sphere is located inside the sphere and not on its surface.

To address this problem, one has to rely on statistical tools that work directly on manifolds in an intrinsic way. We summarize in Sect. 2 the mathematical bases that are needed to properly work on finite dimensional manifolds. Then, we show in Sect. 3 that a consistent set of statistical tools, including mean and covariance matrix analysis, can be developed based on the choice of a Riemannian metric. This algorithmic framework to compute on manifolds is then extended to process fields of geometric features (manifold-valued image). In particular, we show that one can perform interpolation, filtering, isotropic and anisotropic regularization and restoration of missing data (extrapolation or in-painting) on manifold valued images by using generalized weighted means and partial derivative equations (PDEs). Finally, the methodology is exemplified in Sect. 4 with two example applications: the statistical analysis of the anatomic variability of the spine in scoliotic patients, where a set of rigid body transformations is used to model the articulations between the vertebrae; and the modeling of the variability of the brain from a data-set of precisely delineated sulcal lines, where covariance matrices (symmetric positive definite matrices, so-called tensors) are used to describe the individual and joint anatomical variability (Green function) of points in the brain.

2 Mathematical bases of computing on manifolds

Computing on simple manifolds like the 3D sphere or a flat torus (for instance an image with opposite boundary points identified) might seem easy as we can see the geometrical properties (e.g. invariance by rotation or translation) and imagine tricks to alleviate the different problems. However, when it comes to slightly more complex manifolds like tensors, rigid body or affine transformations, without even thinking to infinite dimensional manifolds like spaces of surfaces or diffeomorphisms, computational tricks are much more difficult to find and have to be determined on a case by case basis. The goal of this section is to exemplify with the development of basic but generic statistical tools that the work specific to

each manifold can be limited the determination of a few computational tools derived from a chosen Riemannian metric. These tools will then constitute the basic atoms to build more complex generic algorithms in Sect. 3.

2.1 *The basic structure: the Riemannian metric*

In the geometric framework, one has to separate the topological and differential properties of the manifold from the metric ones. The first ones determine the local structure of a manifold \mathcal{M} by specifying neighboring points and tangent vectors, which allows to differentiate smooth functions on the manifold. The topology also impacts the global structure as it determines if there exists a connected path between two points. However, we need an additional structure to quantify how far away two connected points are: a distance. By restricting to distances which are compatible with the differential structure, we enter into the realm of Riemannian geometry. A *Riemannian metric* is defined by a continuous collection of scalar products $\langle \cdot | \cdot \rangle_p$ (or equivalently quadratic norms $\|\cdot\|_p$) on each tangent space $T_p\mathcal{M}$ at point p of the manifold. Thus, if we consider a curve on the manifold, we can compute at each point its instantaneous speed vector (this operation only involves the differential structure) and its norm to obtain the instantaneous speed (the Riemannian metric is needed for this operation). To compute the length of the curve, this value is integrated as usual along the curve. The distance between two points of a connected Riemannian manifold is the minimum length among the curves joining these points. The curves realizing this minimum are called geodesics. The calculus of variations shows that geodesics are the solutions of a system of second order differential equations depending on the Riemannian metric. In the following, we assume that the manifold is geodesically complete, i.e. that all geodesics can be indefinitely extended. This means that the manifold has neither boundary nor any singular point that we can reach in a finite time. As an important consequence, the Hopf-Rinow-De Rham theorem states that there always exists at least one minimizing geodesic between any two points of the manifold (i.e. whose length is the distance between the two points).

2.2 *Exponential chart*

Let p be a point of the manifold that we consider as a local reference and \vec{v} a vector of the tangent space $T_p\mathcal{M}$ at that point. From the theory of second order differential equations, we know that there exists one and only one geodesic $\gamma_{(p,\vec{v})}(t)$ starting from that point with this tangent vector. This allows to wrap the tangent space onto the manifold, or equivalently to develop the manifold in the tangent space along the geodesics (think of rolling a sphere along its tangent plane at a given point), by mapping to each vector $\vec{v} \in T_p\mathcal{M}$ the point q of the manifold that is reached after

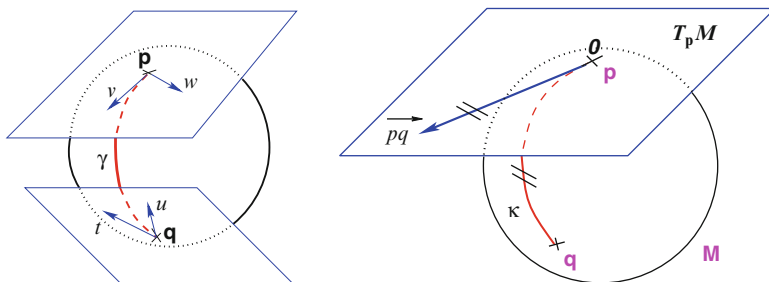


Fig. 1 Left: The tangent planes at points p and q of the sphere \mathcal{S}_2 are different: the vectors v and w of $T_p\mathcal{M}$ cannot be compared to the vectors t and u of $T_q\mathcal{M}$. Thus, it is natural to define the scalar product on each tangent plane. **Right:** The geodesics starting at x are straight lines in the exponential map and the distance along them is conserved

a unit time by the geodesic $\gamma_{(p,\vec{v})}(t)$. This mapping $\text{Exp}_p(\vec{v}) = \gamma_{(p,\vec{v})}(1)$ is called the *exponential map* at point p . Straight lines going through 0 in the tangent space are transformed into geodesics going through point p on the manifold and distances along these lines are conserved (Fig. 1).

The exponential map is defined in the whole tangent space $T_p\mathcal{M}$ (since the manifold is geodesically complete) but it is generally one-to-one only locally around 0 in the tangent space (i.e. around p in the manifold). In the sequel, we denote by $\vec{p}\vec{q} = \text{Log}_p(q)$ the inverse of the exponential map: this is the smallest vector (in norm) such that $q = \text{Exp}_p(\vec{p}\vec{q})$. If we look for the maximal definition domain, we find out that it is a star-shaped domain delimited by a continuous curve C_p called the *tangential cut-locus*. The image of C_p by the exponential map is the cut locus \mathcal{C}_p of point p . This is (the closure of) the set of points where several minimizing geodesics starting from p meet. On the sphere $\mathcal{S}_2(1)$ for instance, the cut locus of a point p is its antipodal point and the tangential cut locus is the circle of radius π .

The exponential and log maps within this domain realize a chart (a local parameterization of the manifold) called *the exponential chart at point p*. It covers all the manifold except the cut locus of the reference point p , which has a null measure. In this chart, geodesics starting from p are straight lines, and the distance from the reference point are conserved. This chart is somehow the “most linear” chart of the manifold with respect to the reference point p . The set of all the exponential charts at each point of the manifold realize an atlas which allows working very easily on the manifold, as we will see in the following.

2.3 Practical implementation

The exponential and logarithmic maps (from now on Exp and Log maps) are obviously different for each manifold and for each metric. Thus they have to be determined and implemented on a case by case basis. Example for rotations, rigid body transformations can be found for the left invariant metric in [60], and examples

Table 1 Re-interpretation of standard operations in a Riemannian manifold

	Euclidean space	Riemannian manifold
Subtraction	$\vec{pq} = q - p$	$\vec{pq} = \text{Log}_p(q)$
Addition	$p = q + \vec{v}$	$q = \text{Exp}_p(\vec{v})$
Distance	$\text{dist}(p, q) = \ q - p\ $	$\text{dist}(p, q) = \ \vec{pq}\ _p$
Mean value (implicit)	$\sum_i (p_i - \bar{p}) = 0$	$\sum_i \vec{pp}_i = 0$
Gradient descent	$p_{t+\varepsilon} = p_t - \varepsilon \overrightarrow{\nabla C}(p_t)$	$p_{t+\varepsilon} = \text{Exp}_{p_t}(-\varepsilon \overrightarrow{\nabla C}(p_t))$
Geodesic interpolation	$p(t) = p_0 + t \overrightarrow{p_0 p_1}$	$p(t) = \text{Exp}_{p_0}(t \overrightarrow{p_0 p_1})$

for tensors in [4, 58]. Exponential charts constitute very powerful atomic functions in terms of implementation on which we will be able to express practically all the geometric operations: the implementation of Log_p and Exp_p is the basis of programming on Riemannian manifolds, as we will see in the following.

In a Euclidean space, the exponential charts are nothing but one orthonormal coordinates system translated at each point: we have in this case $\vec{pq} = \text{Log}_p(q) = q - p$ and $\text{Exp}_p(\vec{v}) = p + \vec{v}$. This example is more than a simple coincidence. In fact, most of the usual operations using additions and subtractions may be reinterpreted in a Riemannian framework using the notion of *bi-point*, an antecedent of vector introduced during the 19th Century. Indeed, vectors are defined as equivalent classes of bi-points in a Euclidean space. This is possible because we have a canonical way (the translation) to compare what happens at two different points. In a Riemannian manifold, we can still compare things locally (by parallel transportation), but not any more globally. This means that each “vector” has to remember at which point of the manifold it is attached, which comes back to a bi-point.

A second way to see the vector \vec{pq} is as a vector of the tangent space at point p . Such a vector may be identified to a point on the manifold using the exponential map $q = \text{Exp}_p(\vec{pq})$. Conversely, the logarithmic map may be used to map almost any bi-point (p, q) into a vector $\vec{pq} = \text{Log}_p(q)$ of $T_p\mathcal{M}$. This reinterpretation of addition and subtraction using logarithmic and exponential maps is very powerful to generalize algorithms working on vector spaces to algorithms on Riemannian manifolds, as illustrated in Table 1 and the in following sections.

2.4 Example of Metrics on Covariance matrices

Let us take an example with positive definite symmetric matrices, called tensors in medical image analysis. They are used for instance to encode the covariance matrix of the Brownian motion (diffusion) of water in Diffusion Tensor Imaging (DTI) [8, 40] or to encode the joint variability at different places (Green function) in shape analysis (see [29, 30, 31] and Sect. 4). They are also widely used in image analysis to guide the segmentation, grouping and motion analysis [16, 47, 73, 74].

The main problem is that the tensor space is a manifold that is not a vector space with the usual additive structure. Indeed, the positive definiteness constraint delimits a convex half-cone in the vector space of symmetric matrices. Thus, convex operations (like the mean) are stable in this space but problems arise with more complex operations. For instance, there is inevitably a point in the image where the time step is not small enough when smoothing fields of tensors with gradient descents, and this results into negative eigenvalues.

To answer that problem, we proposed in [58] to endow the space of tensors with a Riemannian metric invariant by any change of the underlying space coordinates, i.e. invariant under the action of affine transformations on covariance matrices. A few mathematical developments showed that the Exp, Log and distance maps were given with quite simple formulas involving the matrix logarithm exp and log:

$$\begin{aligned}\text{Exp}_{\Sigma}(W) &= \Sigma^{1/2} \exp(\Sigma^{-1/2} W \Sigma^{-1/2}) \Sigma^{1/2} \\ \text{Log}_{\Sigma}(\Lambda) &= \Sigma^{1/2} \log(\Sigma^{-1/2} \Lambda \Sigma^{-1/2}) \Sigma^{1/2} \\ \text{dist}^2(\Sigma, \Lambda) &= \text{Tr}(\log(\Sigma^{-1/2} \Lambda \Sigma^{-1/2})^2)\end{aligned}$$

This metric leads to a very regular Hadamard manifold structure, a kind of hyperbolic space without cut-locus, which simplifies the computations. Tensors with null and infinite eigenvalues are both at an infinite distance of any positive definite symmetric matrix: the cone of positive definite symmetric matrices is changed into a space of “constant” (homogeneous) non-scalar curvature without boundaries. Moreover, there is one and only one geodesic joining any two tensors, the mean of a set of tensors is uniquely defined, and we can even define globally consistent orthonormal coordinate systems of tangent spaces. Thus, the structure we obtain is very close to a vector space, except that the space is curved.

This affine-invariant Riemannian metric derives from affine invariant connections on homogeneous spaces [54]. It has been introduced in statistics to model the geometry of the multivariate normal family (the Fisher information metric) [18, 19, 64] and in simultaneously by many teams in medical image analysis to deal with DTI [9, 33, 42, 52, 58]. In [58], we showed that this metric could be used not only to compute distances between tensors, but also as the basis of a complete computational framework on manifold-valued images as will be detailed in Sect. 3.

By trying to put a Lie group structure on the space of tensors, Vincent Arsigny observed later that the matrix exponential was a diffeomorphism from the space of symmetric matrices to the tensor space. Thus, one can seamlessly transport all the operations defined in the vector space of symmetric matrices to the tensor space [4, 5]. This gives a commutative Lie group structure to the tensors, and the Euclidean metric on symmetric matrices is transformed into a bi-invariant Riemannian metric on the tensor manifold. As geodesics are straight lines in the space of symmetric

matrices, the expression of the Exp, Log and distance maps for the Log-Euclidean metric is easily determined:

$$\begin{aligned}\text{Exp}_{\Sigma}(W) &= \exp(\log(\Sigma) + \partial_W \log(\Sigma)) \\ \text{Log}_{\Sigma}(\Lambda) &= D \exp(\log(\Sigma)) (\log(\Lambda) - \log(\Sigma)) \\ \text{dist}_{LE}^2(\Sigma_1, \Sigma_2) &= \text{Tr}((\log(\Sigma_1) - \log(\Sigma_2))^2)\end{aligned}$$

These formulas look more complex than for the affine invariant metric because they involve the differential of the matrix exponential and logarithm in order to transport tangent vectors from one space to another [59]. However, they are in fact nothing but the transport of the addition and subtraction through the exponential of symmetric matrices. In practice, the log-Euclidean framework consist in taking the logarithm of the tensor data, computing like usual in the Euclidean space of symmetric matrices, and coming back at the end to the tensor space using the exponential [3, 5].

From a theoretical point of view, geodesics through the identity are the same for both log-Euclidean and affine-invariant metrics, but this is not true any more in general at other points of the tensor manifold [4]. A careful comparison of both metrics in practical applications [3, 5] showed that there was very few differences on the results (of the order of 1%) on real DTI images, but that the log-Euclidean computations were 4 to 10 times faster. For other types of applications, like adaptive re-meshing [53], the anisotropy of the tensors can be much larger, which may lead to larger differences. In any case, initializing the iterative optimizations of affine-invariant algorithms with the log-Euclidean result drastically speeds-up the convergence. Important application example of this tensor computing framework were provided in [27, 28] with a statistically grounded estimation and regularization of DTI images. The white matter tractography that was allowed by these methods in clinical DTI images with very poor signal to noise ratios could lead to new clinical indications of DTI, for instance in the spinal chord [23].

3 Statistical Computing on Manifolds

The previous section showed how to derive the atomic Exp and Log maps from a Riemannian metric. We now summarize in this section how one generalizes on this basis many important statistical notions, like the mean, covariance and Principal Component Analysis (PCA), as well as many image processing algorithms like interpolation, diffusion and restoration of missing data (extrapolation). For details about the theory of statistics on Riemannian manifolds in itself, we refer the reader to [56, 57] and reference therein. Manifold-valued image processing is detailed in [58] with the example of tensors.

3.1 First statistical moment: the mean

The Riemannian metric induces an infinitesimal volume element on each tangent space, and thus a reference measure $d\mathcal{M}(p)$ on the manifold that can be used to measure random events on the manifold and to define the probability density function (the function ρ such that $dP(p) = \rho(p)d\mathcal{M}(p)$, if it exists). It is worth noticing that the induced measure $d\mathcal{M}$ represents the notion of *uniformity* according to the chosen Riemannian metric. This automatic derivation of the uniform measure from the metric gives a rather elegant solution to the Bertrand paradox for geometric probabilities [38, 62]. With the probability measure of a random element, we can integrate functions from the manifold to any vector space, thus defining the expected value of this function. However, we generally cannot integrate manifold-valued functions. Thus, one cannot define the mean or expected “value” of a random manifold element that way.

One solution is to rely on a distance-based variational formulation: the Fréchet (resp. Karcher) expected features minimize globally (resp. locally) the variance:

$$\sigma^2(q) = \int \text{dist}(p, q)^2 dP(p) = \frac{1}{n} \sum_{i=1}^n \text{dist}(p_i, q)^2,$$

written respectively in the continuous and discrete forms. One can generalize the variance to a dispersion at order α by changing the L_2 with an α -norm: $\sigma_\alpha(p) = (\int \text{dist}(p, q)^\alpha dP(p))^{1/\alpha}$. The minimizers are called the central Karcher values at order α . For instance, the median is obtained for $\alpha = 1$ and the modes for $\alpha = 0$, exactly like in the vector case. It is worth noticing that the median and the modes are not unique in general in the vector space, and that even the mean may not exist (e.g. for heavy tailed distribution). In Riemannian manifolds, the existence and uniqueness of all central Karcher values is generally not ensured as they are obtained through a minimization procedure. However, for a finite number of discrete samples at a finite distance of each other, which is the practical case in statistics, a mean value always exists and it is unique as soon as the distribution is sufficiently peaked [37, 39].

Local minima may be characterized as particular critical points of the cost function: at Karcher mean points, the gradient of the variance should be null. However, the distance is continuous but not differentiable at cut locus points where several minimizing geodesic meets. For instance, the distance from a point of the sphere to its antipodal point is maximal, but decrease continuously everywhere around it. One can show [56, 57] that the variance is differentiable at all points where the cut locus has a null measure and has gradient: $\overrightarrow{\nabla \sigma^2}(q) = -2 \int \overrightarrow{qp} dP(p) = \frac{-2}{n} \sum_{i=1}^n \overrightarrow{qp}_i$ respectively in the continuous (probabilistic) and discrete (statistical) formulations. In practice, this gradient is well defined for all distributions that have a pdf since the cut locus has a null measure. For discrete samples, the gradient exists if there is

no sample lying exactly on the cut-locus of the current test point. Thus, we end up with the implicit characterization of Karcher mean points as exponential barycenters which was presented in Table 1.

To practically compute the mean value, we proposed in [60] for rigid body transformations and in [56, 57] for the general Riemannian case to use a Gauss-Newton gradient descent algorithm. It essentially alternates the computation of the barycenter in the exponential chart centered at the current estimation of the mean value, and a re-centering step of the chart at the point of the manifold that corresponds to the computed barycenter (geodesic marching step). This gives the Newton iteration: $\bar{p}^{t+1} = \text{Exp}_{\bar{p}^t} \left(\frac{1}{n} \sum_{i=1}^n \overrightarrow{\bar{p}^t p_i} \right)$. One can actually show that its convergence is locally quadratic towards non degenerated critical points [22, 43, 55].

3.2 Covariance matrix and Principal Geodesic Analysis

Once the mean point is determined, using the exponential chart at the mean point is particularly interesting as the random feature is represented by a random vector with null mean in a star-shaped domain. With this representation, there is no difficulty to define the covariance matrix:

$$\Sigma = \int \overrightarrow{\bar{p}q} \cdot \overrightarrow{\bar{p}q}^T dP(q) = \frac{1}{n} \sum_{i=1}^n \overrightarrow{\bar{p}q_i} \cdot \overrightarrow{\bar{p}q_i}^T$$

and potentially higher order moments. This covariance matrix can then be used to defined the Mahalanobis distance between a random and a deterministic feature: $\mu_{(\bar{p}, \Sigma)}(q) = \overrightarrow{\bar{p}q}^T \Sigma^{-1} \overrightarrow{\bar{p}q}$. Interestingly, the expected Mahalanobis distance of a random element is independent of the distribution and is equal to the dimension of the manifold, as in the vector case. This statistical distance can be used as a basis to generalize some statistical tests such as the mahalanobis D^2 test [57].

To analyze the results of a set of measurements in a Euclidean space, one often performs a principal component analysis (PCA). A generalization to Riemannian manifolds called Principal Geodesic Analysis (PGA) was proposed in [32] to analyze shapes based on the medial axis representations (M-reps). The basic idea is to find a low dimensional sub-manifold generated by some geodesic subspaces that best explain the measurements (i.e. such that the squared Riemannian distance from the measurements to that sub-manifold is minimized). Another point of view is to assume that the measurements are generated by a low dimensional Gaussian model. Estimating the model parameters amounts to a covariance analysis in order to find the k -dimensional subspace that best explains the variance. In a Euclidean space, these two definitions correspond thanks to Pythagoras's theorem. However, in the Riemannian setting, geodesic subspaces are generally not orthogonal due to the

curvature. Thus, the two notions differ: while the Riemannian covariance analysis (tangent PCA) can easily be performed in the tangent space of the mean, finding Riemannian sub-manifolds turns out to become a very difficult problem. As a matter of fact, the solution retained by [32] was finally to rely on the covariance analysis.

When the distribution is unimodal and sufficiently peaked, we believe that covariance analysis is anyway much better suited. However, for many problems, the goal is rather to find a sub-manifold on which measurements are more or less uniformly distributed. This is the case for instance for features sampled on a surface or points sampled along a trajectory (time sequences). While the one dimensional case can be tackled by regression [21], the problem for higher dimensional sub-manifolds remains quite open. Some solutions may come from manifold embedding techniques as exemplified for instance in [17].

3.3 Interpolation and filtering as weighted means

One of the important operations in geometric data processing is to interpolate values between known measurements. The standard way to interpolate on a regular lattice is to make a linear combination of samples f_k at integer (lattice) coordinates $k \in \mathbb{Z}^d$: $f(x) = \sum_k w(x - k) f_k$. A typical example is the sinus cardinal interpolation. With the nearest-neighbor, linear (or tri-linear in 3D), and higher order spline interpolations, the kernel is piecewise polynomial, and has a compact support [48, 67]. With normalized weights, this interpolation can be seen as a weighted mean. Thus, it can be generalized in the manifold framework as an optimization problem: the interpolated value $p(x)$ on our feature manifold is the point that minimizes $C(p(x)) = \sum_{i=1}^n w_i(x) \text{dist}^2(p_i, p(x))$. This can easily be solved using the iterative Gauss-Newton scheme proposed for the Karcher mean. The linear interpolation is interesting and can be written explicitly since it is a simple geodesic walking scheme: $p(t) = \text{Exp}_{p_0}(t \overrightarrow{p_0 p_1}) = \text{Exp}_{p_1}((1-t) \overrightarrow{p_1 p_0})$.

Many other operators can be rephrased as weighted means. For instance approximations and convolutions like Gaussian filtering can be viewed as the average of the neighboring values weighted by a (Gaussian) function of their spatial distance. For instance, $\hat{F}(x) = \int_{\mathbb{R}^n} K(u) F(x + u) du$ is the minimizer of $C(\hat{F}) = \int_{\mathbb{R}^n} K(u) \text{dist}^2(F(x + u), \hat{F}(x)) du$. In a Riemannian manifold, this minimization problem is still valid, but instead of a closed-form solution, we have once again a Gauss-Newton iterative gradient descent algorithm to reach the filtered value:

$$\hat{p}^{t+1}(x) = \int_{\mathbb{R}^n} K(u) \text{Log}_{\hat{p}^t(x)}(p(x + u)) du.$$

We can also use anisotropic and non-stationary kernels $K(x, u)$. For instance, it can be modulated by the norm of the derivative of the field in the direction u . We should notice that for a manifold-value field $p(x)$, the directional derivatives $\partial_u p(x)$ is a tangent vector of $T_{p(x)}\mathcal{M}$ which can be practically approximated using finite

“differences” in the exponential chart: $\partial_u p(x) \simeq \text{Log}_{p(x)}(p(x+u)) + O(\|u\|^2)$. However, to measure the norm of this vector, we have to use the Riemannian metric at that point: $\|\partial_u p\|_p$.

3.4 Harmonic diffusion and anisotropic regularization

An alternative to kernel filtering is to consider a regularization criterion that penalizes the spatial variations of the field. A measure of variation is the spatial gradient (the linear form that maps to any spatial direction u the directional derivative $\partial_u p(x)$), which can be robustly computed as the matrix that best approximates the directional derivatives in the neighborhood (e.g. 6, 18 or 26 connectivity in 3D). The simplest criterion based on the gradient is the Harmonic energy

$$\text{Reg}(p) = \frac{1}{2} \int_{\Omega} \|\nabla p(x)\|_{p(x)}^2 dx = \frac{1}{2} \sum_{i=1}^d \int_{\Omega} \|\partial_{x_i} p(x)\|_{p(x)}^2 dx.$$

The Euler-Lagrange equation with Neumann boundary conditions is as usual $\nabla \text{Reg}(p)(x) = -\Delta p(x)$. However, the Laplace-Beltrami operator on the manifold $\Delta p(x)$ is the sum of the usual flat Euclidean second order directional derivatives $\partial_{x_i}^2 p(x)$ in a locally orthogonal system and an additional term due to the curvature of the manifold that distorts the orthonormality of this coordinate system. To practically compute this operator, we proposed in [58] an efficient and general scheme based on the observation that the Christoffel symbols and their derivatives along the geodesics vanish at the origin of the exponential chart. This means that the correction for the curvature is in fact already included: by computing the standard Laplacian *in that specific map*, one gets the directional Laplace-Beltrami operator for free: $\Delta_u p = \text{Log}_{p(x)}(p(x+u)) + \text{Log}_{p(x)}(p(x-u)) + O(\|u\|^4)$. Averaging over all the directions in a neighborhood finally gives a robust and efficient estimation.

A very simple scheme to perform Harmonic diffusion is to use a first order geodesic gradient descent. At each iteration and at each point x , one walks a little bit along the geodesic which start at the current point with the opposite of the gradient of the regularization criterion:

$$p^{t+1}(x) = \text{Exp}_{p^t(x)}(-\varepsilon \Delta p^t(x)) \quad \text{with} \quad \Delta p(x) \propto \sum_{u \in \mathcal{V}} \frac{1}{\|u\|^2} \text{Log}_{p(x)}(p(x+u))$$

In order to filter within homogeneous regions but not across their boundaries, an idea is to penalize the smoothing in the directions where the derivatives are important [35, 61]. This can be realized directly in the discrete implementation of the Laplacian by weighting the directional Laplacian by a decreasing function of the norm $\|\partial_u p\|_p$ of the gradient in that direction. For instance, we used $\Delta_u p = \sum_u c(\|\partial_u p\|_p) \Delta_u p$ with $c(x) = \exp(-x^2/\kappa^2)$ in [58]. As the convergence of

this scheme is not guaranteed (anisotropic regularization “forces” may not derive from a well-posed energy), the problem may be reformulated as the optimization of a ϕ -function of the Riemannian norm of the spatial gradient (a kind of robust M-estimator): $Reg_\phi(p) = \frac{1}{2} \int_\Omega \phi(\|\nabla p(x)\|_{p(x)}) dx$. By choosing an adequate ϕ -function, one can give to the regularization an isotropic or anisotropic behavior [7]. The main difference with a classical Euclidean calculation is that we have to take the curvature into account by using the Laplace-Beltrami operator, and by measuring the length of directional derivatives using the Riemannian metric at the right point [26]. Using $\Psi(x) = \phi'(x)/x$, we get:

$$\nabla Reg_\phi(p) = -\Psi(\|\nabla p\|_p)\Delta p - \sum_{i=1}^d \partial_{x_i} \Psi(\|\nabla p\|_p)\partial_{x_i} p.$$

3.5 Diffusion-based interpolation and extrapolation

The pure diffusion reduces the noise in the data but also the amount of information. Moreover, the total diffusion time that controls the amount of smoothing is difficult to estimate. At an infinite diffusion time, the field will be completely homogeneous. Thus, it is more interesting to consider the data as noisy observations and the regularization as a prior on the spatial regularity of the field. Usually, one assumes a Gaussian noise independent at each position, which leads to a least-squares criterion through a maximum likelihood approach. For a dense data field $q(x)$, the similarity criterion that is added to the regularization criterion is simply $Sim(p) = \int_\Omega \text{dist}^2(p(x), q(x)) dx$. The only difference here is that it uses the Riemannian distance. It simply adds a linear (geodesic) spring $\nabla_p \text{dist}^2(p, q) = -2 \overrightarrow{pq}$ to the global gradient to prevent the regularization from pulling to far away from the original data.

For sparse measures, using directly the maximum likelihood on the observed data leads to deal with Dirac (mass) distributions in the derivatives, which is a problem for the numerical implementation. One solution is to consider the Dirac distribution as the limit of the Gaussian function G_σ when σ goes to zero, which leads to the regularized derivative [58]: $\nabla Sim(x) = -2 \sum_{i=1}^n G_\sigma(x - x_i) \overrightarrow{p(x)p_i}$.

4 Modeling the Anatomy

4.1 A statistical shape model of the scoliotic spine

Now that we have the methodology to work with geometric features, let us see how it can be used to model the anatomy. A first interesting example was proposed by Jonathan Boisvert [12, 13] with a 3D articulated model of the spine. The model



Fig. 2 First (left) and second (right) modes of variation of the statistical spine model depicted at -3 , 0 (mean) and 3 times its standard deviation. Images courtesy of Jonathan Boisvert, Polytechnique School of Montreal, Canada

gathers the relative configurations of the vertebrae along the spinal chord (the parameters are the rigid transforms that superpose neighboring vertebrae) rather than the position and orientation of each vertebra in a global reference frame. As small local motions at one point of the spine may have a large impact of the position at another point, this local representation is better capturing information that may get unnoticed in a global reference frame. However, this requires making statistics on geometric objects (rigid body transformation parameters) rather than on just points.

The statistical model of the spine was established in a population of 307 untreated scoliotic patients. Each vertebra was reconstructed in 3D from anatomical landmarks in bi-planar radiographies. Posture during data acquisition was normalized but individual factors such as the age, sex or type of scoliotic curve were not taken into account. Thus, the statistics capture the anatomical variability inherent to the pathology but also the growth stage. The Fréchet mean and the generalized covariance of the articulated model was then computed. As there are 102 degrees of freedom (5 lumbar and 12 thoracic vertebrae), the analysis of the covariance matrix could hardly be performed by hand. Thus, the most meaningful modes of variation were extracted using a PCA on the tangent plane.

A visual inspection reveals that the first modes had clinical meanings and were explaining curve patterns that are routinely used in different clinical classifications of scoliosis (see [11, 14] for details). For instance, the first mode appears to be associated with the patient growth with a mild thoracic curve (King's type II or III depending on the amplitude of the mode) and the second could be described as a double thoraco-lumbar curve (King's type I), see Fig. 2. A more quantitative analysis showed that there is a statistically significant link between the 4 principal modes and King's classes, although each class is generally linked to a combination of modes rather than only one mode [11].

4.2 Learning Brain Variability from Sulcal Lines

A second interesting shape analysis application is the statistical modeling of the brain variability in a given population of 3D images. In such a process, a first step is to measure the variability of each anatomical position independently by identifying for instance corresponding points among each individual anatomy (structural homologies). This allows us to encode the brain variability by covariance matrices that we call *variability tensors*. The reason why we should not simplify these tensors into simpler scalar values is that there are evidences that structural variations are larger along certain preferred directions [69].

As exemplify in introduction, a hierarchy of anatomical features may be used to abstract the brain anatomy. We chose sulcal lines as they are low dimensional structures easily identified by neuroscientists. Moreover, a certain number of sulcal landmarks consistently appear in all normal individuals and allow a consistent subdivision of the cortex into major lobes and gyri [45]. In the framework of the associated team program Brain-Atlas between Asclepios at INRIA and LONI at UCLA, we use a data-set of sulcal lines manually delineated in 98 subjects by expert neuroanatomists according to a precise protocol¹. We used the 72 sulcal curves that consistently appear in all normal subjects (abusively called *sulci* in the sequel).

To find the corresponding points between the 98 instances of each of the 72 sulci, we proposed in [29, 30] an original methodology which alternatively compute the matches that minimize the distance between the mean curve and the instances, and re-estimates the mean curve from the updated matches. As a result, we obtain the mean sulcal curves, along with a variability tensor which encodes the covariance matrix of the anatomical positions in each subject corresponding to each mean point. To optimally adjust the number of tensor needed to represent the variability information along each sulcus, we proposed a *tensor picking* method. The principle is to approximate our tensor measurements using a linear interpolation in-between N tensors picked along the line. The optimal subset of tensors is determined by optimizing the distance between interpolated and measured tensors along the line. The number of tensors picked along each line is adjusted so that the interpolation error does not exceed a prescribed value. In this process, we used the Riemannian metrics presented in Sect. 2.4 for their very good theoretical properties, and the algorithmic framework developed in Sect. 3.3. Tensor picking is illustrated in Fig. 3. We were able to show that selecting only 366 variability tensors was sufficient to encode the variability of the 72 sulci without a significant loss of accuracy.

The result is a sparse field of tensors, which can naturally be extrapolated to the whole space using the framework described in Sect. 3.5 (Fig. 4). This dense map of tensors was shown to be in good agreement with previous published results: the highly specialized and lateralized areas such as the planum parietale and the temporo-parietal areas consistently show the highest amount of variability.

¹http://www.loni.ucla.edu/~khayashi/Public/medial_surface/

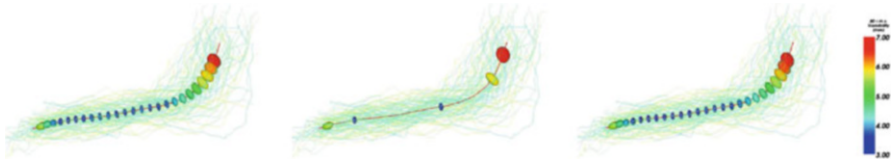


Fig. 3 Measuring variability tensors along the Sylvian Fissure. Left: Covariance matrices (ellipsoids at one standard deviation) are overlaid at regularly sampled spatial positions along the mean sulci. Middle: Tensors selected by our tensor picking operation. Right: Tensors reconstructed by linear interpolation in-between. Notice that only 5 tensors in that case nicely represent the variability of the entire sulcus

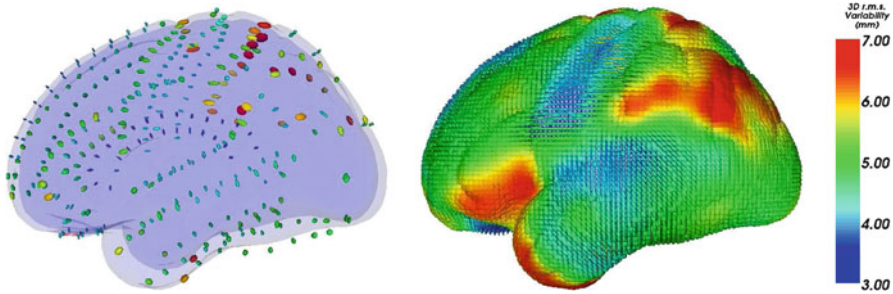


Fig. 4 Variability tensor extrapolation. Left: The 366 tensors retained for our model. Right: Result of the extrapolation. Each point of this average brain shape contains a variability tensor

The lowest amount of variability is consistently found in phylogenetically older areas (e.g. orbitofrontal cortex) and primary cortices that myelinate earliest during development (e.g., primary somatosensory and auditory cortex). However, our variability map gives more than just the amount of variability: we can extract from the tensors the spatial directions where the variability is the greatest at every single anatomical position. We refer the reader to [29, 30] for a more detailed explanation of the method and for the neuroscience interpretation of these results.

Modeling independently the variability at each point may not be sufficient as we may overlook potential statistical relationships between different brain regions. Indeed, long range relationships may arise from common genetic and trophic influences across brain regions (e.g., brain regions that develop together). In our framework, such relationships can be revealed by an analysis of the correlation between spatially close anatomical positions along the lines (neighboring points), but also distant points (e.g., a point and its symmetric counterpart in the opposite hemisphere).

In [31], we studied the correlation between two points \bar{x} and \bar{y} of the mean sulcal lines by canonical correlation analysis. This analysis is based on the total covariance matrix (TCM) of the corresponding points x_i and y_i in each subject anatomy:

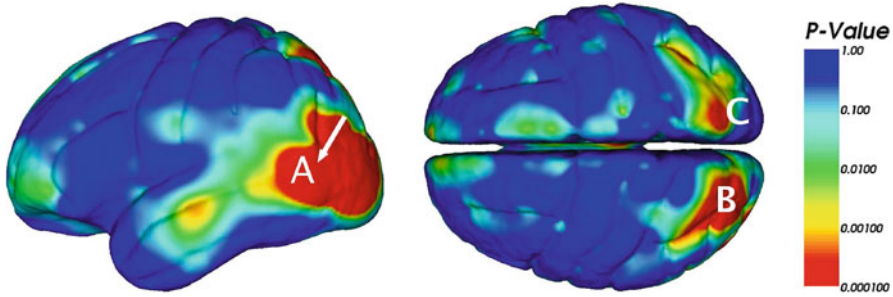


Fig. 5 Map of anatomical correlations. The tip of the superior temporal sulcus (marked A) was picked as a reference point. The map indicates regions which are spotted as correlated with this reference position (hot colors mean correlation). The most correlated points include the parietal sulci (marked B and C), a very interesting neuroscience finding

$$TCM(x, y) = \frac{1}{n-1} \sum_{i=1}^n \begin{pmatrix} x_i - \bar{x} \\ y_i - \bar{y} \end{pmatrix} \begin{pmatrix} x_i - \bar{x} \\ y_i - \bar{y} \end{pmatrix}^T = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_{yy} \end{pmatrix}.$$

The TCM being a 6-dimensional tensor, our Riemannian processing framework naturally allows us to extrapolate this tensor from the sulcal lines to the whole cortex surface. To statistically assess the correlation of the variability at any two anatomical positions, we used the Bartlett-Lawley test which tests the rank of the correlation matrix $\Gamma = \Sigma_{xx}^{-1/2} \Sigma_{xy} \Sigma_{yy}^{-1/2}$. The singular values of this matrix are the correlation coefficients between the spatial directions given by corresponding eigenvector. A rank of at least 1 means that the variability at \bar{x} and at \bar{y} are significantly correlated in at least one direction. To account for multiple comparisons, we used the very conservative Bonferroni correction. In addition to the expected local correlation, results indicates that there was generally a long-range correlations with the symmetric point in the other hemisphere, but also unexpected long-range correlations with other parts of the brain as shown for instance in Fig. 5 with the superior temporal sulcus.

5 Challenges

We have shown in this chapter that the choice of a Riemannian metric and the implementation of a few tools derived from it, namely the Exp and Log maps, provide the bases for building a consistent algorithmic framework to compute on manifolds. In particular, we showed that one can compute consistent statistics, perform interpolation, filtering, isotropic and anisotropic regularization and restoration of missing data.

We also showed that powerful computational models of the anatomy could be built thanks to this Riemannian computing framework. For instance, Sect. 4.1 demonstrates that using a proper non-linear model of the spine allows to find a good

separation of different physiological phenomena such as pathological deformations and normal growth. In this example, using the generic tools of the Riemannian computing framework particularly simplifies both the inception of the experimental setup, its implementation and the exploitation of the results. Sect. 4.2 also proposes a brain variability model that is able to recover the estimation of the sulcal variability with a very low number of parameters. By pushing further this statistical investigation, we showed that the same Riemannian framework could be used to measure and model the anatomical correlations between any two positions. These correlations are very interesting from a neuroscientific point of view since they can reveal factors of dependence between brain regions (like regions that develop or fail to develop together).

However, there are many challenges left open both from the theoretical and application point of views. For instance, it would be necessary to extend the computing framework presented here to infinite dimensional manifolds in order to deal properly with curves, surfaces and diffeomorphisms. For the case of diffeomorphism, we already know how to provide Riemannian metrics for which the geodesics can be computed by optimization [10, 36, 49, 50]. Through the so called EPDiff equation (Euler-Poincaré equation for diffeomorphisms), this optimization framework has been recently rephrased in an exponential/logarithm framework similar to the one developed here [51]. Thus, the basic algorithmic tools are the same, except that optimizing each time to compute the exponential and the logarithm has a deep impact on the computational times. Moreover, the infinite number of dimensions forbids the use of many tools like the probability density functions! For instance, the computation of simple statistics like the mean and the principal component analysis of diffeomorphism raises practical representation problems [24, 71].

From a computational anatomy standpoint, the huge number of degrees of freedom involved in the estimation of the anatomical variability will require to aggregate information coming from many different sources in order to improve the statistical power. As there is no gold standard, we should also be careful that many biases may be hidden in the results. Thus, methods to compare and fuse statistical information coming from many different anatomical features will need to be developed in order to confirm anatomical findings. For the brain variability, one could for instance add to the sulci other cortical landmarks like sulcal ribbons and gyri, the surface of internal structures like the ventricles, the hippocampus or the corpus callosum, or fiber pathways mapped from DTI. These sources of information are individually providing a partial and biased view of the whole variability. Thus, we expect to observe a good agreement in some areas, and complementary measures in other areas. This will most probably lead in a near future to new anatomical findings and more robust medical image analysis applications.

References

1. A. Andrade, F. Kherif, J.-F. Mangin, K. Worsley, A.-L. Paradis, O. Simon, S. Dehaene, and J.-B. Poline. Detection of fMRI activation using cortical surface mapping. *Human Brain Mapping*, 12:79–93, 2001.
2. V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-Euclidean framework for statistics on diffeomorphisms. In *Proc. of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I*, number 4190 in LNCS, pages 924–931, 2–4 October 2006.
3. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple calculus on tensors in the log-Euclidean framework. In J. Duncan and G. Gerig, editors, *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part I*, volume 3749 of LNCS, pages 115–122, Palm Springs, CA, USA, October 26–29, 2005. Springer Verlag.
4. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2006.
5. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, August 2006.
6. J. Ashburner and K. J. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 2000.
7. G. Aubert and P. Kornprobst. *Mathematical problems in image processing - Partial differential equations and the calculus of variations*, volume 147 of *Applied Mathematical Sciences*. Springer, 2001.
8. P. Basser, J. Mattiello, and D. L. Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.
9. P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53:221–225, 2005.
10. M. Beg, M. Miller, A. Trounev, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. Journal of Computer Vision*, 61(2):139–157, 2005.
11. J. Boisvert, F. Chieriet, X. Pennec, N. Ayache, and H. Labelle. A novel framework for the 3D analysis of spine deformation modes. In *Research into Spinal Deformities*, volume 123 of *Studies in Health Technology and Informatics*, pages 176–182, 2006.
12. J. Boisvert, F. Chieriet, X. Pennec, H. Labelle, and N. Ayache. Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Transactions on Medical Imaging*, 27(4):557–568, 2008.
13. J. Boisvert, X. Pennec, N. Ayache, H. Labelle, and F. Chieriet. 3D anatomic variability assesment of the scoliotic spine using statistics on Lie groups. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, pages 750–753, Crystal Gateway Marriott, Arlington, Virginia, USA, April 2006. IEEE.
14. J. Boisvert, X. Pennec, H. Labelle, F. Chieriet, and N. Ayache. Principal spine shape deformation modes using Riemannian geometry and articulated models. In *Proc of the IV Conference on Articulated Motion and Deformable Objects, Andratx, Mallorca, Spain, 11-14 July*, volume 4069 of LNCS, pages 346–355. Springer, 2006. AMDO best paper award 2006.
15. F. Bookstein. *The Measurement of Biological Shape and Shape Change*, volume 24 of *Lecture Notes in Biomathematics*. Springer-Verlag, 1978.
16. T. Brox, J. Weickert, B. Burgeth, and P. Mrázek. Nonlinear structure tensors. *Image and Vision Computing*, 24(1):41–55, 2006.
17. A. Brun. *Manifolds in Image Science and Visualization*. PhD thesis, Linköping University, 2007. Linköping Studies in Science and Technology Dissertations No 1157.
18. J. Burbea and C. Rao. Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. *Journal of Multivariate Analysis*, 12:575–596, 1982.
19. M. Calvo and J. Oller. An explicit solution of information geodesic equations for the multivariate normal model. *Statistics and Decisions*, 9:119–138, 1991.

20. D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, June 1998.
21. B. Davis, P. Fletcher, E. Bullitt, and S. Joshi. Population shape regression from random design data. In *Proc. of ICCV'07*, 2007.
22. J.-P. Dedieu, G. Malajovich, and P. Priouret. Newton method on Riemannian manifolds: Covariant alpha-theory. *IMA Journal of Numerical Analysis*, 23:395–419, 2003.
23. D. Ducreux, P. Fillard, D. Facon, A. Ozanne, J.-F. Lepeintre, J. Renoux, M. Tadié, and P. Lasjaunias. Diffusion tensor magnetic resonance imaging and fiber tracking in spinal cord lesions: Current and future indications. *Neuroimaging Clinics of North America*, 17(1): 137–147, February 2007.
24. S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Measuring brain variability via sulcal lines registration: a diffeomorphic approach. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 4791 of *LNCS*, pages 675–682, Brisbane, Australia, October 2007. Springer.
25. A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.
26. P. Fillard, V. Arsigny, N. Ayache, and X. Pennec. A Riemannian framework for the processing of tensor-valued images. In O. F. Olsen, L. Florak, and A. Kuijper, editors, *Deep Structure, Singularities, and Computer Vision (DSSCV)*, number 3753 in *LNCS*, pages 112–123. Springer Verlag, June 2005.
27. P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, pages 786–789, Crystal Gateway Marriott, Arlington, Virginia, USA, April 2006.
28. P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. *IEEE Transactions on Medical Imaging*, 26(11):1472–1482, Nov. 2007.
29. P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage*, 34(2):639–650, January 2007.
30. P. Fillard, V. Arsigny, X. Pennec, P. M. Thompson, and N. Ayache. Extrapolation of sparse tensor fields: Application to the modeling of brain variability. In G. Christensen and M. Sonka, editors, *Proc. of Information Processing in Medical Imaging 2005 (IPMI'05)*, volume 3565 of *LNCS*, pages 27–38, Glenwood springs, Colorado, USA, July 2005. Springer.
31. P. Fillard, X. Pennec, P. Thompson, and N. Ayache. Evaluating brain anatomical correlations via canonical correlation analysis of sulcal lines. In *Proc. of MICCAI'07 Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation*, Brisbane, Australia, 2007.
32. P. Fletcher, S. Joshi, C. Lu, and S. Pizer. Gaussian distributions on Lie groups and their application to statistical shape analysis. In C. Taylor and A. Noble, editors, *Proc. of Information Processing in Medical Imaging (IPMI'2003)*, volume 2732 of *LNCS*, pages 450–462. Springer, 2003.
33. P. T. Fletcher and S. C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis, ECCV 2004 Workshops CVAMIA and MMBIA, Prague, Czech Republic, May 15, 2004*, volume 3117 of *LNCS*, pages 87–98. Springer, 2004.
34. M. Fleute and S. Lavallée. Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. In Springer, editor, *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496 of *LNCS*, pages 879–887, 1998.
35. G. Gerig, R. Kikinis, O. Kübler, and F. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–232, June 1992.

36. S. C. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Processing*, 9(8):1357–1370, 2000.
37. H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications in Pure and Applied Mathematics*, 30:509–541, 1977.
38. M. Kendall and P. Moran. *Geometrical probability*. Number 10 in Griffin’s statistical monographs and courses. Charles Griffin & Co. Ltd., 1963.
39. W. Kendall. Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence. *Proc. London Math. Soc.*, 61(2):371–406, 1990.
40. D. Le Bihan, J.-F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriet. Diffusion tensor imaging: Concepts and applications. *Journal Magnetic Resonance Imaging*, 13(4):534–546, 2001.
41. G. Le Goualher, E. Procyk, D. Collins, R. Venugopal, C. Barillot, and A. Evans. Automated extraction and variability analysis of sulcal neuroanatomy. *IEEE Transactions on Medical Imaging*, 18(3):206–217, 1999.
42. C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision*, 25(3):423–444, Oct. 2006.
43. R. Mahony and R. Manton. The geometry of the newton method on non-compact lie groups. *Journal of Global Optimization*, 23:309–327, 2002.
44. J.-F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, D. L. Collins, A. C. Evans, and J. Régis. Object-based morphometry of the cerebral cortex. *IEEE Transactions on Medical Imaging*, 23(8):968–982, Aug. 2004.
45. J.-F. Mangin, D. Rivière, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, P. Scifo, T. Ochiai, F. Brunelle, and J. Régis. A framework to study the cortical folding patterns. *NeuroImage*, 23(Supplement 1):S129–S138, 2004.
46. J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci*, 356:1293–1322, 2001.
47. G. Medioni, M.-S. Lee, and C.-K. Tang. *A Computational Framework for Segmentation and Grouping*. Elsevier, 2000.
48. E. Meijering. A chronology of interpolation: From ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, 90(3):319–342, March 2002.
49. M. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Re-view of Biomedical Engineering*, pages 375–405, 2003.
50. M. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1/2):61–84, 2001.
51. M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 2006.
52. M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal of Matrix Analysis and Applications*, 26(3):735–747, 2005.
53. B. Mohammadi, H. Borouchaki, and P. George. Delaunay mesh generation governed by metric specifications. Part II: applications. *Finite Elements in Analysis and Design*, pages 85–109, 1997.
54. K. Nomizu. Invariant affine connections on homogeneous spaces. *American J. of Math.*, 76: 33–65, 1954.
55. B. Owren and B. Welfert. The newton iteration on Lie groups. *BIT Numerical Mathematics*, 40(1):121–145, 2000.

56. X. Pennec. Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements. In A. Cetin, L. Akarun, A. Ertuzun, M. Gurcan, and Y. Yardimci, editors, *Proc. of Nonlinear Signal and Image Processing (NSIP'99)*, volume 1, pages 194–198, June 20–23, Antalya, Turkey, 1999. IEEE-EURASIP.
57. X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, July 2006. A preliminary appeared as INRIA RR-5093, January 2004.
58. X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.
59. X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. Riemannian elasticity: A statistical regularization framework for non-linear registration. In J. Duncan and G. Gerig, editors, *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part II*, volume 3750 of *LNCS*, pages 943–950, Palm Springs, CA, USA, October 26–29, 2005. Springer Verlag.
60. X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, December 1997.
61. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 12(7):629–639, 1990.
62. H. Poincaré. *Calcul des probabilités*. 2nd edition, Paris, 1912.
63. K. Rajamani, S. Joshi, and M. Styner. Bone model morphing for enhanced surgical visualization. In IEEE, editor, *Proc of IEEE Symp. on Biomedical Imaging: Nano to Macro (ISBI) 2004*, volume 2, pages 1255–1258, Apr. 2004.
64. L. Skovgaard. A Riemannian geometry of the multivariate normal model. *Scand. J. Statistics*, 11:211–223, 1984.
65. G. Subsol, J.-P. Thirion, and N. Ayache. A scheme for automatically building 3D morphometric anatomical atlases: application to a skull atlas. *Medical Image Analysis*, 2(1):37–60, 1998.
66. J. Talairach and P. Tournoux. *Co-Planar Stereotaxic Atlas of the Human Brain: 3-dimensional Proportional System : an Approach to Cerebral Imaging*. Thieme Medical Publishers, New York, 1988.
67. P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758, July 2000.
68. P. Thompson, D. MacDonald, M. Mega, C. Holmes, A. Evans, and A. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21(4):567–581, 1977.
69. P. Thompson, M. Mega, R. Woods, C. Zoumalan, C. Lindshield, R. Blanton, J. Moussai, C. Holmes, J. Cummings, and A. Toga. Cortical change in alzheimer's disease detected with a disease-specific population-based brain atlas. *Cerebral Cortex*, 11(1):1–16, January 2001.
70. A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.
71. M. Vaillant, M. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23(Supp. 1):S161–S169, 2004.
72. M. Vaillant, A. Qiu, J. Glaunès, and M. Miller. Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus. *NeuroImage*, 34(3):1149–1159, 2007.
73. J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix-valued images. In M. Nashed and O. Scherzer, editors, *Inverse Problems, Image Analysis, and Medical Imaging*, volume 313 of *Contemporary Mathematics*, pages 251–268, Providence, 2002. AMS.
74. J. Weickert and H. Hagen, editors. *Visualization and Processing of Tensor Fields*. Mathematics and Visualization. Springer, 2006.

Building Patient-Specific Physical and Physiological Computational Models from Medical Images

H. Delingette and N. Ayache

Abstract We describe a hierarchy of computational models of the human body operating at the geometrical, physical and physiological levels. Those models can be coupled with medical images which play a crucial role in the diagnosis, planning, control and follow-up of therapy. In this paper, we discuss the issue of building patient-specific physical and physiological models from macroscopic observations extracted from medical images. We illustrate the topic of model personalization with concrete examples in brain shift modeling, hepatic surgery simulation, cardiac and tumor growth modeling. We conclude this article with scientific perspectives.

1 Introduction

Computational models of the human body [3] aim at reproducing the geometrical, physical and physiological properties of human organs and systems at various scales (see Fig. 1). This is an emerging and rapidly progressing area of research that is driven by a better understanding of the physical and physiological processes involved but also by more efficient computational tools (either software or hardware) for their realistic numerical simulation. The purpose of this paper is to show how medical imaging plays a growing role in the development of those models. Indeed, medical imaging provides macroscopic observations of the human anatomy and its function for a wide number of patients. It can serve to personalize those computational models, *i.e.* to choose a specific set of parameters that best corresponds to the studied patient.

Before entering the issue of creating patient specific models, it is useful to structure those models into a hierarchy of three main levels [16], the knowledge

H. Delingette (✉) • N. Ayache
Inria, 2004 Route des Lucioles, Sophia Antipolis 06600, France
e-mail: Herve.Delingette@inria.fr

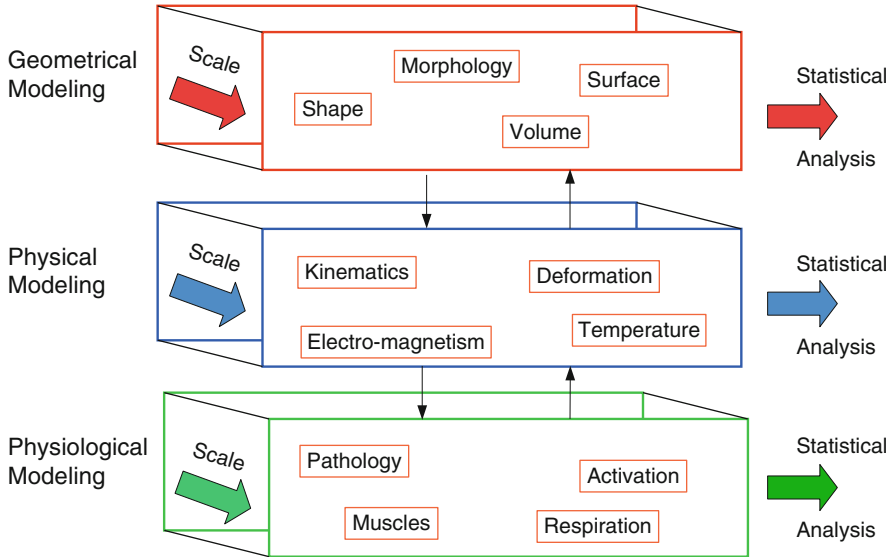


Fig. 1 Hierarchy of computational models of the human body (from [15])

of the lower level being required for the implementation of the upper level. The first level is mainly *geometrical* and addresses the construction of digital static descriptions of the anatomy, often based on medical imagery. The techniques for segmenting and reconstructing anatomical and pathological structures from medical images have been developed for the past 15 years and have brought many advances in several medical fields including computer-aided diagnosis, therapy planning, image-guided interventions, drug delivery, *etc.* A distinctive achievement of computational anatomy has been carried out by the “Visible Human Project” [1] which provided the first digital multimodal anatomical representation of the full human body.

A second level of modeling describes the *physical* properties of the human body, involving for instance the biomechanical behavior of various tissues, organs, vessels, muscles or bone structures [22].

A third level of modeling describes the functions of the major biological systems [24, 38] (e.g. cardiovascular [5, 44], respiratory [25], digestive, hormonal, muscular, central or peripheral nervous system, *etc.*) or some pathological metabolism (e.g. evolution of inflammatory or cancerous lesions [48], formation of vessel stenoses [7, 42], *etc.*). Such physiological models often include reactive mechanisms while physical models only provide a passive description of tissues and structures.

There is an additional dimension associated with each level: the scale at which the anatomical, physical or physiological structure is described. With the development of new imaging modalities, it is now possible to observe the shape or function of most structures at the macroscopic (tissue), microscopic (cellular) levels and

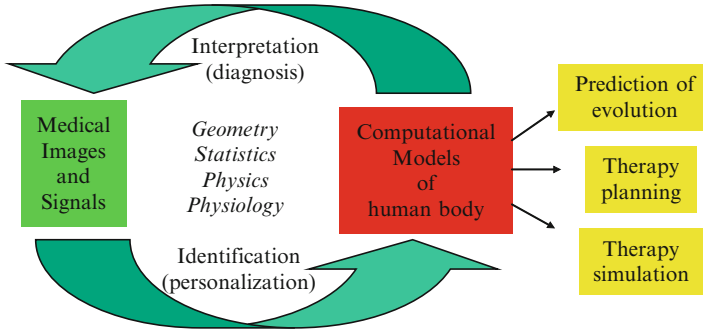


Fig. 2 This diagram shows how computational models can be coupled with medical images and signals to build personalized models and use them in clinical applications [4]

even in such cases to reveal the metabolic activity at the nanoscopic (molecular) scale. Coupled with those multiscale observations are new generations of multiscale computational models [24, 38]

Furthermore, each model is specified by a number of parameters (*e.g.* material stiffness or electrical conductivity for a physical model) and a related task consists in finding a set of those parameters that produces the best agreement between the simulated processes (deformation, activation,...) and the observed ones. The techniques consisting in finding the patient-specific parameters of a dynamic model typically require to solve an inverse problem and are sometimes called “data assimilation” techniques in the field of oceanography or climatology. As illustrated in Fig. 2, the personalization of a model from medical images or signals is often a requirement for using its predictive power in clinical applications such as therapy planning or simulation. The personalized models may also be used as advanced image processing tools that can provide a decision support system with additional physical or physical parameters relevant for establishing a diagnosis.

Finally, the ability to recover patient-specific parameters leads to the variability study of those parameters across a given population. Thus, statistical modeling of those computational models can be seen as an orthogonal modeling activity that aims for instance at finding the local or global similarities and dissimilarities of a structure or a function between two populations [18, 34, 49]. Statistical findings may also be used to calibrate, refine or constrain [11] a given model. At the basis of this activity is the growing availability of large databases of subjects and patients including biomedical signals and images as well as genetic information.

In the next sections, following the proposed hierarchy, we describe a number of practical cases involving the personalization of computational models before proposing some perspectives and challenges for future research.

2 Patient Specific Biomechanical/Physical modeling

Anatomical models only provide a static geometrical representation of the patient anatomy and do not take into account the deformation of soft tissue that may occur before or during therapy. To address this issue, it is necessary to add a biomechanical model that can estimate soft tissue deformations under the application of known forces or displacements. The additional complexity of modeling may be used to improve the pre-operative planning of a therapy [23, 43] or to provide advanced surgical gesture training.

Personalization of biomechanical models mainly involves the estimation of material parameters and boundary conditions from a time series of images, before and after deformation. One should stress the importance of specifying the proper boundary conditions (constraining the displacement of some nodes at the boundary of the structure) when modeling soft tissue deformation since it has a key effect of the displacement of neighboring nodes.

We briefly discuss below two limit cases for the personalization of biomechanical models: small deformations involved in brain-shift during neurosurgery and large deformations involved in liver surgery simulation.

2.1 *Brain-Shift modeling in neurosurgery*

The brain shift that occurs during a neuro-surgical intervention is the main source of intra-operative localization inaccuracies of pathologies (cerebral tumors, . . .). Indeed, a neurosurgeon establishes the surgical plan based on a pre-operative MR image: any non-rigid motion of the brain between the pre-operative and the intra-operative configuration may lead to an error in the localization of the target. To model the brain motion after opening the dura, a number of authors [17, 32] have made the hypothesis that the loss of cerebro-spinal fluid causes a pressure field along the gravity direction (Archimedes principle). Furthermore, anatomical constraints (falx cerebri, skull) of the deformation field can be enforced with a biomechanical model of the brain discretized as a tetrahedral mesh since the relevant anatomical information can be extracted from MR images and enforced on the mesh.

This clinical problem has motivated the study of the biomechanical behavior of the brain. For instance, Miller [33] has proposed a rheological model for swine brains valid for large displacements. However, in most cases, authors have relied on linear elastic models to extrapolate displacement fields [17, 47] from the cortex surface. Similarly, partial validation of brain shift models has been carried out [10] with a linear elastic model by comparing computed displacements with those observed from intra-operative MR images.

The fairly good predictive power of those simplified models show that a quasi-incompressible (Poisson ratio close to 0.5) linear elastic model is a good choice for simulating the small displacements induced by the brain shift. This is a

sensible result since any non-linear material can be approximated as a linear elastic material for sufficiently small displacements. Another important point that makes the personalization of those biomechanical models less difficult is the fact they are often used to predict displacements from given imposed displacements. In such cases, the knowledge of the Young Modulus is irrelevant and only the Poisson ratio and the boundary conditions must be chosen properly.

2.2 *Hepatic Surgery Simulation*

Surgery simulation aims at reproducing the visual and haptic senses experienced by a surgeon during a surgical procedure, through the use of computer and robotics systems. The medical scope of this technology is linked with the development of minimally invasive techniques especially videoscopic surgery (endoscopy, laparoscopy,...) and possibly telesurgery.

By creating patient-specific models, surgery simulation allows surgeons to verify, optimize and rehearse the surgical strategy of a procedure on a given patient.

However, an important issue for patient-specific soft-tissue models is the estimation of their material properties. Such parameters may be the Young Modulus and Poisson ratios for linear elastic materials or other stiffness parameters for general hyperelastic materials. There are three different sources of rheological data to estimate those parameters: *ex-vivo testing* where a sample of a tissue is positioned inside a testing rig [39]; *in-vivo testing* where a specific force and position sensing device is introduced inside the abdomen to perform indentation [8,37]; *Image-based elastometry* from ultrasound, Magnetic Resonance Elastometry [26, 31] or CT-scan imaging.

There is no consensus on which method is best suited to recover meaningful material parameters, each one having its limitation. For instance *ex-vivo testing* may not be relevant because of the swelling or drying of the tissue. *In-vivo* experiments should also be considered with caution because the response may be location-dependent (linked to specific boundary conditions or non-homogeneity of the material) and the influence of the loading tool caliper on the deformation may not be well understood. Finally elastometry commonly assumes that the tissue undergoes small displacements and need to be thoroughly calibrated.

Thus, when assessing the mechanical parameters of the liver, several authors have reported widely varying parameters [15]. It is especially difficult in the case of the liver because it undergoes large displacements and its perfusion affects deeply its rheology (the liver receives one fifth of the total blood flow at any time). In fact, trying to estimate the liver Young Modulus is prone to large errors because the liver response largely depends on the speed at which the pressure was applied. One can expect to obtain meaningful material parameters only if this material is highly viscoelastic such as the one proposed by Kerdork *et al.* [27].



Fig. 3 (Left) View of the simulated hepatic resection involving linear-elastic materials (from [12, 13, 15] and [20, 21]); (Right) A force feedback system suited for surgery simulation

Furthermore, in a surgical simulator there are strict real-time constraints for simulating soft tissue deformations. For instance, the required refresh rate for user interaction is 30 Hz for visual feedback and more than 500 Hz for force-feedback, although the latter constraint can be alleviated with the addition of a local haptic model [20]. This implies that specific optimisations of soft tissue deformation must be devised either in the form of precomputations [15], or dedicated data structures [40] or multigrid [14] algorithms.

Thus, providing patient-specific soft tissue deformation suitable for surgery simulation remains largely an open problem that encompasses two separate issues. The former one corresponds to the proposition of a physically realistic constitutive material (with viscoelastic and non-linear behaviours) that is suitable for real-time computation. The latter one consists in finding reliable and non-invasive ways to estimate those material parameters from in-vivo data.

3 Patient Specific Physiological Modeling

To model the active properties of living tissues and the dynamic nature of normal or pathological evolving processes, it is necessary to introduce physiological models of the human body. We illustrate the personalization of those models with two examples related to the modeling of the electro-mechanical activity of the heart and the growth of brain tumors.

3.1 Cardiac modeling

During the past 15 years, a scientific INRIA consortium¹ has developed an electro-mechanical model of the cardiac ventricles for medical purposes. The model reproduces the electrical depolarization and repolarization of the cardiac tissues through a set of macroscopic reaction-diffusion equations initially proposed by Fitzugh and Nagumo [19] and further refined by Aliev and Panvilov [2].

$$\begin{aligned}\varepsilon^2 \partial_t u &= \varepsilon \operatorname{div}(D \nabla(u)) + ku(1-u)(u-a) - uz \\ \partial_t z &= -(ku(u-a-1) + z)\end{aligned}\quad (1)$$

where u is the action potential averaged within a volume element of the cardiac tissue, z is the repolarization variable, D is the electrical conductivity tensor, ε is a numerical constant, k controls the electrical reaction and a controls the action potential duration.

The electrical activity can be synchronized with the actual ECG (electrocardiogram) of the patient and creates a mechanical contraction followed by an active relaxation which are modeled by a set of partial differential equations initially proposed by Bestel, Clément and Sorine [6]. The average direction of the myocardium fibers is also integrated into this model (for instance through the conductivity tensor D), since it plays an important role in the anisotropic propagation of both the electrical and mechanical excitations.

This electromechanical model of the heart includes several parameters related to the heart anatomy (fiber orientation), electrophysiology (electrical conductivity), blood flow (preload and afterload) or cardiac mechanics (passive stiffness, contractility). It was shown in [44] that this model could be interactively adjusted to the actual geometrical, mechanical or electrical properties of patient's heart through the use of conventional or tagged MR images and some *in vivo* electrophysiological measurements.

However it is important to make the personalization as automatic as possible in order to predict in the most objective way the effect of a therapy or the evolution of a pathology [45, 46]. This parameter estimation is typically an inverse problem that can be formulated as the minimization of a functional measuring the discrepancy between simulated and observed quantities: the “best” set of parameters specific to a given patient observation is the one that minimizes the functional. Not all parameters can be identified however since several combinations of parameters may lead to the same simulation. For instance, in Eq. (1), the speed of the depolarization propagation is governed by the product between the conductivity and the reaction term: $D \times k$.

¹cf. CardioSense3D URL <http://www-sop.inria.fr/CardioSense3D/>

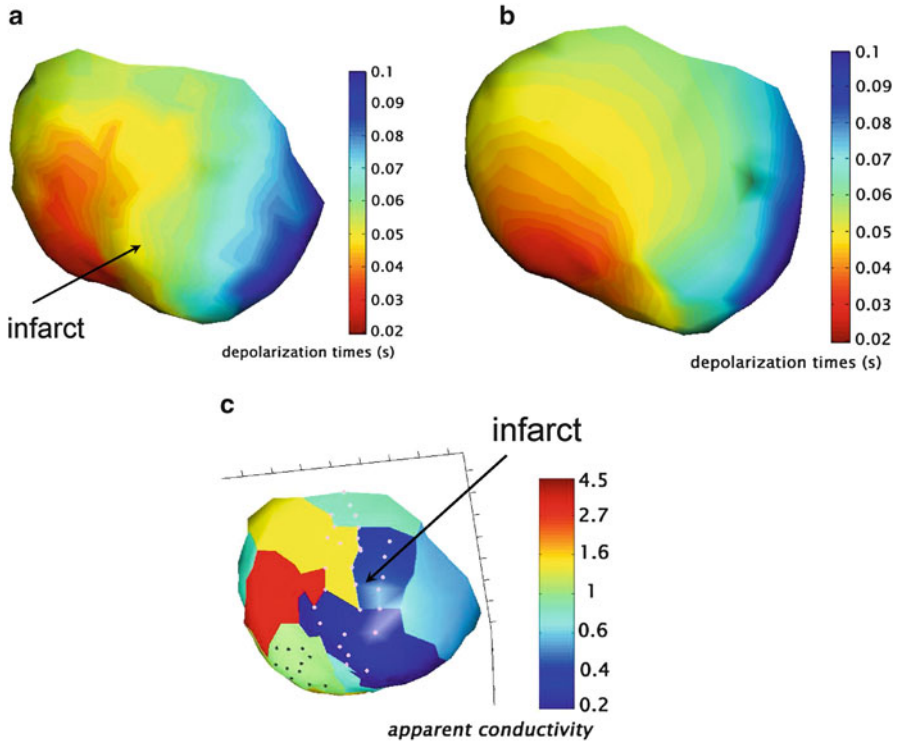


Fig. 4 (a) Isochrone map of the epicardium measured on a canine heart. The location of the infarcted zone is shown; (b) Simulated isochrones after the estimation of global and regional parameters; (c) Map of the apparent conductivity, the conductivity being constant for each region. There exists a good correlation between regions of low conductivity and infarcted regions [36]

This problem of identifiability has been tackled in [36] by estimating *apparent conductivities* instead of physical ones from isochrones maps (maps of the epicardium describing the time at which the electrical signal reaches a given point). This is performed by estimating a global value of the reaction term D while finding regional variations of the electrical conductivities (see Fig. 4) around a reference value. Minimizing the difference between simulated and observed quantities is numerically challenging because the gradient of the functional cannot be computed in a closed form. Instead, various numerical schemes can be used to avoid an exhaustive search of parameters. For instance, in [36], the causality of the propagation was assumed to break a multivariate optimisation into a series of one dimensional optimization of each region conductivity. Furthermore, the Brent algorithm allowed to minimize the functional without estimating its gradient.

To build patient specific models of the cardiac mechanics and estimate the contractility of the myocardium, data assimilation techniques based on variational [45] or sequential [35] approaches can be used assuming for instance that the speed

of material points are known from the analysis of time series of medical images. A common difficulty in those approaches lies in the large size of the state vector and associated covariance matrices due to the addition of the parameters to be estimated.

3.2 Tumor growth

The second physiological model is related to the modeling of the growth of brain tumors. A joint action between INRIA, a Nice Hospital (Centre Antoine Lacassagne), the Brigham and Women's hospital and MIT (Boston) has led to the development of a three stage model [9].

The first stage includes the geometrical model of a patient's head, including the skull, the brain parenchyma (grey and white matter) and the Cerebro-Spinal Fluid (CSF). The shape of each region is acquired through a conventional MR exam. In addition, the direction of the main white matter fibers is also acquired either through Diffusion Tensor Imaging (DTI), or using average directions provided by a brain atlas. A second stage includes the modeling of the biomechanical properties of the brain tissues. Because we are considering small deformations only, this model is linear elastic and implements the boundary conditions imposed by the bony (skull) and fluid (ventricles) structures surrounding the brain parenchyma. The third stage is an evolution model of the tumoral tissues, which is based on a set of reaction-diffusion equations describing the proliferation and diffusion of malignant cells [48].

An original point is the coupling of the third level with the previous two levels in order to create a realistic deformation of the brain tissues (also called *mass effect*) produced by the tumor growth. The direction of the white matter fibers plays an important role in the highly anisotropic diffusion of cancerous cells (see Fig. 5).

In order to predict the tumor evolution beyond the time of image acquisition, the parameters controlling the tumor growth must be estimated from T2 weighted magnetic resonance images (MRI). Those patient-specific parameters include the proliferation rate ρ , the diffusion coefficients d_g and d_w in the grey and white matter respectively. However the proliferation rate cannot be determined independently from the diffusion coefficients by just looking at the motion of the tumor front. That is why our work in tumor growth parameter identification [28] has focused on the estimation of apparent diffusion coefficients assuming a known proliferation rate. In order to speed-up the computation, the tumor growth is simulated with an anisotropic fast marching method [29] based on an Eikonal approximation of the reaction-diffusion equation. The patient-specific parameters are estimated [28] by minimizing the symmetric distance between the segmented tumor front and the simulated one (see Fig. 5). The minimization is performed by using an unconstrained optimization algorithm [41] that does not require derivatives of the objective function. The average symmetric distance between the two fronts are less than $0.1mm$ which seem to imply that the tumor growth model captures reasonably well the physiological process.

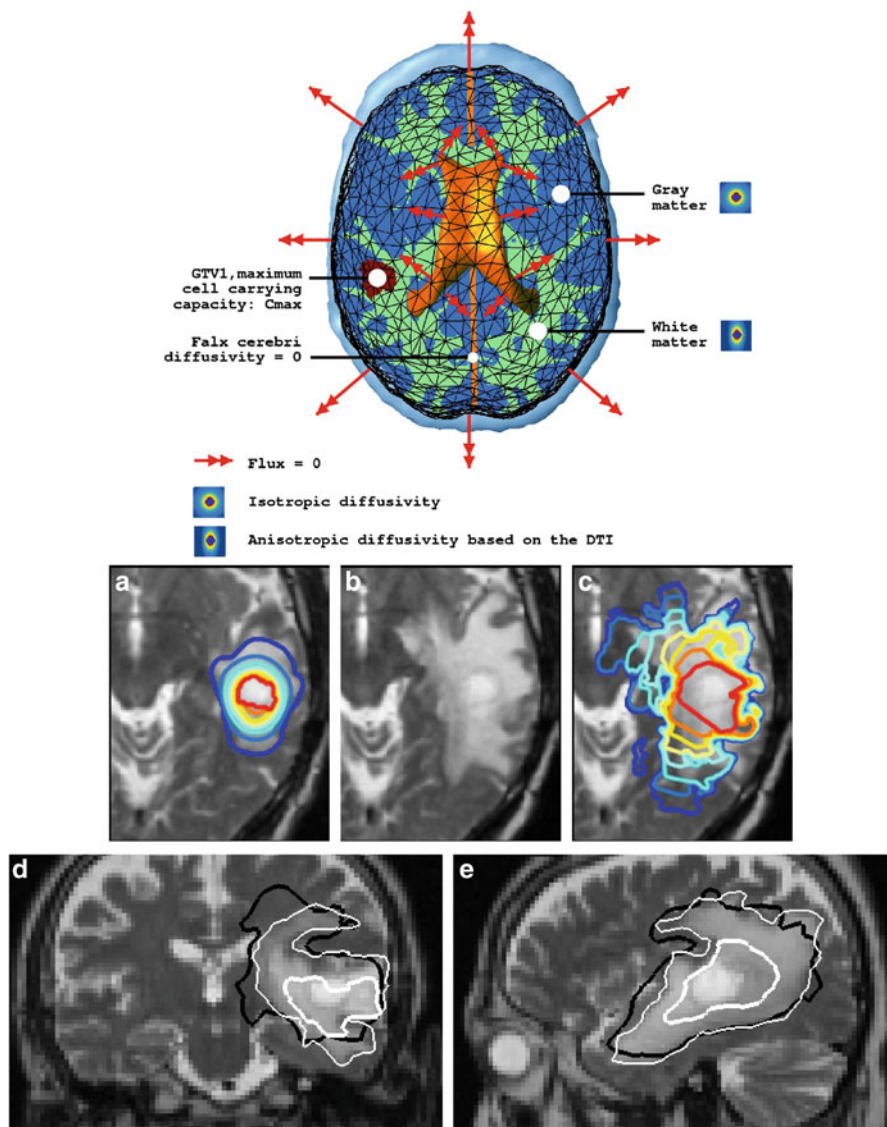


Fig. 5 (Top) Overview of the evolution model of cancerous cells which takes into account the anisotropic diffusion along white matter fibers; (Bottom) Result of the tumor growth simulation on a brain slice; (a) Initial MR T2 image of the patient with lines of constant tumor density; (b) View of the corresponding MR T2 slice (after rigid registration) six months later; (c) Lines of constant tumor density predicted by the tumor growth model [9] after 6 months of evolution. (d) and (e) Patient specific tumor growth modeling in coronal and sagittal views. Thick white contours correspond to the segmented initial tumor front while the thin white contours are the segmented front 270 days later. Thick black contours are the simulated front after optimizing the diffusion coefficients : $d_w = 0.55mm^2/day$, $d_g = 2.7 \times 10^{-3}mm^2/day$ (from [28])

4 Summary and Perspectives

We have shown in this article some examples of patient-specific physical and physiological models resulting from their coupling with medical images. In fact those models should be seen as a new generation of medical image analysis tools that go beyond the geometrical principles used in the already existing tools (image segmentation, image registration). An important aspect is the possibility offered by these personalized models to actually fuse the geometrical, physical, and physiological information necessary to provide a thorough and reliable analysis of the complex and multimodal biomedical signals acquired on each patient, possibly at various scales.

We list below some research topics that should open new perspectives in combining physical and physiological models with medical images:

- *Soft Tissue Modeling.* The liver example presented in Sect. 2.2 illustrates the need to develop sophisticated *ex-vivo* and *in-vivo* indentation devices in order to provide a better understanding and new mathematical models of the mechanical behavior of human organs. Without developing realistic macroscopic biomechanical models, it is impossible to reliably recover patient specific material parameters useful for surgery simulation or therapy planning.
- *Respiratory and cardiac modeling.* In addition to providing better understanding of pulmonary and cardio-vascular diseases, patient-specific models of the lungs and the heart can be helpful to compensate the motion artefacts created when imaging the thorax or the abdomen.
- *Parameter estimation.* Inverting realistic computational models from observations still remains an open problem due to the issue of identifiability and the curse of dimensionality associated with sequential or variational approaches. A reasonable objective for parameter identification would be to provide a probability density function associated with each parameter in order to take into account the uncertainty of the inversion.
- *Statistical Analysis.* The development of large databases of medical images should further improve the robustness and accuracy of the previously discussed computational models, and therefore the performances of image-guided intervention or simulation systems.
- *Microscopic Imaging.* New *in vivo* microendoscopy techniques [30,50] providing structural and functional information on the tissues at the cellular level should also open new avenues for building patient specific models.

References

1. M. J. Ackerman. The visible human project. *Proceedings of the IEEE : Special Issue on Surgery Simulation*, 86(3):504–511, Mar. 1998.
2. R. Aliev and A. Panfilov. A simple two-variable model of cardiac excitation. *Chaos, Solitons & Fractals*, 7(3):293–301, 1996.

3. N. Ayache, editor. *Computational Models for the Human Body*. Handbook of Numerical Analysis (Ph. Ciarlet series editor). Elsevier, 2004. 670 pages.
4. N. Ayache, O. Clatz, H. Delingette, G. Malandain, X. Pennec, and M. Sermesant. Asclepios: a research project-team at inria for the analysis and simulation of biomedical images. In *From semantics to computer science: essays in honor of Gilles Kahn*. Cambridge University Press, 2008.
5. M. Belik, T. Usyk, and A. McCulloch. Computational methods for cardiac electrophysiology. In N. Ayache, editor, *Computational Models for the Human Body*, pages 129–187. Elsevier, 2004.
6. J. Bestel, F. Clément, and M. Sorine. A biomechanical model of muscle contraction. In *Proc. of MICCAI'01*, volume 2208, pages 1159–1161. Springer, 2001.
7. J.-D. Boissonnat, R. Chaine, P. Frey, G. Malandain, F. Nicoud, S. Salmon, E. Saltel, and M. Thiriet. From arteriographies to computational flow in saccular aneurisms: the INRIA experience. *Medical Image Analysis*, 9(2):133–143, Apr. 2005.
8. T. Chanthasopeephan, J. P. Desai, and A. Lau. Measuring forces in liver cutting: New equipment and experimental results. *Annals of Biomedical Engineering*, 31(11):1372–1382, 2003.
9. O. Clatz, P. Bondiau, H. Delingette, G. Malandain, M. Sermesant, S. K. Warfield, and N. Ayache. In silico tumor growth: Application to glioblastomas. In *Proc. of MICCAI 2004*, volume 3217 of *LNCS*, pages 337–345. Springer Verlag, September 2004.
10. O. Clatz, H. Delingette, E. Bardinet, D. Dormont, and N. Ayache. Patient specific biomechanical model of the brain: Application to parkinson's disease procedure. In N. Ayache and H. Delingette, editors, *International Symposium on Surgery Simulation and Soft Tissue Modeling (IS4TM'03)*, volume 2673, pages 321–331. Springer-Verlag, 2003.
11. T. Cootes, C. Taylor, A. Lanitis, D. Cooper, and J. Graham. Building and using flexible models incorporating grey-level information. In *Proc. of the Int. Conf. on Computer Vision (ICCV'93)*, pages 242–245, 1993.
12. S. Cotin, H. Delingette, and N. Ayache. Real-time elastic deformations of soft tissues for surgery simulation. *IEEE Transactions On Visualization and Computer Graphics*, 5(1):62–73, January-March 1999.
13. S. Cotin, H. Delingette, and N. Ayache. A hybrid elastic model allowing real-time cutting, deformations and force-feedback for surgery training and simulation. *The Visual Computer*, 16(8):437–452, 2000.
14. G. DeBunne, M. Desbrun, M.-P. Cani, and A. H. Barr. Dynamic real-time deformations using space and time adaptive sampling. *Computer Graphics Proceedings*, Aug 2001. Proceedings of SIGGRAPH'01.
15. H. Delingette and N. Ayache. Soft tissue modeling for surgery simulation. In N. Ayache, editor, *Computational Models for the Human Body*, Handbook of Numerical Analysis (Ed : Ph. Ciarlet), pages 453–550. Elsevier, 2004.
16. H. Delingette, X. Pennec, L. Soler, J. Marescaux, and N. Ayache. Computational models for image guided, robot-assisted and simulated medical interventions. *Proceedings of the IEEE*, 94(9):1678– 1688, September 2006.
17. M. Ferrant, A. Nabavi, B. Macq, P. Black, F. Jolesz, R. Kikinis, and S. Warfield. Serial registration of intraoperative MR images of the brain. *Medical Image Analysis*, 2002.
18. P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage*, 34(2):639–650, January 2007.
19. R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1:445–466, 1961.
20. C. Forest, H. Delingette, and N. Ayache. Surface contact and reaction force models for laparoscopic simulation. In *International Symposium on Medical Simulation*, volume 3078 of *LNCS*, pages 168–176. Springer-Verlag, June 2004.
21. C. Forest., H. Delingette, and N. Ayache. Removing tetrahedra from manifold tetrahedralisation : application to real-time surgical simulation. *Medical Image Analysis*, 9(2):113–122, Apr. 2005.

22. E. Haug, H.-Y. Choi, S. Robin, and M. Beaugonin. Human models for crash and impact simulation. In N. Ayache, editor, *Computational Models for the Human Body*, pages 231–452. Elsevier, 2004.
23. D. Hawkes, D. Barratt, J. Blackall, C. Chan, P. Edwards, K. Rhode, G. Penney, J. McClelland, and D. Hill. Tissue deformation and shape models in image-guided interventions: a discussion paper. *Medical Image Analysis*, 9(2):163–175, Apr. 2005.
24. P. Hunter and T. Borg. Integration from proteins to organs: the Physiome project. *Nature Reviews Molecular Cell Biology*, 4:237–243, 2003.
25. J. Kaye, F. Primiano, and D. Metaxas. A 3D virtual environment for modeling mechanical cardiopulmonary interactions. *Medical Image Analysis*, 2(2):1–26, 1997.
26. A. E. Kerdok, S. M. Cotin, M. P. Ottensmeyer, A. M. Galea, R. D. Howe, and S. L. Dawson. Truth Cube: Establishing Physical Standards for Real Time Soft Tissue Simulation. *Medical Image Analysis*, 7:283–291, 2003.
27. A. E. Kerdok, M. P. Ottensmeyer, and R. D. Howe. Effects of perfusion on the viscoelastic characteristics of liver. *Journal of Biomechanics*, 39(12):2221–2231, 2006.
28. E. Konukoglu, O. Clatz, P.-Y. Bondiaou, M. Sermesant, H. Delingette, and N. Ayache. Towards an identification of tumor growth parameters from time series of images. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 4791 of *LNCS*, pages 549–556, Brisbane, Australia, October 2007. Springer.
29. E. Konukoglu, M. Sermesant, O. Clatz, J.-M. Peyrat, H. Delingette, and N. Ayache. A recursive anisotropic fast marching approach to reaction diffusion equation: Application to tumor growth modeling. In *Proceedings of the 20th International Conference on Information Processing in Medical Imaging (IPMI'07)*, volume 4584 of *LNCS*, pages 686–699, 2-6 July 2007.
30. C. MacAulay, P. Lane, and R. Richards-Kortum. In vivo pathology: microendoscopy as a new endoscopic imaging modality. *Gastrointestinal Endoscopy Clinics of North America*, 14: 595–620, 2004.
31. A. Manduca, T. E. Oliphant, M. A. Dresner, J. L. Mahowald, S. A. Kruse, E. Amromin, J. P. Felmlee, J. F. Greenleaf, and R. L. Ehman. Magnetic resonance elastography: Non-invasive mapping of tissue elasticity. *Medical Image Analysis*, 5(4):237–254, Dec. 2001.
32. M. Miga, K. Paulsen, J. Lemry, F. Kennedy, S. Eisner, A. Hartov, and D. Roberts. Model-updated image guidance: Initial clinical experience with gravity-induced brain deformation. *IEEE Transactions on Medical Imaging*, 18(10):866–874, 1999.
33. K. Miller. Constitutive modelling of abdominal organs. *Journal of Biomechanics*, 33(3): 367–373, 2000.
34. M. I. Miller. Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage*, 23(Supplement 1):S19–S33, 2004. Special Issue : Mathematics in Brain Imaging.
35. P. Moireau, D. Chapelle, and L. T. P. Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 197:659–677, 2008.
36. V. Moreau-Villéger, H. Delingette, M. Sermesant, H. Ashikaga, O. Faris, E. McVeigh, and N. Ayache. Building maps of local apparent conductivity of the epicardium with a 2D electrophysiological model of the heart. *IEEE Transactions on Biomedical Engineering*, 53(8):1457–1466, Aug. 2006.
37. A. Nava, E. Mazza, F. Kleineremann, N. Avis, and J. McClure. Determination of the mechanical properties of soft human tissues through aspiration experiments. In *Proc. of Conference on Medical Robotics, Imaging And Computer Assisted Surgery: MICCAI 2003*, LNCS, Montreal, Canada, Nov. 2003.
38. D. Noble. Modeling the Heart, from genes to cells to the whole organ. *Science*, 295:1678–1682, 2002.
39. M. P. Ottensmeyer, A. E. Kerdok, R. D. Howe, and S. L. Dawson. The effects of testing environment on the viscoelastic properties of soft tissues. In *International Symposium on Medical Simulation*, pages 9–18, June 2004.

40. G. Picinbono, H. Delingette, and N. Ayache. Non-Linear Anisotropic Elasticity for Real-Time Surgery Simulation. *Graphical Models*, 65(5):305–321, Sept. 2003.
41. M. Powell. Uobyqa : unconstrained optimization by quadratic approximation. *Mathematical Programming*, 92(3):555–582, May 2002.
42. A. Quarteroni and L. Formaggia. Mathematical modeling and numerical simulation of the cardiovascular system. In N. Ayache, editor, *Computational Models for the Human Body*, pages 3–128. Elsevier, 2004.
43. J. Schnabel, C. Tanner, A. Castellano-Smith, A. Degenhard, M. Leach, D. Hose, D. Hill, and D. Hawkes. Validation of non-rigid image registration using finite element methods: application to breast MR images. *IEEE Trans. Medical Imaging*, 22(2):238–247, 2003.
44. M. Sermesant, H. Delingette, and N. Ayache. An electromechanical model of the heart for image analysis and simulation. *IEEE Transactions on Medical Imaging*, 25(5):612–625, 2006.
45. M. Sermesant, P. Moireau, O. Camara, J. Sainte-Marie, R. Andriantsimiavona, R. Cimrman, D. G. Hill, D. Chapelle, and R. Razavi. Cardiac function estimation from mri using a heart model and data assimilation: Advances and difficulties. In *Functional Imaging and Modeling of the Heart (FIMH'05)*, pages 325–337, 2005.
46. M. Sermesant, K. Rhode, A. Anjorin, S. Hedge, G. Sanchez-Ortiz, D. Rueckert, P. Lambiase, C. Bucknall, D. Hill, and R. Razavi. Simulation of the electromechanical activity of the heart using xmr interventional imaging. In *Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery: MICCAI 2004*, pages 786–794, Oct. 2004.
47. O. Skrinjar, A. Nabavi, and J. Duncan. Model-driven brain shift compensation. *Medical Image Analysis*, 6(4):361–373, 2002.
48. K. Swanson, E. Alvord, and J. Murray. Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *British Journal of Cancer*, 86(1):14–18, 2002.
49. P. M. Thompson, M. I. Miller, J. T. Ratnanather, R. A. Poldrack, and T. E. Nichols. Guest Editorial. *NeuroImage*, 23(Supplement 1):S1, 2004. Special Issue : Mathematics in Brain Imaging.
50. T. Vercauteren, N. Ayache, N. Savoie, G. Malandain, and A. Perchant. Processing of in vivo fibered confocal microscopy video sequences. In J. Rittscher, R. Machiraju, and S. T. C. Wong, editors, *Microscopic Image Analysis for Life Science Applications*. Chapter 19, pages 441–463, Artech House, 2008.

Constructing a Patient-Specific Model Heart from CT Data

D.M. McQueen, T. O'Donnell, B.E. Griffith, and C.S. Peskin

Abstract The goal of our work is to predict the patterns of blood flow in a model of the human heart using the Immersed Boundary method. In this method, fluid is moved by forces associated with the deformation of flexible boundaries which are immersed in, and interacting with, the fluid. In the present work the boundary is comprised of the muscular walls and valve leaflets of the heart. The method benefits by having an anatomically correct model of the heart. This report describes the construction of a model based on CT data from a particular individual, opening up the possibility of simulating interventions in an individual for clinical purposes.

D.M. McQueen (✉)

Department of Mathematics, University of North Carolina at Chapel Hill, Phillips Hall, Campus Box 3250, University of North Carolina, Chapel Hill, NC 27599

e-mail: mcqueen@cims.nyu.edu

T. O'Donnell

Siemens Medical Solutions, Malvern, PA

e-mail: tom.odonnell@siemens.com

B.E. Griffith

Leon H. Charney Division of Cardiology, Department of Medicine, New York University School of Medicine, 522 First Avenue, New York, NY 10016, USA

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

e-mail: griffith@cims.nyu.edu

C.S. Peskin

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

e-mail: peskin@cims.nyu.edu

1 Introduction

We wish to compute blood flow in the chambers of a model of the human heart. For spatial scales on the order of the sizes of the chambers of the heart or the great vessels, the motion of blood is well-described by the Navier-Stokes equations. Solution of the Navier-Stokes equations requires specifying conditions on the boundary. This can be a challenging requirement in the computation of blood flow within the chambers of the heart. The valve leaflets comprise an important part of the boundary, and the motion of those leaflets cannot be specified in advance. The valve leaflets and the surrounding fluid (blood) form a coupled system; the motion of the leaflets and the motion of the blood must be computed simultaneously. Although less obviously so, this is equally true for the motion of the heart walls.

We have developed a numerical method, the Immersed Boundary (IB) method, which simultaneously computes the motion of a fluid and the motion of an elastic boundary immersed in, and interacting with, that fluid. In this method, the fluid is represented by Eulerian velocities and pressures which are stored on a regular three-dimensional computational lattice. The boundary is represented by elastic structures which are free to move continuously in the space sampled by the computational lattice. The essence of the method is to replace the elastic boundary by the forces which result from its deformations. These forces are applied to the lattice in the neighborhood of the elastic boundary with the aid of a numerical approximation to the Dirac delta function. The fluid moves under the action of this body force field. The numerical delta function is then used again, to interpolate the newly computed lattice velocities to the locations of the boundary, and then the boundary is moved at the interpolated velocity to a new location (this is the no-slip condition). The process of calculating forces, computing fluid motion and moving the boundary is repeated cyclically in a time-stepping procedure with a suitably chosen time step. Neither the fluid motion nor the boundary motion is an input: both motions are outputs. The inputs are physical properties of the fluid, the elastic properties of the boundary (which may be time-dependent), and the initial geometry of the boundary. A systematic description of the IB method can be found in [4], and its application to the heart is illustrated in [2].

The myocardial muscle fibers supply the force which moves the heart and the blood. Computation of blood flow within the chambers of the heart by the IB method would benefit by the availability of an anatomically correct model of the cardiac muscle fibers. We have previously described a somewhat idealized model of the fiber anatomy of the heart [1], strongly influenced by the dissections of Thomas [8]. In the idealized model the heart at end-systole was comprised of a collection of conical surfaces; muscle fibers were represented by geodesic paths on these surfaces, each geodesic beginning on one valve ring and ending on a (possibly different) valve ring. Representing muscle fibers by geodesics was motivated by the observations of Streeter, et al. [6]. The use of conical surfaces makes computation of the geodesics straightforward, since a conical surface can be unrolled onto a plane, but it also gives rise to some unnatural anatomical shapes, most notably in the neighborhood of the apex and in the neighborhood of the valve rings. Since the flow patterns in

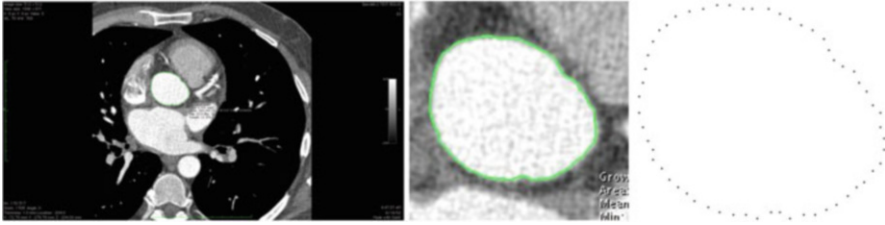


Fig. 1 CT data from a patient with congestive heart failure visualized using the Osirix DICOM viewer. The middle and right panels show the segmentation of the ascending aorta (enlarged)

the chambers of the heart are influenced by the shapes of the chambers, it would be unrealistic to use the conical-surface model to study the blood flow in any particular human.

A major (and ambitious) goal of our research is to provide a tool that would permit a cardiologist to replicate a patient's disease state in a computer model and to study how a proposed intervention, such as surgery, changes the behavior of the model, as a guide to how that intervention might change the behavior of the patient's heart. A vital component of such a tool would be an accurate model of the patient's cardiac anatomy, that is, a patient-specific heart model. In the following we describe the techniques by which we produce a computer model of the heart based on measurements from a particular individual.

2 Computed Tomography Data

The starting point of our construction is a computed tomography (CT) data set from a patient with congestive heart failure, obtained by Arthur E. Stillman and Randolph M. Setser at the Cleveland Clinic. A sample plane slice of this data set is shown in the left panel of Fig. 1. The data set consists of about 300 such slices, 0.5 mm apart. Cross-sections of the heart and the nearby great vessels from this data set are segmented by hand. The right panel of Fig. 1 shows a sample segmentation of the ascending aorta.

For technical reasons, the IB method requires that neighboring points on the boundary be spaced apart no more than $1/2$ of the computational lattice meshwidth in order for the boundary not to leak. Consequently, the cross-section perimeters produced by segmentation are re-discretized so that interpoint distances are slightly less than $1/2$ of the intended lattice meshwidth and are uniform around the perimeter. All subsequent references here to discretized boundaries should be understood to mean boundaries re-discretized in this way.

Figure 2(upper) shows oblique views of two neighboring cross-sections of the ascending aorta just below the aortic arch. The aortic surface between these two perimeters can be defined by triangulation. We find a good triangulation by the following heuristic procedure. On each perimeter choose any point to be the first

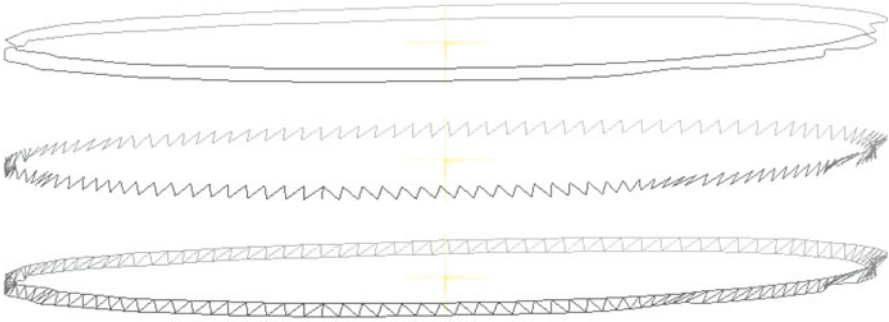


Fig. 2 Upper panel: oblique view of two neighboring cross-sections of the ascending aorta just below the aortic arch; Middle panel: set of edges having minimal aggregate length joining points of the discretized cross-sections; Lower panel: triangulation of the surface between the cross-sections

point and then compute the aggregate arclength to each subsequent point. The perimeter is periodic, so the last point is also the first point. Normalize so that the arclength from first point to last point is 1.0. Logically, the points from both perimeters could be thought of as coexisting on a line segment of length 1.0. We are going to connect points of one perimeter to points of the other, and it is convenient to maintain a pointer on each perimeter to the “current point”, which is the last point on that perimeter which has been connected to another point. Identify either perimeter as perimeter one and the other perimeter as perimeter two. Connect the current point of perimeter one to each point of perimeter two which has an arclength greater than that of the current point of perimeter two and which also has an arclength less than or equal to the arclength of the next point on perimeter one (or arclength = 1.0 if the current point of perimeter one is its last point). With each new connection, the current point of perimeter two is incremented. After making these connections, interchange perimeter identities, so the perimeter which had identity one now has identity two, and vice-versa. The cycle of connecting points and interchanging identities continues until the two last points are connected. These connections define the edges of a triangulation whose other edges lie on the perimeters. The aggregate length of the inter-perimeter edges depends on which points are chosen as the first points. We test all possible pairs of first points and choose the pair which results in the shortest aggregate length of the inter-perimeter edges. Figure 2(middle) shows this set of inter-perimeter edges for the perimeters in Fig. 2(upper). Figure 2(lower) shows the resulting surface triangulation.

Applying this procedure to the entire data set results in triangulated surfaces for the major anatomical structures at end-systole: aorta and pulmonary artery; superior and inferior vena cava and pulmonary veins; right atrium; left atrium and appendage; right ventricle; left ventricular endocardium; left ventricular epicardium; the epicardium of the entire heart. The valves are produced by a procedure described later. In addition, we require a surface in the left ventricle, midway between the endocardium and epicardium. For each level at which there are both endocardial and epicardial cross-sections, a midwall cross-section can be constructed by averaging

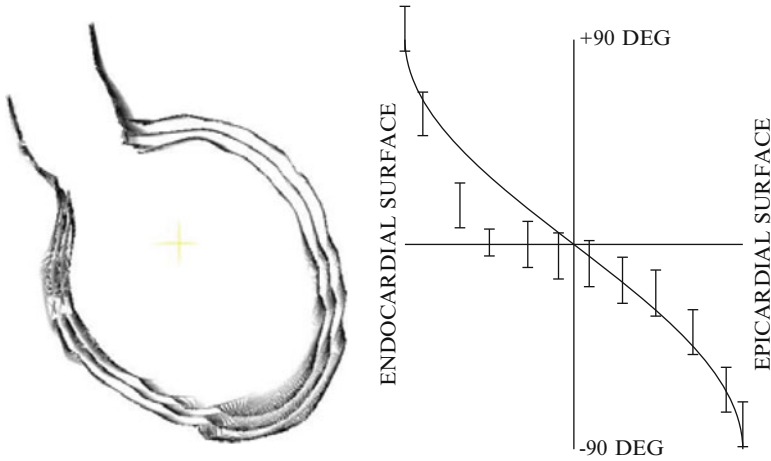


Fig. 3 Left panel: left ventricular endocardial, midwall and epicardial triangulated surfaces based on CT data. Right panel: fiber-angle distribution in the LV wall. Continuous curve is predicted by the asymptotic theory of Peskin [3], and vertical bars represent the observations of Streeter, et al. (ref. [7], Fig. 4, curve a) plotted with a tolerance of $\pm 10^\circ$. Horizontal axis is radial distance through the wall with midwall at zero and epicardial and endocardial surfaces as indicated. Vertical axis is angle between cardiac fibers and a plane perpendicular to the axis of the left ventricle along line segments connecting the two cross-sections, drawn normal to the endocardial cross-section. In all cases the region of the apex is approximated by the plane of the most apical cross-section of the data set. Figure 3(L) shows the three surfaces (endocardial, midwall, epicardial) of the left ventricle (LV) with the front and rear clipped away to improve visibility.

3 Construction of Model Heart Fibers

The model heart is comprised of three types of structures in each of which fibers are constructed using variations on the same general theme: geodesics on surfaces. The three types of structure are: thick-walled (the LV), thin-walled (all other chambers and the great vessels), and valvular.

3.1 Model Left Ventricular Muscle Fibers

At the present time the technology for directly imaging cardiac muscle fibers (e.g., diffusion tensor MRI) is not sufficiently developed for our purposes. Instead, we approximate the muscle fibers in the left ventricle using a method motivated again by the observations of Streeter, et al. [6] and by an asymptotic analysis of Peskin [3]. From measurements on the hearts of macaques, and treating the left ventricle as a nest of ellipsoidal surfaces of revolution, Streeter observed that muscle fibers in the left ventricle follow trajectories that are approximately geodesic paths on those

surfaces. Using an asymptotic approach, also treating the wall of the left ventricle as a nest of surfaces (but not necessarily ellipsoids) of revolution, Peskin was able to derive the relation between the angle made by the geodesics as a function of their distance from the midwall surface. The angle is measured relative to the latitude lines of an appropriate coordinate system constructed on the surface of revolution. Figure 3(R) (redrawn from [3]) shows the relation between angle and distance from the midwall surface.

We treat the midwall surface shown in Fig. 3(L) as if it were a surface of revolution, even though it is not. A coordinate system is constructed on this surface consisting of geodesic curves which radiate out from the apex and terminate at the upper cross-sections of the data set, above the mitral or aortic valve ring. These coordinates can be thought of as lines of longitude. Lines of latitude are constructed by joining points of equal arclength measured from the apex along the lines of longitude. It is an interesting theorem of differential geometry that the latitudes and longitudes so constructed form an orthogonal net, even when the surface on which the above construction is done is not a surface of revolution. Because the midwall surface is not in fact a surface of revolution, some of these longitude lines intersect. When a pair of longitude lines intersects, the longer of the longitudes is terminated at the intersection. This insures that when following a line of constant latitude there is a monotonic change in longitude.

Following Streeter's observation, and the theoretical curve of Fig. 3(R), muscle fibers are parallel to the lines of latitude on the midwall surface. We construct families of model fibers by computing geodesic paths on the CT scan midwall surface, starting at locations equally spaced on, and initially parallel to, a line of constant latitude. Geodesics paths are computed in both directions (increasing longitude and decreasing longitude), terminating whenever a valve ring is encountered. When a geodesic crosses a longitude line, its angle with respect to the latitude line is calculated, and the intersection point is projected toward the epicardial or endocardial surface by the distance given by the theoretical curve of Fig. 3(R). In this way model muscle fibers fill the space between the epicardial and endocardial surfaces, and have an angular distribution through the wall which matches the observed distribution.

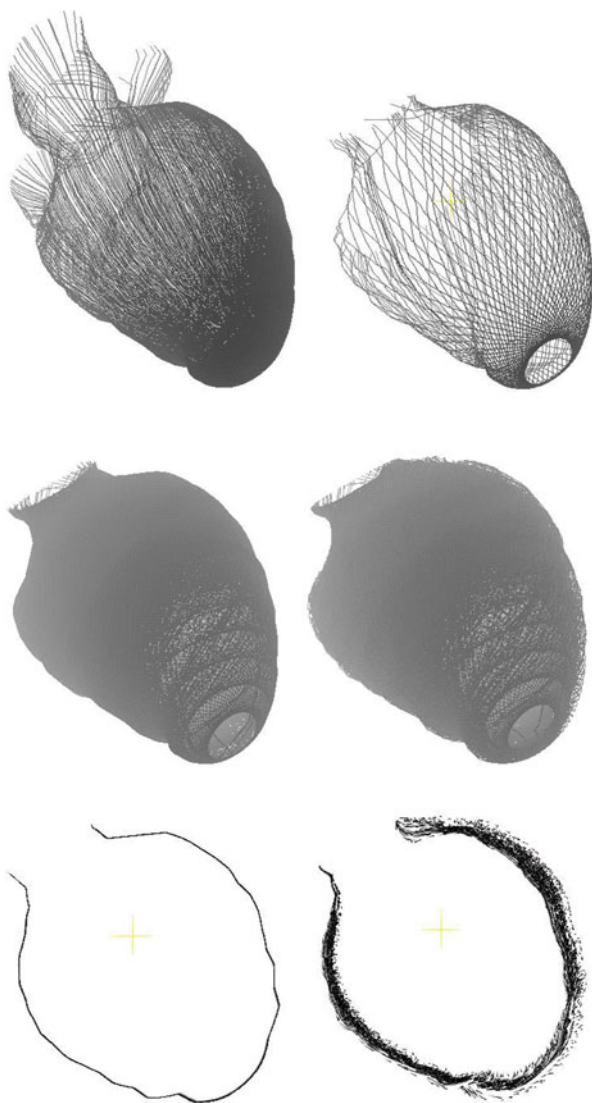
How are geodesics constructed? Recall that the surface is triangulated. Each triangle in such a structure "knows" which triangles are its neighbors. Each triangle shares each of its edges either with another triangle or with nothing (in the case of edges on the upper end of the CT scan). No triangle shares any two of its edges with the same other triangle. It is straightforward to construct a map that indicates the other triangles with which any triangle shares edges. A triangle is a plane figure, so on each triangle a local coordinate system may be constructed having one unit vector normal to the plane of the triangle and two unit tangent vectors in the plane of the triangle. If a line is drawn from a point within any triangle in a known direction in the plane of the triangle, its intersection with one of the edges of the triangle is easily calculated. From these considerations geodesic curves on the surface are constructed as follows: draw a straight line or ray emanating from a point on a latitude line in one direction along the latitude line. Call the triangle containing the

starting point “the current triangle”. The drawn ray, straight and in the plane of the triangle, is a geodesic of the triangle by definition. Calculate the intersection point of the drawn ray with an edge of the current triangle (there can be only one such point). Determine which triangle shares that edge with the current triangle, and call that triangle “the next triangle”. The vector in the direction of the drawn ray intersecting the edge can be decomposed into two components, along the edge and normal to the edge within the current triangle, and then recomposed along the edge and normal to the edge within the next triangle. The component along the edge is unchanged since the edge is shared; the component which was normal to the edge within the current triangle retains its magnitude but changes its direction to be normal to the edge within the next triangle. The ray drawn in the next triangle line has a known starting location (on the edge) and a known direction. Rename the next triangle as the current triangle, and repeat the drawing process just described. The result is a piecewise linear geodesic path on a triangulated surface. Note that the path is not necessarily the globally shortest path between the endpoints, but *is* the shortest path given the particular starting direction, which is sufficient for a geodesic path.

Figure 4 (upper row) shows the lines of longitude on the midwall surface (left panel), and a family of midwall-surface geodesics starting on a latitude line near the apex (right panel). Notice that even though geodesics are initially perpendicular to longitude lines (near the apex), they become more longitudinal as they rise (toward the base). Figure 4 (middle row, left panel) shows several families of geodesics on the midwall surface, each family arising from one of a set of regularly spaced latitude lines.

We now have a midwall surface covered by geodesic fibers. The density of coverage can be increased (or decreased) by starting more (or fewer) geodesics on any particular latitude line, or by having more (or fewer) latitude lines from which geodesics are launched. Choosing an appropriate density of coverage is discussed later. During the process of construction the angle any geodesic makes when it intersects any longitude line can be tracked. The major requirement is a table which lists the coordinates of the end points of the longitude line segments crossing any triangle on the midwall surface. The interior of the LV wall can now be populated with muscle fibers as follows. Recall that Fig. 3(R) shows fiber angle as a function of depth. Simply inverting this gives depth as a function of fiber angle. For each intersection of a geodesic with a longitude line, the intersection point can be “inflated” along the normal to the midwall surface triangle by an amount given by its angle with respect to the latitude line. For this purpose the distance in Fig. 3(R) is taken to be the fraction of the maximum possible distance along the normal. Since every surface is defined as a set of triangles, this inflation process requires finding the intersection of a midwall-triangle normal with an epicardial or endocardial surface triangle. There are several thousand such triangles, and hundreds of intersection points on each of several thousand geodesics. A brute force search for intersections of normals and triangles would be quite time-consuming. To improve efficiency we construct a table which lists the endocardial and epicardial surface triangles intersected by a normal at the centroid of each midwall surface triangle. For a normal starting at some other point (not the centroid), this table

Fig. 4 Upper left panel shows midwall-surface longitude lines; upper right panel shows a family of midwall-surface geodesics arising from a latitude line near the apex; middle left panel shows families of midwall-surface geodesics arising from several regularly spaced latitude lines; middle right panel shows fibers inflated from the midwall-surface geodesics to fill the LV wall; lower left panel shows a thin cross-section through the midwall-surface geodesics; lower right panel shows a thin cross-section through the inflated fibers



indicates the triangle which is probably intersected by the normal, and, if not, is a reasonable place from which to start searching for the intersected triangle. (Recall that each triangle knows its neighbors; a nearest neighbor flooding type of search proves to be effective here.) Fig. 4 (middle row, right panel) shows the muscle fibers resulting from inflating the geodesics shown to its left. Figure 4 (lower row) shows thin cross-sections through the entire collection of geodesics on the midwall surface (left panel) and the entire collection of inflated fibers in the LV wall (right panel).

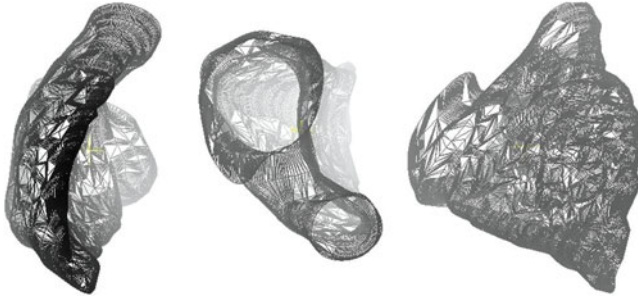


Fig. 5 Triangulated right ventricular surface, three different points of view. The right panel uses the same point of view as in Fig. 4

3.2 *Model Thin-Walled Chambers*

All the other chambers of the heart are treated as thin-walled surfaces covered with geodesics using the same general approach as used in the midwall surface of the LV. For each chamber the particular geodesics selected are intended to represent fibers observed experimentally. Figure 5 shows the triangulated surface resulting from segmentation of the right ventricle (RV). Figure 6 shows the computed geodesics on the model RV surface. There are two families: first, fibers which radiate out from the RV apex (“lines of longitude”) and second, fibers which are approximately orthogonal to the first on the septal surface. These particular families were suggested by the dissections of Thomas [8].

Figure 7 shows the triangulated surfaces of the model atria. There are three atrial surfaces in the model: left atrium, right atrium and a combined atrial surface which serves to hold the two atrial chambers together. The left atrial surface includes the appendage; the right atrial surface does not. The combined atrial surface is constructed by joining the left and right surfaces with bridging surfaces consisting of a small number of triangles (that is, the combined atrial surface is not the result of another segmentation). All atrial surfaces are considered to be thin-walled. Geodesic paths are constructed on each of these three surfaces. The resultant muscle fibers are shown in Fig. 8. Two families of fibers are used here, one to represent the pectinate muscles and one to represent the interatrial band.

3.3 *Model Great Vessels*

It is our intention initially to use the great vessels as a means to anchor the model heart, that is, points on the great vessel walls will be connected to fixed points in space with springs of appropriate stiffness, as if the vessels were surrounded and constrained by the tissues of the body. Hence, it is unnecessary to construct paths which represent fibers in the vessels. It is sufficient to represent the vessels as surfaces triangulated finely enough that there are no leaks. Triangles with edges

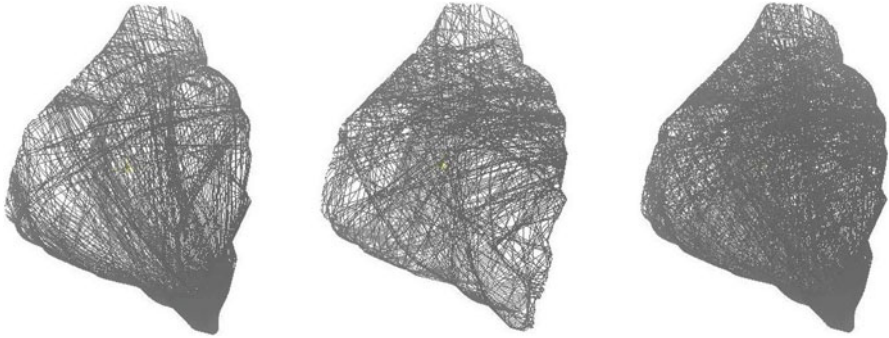


Fig. 6 Computed geodesics on the model RV surface. Left panel: family 1, radiating from the apex (longitude lines). Middle panel: family 2, perpendicular to a longitude line on the septum. Right panel: both families

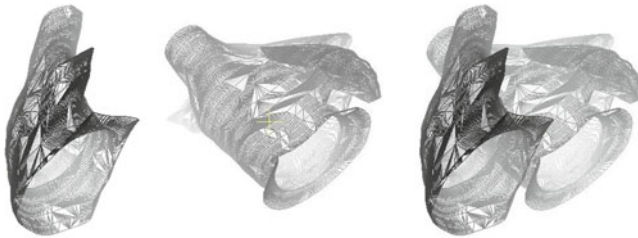


Fig. 7 Triangulated surfaces of the model atria. Left panel: right atrium; Middle panel: left atrium; Right panel: combined left and right atrium

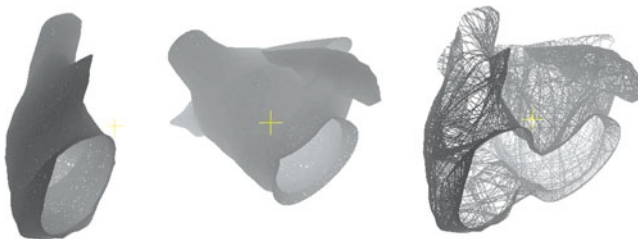


Fig. 8 Muscle fibers of the model atria. Left panel: right atrium; Middle panel: left atrium; Right panel: combined left and right atrium

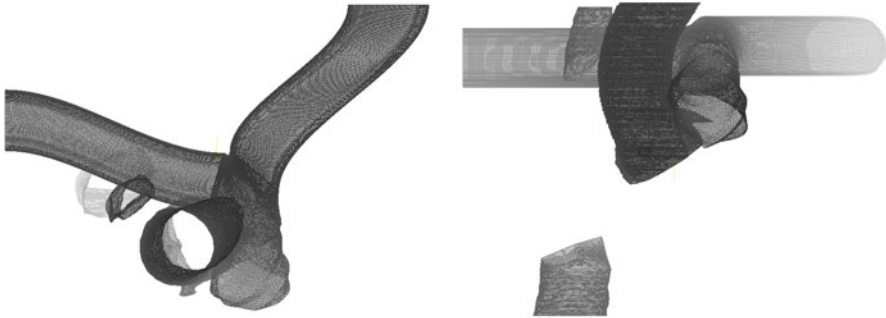


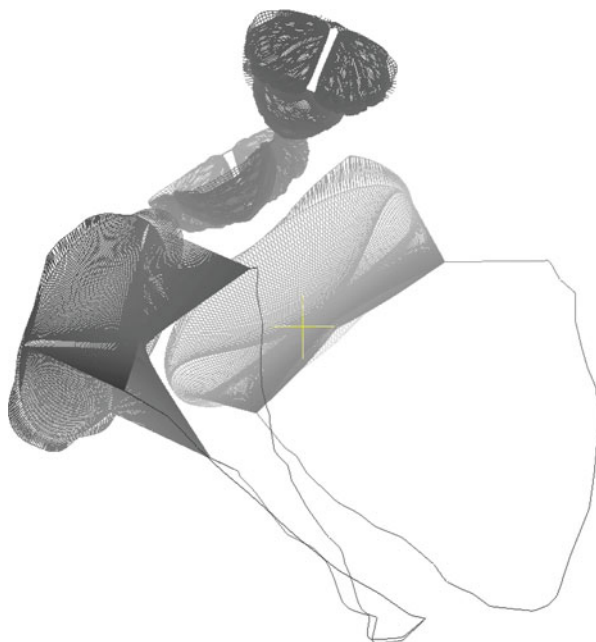
Fig. 9 Model great vessels. Left panel is a top view in which extensions of the left and right pulmonary artery branches to the edges of the computational domain are clearly shown; right panel is a front view in which the aorta is to viewer's left of the pulmonary trunk, the superior vena cava is to viewer's left of the aorta, and the inferior vena cava is at the bottom

on the order of $1/2$ of the computational lattice meshwidth are expected to be adequate for this purpose. Figure 9 shows the triangulated surfaces for the inferior and superior vena cava, the aorta, and the pulmonary artery. The data set ends just below the arch of the aorta. In the figure vessels have been extended to the edges of the computational domain where boundary conditions representing the circulation (e.g., windkessels) will be imposed.

3.4 Valves

The CT data set from which we are constructing the model does not include any of the four valves. The valves we intend to use, initially, are modified versions of the valves used in the conical surface model. Figure 10 shows these (modified) valves. The mitral valve consists of fibers lying on a surface which smoothly interpolates between a line joining the tips of the papillary muscles and a circle in the plane of the valve ring. This is a surface of the type described in [1] which interpolates between two ellipses at different heights. The line joining the tips of the papillary muscles is an ellipse of eccentricity 1.0, and the circle in the plane of the valve ring is an ellipse of eccentricity 0.0. Fibers fan out from each papillary tip and contribute to each leaflet. Where fibers from different papillary tips cross, a fabric which constitutes the leaflet is formed. Between the papillary tip and the leaflet, each fiber forms one of the chordae tendineae. The construction of the tricuspid valve is similar, except that there are three papillary tips and three leaflets. Again, fibers from each papillary tip fan out to contribute to the two leaflets nearest the tip. Where fibers from different tips cross, a leaflet fabric is formed. Between the

Fig. 10 Model valves in their correct anatomical relations. In the figure, the pulmonic valve is uppermost, with the aortic valve just below and to its left. The tricuspid valve is the leftmost valve, with the mitral valve to its right. (All directions are in the viewers framework.)



tip and the leaflet, each fiber is one of the chordae. Each valve ring in the model is constructed as the intersection of a plane with an appropriate triangulated surface. The triangulation is left unchanged, but whenever a geodesic curve on a triangulated surface intersects a valve ring, the geodesic is terminated at the point of intersection. It is in the nature of the data set and the segmentation that the valve rings on any triangulated surface are unlikely to be perfectly circular. Collars are inserted in order to join the circular parts of the model valves with the non-circular valve rings. These collars are visible in Fig. 10.

The model aortic and pulmonic valves result from the solution of a partial differential equation which describes the equilibrium of a single family of fibers supporting a hydrostatic pressure load [5]. As in the case of the inflow valves, the outflow valve rings are not circular, and collars are inserted. The collars for the outflow valves are significantly smaller than those for the inflow valves. Except for the collars, the shapes of the aortic and pulmonic valves are identical. In the current construction the diameter of the pulmonic valve is approximately 11% smaller than that of the aortic valve.

All the valves shown in Fig. 10 are in their closed configurations. Nonetheless there is a small gap between neighboring leaflets in each valve. This gap is required in order for the valves to open. In the IB method, the velocity of any boundary point is interpolated from the velocities stored on the surrounding computational lattice. Any two boundary points with identical locations would therefore have the same velocity and would always stay together. If there were no gap between valve leaflets, points on neighboring leaflets would move with the same velocity and the leaflets

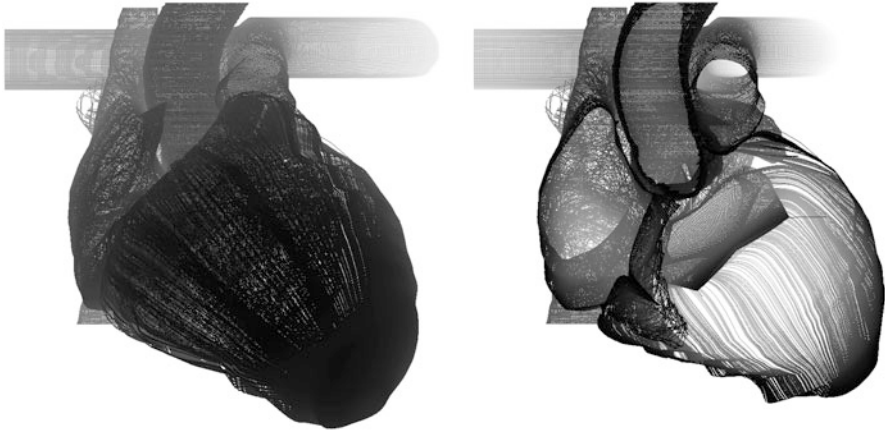


Fig. 11 Left panel: entire heart model; Right panel: interior of heart model

would be unable to open. The gap must be sufficiently large that neighboring leaflets can separate on opening, but not so large that the closed valve leaks. In practice, the leaflet gap that works best is found by trial and error.

4 Entire Heart Model

Figure 11(Left) shows the entire model heart composed of all the pieces described above. The left and right ventricles are contained in an investment layer composed of longitudinal fibers radiating from the apex of the model heart. At this point in the construction the great vessels all have open ends. In our earlier (conical surface) model, the great vessels were capped off, and sources or sinks as appropriate were inserted within the caps to represent the portions of the circulatory system not modeled in detail. Although such caps have been constructed for the model described here, we are currently investigating other approaches for treating inflow and outflow, such as connecting the great vessels to circulation models (e.g. windkessels) on the boundaries of the domain. Figure 11(Right) shows the interior of the heart model by clipping away portions of the heart wall.

5 Selecting Boundary Point Density

We have previously remarked that the IB method requires neighboring points on the boundary to be spaced apart $1/2$ the meshwidth of the computational lattice in order for the boundary to not leak. The meaning of this spacing requirement is clear for a closed 2D curve. In three dimensions the situation is more complicated. It is easy to

imagine, in 3D, that every point is within $1/2$ meshwidth of some other point and that there are gaping holes in the structure through which fluid could pass.

We have adopted a strategy based on flooding (our favorite technique) for determining if a 3D structure is leak-proof on some target computational lattice resolution, that is, whether the boundary has been discretized appropriately not to leak in a particular computational lattice.

Flooding is a method for estimating the enclosed area within a curve (in 2D) or volume within a surface (in 3D). It is most easily described for the 2D setting and the extrapolation to 3D is obvious. Consider a closed 2D discretized curve. Construct a Cartesian grid covering a rectangular domain which includes the 2D curve. Mark with a 1 the corner points of every grid cell containing any point of the curve, mark with a 1 the grid points on the edges of the grid, and unmark (with a 0) all other grid points. The marks on the grid in the immediate neighborhood of the 2D curve form a fat stair-step representation of the 2D curve. To begin flooding, locate an unmarked grid point in the interior of the 2D curve, mark that point with a 1 and place the coordinates of that grid point on a *queue*. Repeat the following process until all the items in the queue have been examined:

(1) Examine the grid points which are neighbors of the grid point at the head of the queue (at most 8 neighbors in 2D). Each neighbor which is found to be unmarked is then marked and added to the tail of the queue.

(2) Advance the head pointer to the next grid point on the queue.

This process finishes when the head pointer advances beyond the tail of the queue. The portion of the grid interior to the 2D curve will be marked (i.e., the interior of the 2D curve will have been “flooded”) and the number of items placed on the queue is the number of grid points in the interior of the 2D curve. If each such grid point is treated as the lower left-hand corner of a grid cell, the number of such grid points estimates the area enclosed by the 2D curve.

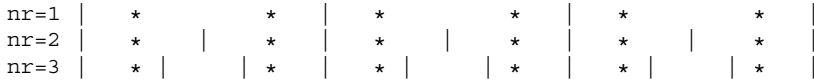
A difficulty arises if the distance between neighboring points on the curve is large compared with the grid meshwidth. In this case there will be a gap in the stair-step representation of the 2D curve, and the flooding will not be confined to the interior of the curve, but will leak into the exterior, encompassing virtually the entire grid. A remedy for this difficulty would be to interpolate points onto the boundary at a density appropriate for the grid resolution. For our present purposes we wish to consider a fixed discretization of the boundary and to use the onset of leaking when the grid is made finer as an indicator of when that discretization of the boundary is leak-proof.

Consider the following 1D model problem in which

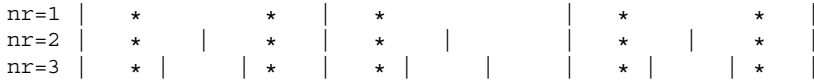
```
'|' indicates the locations of lines in the grid
'*' indicates boundary points
nr how fine the flooding grid is w.r.t. the target grid
nr=1 means flooding grid is same resolution as target grid
```

In the picture below boundary points are spaced at $1/2$ of the target grid meshwidth ($nr = 1$). This is the spacing of boundary points which the IB method sees as leak-proof. The flooding grid is initialized by marking the grid on either side of a boundary point. Notice that in all 3 cases ($nr = 1, 2, 3$) there is a sufficient

density of boundary points that every grid line ends up being marked, even though there is *not* a boundary point between every pair of neighboring grid lines.



Now suppose one boundary point is absent producing a local failure of the ‘‘1/2 of the target grid meshwidth’’ rule.



For $nr = 1$ or $nr = 2$ there is still sufficient boundary point density to mark every grid line, but at $nr = 3$ there is one grid line that does not get marked because the cells on either side of it are empty. This would permit a flood of the interior (say the space above the line) to pass through the line to the exterior (the space below the line). We take this to mean that, if a 3D structure made of points does not leak when the flooding grid is refined by a factor of 3 relative to the target grid, the 3D structure has a sufficient density of points for the IB method with the target grid resolution.

Acknowledgements The authors are grateful to Arthur E. Stillman, M.D., Ph.D., and Randolph M. Setser, D.Sc. of The Cleveland Clinic Foundation, Cleveland, Ohio for providing the CT images on which this work was based. We are also deeply grateful to ‘‘Mr. C.’’, the patient whose heart was imaged.

We thank Nikos Paragios for organizing the collaboration between the Cleveland Clinic, Siemens Corporate Research and NYU that made possible the present work.

References

1. D. M. McQueen and C. S. Peskin. A three-dimensional computer model of the human heart for studying cardiac fluid dynamics. *Computer Graphics*, 34:56–60, 2000.
2. D. M. McQueen and C. S. Peskin. Heart simulation by an immersed boundary method with formal second-order accuracy and reduced numerical viscosity. In H. Aref and J. Phillips, editors, *Mechanics for a New Millennium, Proceedings of the International Conference on Theoretical and Applied Mechanics (ICTAM) 2000*, pages 429–444. Kluwer Academic Publishers, 2001.
3. C. S. Peskin. Fiber-architecture of the left ventricular wall: an asymptotic analysis. *Commun. Pure and Appl. Math.*, 42:79–113, 1989.
4. C. S. Peskin. The immersed boundary method. *Acta Numerica*, 11:479–517, 2002.
5. C. S. Peskin and D. M. McQueen. Mechanical equilibrium determines the fractal fiber architecture of the aortic heart valve leaflets. *Am J. Physiol.*, 266:H319–H328, 1994.
6. D. D. Streeter, W. E. Powers, A. Ross, and F. Torrent-Guasp. Three-dimensional fiber orientation in the mammalian left ventricular wall. In J. Baan, A. Noordergraaf, and J. Raines, editors, *Cardiovascular System Dynamics*, pages 73–84. MIT Press, 1978.
7. D. D. Streeter, H. M. Spotnitz, D. P. Patel, J. Ross, and E. H. Sonnenblick. Fiber orientation in the canine left ventricle during diastole and systole. *Circ. Res.*, 24:339–347, 1969.
8. C. E. Thomas. The muscular architecture of the ventricles of hog and dog hearts. *Am. J. Anat.*, 101:17–57, 1957.

Image-based haemodynamics simulation in intracranial aneurysms

A.G. Radaelli, H. Bogunović, M.C. Villa Uriol,
J.R. Cebral, and A.F. Frangi

Abstract Image-based haemodynamics simulation is a computational technique that combines patient-specific vascular modeling from medical images with Computational Fluid Dynamics techniques to approximate the complex blood flow characteristics of healthy and diseased vessels. Advances in image quality, algorithmic sophistication and computing power are enabling the introduction of such technology not only as a biomedical research tool but also for clinical practice. In particular, the interaction between haemodynamical forces and arterial wall biology is believed to play an important role in the formation, growth and, eventually, rupture of intracranial aneurysms. Due to the absence of ground truth image modalities to measure blood flow, image-based haemodynamics simulation represents an attractive tool to provide insight into the haemodynamics characteristics of intracranial aneurysms. In this chapter, we provide an overview of the main components of this technique, illustrate recent efforts in its validation and sensitivity analysis and discuss preliminary clinical studies and future research directions.

A.G. Radaelli

CISTIB Centre for Computational Imaging & Modelling in Biomedicine, Universitat Pompeu Fabra, c/ Tàrrer 122-140, E08018 Barcelona, Spain

H. Bogunović

CISTIB Centre for Computational Imaging & Modelling in Biomedicine, University of Sheffield, c/ Tàrrer 122-140, E08018 Barcelona, Spain

M.C. Villa-Uriol • A.F. Frangi (✉)

CISTIB Centre for Computational Imaging & Modelling in Biomedicine, University of Sheffield, Sir Frederick Mappin Building, Sheffield, S1 3JD, UK

e-mail: a.frangi@sheffield.ac.uk

J.R. Cebral

Center for Computational Fluid Dynamics, School of Physics, Astronomy and Computational Sciences, George Mason University, Planetary Hall, Room 103A, 4400 University Drive, MSN 3F3, Fairfax, VA 22030, USA

1 Introduction

Intracranial aneurysms are pathological, localized dilatations of the cerebral arteries. They have been reported to affect around 1-6% of the population and the major complication is their rupture, which causes subarachnoid haemorrhage (SAH). Although the annual risk of rupture is as low as 0.7%, aneurysm rupture has exceedingly high mortality (50%) and morbidity (20%) rates [34, 66, 70]. Intracranial aneurysms may be treated either by surgical clipping or endo vascular procedures, performed primarily by coil embolization, whose aim is to disconnect the aneurysm from the cerebral circulation and re-establish a physiological flow passage through the parent artery. However, post-treatment complications and failure have been reported to even potentially outweigh the risk of rupture, thus clinicians are often left with the fundamental question whether to treat or not [52, 69]. Currently, clinical risk assessment is based on the combination of patient-level and environmental risk factors, aneurysm location and simple descriptors of aneurysmal size and shape such as aneurysmal volume and aspect ratio. Intracranial aneurysms frequently occur at or near arterial bifurcations in the circle of Willis, which suggests an important role for haemodynamical stresses [37]. Factors such as wall shear stress (WSS), mural stress, impingement force, flow rate and residence time have been suggested to influence the initiation, progression and rupture of intracranial aneurysms, although the exact mechanisms not only remain unknown but also are thought to differ at various phases of their natural history [7, 26]. Abnormal haemodynamics has also been indicated to influence the failure of surgical and endo vascular treatment [27, 39].

Insight into the haemodynamics of a particular individual may allow for improved diagnostic and treatment planning. With this goal in mind, several authors have initially analyzed the fundamental haemodynamical properties of blood flow in idealized models of intracranial aneurysms either experimentally in phantom models [40] or numerically using 2D and 3D Computational Fluid Dynamics (CFD) [24, 33]. Due to the strong dependence of fluid flow on the geometrical configuration of branching vessels and on the aneurysm shape and neck size, a personalized haemodynamical description was found to rely on the patient-specific geometry [67]. A variety of image processing techniques have been then applied to retrieve patient-specific flow measurements directly from medical images [54, 55, 63]. In particular, quantitative measurements of blood velocity and volumetric flow rates are provided by time-resolved 2D or, more recently, 3D phase contrast (PC) magnetic resonance angiography (MRA) [72], although detailed information on blood velocity is undermined by limited resolution and the effectiveness of the methodology is still unproven for intracranial aneurysms. On the other hand, current image acquisition techniques provide accurate measurements of vascular anatomy. Realistic experimental phantoms, typically made of a translucent polymer, may be casted starting from vascular models extracted by segmentation of the medical images. The phantom is then connected to a pulsatile pump reproducing physiological flow waveforms. Imaging techniques such as Laser Doppler Velocimetry

(LDV) or Particle Image Velocimetry (PIV) are then applied to visualize and measure the velocity field [59, 60]. A detailed measurement is however quite time consuming and is generally not achievable close to the wall, thus hampering the calculation of important haemodynamical variables such as WSS. In addition, the manufacturing process is not trivial for small arteries and small aneurysms and would be impractical for a systematic patient-specific analysis.

Instead, image-based haemodynamics simulation may represent an attractive tool to provide a description of patient-specific haemodynamical variables with compelling detail. The basic principle of this technique is to obtain accurate patient-specific vascular models from medical images and apply CFD techniques to reconstruct the time-resolved blood velocity and pressure fields starting from flow and/or pressure conditions prescribed at the boundaries of the vascular domain. The availability of high-resolution scanners, fast modeling algorithms and powerful computational resources has been key to the widespread adoption of image-based haemodynamics simulation techniques in biomedical research. In the next sections, the fundamental components of an image-based haemodynamics simulation workflow will be illustrated. The significance of assumptions, approximations and modeling parameters will be considered and efforts in the validation of the simulation results will be discussed. Current application of state-of-the-art technologies for the understanding of aneurysm rupture will be further addressed and directions for future research and development will conclude the chapter.

2 Workflow overview

The application of image-based haemodynamics simulation technologies to cerebral vessels and intracranial aneurysms involves the cross-disciplinary integration of imaging, modeling and simulation techniques [58]. The main challenges are the development of an efficient pipeline for the generation of patient-specific anatomical models from medical images and the establishment of a sustainable level of realism of the haemodynamics simulation model. In addition, due to the need of regular patient monitoring, the processing pipeline has to be applicable to a variety of angiographic imaging modalities, including Computed Tomography Angiography (CTA), 3D Rotational Angiography (3DRA) and MRA. Our groups have recently proposed an efficient pipeline for the modeling of blood flow patterns in intracranial aneurysms from 3DRA and CTA images, with the eventual possibility of performing CFD simulation in large samples of aneurysm models (and with other image modalities, e.g. MRA) and of investigating the association of specific haemodynamics variables with clinical events such as rupture [12]. Similar approaches have been adopted by other groups in [13, 17, 29, 36, 57, 68].

A sketch of the image-based haemodynamics simulation pipeline is shown in Fig. 1. The pipeline is composed of five key components: 1) medical image acquisition; 2) anatomical modeling; 3) computational grid generation; 4) specification of boundary conditions; 5) CFD simulation and visualization. Medical

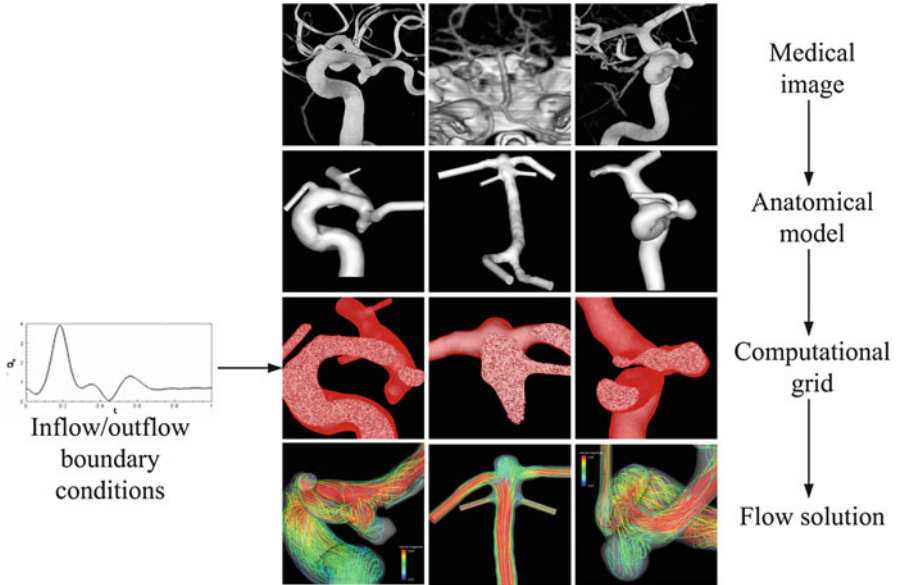


Fig. 1 Sketch of the haemodynamics simulation pipeline

image acquisition requires the identification of protocols providing good contrast between vessels and other tissues. At the same time, it is also important to adopt an accurate anatomical modeling strategy able to reconstruct both the morphology of the vessels of the Circle of Willis and of the aneurysm(s), which are both complex and a priori unknown. Anatomical modeling comprises image segmentation and mesh processing. Image segmentation here refers to the extraction of the boundaries of vascular objects from a 3D image. The outcome is a binarization of the image in voxels belonging either to the vessels or to the background (including other tissues) or a distance transform image where voxels are given values representing the Euclidean distance to the vascular surface. Mesh processing algorithms involve first the extraction of a surface triangulation representing the vascular wall and then some geometrical and topological corrections that lead to a smooth representation of the vessels of interest.

The anatomical model is used as a support surface to generate a computational grid that divides the complex arterial domain into a finite number of smaller polyhedral elements such as tetrahedra or hexahedra. Before simulating the blood flow through the geometrical domain, the nodal velocities or pressure at the geometric boundaries of the model must be specified. Blood is approximated as a Newtonian fluid and vessel walls are typically assumed rigid. Blood flow is mathematically modeled by the unsteady Navier-Stokes equations for an incompressible fluid. These equations are converted into a matrix of discrete equations whose solution yields the unknown velocity and pressure at the nodes of the entire volumetric grid. The velocity field is then processed to facilitate the analysis of the complex unsteady

flow characteristics, which includes simple operations such as derivation or time averaging to extract clinically relevant quantities such as WSS magnitude and gradient or the oscillatory shear index (OSI).

2.1 Image acquisition

At present, there are no standard guidelines for the evaluation of image quality for haemodynamics simulation. This often requires a certain familiarity with the flow realization across the vessels of the Circle of Willis, an understanding of the geometrical features prone to affect blood flow and a basic knowledge of the main components of the processing pipeline. It is generally accepted that 3DRA images provide both higher resolution ($\sim 0.15mm$) and better contrast between the signal intensity of vascular lumen and background when compared to CTA and MRA. On the other hand, 3DRA images have been reported to occasionally present artefacts leading to the underestimation of aneurysm dimensions in specific locations of the Circle of Willis [35] or to the appearance of pseudo stenosis of parent vessels [20, 32]. Due to the profound influence of parent vasculature on aneurysmal haemodynamics, the absence of artefacts has to be ensured not only in the aneurysm but also for all the vessels influencing the realization of the aneurysm haemodynamics. While providing important information on the aneurysm environment in the brain, CTA images offer instead lower resolution ($\sim 0.4mm$) and the ranges of intensities of vessels and bone overlap. These conditions may limit an accurate segmentation of both small aneurysms and vessels crossing the skull base, where vascular and bone structures are very close. Similar resolution is achieved with TOF MRA, which has the disadvantage of sensitivity to metal implants such as endo vascular devices. In addition, progressive saturation may occur as the spins penetrate into the imaging volume and for disturbed or slow flow. The result is the appearance of signal voids in the vascular lumen or blurred regions across the vascular boundaries, at bifurcations and inside aneurysms, where complex flow patterns may be expected. Despite these limitations, MRA is potentially the most suited imaging technique for haemodynamics simulation as it can derive all the necessary boundary conditions from a single non-invasive imaging session by combining TOF and PC MRA acquisitions.

2.2 Image segmentation

For adapt to the topological changes common in the cerebral vasculature, the segmentation of cerebral vessels may be performed using the geometric deformable model technique within the level set framework [45]. This technique describes the evolution of a surface model within the image domain that deforms under the action of internal (curvature and smoothness) and external (image gradient and

other features) forces to recover the unknown shape of the vessels. The surface is represented implicitly as a zero level set of a distance transform image whose size corresponds to the original medical image. Information on image gradient drives the evolution of the model towards the locus of maximum intensity variation across the vascular boundaries [9]. Due to limited resolution and/or image artefacts, the gradient information may be discontinuous or weak across the vascular boundaries, thus leading to boundary leakage. To overcome this limitation, the Geodesic Active Regions (GAR) [48] technique combines image gradient maps with statistical region-based information. The region-based information is presented in the form of a probability image map, that contains the probability of each image voxels to belong to a certain region (or tissue) R . The estimated probability value can be interpreted as a conditional probability, $P(x \in R \mid \mathbf{f}(x))$, where x is a point in the image domain and $\mathbf{f}(x)$ is the feature vector used to characterize the tissue at the point x . The regional descriptors k for each region are then based on the corresponding probability map. The equation that drives the evolution of the surface is expressed as:

$$\Phi_t + \zeta(k_{out} - k_{in}) \mid \nabla \Phi \mid - \eta(\varepsilon g K_m \mid \nabla \Phi \mid + \nabla g \cdot \nabla \Phi) = 0 \quad (1)$$

where Φ is an implicit function whose zero level set at any time t of the evolution represents the vascular surface, k_{out} and k_{in} are the descriptors of the inner and outer tissues with respect to the vascular lumen, K_m is the mean curvature of the surface, ε is a parameter that controls the contribution of the curvature to the evolution, while ζ and η control the influence of region-based and boundary information, respectively. We have proposed to create features as vectors of differential invariants [64] up to the second order, which are computed at multiple scales to provide a richer description of the different tissues [30]. These feature vectors are expected to be able to differentiate between tissues that cover overlapping image intensity ranges but present different shapes, such as vascular and bone tissues in CTA images. The set of feature vectors belonging to a specific region are first learned in a supervised fashion and then the probability of each image voxel to belong to a particular tissue is estimated using a non-parametric technique. Such approach is particularly suitable for multimodal vessel segmentation as for each modality the vectors of features corresponding to each tissue can be learned from its own training set. Additionally, the initialization process does not require user intervention since it can be obtained by thresholding the probability map corresponding to the vessel region.

2.3 Mesh processing

The generation of numerical grids for flow simulation requires a watertight description of the vascular surface. This can be obtained by automatically applying the method of the marching cubes (or tetrahedra) [6] to the distance transform image obtained from the segmentation process. A sequence of global and local operations

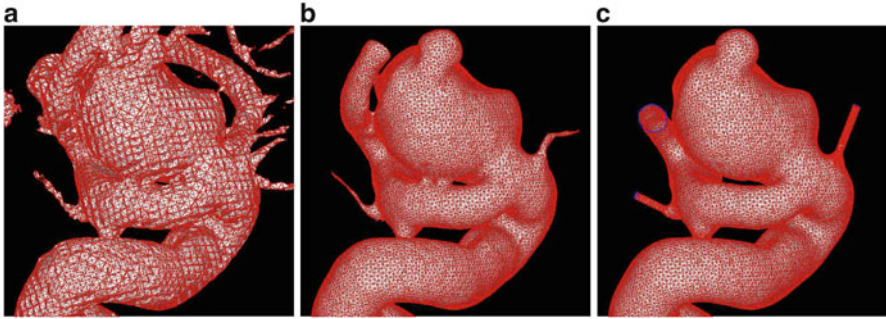


Fig. 2 Mesh processing steps. (a) initial model extracted using the marching tetrahedra technique; (b) smooth model after the application of the Taubin smoothing technique and a gross clipping to discard unwanted vessels; (c) final model after mesh optimization, editing, clipping and extrusion

is then applied to either improve the quality of the resulting triangulation or to correct geometrical and topological irregularities, which include the fusion of the surface with touching vessels or bones (in CTA only) and over- or under-estimation of the aneurysm neck due to low image contrast and/or resolution. Improvements in mesh quality are obtained by the use of automatic mesh smoothing and optimization algorithms. In particular, the Taubin algorithm [61], which is a volume preserving smoothing technique, is employed to remove bumps and sharp corners, although there is no guarantee that some triangles do not become distorted or too small. Therefore, some mesh optimization operations, including edge-collapsing and side swapping or the removal of degenerate, stretched, repeated and small triangles [12], are applied to the whole triangulation.

In the presence of residual geometrical or topological irregularities, a set of interactive tools are adopted to efficiently edit the triangulation, locally remove or smooth a single or a group of elements and close holes while preserving the general topology and smoothness of the vascular surface [5]. The vessels are then interactively cut perpendicularly to the axis. Each operation is carefully considered to minimize the sensitivity of the results to parent vessel modeling and, at the same time, obtain a workable model size for further grid generation and numerical solution. In addition, the boundary surfaces are extruded along the direction of the axis to minimize the effects of boundary conditions on the realization of blood flow. Following work presented in [19, 25, 44], the length of each extension is typically 5 to 10 times the diameter of the associated vessel.

3DRA acquisition achieves a full description of the Circle of Willis only if multiple injections are performed at the same time. As this carries some risks, Castro [11] applied a surface merging technique to fuse models constructed from different 3DRA images into a single watertight model including all possible avenues of flow.

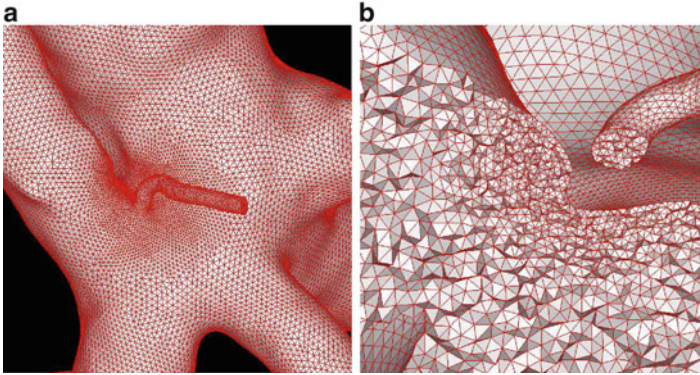


Fig. 3 (a) Surface grid with increased mesh resolution in correspondence to a small branching vessel near the aneurysm neck. (b) Detail of the corresponding volumetric grid

This approach allows to construct complex arterial networks and to investigate the haemodynamics environment of aneurysms that are fed by vessels typically captured in separate 3DRA images.

2.4 Grid generation

The process of generation of a finite element grid for CFD simulation typically starts with the generation of a higher quality surface grid. Although mesh processing operations lead to an appropriate representation of the vascular surface, the quality and the size distribution of the triangulation is not adequate for CFD simulation. New surface grids can be generated starting from an analytical representation of the surface or directly from the surface triangulation obtained after mesh processing. In the former case, the analytical representation of the surface can be obtained using parametric patches or implicit functions of global support [50]. Our approach instead adopts an advancing front method that places newly created points on the original surface triangulation by linear or quadratic interpolation [41]. This process provides elements of higher quality and a more uniform distribution of the element size across the model. This surface mesh is then employed as a support for the generation of tetrahedral elements inside the anatomical model using the advancing front method [42]. Other approaches include Delaunay and octree meshing, but we refer to the literature for further details on the progress in this field [47].

The distribution of element sizes is typically prescribed to obtain a minimum number of elements across the smallest vessel. Adaptive background grids can also be used and interactively specified to increase the mesh resolution in regions where the anatomical model has a large surface curvature. The number of tetrahedral elements in our models typically varies from 1.5 to 4 millions (average element size of 0.1 to 0.15 mm). Details of the surface and volumetric grids generated with our approach are provided in Fig. 3.

2.5 Specification of boundary conditions

The numerical solution of the equations governing the fluid flow requires the specification of the fluid behavior at the boundaries of the computational domain, that account for the unmodeled part of the vascular network. The assumption of non-moving walls implies that the value of the velocity at the surface nodes is null. Physiological boundary conditions are instead derived for the inlets and outlets of the model. These can be obtained from 2D PC MRA images for the main branches of the Circle of Willis, especially if contributing to the realization of the flow in the aneurysm. Flow rate curves are obtained at each location across one cardiac cycle, but due to resolution constraints the technique does not offer an adequate description of the velocity profile. The flow rate curves are therefore decomposed into their Fourier modes and the velocity profile is analytically computed from the Womersley solution [62].

It must be noted that patient-specific flow rate measurements are seldom available and reference data obtained from normal volunteers are instead employed. Ford [21] suggested that the use of reference waveforms may be appropriate, although the time-average flow rate should be scaled to coincide with the patient's measured one. If time-averaged flow rates are not available, an approach could be to scale the reference waveforms by a factor depending on vascular dimensions or so that the time-average flow rate corresponds to physiological shear stress values. In the case reference waveforms are also not available, which is the typical scenario for small outflow branches, traction free (zero-pressure) boundary conditions may be applied, assuming that the flow division among the arterial branches is determined by the geometry of the anatomical model. As flow divisions are actually determined by the impedance of the distal arterial tree, more sophisticated strategies have been proposed to integrate vascular bed models of the brain [15] or 1D models of the whole cardiovascular system [3] into the simulation model.

2.6 CFD simulation

Blood is a suspension of particles (e.g. red/white cells and platelets) immersed into an aqueous fluid (plasma). However, in medium and large sized vessels (diameter $> 1mm$) blood may be considered a homogenous, incompressible, viscous fluid with constant density. Blood flow is here governed by the Navier-Stokes equations for incompressible fluids. Starting from an initial condition at $t = 0$, these equations require the solution of a system of partial differential equations (PDE) for $t > 0$, that express the conservation of mass and linear momentum, respectively, as

$$\nabla \cdot \mathbf{v} = 0 \quad (2)$$

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p + \nabla \cdot \boldsymbol{\tau} + f \quad (3)$$

where ρ is the density, $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ is the velocity field, p is the pressure, τ is the deviatoric stress tensor and f is a source term that accounts for additional contributions such as gravity, which is however generally not considered in our case. The momentum equation is a vector equation and is therefore composed of three differential equations. The role of the pressure term is peculiar for an incompressible fluid as it is not related to density by an equation of state but is obtained from the divergence constraint. The pressure establishes itself instantaneously and must therefore be integrated implicitly. For a viscous and isotropic fluid, the stress/strain rate relationship (although a tensor relation) is usually expressed as an algebraic equation of the form $\tau = \mu \dot{\gamma}$, where μ is the viscosity, while the strain-rate $\dot{\gamma}$ is defined as the second invariant of the strain-rate tensor and depends on first derivatives of the velocity. The system of equations is closed by providing a constitutive law for the local viscosity of the fluid. The simplest rheological model is a Newtonian fluid that assumes a constant viscosity $\mu = \mu_0$ and a linear relationship between stress and strain. Typical values used for blood are $\rho = 10^3 \text{ kg/m}^3$ and $\mu = 4.0 \text{ cP}$. As the caliber of cerebral vessels quickly decreases with the degree of branching and low shear regions are common in intracranial aneurysms, more accurate non-Newtonian models such as power law, Casson's or based on Carreau's law may be used to introduce the shear thinning behavior of blood [38].

The solution of this system of equations requires the specification of boundary conditions (see Sect. 2.5) and of initial conditions $\mathbf{v}(\mathbf{x}, 0) = \mathbf{v}_0(\mathbf{x})$ the value of $\mathbf{v}_0(\mathbf{x})$ is usually chosen arbitrarily, thus the numerical solution of the system of equations may produce initial transients to recover, typically after two or three cardiac cycles, from the incorrect initial data.

The Navier-Stokes equations admit analytical solutions only in very simple conditions such as for a laminar flow through a cylinder (Poiseuille's law), thus we need to resort to numerical techniques to find an approximation of the mathematical problem. The most popular techniques in haemodynamics simulation involve the reformulation and approximation of the system of PDEs in a finite number of equations, which are then numerically solved on computational grids (see Sect. 2.4). The numerical solution of the unknown \mathbf{v} , p requires a spatial discretization and a temporal advancement of the approximation \mathbf{v}_h , p_h . Popular spatial discretization techniques are the finite volume (FV) [31] and finite element (FE) [51] methods. The FV method adopts an integral formulation of the PDEs, called the divergence or conservation form, leading to a system of equations where the approximation \mathbf{v}_h , p_h of \mathbf{v} , p at each element of the grid depends only on the value of \mathbf{v}_h , p_h at neighboring elements, which are thus used as a control volumes of a piecewise linear (or non-linear) approximation. The FE method instead is based on the Galerkin approximation method, which turns the system of PDEs into its weak variational form. The approximation \mathbf{v}_h , p_h is searched in functional subspaces and is expressed as a linear combination of a finite number of basis functions weighted by the values of \mathbf{v} and p at the nodes of the computational grid. The basis functions are typically high-order polynomials, whose degree determines the order of accuracy of the method. In addition, the basis functions have small support so that the value of \mathbf{v}_h , p_h at each node only depends on the values of \mathbf{v}_h , p_h in a limited region of

the grid. Thus, as the finite number of equations is subsequently turned into a linear system $\mathbf{A}\mathbf{u} = \mathbf{b}$, this has the implication that the stiffness matrix \mathbf{A} is sparse and that smaller memory requirements and more efficient solutions can be achieved. Similar considerations can be extended to the FV method. In addition, clearly for both methods the finer the grid the more accurate the solution, but also the higher the computational costs. Higher accuracy could also be achieved by increasing the degree of the polynomial basis as in the spectral elements method [49].

For advance the solution in time, the two main approaches involve either an implicit or an explicit time integration of the Navier-Stokes equations. Implicit schemes involve an update of the solution at each time t_n considering also the unknowns at the new time t_{n+1} , such as for the backward Euler or the Crank-Nicholson schemes, and typically guarantee stability regardless the time step size. Implicit schemes involve the iterative solution of a linear system of equations, which is not required for explicit schemes, where the update at t_{n+1} is directly obtained from the values of \mathbf{v} , p at t_n . On the other hand, explicit schemes, such as the forward Euler method, are subject to a stability condition, which imposes limits to the time step size. More details on the numerical solution of the Navier-Stokes equations can be found in [43, 65, 71]. Novel techniques that are gaining ground in computational haemodynamics include spectral/hp element methods [56] and Lattice-Boltzmann methods [28].

2.7 Visualization and data extraction

The solution of the Navier-Stokes equations provides the values of the velocity and pressure fields at the nodes of the computational grid for all the time steps. A number of techniques may be adopted to investigate the blood flow structure across the vascular domain and to reduce the large amount of data to a set of meaningful quantities describing both the blood flow patterns and the fluid stresses on the vascular wall. In the first case, typical visualization techniques involve the extraction of particle trajectories and velocity streamlines (Fig. 4(a)), or the projection of the three-dimensional vectorial field onto 2D cut planes placed within the vascular domain (Fig. 4(b)).

The visualization of fluid stresses on the surface requires post-processing operations applied to the CFD solution. In particular, WSS represents the viscous frictional force of blood that acts parallel to the vessel wall. The OSI instead measures the degree of angular deviation of the shear stress force with respect to the mean shear stress during a cardiac cycle. The WSS is represented by a 3D vector field lying on tangential planes to the vascular surface at each node and at each time. The WSS magnitude is commonly time-averaged over a cardiac cycle to investigate

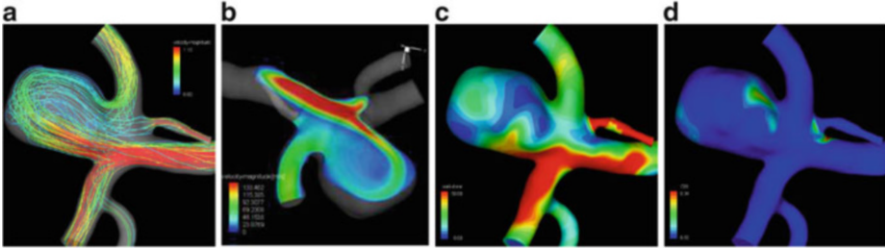


Fig. 4 Visualization of haemodynamics simulation results. (a) Streamlines color-coded by velocity magnitude at peak systole; (b) Velocity magnitude on a cut-plane through the aneurysm at peak systole; (c) Time-averaged WSS distribution; (d) OSI distribution

its distribution across the aneurysm region. The equations used for the calculation of time-averaged WSS ($\bar{\tau}$) and OSI have the form:

$$\bar{\tau} = \int \tau dt = \int \sigma \cdot \mathbf{n} dt \quad (4)$$

$$OSI = \frac{1}{2} \left(1 - \frac{|\bar{\tau}|}{|\tau|} \right) \quad (5)$$

where σ is the strain rate tensor and \mathbf{n} is the surface normal. Surface color maps of time-averaged WSS and the OSI are shown in Fig. 4(c) and (d), respectively.

3 Sensitivity analysis

Blood is a non-Newtonian fluid that circulates in vessels with distensible walls and in a patient-specific pulsatile regime that may change over time. Sensitivity analysis aims at both assessing the variability of the simulation to assumptions, approximations and uncertainties in the modeling process and at potentially providing confidence intervals to the simulation output. Our groups have recently conducted a sensitivity analysis in four models of intracranial aneurysms [12] selected from a database of 40 anatomical models obtained from CTA and 3DRA images. The results showed that a qualitative characterization of the intra-aneurysmal flow patterns is robust to variations to the mean input flow, the outflow division, the viscosity model or to the grid resolution, while it strongly depends on the geometry of the parent vessel and on the shape and neck size of the aneurysm. Conversely, an accurate quantification of the intra-aneurysmal flow patterns and haemodynamical forces may require the availability of patient-specific flow boundary conditions and the use of a non-Newtonian rheology model. Without patient-specific information, it was concluded that sensitivity analyzes should be routinely performed and that a major effort should be paid to obtain accurate geometrical models. Similar observations were reported in [10, 11, 68].

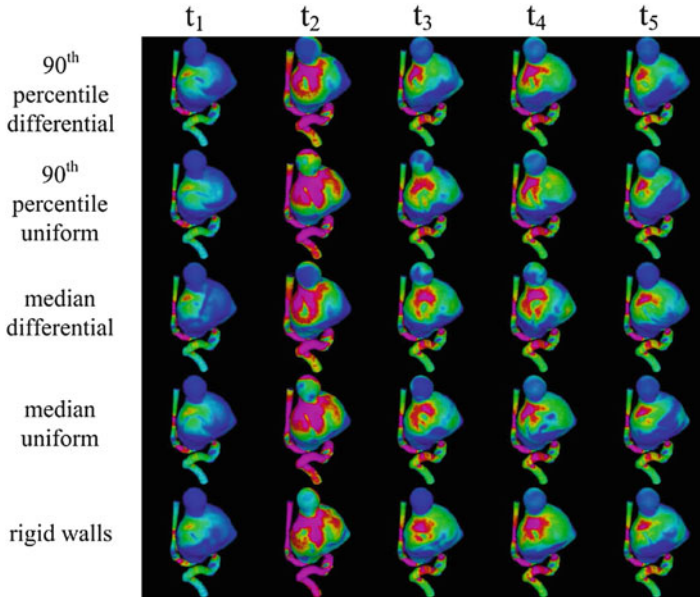


Fig. 5 Maps of WSS for different distributions of wall displacement at some time points over the cardiac cycle

Wall compliance has been indicated as an important factor altering the local haemodynamics. In [18], we have introduced a method based on 2D non-rigid registration to estimate the wall motion of aneurysm dome and neighboring vessels from dynamic biplane Digital Subtraction Angiography (DSA). The method was extended in [46] by post-processing the recovered motion in the Fourier domain and further applied to cases providing higher acquisition frame rate. Wall motion curves were compared to blood pressure waveforms confirming that the measured values corresponded to real displacement. Wall compliance was then integrated in flow simulations by imposing the measured wall displacement directly to the 3D mesh, thus circumventing the difficulties in estimating personalized elasticity properties in a fluid-solid interaction approach. Due to the two-dimensional nature of DSA images, some motion patterns were imposed to the flow simulations and compared to a rigid wall condition. It was observed that the area of the aneurysm under elevated WSS with respect to the average WSS in the proximal parent vessel, the contribution to the total shear force of this region, and the shear force concentration factor were relatively unaffected by the wall motion. In addition, changing the amplitude of the wall motion or imposing differential rather than uniform deformations (Fig. 5) did not have a considerable effect on these variables, although the impact of the real three-dimensional displacement field (including the possible angular movement of parent arteries) should be considered in further studies.

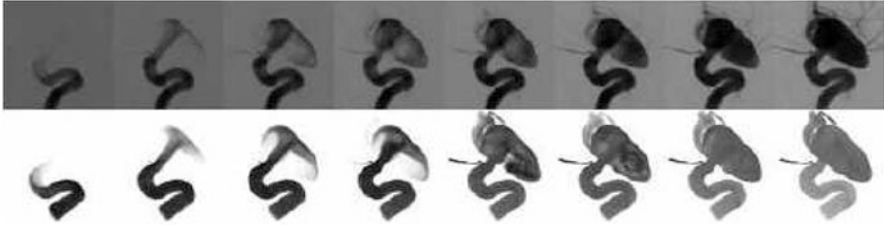


Fig. 6 Conventional (top row) and virtual (bottom row) angiograms during the filling phase of an aneurysm

4 Validation

The introduction of blood flow simulation technology into routine clinical practice will eventually require its validation against ground truth *in vivo* flow measurements. Due to the absence of *in vivo* ground truth data, some authors have compared simulations results with experimental flow measurements in anatomically realistic replicas [22], while others have confronted their ability to reproduce the gross haemodynamical features observed during routine clinical examinations. In particular, Cebal [14] compared the flow structure described by signal isointensity surfaces in TOF MRA images used in [55] with high flow velocity isosurfaces extracted from haemodynamics simulation data obtained starting from the same MRA images, finding good agreement in the aneurysmal inflow region. In two separate works, Ford [23] and Calamante [8] presented a strategy to simulate the transport of the contrast agent from the cerebral vessels through to an intracranial aneurysm and generate visualizations similar to conventional angiography. This technique is called virtual angiography and allows to qualitatively compare simulation results with high frame rate angiographic images routinely acquired during treatment. We have recently applied this technique to assess the reliability of haemodynamics simulation in three aneurysm models extracted from 3DRA images [16]. The approximated time-dependent velocity fields were used to simulate the transport of a contrast agent by solving the transport (or advection-diffusion) equation using an implicit finite element formulation on unstructured grids. Virtual angiograms were then constructed by volume rendering of the simulated contrast concentration field. As depicted in Fig. 6, the virtual angiograms showed good agreement with the conventional angiograms. Analogous size and orientation of the inflow jet, regions of flow impaction, major intra-aneurysmal vortices and regions of outflow were observed in both the conventional and virtual angiograms. Similar conclusions were drawn by Ford [23], thus supporting the ability of patient-specific image-based computational models to realistically reproduce the major intra-aneurysmal flow structures observed with conventional angiography.

5 Current applications and future work

Image-based haemodynamics simulation is a computational technique that combines image processing, anatomical modeling and CFD simulation to recover personalized vascular geometry and approximate blood flow characteristics starting from medical images. Advances in image quality, algorithmic sophistication and computing power are enabling the introduction of such technology not only as a biomedical research tool but also for clinical evaluation. However, due to the absence of ground truth data, it still remains unclear how accurately current combined image processing and modeling techniques truly recover the *in vivo* geometry and what degree of geometrical and numerical accuracy is required to obtain a clinically meaningful prediction. In particular, the identification of diagnostic and prognostic flow descriptors is a first necessary step to further establish the allowed range of variability of the simulation results. Towards this goal, some investigators have initiated international projects whose aim is to provide comprehensive models of human physiology and pathology starting from biochemical pathways through to cells, tissues and organs [1, 2]. A complementary approach is instead to generate hypotheses on relevant flow descriptors of the disease from observations of clinically measured and simulated data. Although still in its infancy, this top-down methodology is being used to investigate the possible association between intracranial aneurysmal haemodynamical characteristics and rupture in samples of intracranial aneurysm models. For example, Cebal [13] reported that aneurysms with complex or unstable flow patterns, small impingement regions and small jet sizes may be more likely to experience rupture. Radaelli [53], analyzing pairs of ruptured and unruptured mirror aneurysms, observed that wide angles of bifurcation, longer and highly curved parent vessels, irregular aneurysmal WSS distribution and small impingement jets may be more common in ruptured aneurysms. From observations on the location of high and low WSS regions, Shojima [57] suggested that high WSS may have an important role in the initiation of the disease, while low WSS may facilitate the growing phase and trigger the rupture event.

Further work towards the creation of larger datasets of aneurysm models has the potential to facilitate the execution of clinical trials and help the evaluation of novel endo vascular devices. Our immediate efforts are directed to the development of a highly controlled and interactive environment to facilitate the use of our software routinely in clinical settings by non-expert users. An appropriate integration into a clinical workstation will also require the availability of multimodal registration and fusion techniques to merge the information provided by different image modalities and to facilitate the specification of boundary conditions. It is clear that progress in modeling and computing facilities need to be accompanied with stricter requirements in image quality. It can also be foreseen that improvements in acquisition systems may eventually lead to more reliable *in vivo* descriptions of aneurysm haemodynamics. Depending on the level of detail, these measurements

may be used to offer accurate flow boundary conditions, for the validation of simulation techniques or directly for diagnostic purposes. Image-based haemodynamics simulation could then be used in combination with *in vivo* measurements to offer prognostic information. Among our current interests is to progress in the implementation of effective techniques to simulate blood flow around endo vascular devices [4] and to further integrate models of blood clotting and tissue perfusion to study the association between haemodynamics and treatment failure.

Among the new ventures ahead of investigators working in the field, we could also cite the integration of haemodynamics information with imaging data capturing cellular activities. This process has the potential to enable both the investigation of the correlation between haemodynamics and biomarkers expressed at the cellular level and the integration of image-based haemodynamics simulation with predictive models developed within the framework of the STEP (A Strategy Towards the Europhysiome) roadmap initiative [1, 2].

References

1. STEP: A strategy towards the EuroPhysiome. <http://www.europhysiome.org>.
2. Towards virtual physiological human: multilevel modelling and simulation of the human anatomy and physiology. White Paper DG INFSO & DG JRC, 2005.
3. J. Alastruey, K. Parker, J. Peiró, S. Byrd, and S. Sherwin. Modelling the circle of Willis to assess the effects of anatomical variations and occlusions on cerebral flows. *Journal of Biomechanics*, 40(8):1794–1805, 2007.
4. S. Appanaboyina, F. Mut, R. Lohner, C. Putman, and J. Cebral. Computational fluid dynamics of stented intracranial aneurysms using adaptive embedded unstructured grids. *International Journal for Numerical Methods in Fluids*, In press, 77, 2007.
5. M. Attene and B. Falcidieno. ReMESH: An Interactive Environment to Edit and Repair Triangle Meshes. *SMI06: Proceedings of the IEEE International Conference on Shape Modeling and Applications 2006 (SMI06)*.
6. J. Bloomenthal. An implicit surface polygonizer. *Graphics Gems IV*, pages 324–349, 1994.
7. A. Burlison and V. Turitto. Identification of quantifiable hemodynamic factors in the assessment of cerebral aneurysm behavior. On behalf of the Subcommittee on Biorheology of the Scientific and Standardization Committee of the ISTH. *Thromb Haemost*, 76(1):118–23, 1996.
8. F. Calamante, P. Yim, and J. Cebral. Estimation of bolus dispersion effects in perfusion MRI using image-based computational fluid dynamics. *Neuroimage*, 19(2):341–353, 2003.
9. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
10. M. Castro, C. Putman, and J. Cebral. Computational Fluid Dynamics Modeling of Intracranial Aneurysms: Effects of Parent Artery Segmentation on Intra-Aneurysmal Hemodynamics. *American Journal of Neuroradiology*, 27(8):1703–1709, 2006.
11. M. Castro, C. Putman, and J. Cebral. Patient-Specific Computational Modeling of Cerebral Aneurysms With Multiple Avenues of Flow From 3D Rotational Angiography Images. *Academic Radiology*, 13(7):811–821, 2006.
12. J. Cebral, M. Castro, S. Appanaboyina, C. Putman, D. Millan, and A. Frangi. Efficient Pipeline for Image-Based Patient-Specific Analysis of Cerebral Aneurysm Hemodynamics: Technique and Sensitivity. *IEEE Transactions on Medical Imaging*, 24(4):457–467, 2005.

13. J. Cebral, M. Castro, J. Burgess, R. Pergolizzi, M. Sheridan, and C. Putman. Characterization of Cerebral Aneurysms for Assessing Risk of Rupture By Using Patient-Specific Computational Hemodynamics Models. *American Journal of Neuroradiology*, 26(10):2550–2559, 2005.
14. J. Cebral, M. Castro, T. Satoh, and J. Burgess. Evaluation of image-based CFD models of cerebral aneurysms using MRI. *ISMRM Flow Motion Workshop, Zurich, Switzerland, July*, pages 11–13, 2004.
15. J. Cebral, R. Lohner, and J. Burgess. Computer Simulation of Cerebral Artery Clipping: Relevance to Aneurysm Neuro-Surgery Planning. *Proc. ECCOMAS*, pages 11–14, 2000.
16. J. Cebral, A. Radaelli, A. Frangi, and C. Putman. Qualitative comparison of intra-aneurysmal flow structures determined from conventional and virtual angiograms. *Proceedings of SPIE*, 6511:65111E, 2007.
17. I. Chatziprodromou, V. Butty, V. Makhijani, D. Poulidakos, and Y. Ventikos. Pulsatile Blood Flow in Anatomically Accurate Vessels with Multiple Aneurysms: A Medical Intervention Planning Application of Computational Haemodynamics. *Flow, Turbulence and Combustion*, 71(1):333–346, 2003.
18. L. Dempere-Marco, E. Oubel, M. Castro, C. Putman, A. Frangi, and J. Cebral. CFD Analysis Incorporating the Influence of Wall Motion: Application to Intracranial Aneurysms. *Lecture notes in computer science*, 4191:438–445, 2006.
19. D. Doorly, S. Sherwin, P. Franke, and J. Peiró. Vortical Flow Structure Identification and Flow Transport in Arteries. *Computer Methods in Biomechanics and Biomedical Engineering*, 5(3):261–273, 2002.
20. U. Ernemann, E. Gronewaller, F. Duffner, O. Guervit, J. Claassen, and M. Skalej. Influence of Geometric and Hemodynamic Parameters on Aneurysm Visualization during Three-Dimensional Rotational Angiography: An in Vitro Study. *American Journal of Neuroradiology*, 24(4):597–603, 2003.
21. M. Ford, N. Alperin, S. Lee, D. Holdsworth, and D. Steinman. Characterization of volumetric flow rate waveforms in the normal internal carotid and vertebral arteries. *Physiological Measurement*, 26(4):477–488, 2005.
22. M. Ford, H. Nikolov, J. Milner, S. Lownie, E. DeMont, W. Kalata, F. Loth, D. Holdsworth, and D. Steinman. PIV-Measured Versus CFD-Predicted Flow Dynamics in Anatomically-Realistic Cerebral Aneurysm Models. *Journal of Biomechanical Engineering, In press*, 2007.
23. M. Ford, G. Stuhne, H. Nikolov, D. Habets, S. Lownie, D. Holdsworth, and D. Steinman. Virtual angiography for visualization and validation of computational models of aneurysm hemodynamics. *IEEE Transactions on Medical Imaging*, 24(12):1586–1592, 2005.
24. G. Foutrakis, H. Yonas, and R. Scلابassi. Saccular Aneurysm Formation in Curved and Bifurcating Arteries. *American Journal of Neuroradiology*, 20(7):1309–1317, 1999.
25. P. T. J. Franke. *Blood flow and transport in artificial devices*. PhD thesis, Imperial College London, 2002.
26. S. Glagov, C. Zarins, D. Giddens, and D. Ku. Hemodynamics and atherosclerosis: insights and perspectives gained from studies of human arteries. *Archives of pathology & laboratory medicine(1976)*, 112(10):1018–1031, 1988.
27. Y. Gobin, J. Counord, P. Flaud, and J. Duffaux. In vitro study of haemodynamics in a giant saccular aneurysm model: influence of flow dynamics in the parent vessel and effects of coil embolisation. *Neuroradiology*, 36(7):530–536, 1994.
28. S. Harrison, J. Bernsdorf, D. Hose, and P. Lawford. Development of a Lattice Boltzmann Framework for Numerical Simulation of Thrombosis. *International Journal of Modern Physics C*, 18(04):483–491, 2007.
29. T. Hassan, E. Timofeev, T. Saito, H. Shimizu, M. Ezura, T. Tominaga, A. Takahashi, and K. Takayama. Computational Replicas: Anatomic Reconstructions of Cerebral Vessels as Volume Numerical Grids at Three-Dimensional Angiography. *American Journal of Neuroradiology*, 25(8):1356–1365, 752.
30. M. Hernandez and A. Frangi. Non-parametric geodesic active regions: Method and evaluation for cerebral aneurysms segmentation in 3DRA and CTA. *Medical Image Analysis*, 11(3): 224–241, 2007.

31. T. Hino. Computation of Free Surface Flow Around an Advancing Ship by the Navier-Stokes Equations. *Proceedings: Fifth International Conference on Numerical Ship Hydrodynamics. Hiroshima, Japan*, 1989.
32. T. Hirai, Y. Korogi, K. Ono, M. Yamura, S. Uemura, and Y. Yamashita. Pseudostenosis Phenomenon at Volume-rendered Three-dimensional Digital Angiography of Intracranial Arteries: Frequency, Location, and Effect on Image Evaluation. *Radiology*, 232:882–887, 2004.
33. Y. Hoi, H. Meng, S. Woodward, B. Bendok, R. Hanel, L. Guterman, and L. Hopkins. Effects of arterial geometry on aneurysm growth: three-dimensional computational fluid dynamics study. *J Neurosurg*, 101(4):676–681, 2004.
34. J. Hop, G. Rinkel, A. Algra, and J. van Gijn. Case-Fatality Rates and Functional Outcome After Subarachnoid Hemorrhage A Systematic Review. *Stroke*, 28(3):660–664, 1997.
35. L. Jou, A. Mohamed, D. Lee, and M. Mawad. 3D Rotational Digital Subtraction Angiography May Underestimate Intracranial Aneurysms: Findings from Two Basilar Aneurysms. *American Journal of Neuroradiology*, 28(9):1690–1692, 2007.
36. L. Jou, C. Quick, W. Young, M. Lawton, R. Higashida, A. Martin, and D. Saloner. Computational Approach to Quantifying Hemodynamic Forces in Giant Cerebral Aneurysms. *American Journal of Neuroradiology*, 24(9):1804–1810, 168.
37. K. Kayembe, M. Sasahara, and F. Hazama. Cerebral aneurysms and variations in the circle of Willis. *Stroke*, 15(5):846–850, 1984.
38. S. Lee and D. Steinman. On the Relative Importance of Rheology for Image-Based CFD Models of the Carotid Bifurcation. *Journal of Biomechanical Engineering*, 129:273–278, 2007.
39. B. Lieber and M. Gounis. The physics of endoluminal stenting in the treatment of cerebrovascular aneurysms. *Neurological Research*, 24:33–42, 2002.
40. T. Liou and S. Liou. A review of in vitro studies of hemodynamic characteristics in terminal and lateral aneurysm models. *Proc. Natl. Sci. Counc. ROC (B)*, 23(4):133–148, 1999.
41. R. Löhner. Regridding Surface Triangulations. *Journal of Computational Physics*, 126(1):1–10, 1996.
42. R. Löhner. Automatic unstructured grid generators. *Finite Elements in Analysis & Design*, 25(1-2):111–134, 1997.
43. R. Löhner. Renumbering strategies for unstructured-grid solvers operating on shared-memory, cache-based parallel machines. *Computer Methods in Applied Mechanics and Engineering*, 163(1-4):95–109, 1998.
44. J. Myers, J. Moore, M. Ojha, K. Johnston, and C. Ethier. Factors Influencing Blood Flow Patterns in the Human Right Coronary Artery. *Annals of Biomedical Engineering*, 29(2):109–120, 2001.
45. S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79(1):12–49, 1988.
46. E. Oubel, M. De Craene, C. Putman, J. Cebral, and A. Frangi. Analysis of intracranial aneurysm wall motion and its effects on hemodynamic patterns. *Proceedings of SPIE*, 6511:65112A, 2007.
47. S. Owen. A survey of unstructured mesh generation technology. *7th International Meshing Roundtable*, 3(6), 1998.
48. N. Paragios and R. Deriche. Geodesic Active Regions and Level Set Methods for Supervised Texture Segmentation. *International Journal of Computer Vision*, 46(3):223–247, 2002.
49. A. Patera. A spectral element method for fluid dynamics-Laminar flow in a channel expansion. *Journal of Computational Physics*, 54:468–488, 1984.
50. J. Peiro, L. Formaggia, M. Gazzola, A. Radaelli, and V. Rigamonti. Shape reconstruction from medical images and quality mesh generation via implicit surfaces. *International Journal for Numerical Methods in Fluids*, 53(8):1339–1360, 2007.
51. A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*. Springer, Heidelberg, Germany, 2000.
52. T. Raaymakers, G. Rinkel, M. Limburg, and A. Algra. Mortality and Morbidity of Surgery for Unruptured Intracranial Aneurysms A Meta-Analysis. *Stroke*, 29(8):1531–1538, 1998.

53. A. Radaelli, T. Martínez, E. Díaz, X. Mellado, M. Castro, C. Putman, L. Guimaraens, J. Cebal, and A. Frangi. Combined clinical and computational information in complex cerebral aneurysms: application to mirror cerebral aneurysms. *Proceedings of SPIE*, 6511:65111F, 2007.
54. T. Satoh, C. Ekino, and C. Ohsako. Transluminal color-coded three-dimensional magnetic resonance angiography for visualization of signal Intensity distribution pattern within an unruptured cerebral aneurysm: preliminarily assessment with anterior communicating artery aneurysms. *Neuroradiology*, 46(8):628–634, 2004.
55. T. Satoh, K. Onoda, and S. Tsuchimoto. Visualization of Intraaneurysmal Flow Patterns with Transluminal Flow Images of 3D MR Angiograms in Conjunction with Aneurysmal Configurations. *American Journal of Neuroradiology*, 24(7):1436–1445, 2004.
56. S. Sherwin, J. Peiró, O. Shah, G. Karamanos, and D. Doorly. Computational haemodynamics: geometry and non-newtonian modelling using spectral/hp element methods. *Computing and Visualization in Science*, 3(1):77–83, 2000.
57. M. Shojima, M. Oshima, K. Takagi, R. Torii, M. Hayakawa, K. Katada, A. Morita, and T. Kirino. Magnitude and Role of Wall Shear Stress on Cerebral Aneurysm Computational Fluid Dynamic Study of 20 Middle Cerebral Artery Aneurysms. *Stroke*, 35(11):2500–2505, 2004.
58. D. Steinman, J. Milner, C. Norley, S. Lownie, and D. Holdsworth. Image-Based Computational Simulation of Flow Dynamics in a Giant Intracranial Aneurysm. *American Journal of Neuroradiology*, 24(4):559–566, 2003.
59. S. Tateshima, Y. Murayama, J. Villablanca, T. Morino, K. Nomura, K. Tanishita, and F. Vinuela. In Vitro Measurement of Fluid-Induced Wall Shear Stress in Unruptured Cerebral Aneurysms Harboring Blebs. *Stroke*, 34(1):187–192, 2003.
60. S. Tateshima, K. Tanishita, H. Omura, J. Villablanca, and F. Vinuela. Intra-Aneurysmal Hemodynamics during the Growth of an Unruptured Aneurysm: In Vitro Study Using Longitudinal CT Angiogram Database. *American Journal of Neuroradiology*, 28(4):622, 2007.
61. G. Taubin. A signal processing approach to fair surface design. *Computer Graphics*, 29: 351–358, 1995.
62. C. Taylor, T. Hughes, and C. Zarins. Finite element modeling of blood flow in arteries. *Computer Methods in Applied Mechanics and Engineering*, 158(1-2):155–196, 1998.
63. H. Tenjin. Evaluation of intraaneurysmal blood velocity by time-density curve analysis and digital subtraction angiography. *American Journal of Neuroradiology*, 19(7):1303–1307, 1998.
64. B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever. Scale space: its natural operators and differential invariants. *Information Processing in Medical Imaging: 12th International Conference, IPMI'91, Wye, UK, July 7-12, 1991: Proceedings*, 1991.
65. F. Thomasset. *Implementation of Finite Element Methods for Navier-Stokes Equations*. Springer-Verlag series in computational physics, 1981.
66. F. Tomasello, D. D'Avella, F. Salpietro, and M. Longo. Asymptomatic aneurysms. Literature meta-analysis and indications for treatment. *J Neurosurg Sci*, 42(1 Suppl 1):47–51, 1998.
67. A. Valencia. Flow dynamics in Models of Intracranial Terminal Aneurysms. *Mech Chem Biosystems*, 1:221–231, 2004.
68. A. Valencia, A. Zarate, M. Galvez, and L. Badilla. Non-Newtonian blood ow dynamics in a right internal carotid artery with a saccular aneurysm. *International Journal of Numerical Methods in Fluids*, 50:751–764, 2006.
69. W. van Rooij and M. Sluzewski. Procedural Morbidity and Mortality of Elective Coil Treatment of Unruptured Intracranial Aneurysms. *American Journal of Neuroradiology*, 27(8): 1678–1680, 2006.
70. J. Wardlaw and P. White. The detection and management of unruptured intracranial aneurysms. *Brain*, 123(2):205–221, 2000.
71. P. Wesseling. *Principles of Computational Fluid Dynamics*. Springer, 2001.
72. S. Wetzel, S. Meckel, A. Frydrychowicz, L. Bonati, E. Radue, K. Scheffler, J. Hennig, and M. Markl. In Vivo Assessment and Visualization of Intracranial Arterial Hemodynamics with Flow-Sensitized 4D MR Imaging at 3T. *American Journal of Neuroradiology*, 28(3):433, 2007.

Part III
Biomedical Perception

Atlas-based Segmentation

M. Bach Cuadra, V. Duay, and J.-Ph. Thiran

Abstract Image segmentation is a main task in many medical applications such as surgical or radiation therapy planning, automatic labelling of anatomical structures or morphological and morphometrical studies. Segmentation in medical imaging is however challenging because of problems linked to low contrast images, fuzzy object-contours, similar intensities with adjacent objects of interest, etc. Using *prior knowledge* can help in the segmentation task. A widely used method consists to extract this prior knowledge from a *reference image* often called *atlas*. We review in this chapter the existing approaches for atlas-based segmentation in medical imaging and we focus on those based on a volume registration method. We present the problem of using atlas information for pathological image analysis and we propose our solution for atlas-based segmentation in MR image of the brain when large space-occupying lesions are present. Finally, we present the new research directions that aim at overcome current limitations of atlas-based segmentation approaches based on registration only.

M. Bach Cuadra (✉)

Department of Radiology, Center for Biomedical Imaging, Lausanne University Center (CHUV) and University of Lausanne (UNIL), Rue du Bugnon 46, BH 08-079, Lausanne, Vaud 1011, Switzerland
e-mail: meritxell.bachcuadra@unil.ch

V. Duay

Department of Industrial Technology, University of Applied Sciences Western Switzerland Technology, Architecture and Landscape, Rue de la Prairie 4, Geneva 1202, Switzerland
e-mail: valerie.duay@hesge.ch

J.-P. Thiran

Signal processing Laboratory (LTS5), Swiss Federal Institute of Technology Lausanne (EPFL), EPFL-STI-IEL-LTS5, Station 11, Lausanne 1015, Switzerland
e-mail: jp.thiran@epfl.ch

1 Introduction

A key research area in medical image analysis is *image segmentation*. Image segmentation aims at extracting *objects of interest*¹ in images or video sequences. These objects are *image structures* containing relevant information for a given application. Some examples of medical applications implying a segmentation task are surgical or radiation therapy planning [48, 107], automatic labelling of anatomical structures [5, 29] or morphological and morphometrical studies [49].

However, image segmentation is not an easy task. In many cases, the contours of the objects of interest are difficult to delineate, even manually. The problems linked to segmentation are often due to low contrast, fuzzy contours or too similar intensities with adjacent objects. Fortunately, using *prior knowledge* can help in the segmentation problem. A widely used method consists to extract this prior knowledge from a *reference image* often called *atlas*. The goal of the atlas is to describe the image to segment much like a map would describe the components of a geographical area. In medical imaging, an atlas gives an estimation of the object position in the image. This spatial information permits to save a lot of processing time in the localization of the objects to extract and it allows distinguishing the objects of interest from other objects with similar features. Moreover, an atlas can also bring information that concerns the texture and the shape of the researched objects. This information allows detecting the object contour in the area of interest. Finally, the atlas can point out the features of adjacent objects. This last information allows to better distinguishing the objects of interest from their neighborhood.

Until recently, atlases were paper-based. The development of digital image processing techniques allowed the creation of digital versions of these atlases. Digital atlases have an increased potential: they provide a lot of details and may be used in a number of embedded software-based or computer-based applications (such as computer assisted diagnosis, planning and guidance of surgical procedures, statistical studies, etc.). While geographical atlases represent one region only, anatomical atlas can characterize a group of individuals. This is due to the consistence between anatomical structures of a same type. Biological images, especially medical images, are thus particularly well suited for atlas-based segmentation methods. Digitized atlases have proven their success in medical image analysis notably in atlas-based segmentation of the head and neck [23, 95], the heart [81], the prostate [117] or the knee [55]. However, the large majority of the works done so far on digitized medical atlases concern the human cerebral anatomy. This is why we have mainly focused the brief digitized atlases survey presented here after on human brain.

¹Objects of interest can be for instance points, lines, surfaces or volumes.

1.1 Digitized Atlases of Human Brain

Determinist atlases. First attempts of atlas construction of the human brain are based on a single subject. This type of atlas is called *single-subject based atlas* or *determinist atlas*. It often corresponds to a reference system or volume image that has been selected among a data set to be representative of the objects to segment in other images (average size, shapes or intensity).

In medicine, the pioneer atlas reference system is the Talairach atlas [97] that was proposed to identify deep brain structures in stereotaxic coordinates. The first determinist digital atlas was proposed by the Visible Human Project of the National Library of Medicine [72]. The goal of this project was the creation of complete and detailed three-dimensional anatomical representations of the normal male and female human bodies. These representations were obtained from the acquisition of transverse CT, MR and cryosection high resolution images of representative male and female cadavers. Also derived from a digitized cryosectioned human brain, the Karolinska Institute and Hospital, Stockholm, created a *Computerized Brain Atlas (CBA)* that was designed for display and analysis of tomographic brain images. The atlas includes the brain surface, the ventricular system and about 400 structures and all Brodmann areas [46, 99].

But the large majority of the determinist atlases are created directly from imaging acquisition. For instance, the *digital brain atlas* developed by the Harvard Medical School [58] is based on a 3D MR digitized atlas of the human brain to visualize spatially complex structures. Also, the *digital brain phantom* from McConell Brain Imaging Center [21] is based on 27 high-resolution scans of the same individual. Its average resulted in a high-resolution (1mm isotropic voxels) brain atlas with an increased signal-to-noise ratio. This template is the reference data within the BrainWeb simulator [68]. Bajcsy et al. in [9, 11] created an artificial CT anatomical volume based on the stained slices of a dead soldier brain of a 31 year old normal male (from the so-called Yakovlev Collection).

Statistical atlases. Atlases based on a single-subject are however not representative of the diversity human anatomy. To better characterize the variability of the anatomical structures, atlases have been constructed based on a population. This second type of atlases is called *population-based* or *statistical atlas*. Such atlases are in continuous evolution since new images can be easily incorporated. Moreover, the population that a statistical atlas represents can be easily subdivided into groups according to specific criteria (age, sex, handedness, etc.). Initial population-based atlases were based in Talairach space [51, 111] Later, to compensate for implicit limitations of Talairach space based atlases such as poor resolution across slices (from 3 to 4 mm), population-based atlases from MR images were proposed. A composite MRI data set was constructed by Evans et al. [39] from several hundreds of normal subjects (239 males and 66 females of 23.4 ± 4.1 years old). All the scans were first individually registered into the Talairach coordinate system. Then, they were intensity normalized and, finally, all the scans were averaged voxel-by-voxel and probabilistic maps for brain tissue were created. The same procedure

for constructing an average brain was used later by the International Consortium for Brain Mapping (ICBM) to 152 brains [60]. Recently, *disease-based atlases* [38, 67] have been created. They represent subgroups of some disease instead of using a healthy group of subjects. Such atlases provide the way to examine the history and evolution (due to natural disease evolution or reaction to clinical treatment) of a specific disease.

Important questions arise when generating population-based atlases such as the selection of a reference space or the registration method for the data alignment. Many researchers have proposed new strategies to create unbiased average templates and multi-subject registration as in [12, 23, 31, 47, 54, 63, 80, 118].

As conclusion, the term atlas is rather used in the literature to designate a reference image generally composed of an intensity image, *the intensity atlas* and a labeled image, *the labeled atlas*). However, as in [2], active shape models [25] or active appearance models [26] can also be considered as atlases since they bring spatial prior knowledge in a segmentation process.

2 Atlas Driven Segmentation

To exploit the information given by an atlas in a segmentation task, we need to register it with the target image. Global image registration is usually applied first in order to compensate for the difference of position, orientation and possibly size between both images. It can consist in a Talaraich registration (piecewise affine registration from anterior and posterior commissures), or an affine registration (12 parameters). Then, when more precise localization or delineation is needed, deformations with a higher degree of freedom should be applied. In fact, this second step aims to compensate for the anatomical variability between both images. Thus, atlas-based segmentation is most often considered as a registration problem only. However, many examples in the literature show that the atlas-based segmentation problem can also include a statistical segmentation or a contour morphing problem. Let us describe here after these three different approaches of *atlas-based segmentation*. Note that we will focus on this one based on registration.

2.1 A Statistical Classification Approach

Statistical methods usually solve the classification problem by either assigning a class label to a voxel or by the estimation of the relative amounts of the various class types within a voxel [77, 96]. Such decision is often taken in a Bayesian framework where *prior probability* is used to model the spatial coherence of the images and *conditional probability* models the image intensity distribution for each class [4, 90, 102, 109].

In such probabilistic framework, prior probabilities from statistical atlas are easily incorporated as shown in [41, 73, 84, 103]. Thus, statistical classification approaches can be considered as atlas-based segmentation approaches when they incorporate spatial probabilistic information from an atlas. (*probably also refer here to previous chapter on statistical classification*).

2.2 A Contour Morphing Approach

We understand atlas-based segmentation as a morphing problem when the atlas contours selected by the labeled atlas are directly deformed without extracting a dense deformation field in the whole image [1, 8, 32, 35]. The contour matching techniques that has attracted the most attention in atlas-based segmentation are the active contour models [56].

The atlas-based segmentation process via active contours is the following. After global matching of the atlas with the image to segment, and also sometimes an approximative local registration to bring the atlas contour closer to the target contour, the active contour segments the target image by combining the prior information given by the deformed labeled atlas and the intensity atlas. The advantage of this technique is its low computational cost. This is due to the fact that it computes the deformation of the contours of interest only, not of the whole image. Also, as it deforms the contours, it is directly designed to use local contour-based information. However, with this method, the segmentation of several contours, open contours or not visible contours is not as straightforward as with the atlas-based segmentation via a registration approach presented in the next section.

2.3 A Registration Approach

The most known atlas-based segmentation approach consists to reduce the segmentation problem to an image registration problem (see [85] for a recent survey). To segment a new image, a dense deformation field that puts the atlas into a point-to-point spatial correspondence with the target image is first computed. This transformation is then used to project the labels assigned to structures from the atlas onto the target image to be segmented. As mentioned above, the atlas is first globally registered to the volume of interest and in a second step, a more local registration is applied to compensate for the variability between both images. The main advantage of this approach is that the dense deformation field, interpolated on the whole image from the registration of visible image features, permits one to easily estimate, in the target image, the position of structures with fuzzy or not visible contours. Moreover, this approach permits one to segment simultaneously several contours of any types (closed, open, connected or disconnected).

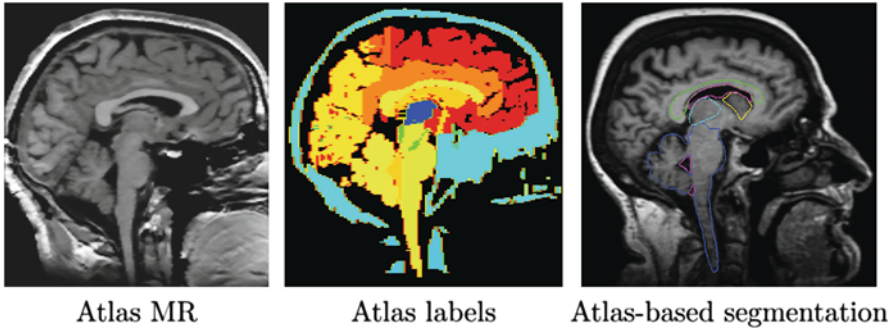


Fig. 1 Atlas MRI and labeled image from Harvard Medical School [58]. Atlas-based segmentation results (ventricles, corpus callosum, thalamus, trunk and cerebellum) using a tandem of affine registration and Demons algorithm as proposed in [5]

In fact, a wide range of image registration techniques that allow to deform a given atlas to a subject have been developed over the last 20 years [5, 10, 14, 18, 22, 29, 43, 53, 86, 98]. It is though beyond the scope of this chapter to review all them and we refer the interested reader to [66, 85, 100]. Mathematically, they formulate the registration as an optimization problem where two images, the intensity atlas and the *target* image to be segmented are matched to each other. Then, the space of allowed transformations (parametric or non-parametric), the similarity measure and the optimization strategy are defined. Local transformations must usually add to the similarity criteria a *regularization* term to ensure the smoothness of the transformation while global transformations, defined by an analytical expression, explicitly fulfill such *physical* constraint. However, a drawback of image registration methods is that they often lead to a compromise between the accuracy of the registration and the smoothness of the deformation. When at some places the registration is not accurate enough, a widely-used solution is to globally or locally allow more variability in the registration model in order to obtain more local deformation, but with the risk of creating irregularities in the deformation field. Also, this does not assure that the desired level of precision will be obtained. To cope with this problem, local constraints should be included in the registration process. Novel types of atlas registration algorithm that aim to overcome these limitations are presented in Sect. 4.

As for image registration, the assessment of atlas-based segmentation is still an open question. First, the atlas must represent the subject anatomy as well as possible. Then, the accuracy of the registration algorithm in capturing the anatomical variability between the atlas and the subject is evaluated. It is though not evident to define a quantitative *accuracy* measure for validation in the particular case of inter-subject non-rigid registration. In [108], Warfield et al. proposed a binary minimum entropy criterion that allows the identification of an intrinsic coordinate system of the subjects under study. Schnabel et al. [89] propose a biomechanical *gold standard* of non-rigid registration based on finite element methods. Actually, validation of

image registration is often *application-based* and, usually, the evaluation process is only presented as a subsection of the work. However, some comparative studies are published concerning the evaluation of different non-rigid registration techniques. For instance, West et al. [110] present a survey and comparison of multimodal registration techniques. The same is done by Hellier et al. [50] for inter-subject image registration. Recently, Dawant et al. in [30] and Sanchez et al. in [88] proposed a validation study of atlas-based segmentation technique in the framework of subthalamic nuclei localization for the Parkinson surgery.

As mentioned above, a main assumption of the atlas-based segmentation process using registration is that the images to segment are consistent with the atlas. However, possible inconsistencies between two images to be registered can exist. We can distinguished two types of inconsistencies: the intensity-based and the content-based. Trying to find a point to point correspondence between inconsistent images is a challenging task. In the literature several methods have been proposed to remove these inconsistencies. For example, different intensity ranges or the presence of a bias field in the intensity of the target images can be corrected in a pre-processing step by an histogram equalization [91] or a bias correction algorithm [61]. Some similarity measures like gradient-based [15], joint-entropy [20, 94] or mutual information [65, 82, 106] permit the registration of images from different modalities. For content-based inconsistencies, geometrical models have been proposed either to introduce the inconsistent object in the atlas [6, 7, 83] or to force corresponding objects to match [64].

3 Atlas-based Segmentation in Pathological Cases

Atlas-based segmentation has been of limited use in presence of large space-occupying lesions. In fact, brain deformations induced by such lesions may dramatically shift and deform functionally important brain structures. In [7, 83] we presented the problem of inter-subject registration of MR images with large tumors, inducing a significant shift of surrounding anatomical structures. In such pathological cases, the goal of the registration becomes even more complex: it not only tries to capture the normal anatomical variability between subjects but also the deformation induced by the pathology. Moreover, the anatomical meaningful correspondence assumption done in the atlas-based segmentation paradigm is usually strongly violated since voxels located inside the damaged area have no correspondence to the atlas. However, precise segmentation of functionally important brain structures would provide useful information for therapeutic consideration of space-occupying lesions, including surgical, radio-surgical, and radiotherapeutic planning, in order to increase treatment efficiency and minimize neurological damage.

Two early works related to atlas-based segmentation in presence of space-occupying tumors were published in the late 90s. Kyriacou et al. [59] proposed a biomechanical model of the brain using a finite-element method. On the other hand, Dawant et al. [28] relied on a simpler approach based on optical-flow - Thirion's

demons algorithm [98] - for both tumor growth modeling and atlas matching deformation. Their solution was called *seeded atlas deformation (SAD)*, as they put a *seed* with the same intensity properties as the lesion in the atlas image, and then computed the non-rigid registration. Other methods [36, 92, 93] locally adapted the elasticity of the transformation, rather than modeling the deformation induced by the tumor, in a way that large deformations induced by the tumor can be captured. Recently, Nowinski et al [74] proposed to use a Talairach registration followed by a three-dimensional nonlinear tumor deformation based on a geometric assumption, as in [3, 7], that the tumor compresses its surrounding tissues radially. A more sophisticated model of lesion growth was proposed by Mohamed et al. [71] based on 3D biomechanical finite element model.

Our proposed solution [7, 83] improved the *SAD*: instead of applying the nonlinear registration algorithm to the whole image, a specific model of tumor growth inside the tumor area was proposed, which assumed the tumor growth radial from a single voxel seed. Demons algorithm [98] was used outside the tumor area and the displacement vector field was regularized by an adaptive Gaussian filter to avoid possible discontinuities. Note that this method does not apply to infiltrating tumors or take into account the presence of the edema.

3.1 Methodology

Our approach is called *model of lesion growth (MLG)*, and it works in 4 steps. First, an affine transformation is applied to the brain atlas in order to globally match the patient's volume [27]. Second, the lesion is automatically segmented [57, 107]. Third, the atlas is manually seeded with a voxel synthetic lesion placed on the estimated origin of the patient's lesion. At this point, the affine registration ensures that the small displacement assumption is respected in the region of the brain that is far from the tumor. Meanwhile, the segmentation of the tumor volume and the manual selection of the tumor seed provides an adequate model for the tumor and its influence on immediately surrounding tissues. Fourth, the proposed non-rigid deformation method distinguishes between those two areas fixed from the segmentation of the lesion. The instantaneous displacement field is computed as:

$$\vec{d}^{i+1} = \vec{D}^i + \Delta t \cdot \vec{v}^{i+1}, \quad (1)$$

where $\Delta t=1$, \vec{D}^i is the current total displacement field. Outside the tumor, the instantaneous force \vec{v}^{i+1} (velocity) for each demon point \vec{p} at iteration $i+1$, is computed as

$$\vec{v}^{i+1} = \frac{(g(\vec{p} + \vec{D}^i(\vec{p})) - f(\vec{p}))\vec{\nabla} f(\vec{p})}{|\vec{\nabla} g(\vec{p})|^2 + (g(\vec{p} + \vec{D}^i(\vec{p})) - f(\vec{p}))^2}, \quad (2)$$

where $f(\cdot)$ and $g(\cdot)$ are the image intensities. Thus, there is a displacement in the direction of the gradient provided by both a difference in image intensities and a reference image gradient different from zero. Note that (2) is asymmetrical, that is, it gives different results depending on which image is chosen as the reference and which is chosen to be floating. As proposed by Thirion [98], *bijectivity* is ensured by computing at each iteration both the *direct* deformation field (\vec{d}_{direct} , from Eqs. (1) and (2)) and the *inverse* deformation field ($\vec{d}_{inverse}$, also from Eqs. (1) and (2) but replacing f instead of g and viceversa). Then, a residual vector field $\vec{R} = \vec{d}_{direct} + \vec{d}_{inverse}$ is equally distributed onto the two deformation fields:

$$\begin{aligned}\vec{D}_{direct}^{*i+1} &= \vec{d}_{direct}^{i+1} - \frac{\vec{R}}{2}, \\ \vec{D}_{inverse}^{*i+1} &= \vec{d}_{inverse}^{i+1} - \frac{\vec{R}}{2},\end{aligned}\quad (3)$$

Inside the tumor, the tumor growth model assumes a *radial* growth of the tumor from the tumor seed, i.e.

$$\vec{v}_{lesion}^{i+1}(\vec{p}) = \frac{\overrightarrow{DM}_{seed}}{N_{it}}, \quad (4)$$

where \vec{v}_{lesion} is the instantaneous velocity inside the lesion area, $\overrightarrow{DM}_{seed}$ is the distance from the corresponding point \vec{p} to the seed, and N_{it} is the number of iterations of the deformation algorithm that have to be performed. Then, the deformation field \vec{d}_{lesion}^{i+1} is computed similarly as in (2). The bijectivity inside the lesion area is ensured by forcing $\vec{d}_{direct} = -\vec{d}_{inverse}$. This model allows the points inside the lesion area to converge towards the seed voxel², while remaining simple and allowing any number of iterations to take place outside the tumor volume.

The displacement vector computed at every voxel using either the demons force (1) or the tumor growth model (4) is regularized by an adaptive Gaussian filter to avoid possible discontinuities

$$\vec{D}_\alpha = \vec{D}_\alpha^{*i+1} \circ G(\sigma), \quad (5)$$

where \vec{D}^* is the deformation field at the current iteration, α refers to *direct* and *inverse*, $G(\sigma)$ is the Gaussian filter with standard deviation σ , and \vec{D} is the regularized deformation field that will be used in Eq. (1) the next iteration. In fact, three different σ areas are considered: inside the lesion area, close to the lesion (within 10 mm of the tumor) where large deformations occur, and the rest of the brain. Smoothing is not necessary inside the lesion because the vector field induced by (4) is highly regular and the continuity is ensured. So, $\sigma = 0$ inside

²Note that the vector field points the origin, and not the destiny, of a voxel.

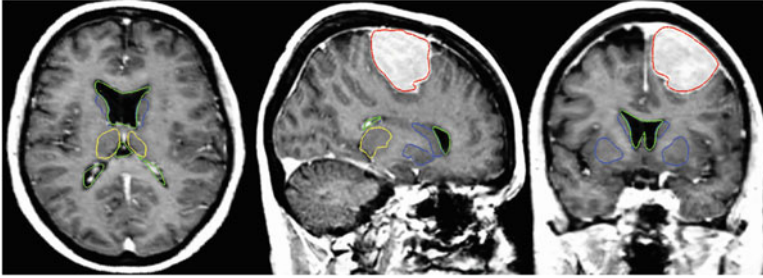


Fig. 2 Segmentation results after applying the *MLG* algorithm. Displayed structures are: tumor (red), ventricles (green), thalamus (yellow), and central nuclei (blue)

the lesion area. In the region close to the tumor (including the tumor contour) there are large deformations due to the tumor growth. Then, it is necessary to allow large elasticity, i.e. σ should have a small value, typically 0.5 mm. In the rest of the brain, deformations are smaller, due primarily to inter-patient anatomical variability. So, a larger σ proves to be better, as it simulates a more rigid transformation. Previous studies [3] suggest that a typical σ to match two healthy brains is about 0.5 mm and 1 mm. In what follows, $\sigma = 0.8$ mm is used. The number of iterations is arbitrarily fixed to $256 + 128 + 32 + 16$ from low to high resolution scale and it stops at the end of the iterations. The algorithm is implemented in a multiscale way: a first match is made with downsampled images and the resulting transformation is upsampled to initialize the next match with finer image resolution.

3.2 Results

The result after applying these steps is a deformed brain atlas in which a tumor has grown from an initial seed, causing displacement and deformation to the surrounding tissues. After this, structures and substructures from the brain atlas may be projected to the patient's image. This is illustrated in Fig. 2. The need of a correct estimation of the tumor growth in order to obtain a good final segmentation of the structures directly displaced and deformed by the lesion is proven in [7]. As illustrated in Fig. 3, without an explicit model of tumor growth, SAD cannot grow the small seed until the final tumor size. However, it seems that a good deformation has been obtained in the rest of the brain. Thus, since we are interested in the deep brain structures and not in the tumor itself, the need of simulating the lesion growth could be questionable. Let us then compare the accuracy of the segmentation results of the ventricles, the thalamus and the central nuclei for both approaches. The obtained results, Fig. 4 show that the *MLG* performs clearly better in the case of the structures near the tumor (thalamus and central nuclei). The most critical

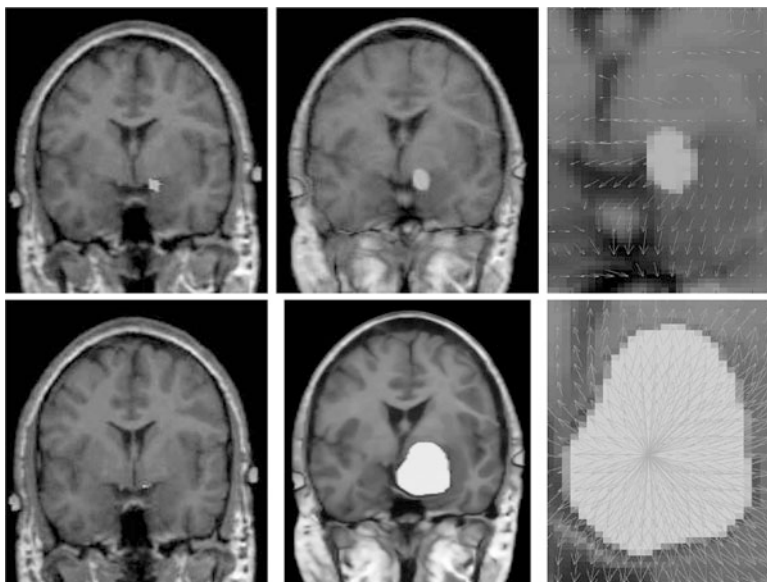


Fig. 3 Top and bottom row, without and with model of lesion growth, respectively. First column, seeded atlas (small seed on top and one voxel seed on bottom). Second column, deformation of seeded atlas. Third column, deformation field in the tumor area

structure is the central nuclei (MLG in green and SAD in red) since it is initially placed inside the tumor area. In this case, SAD method fails because the central nuclei segmentation is placed inside the tumor area. On the contrary, MLG pushes the central nuclei out of the tumor region and it obtains a better segmentation.

3.3 Limitations and future work

Our proposed method in [7, 83] increases the robustness of previous existing methods but other limitations arise: the placement of the seed needs expertise and previous segmentation of the lesion is still necessary. Also, since the Demons algorithm [98] uses the sum of the mean squared differences of intensities as similarity measure, the contrast agent (often present in MR scans of such pathologies) induced some errors in the deformation field. Recently, we proposed in [6] a Mutual Information (MI) flow algorithm combined with the same radial growth model presented here. This approach is in fact very similar to the one presented in [7, 83] but MI flow algorithm has proven its robustness to intensity inconsistencies between the atlas and the patient images.

Other limitations persist and they are shared by all voxel-based methods since they often lead to a compromise between the accuracy of the resulting segmentation

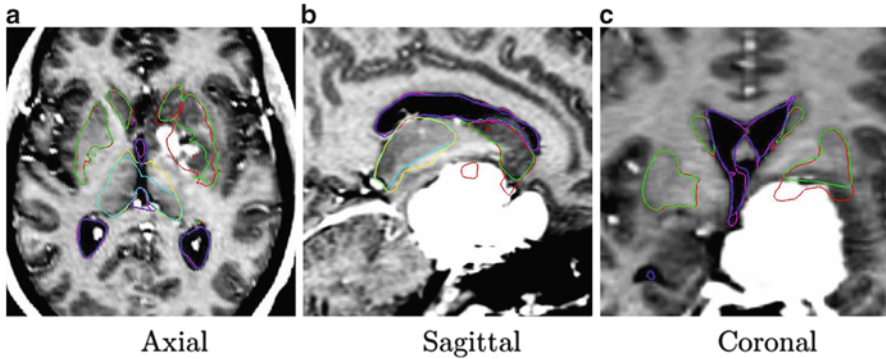


Fig. 4 Importance of the tumor growth model: without (SAD) and with (MLG) model of lesion growth, respectively. Segmented structures: ventricles (MLG in blue and SAD in magenta), thalamus (MLG in cyan and SAD in yellow), and central nuclei (MLG in green and SAD in red)

and the smoothness of the transformation. In our solutions, the regularization of the deformation field is done by an adaptive Gaussian filtering and three different elasticities are allowed according to the distance to the tumor: inside the tumor the radial force ensures the regularity, near the tumor large deformability is allowed and far from the lesion elasticity modeling normal anatomy deformations is applied. A more realistic model of deformability would be to consider not only the distance to the tumor but also the implicit elasticity of some structures such the ventricles, midsagittal plane or skull. This could be included as proposed by Duay et al. [36] by considering that different anatomical structures in the atlas have different elasticity parameters σ but the estimation of these elasticity parameters remains not straightforward. However, even if a better deformability is modeled, voxel-based method still miss local registration constraints that could help selected atlas contours to match their corresponding contours in the patient image.

To cope with this problem, more local constraints have to be included in the atlas registration process for instance by incorporating local statistical measures in the registration process as proposed recently by Commowick et al. in [24]. In our opinion, among the different techniques proposed so far for image analysis, the active contour framework is particularly well suited to define and implement local constraints but it was initially designed for image segmentation. Recently, new algorithms including active contour constraints in a registration process have been proposed. These algorithms perform in fact a registration and segmentation tasks jointly (see Sect. 4.1). We proposed in [34, 35, 37, 45] such type of algorithm and we show in Fig. 5 some preliminary results with in a tumor growth application. The atlas and the patient images are respectively shown in Figs. 5(a) and 5(b). These images correspond to 2D slices extracted from 3D brain MR images. A one-voxel seed (shown in Fig. 5(a) by a red point) has been inserted inside the atlas to model the tumor growth. The difference between this model and the tumor growth model presented above is that this seed corresponds to the initial position of an active

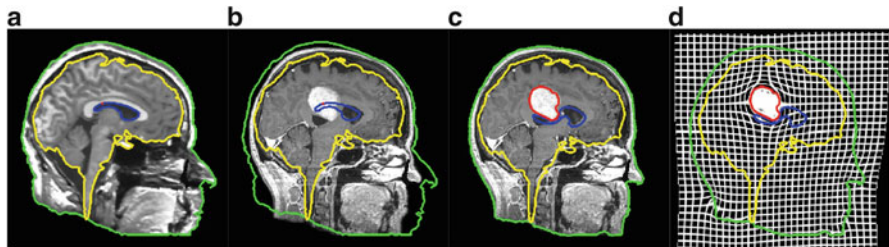


Fig. 5 Active contour-based registration of an atlas on a brain MR image presenting a large occupying tumor. **(a)** Intensity atlas [58] with objects of interest (the head in green, the brain in yellow, the ventricles in blue, and the tumor one-voxel seed in red). **(b)** Atlas contours superimposed to the patient image. **(c)** Results of the joint segmentation and registration driven by the external contour of the head and the tumor contour. **(d)** Corresponding deformation field

contour. This active contour is going to segment the tumor of the patient image during the registration process. Thus, with this model, the pre-segmentation of the patient image is not anymore required. Contours of the objects of interest selected in the atlas (the head in green, the brain in yellow, the ventricles in blue and the tumor in red) are superimposed to all images. Our active contour-based algorithm permits to select the atlas contours that will drive its registration. In this case, the registration was performed following the registration of the head contour and the tumor growth only. The rest of the image just follow the deformation interpolated from the displacement of the selected contours. Fig. 5(c) shows the segmentation result obtained after this type of registration and Fig. 5(d) shows the corresponding deformation field. We can see that the registration of the selected green and red contours are bringing the yellow and blue contours closer to their target contours. This *object-based registration* process points out the spatial dependence that exists between anatomical structures. Such spatial dependence can be exploited in a hierarchical atlas as described in Sect. 4.2.

The registration framework we propose allows not only to easily include object-based registration constraints but also to select different type of segmentation forces derived from the active contour framework for the registration of different structures in the atlas. Like the biomechanical methods, this algorithm can also be seen as a surface-based registration method. Its first difference is that it computes the deformation based on the image and not on mechanical or biological laws. This implies that its accuracy does not depend on a good physical modeling of the tissues. Above all, as mentioned above, its main advantage is that it does not need a pre segmentation of the patient image. The contours defined in the atlas evolve following an energy functional specially defined to be minimal when they have reached the desired object contours, as in the active contour-based segmentation framework [76]. Unfortunately, the main limitation of the surface to surface registration algorithm remains. As the deformation is only based on contours of interest, the probability of registration errors increases more we are far from these contours. As far as we know, most of the existing methods for registration of images

with space-occupying tumors are either surface-based [59, 62, 71, 116] or voxel-based [7, 13, 28, 36, 83, 93] approaches. However, in our opinion it is worth to study how to combine the advantages of both approaches.

4 Discussion

In this chapter, we saw that the atlas-based segmentation has become a standard paradigm for exploiting spatial prior knowledge in medical image segmentation [16]. The labeled atlas aims to delineate objects of interest in the intensity atlas. The atlas-based segmentation process consists to deform the selected atlas objects in order to better align them with their corresponding objects in the patient image to be segmented. To perform this task, we have distinguished two types of approaches in the literature.

The most known approach reduces the segmentation problem in an image registration problem [5, 10, 14, 18, 22, 29, 43, 53, 86, 98]. Its main advantage is that the deformation field computed from the registration of visible contours allows to easily estimate the position of regions with fuzzy contours or without visible contours, as for instance the subthalamic nuclei (STN) targeting for Parkinson disease [30, 88]. However, these registration algorithms do not exploit the *object-based information* that could be obtained by combining the intensity and the labeled atlas. This object-based information contains an estimation of the objects position in the patient image, the description of their shape and texture, and the features of their adjacent regions. Therefore, usual atlas-based segmentation methods via registration often miss local constraints to improve the accuracy of the delineation of some objects. Besides, this technique assumes that a point to point correspondence exists between the atlas and the image to be segmented. Then, the presence of inconsistencies can generate registration errors, and consequently segmentation errors, if special schemes are not used (see Sect. 3).

As mentioned in Sect. 2.2, contours in the labeled atlas can be directly deformed without need of a geometrical deformation [8, 19, 115]. The contour morphing technique that has attracted the most attention to date is the *active contour segmentation model* [56]. Its advantage is that it can exploit the image information directly linked to the object to be delineated. Therefore, active contour models are often able to extract objects where the atlas-based segmentation method by registration fails (see an example on the cerebellum segmentation in [32]). Moreover, this segmentation method allows extracting only the objects that are consistent in the image. Thus, it usually does not need special model to solve the problem of inconsistencies. On the other hand, this segmentation method is very sensitive to the initial position of the atlas contour: the closer to the contours to be detected, the more robust the active contour-based segmentation will be. Besides, this segmentation technique needs prior shape models to be able to estimate the position of regions with fuzzy or without visible contours. Note that these shape models are in fact atlases that are registered with the patient image during the process to incorporate prior knowledge

in the active contour segmentation. See [17, 75, 78, 79, 104] for examples of active contour segmentation of medical images with shape priors.

Biomechanical models such as in [40, 42, 59, 69] can be seen as object-based registration since they are based on selected image objects. They track key surfaces of objects, as for instance the cortical surface or the lateral ventricles of the brain, to propagate then the surface displacements through the entire volume guided by the prior biomechanical knowledge about the deformability of anatomical structures. One drawback of this type of object-based registration method, compared to image-based registration methods as [98] or [86], is that it remains difficult to accurately estimate all the forces that can interact with the model. The level of accuracy stays thus highly dependent on the number of surfaces tracked.

4.1 Atlas-based Segmentation Approaches combining Registration and Segmentation

Novel atlas-based segmentation methods aim at combining segmentation and registration to include more local constraints in the matching of the atlas contours (see [2] for a recent survey on methods combining registration to level set-based segmentation). We distinguish here two types of approaches: the methods that performs a registration and segmentation task in two successive steps and the methods that perform these two tasks jointly.

In a two steps registration and segmentation method, one approach consists to assist the registration process with the solution of a segmentation task (see for example [87, 98]). Indeed, the pre-segmentation of the intensity atlas and the patient image permits to avoid problems linked to intensity differences between the images to register (due to different acquisition conditions, different modalities or the presence of noise) but also to isolate the image regions that are consistent and relevant for the registration task. This contributes to make the registration process more robust. A second approach consists to first globally register the objects contours selected in the atlas to bring them as close as possible to their corresponding contours in the patient image. Then, a deformable model is used to match the atlas objects to their target contours [8, 32]. The drawback to combine both approaches in two successive steps is that this does not permit to compute a dense deformation field tracking the overall atlas contour deformation.

In a joint registration-segmentation method, one approach consists to track the deformation of the atlas contours modeled by a level set function during an active contour-based segmentation process [35, 44, 45, 105]. This method permits to apply segmentation constraints directly in the deformation field calculation. A second approach consists to extract the same objects in both images and to compute the deformation regarding the contours of these extracted objects. In this method, the segmentation by active contours permits to remove the possible inconsistencies between both images. Conversely, the combination of the information of both images by the registration process incorporates prior knowledge in the segmentation

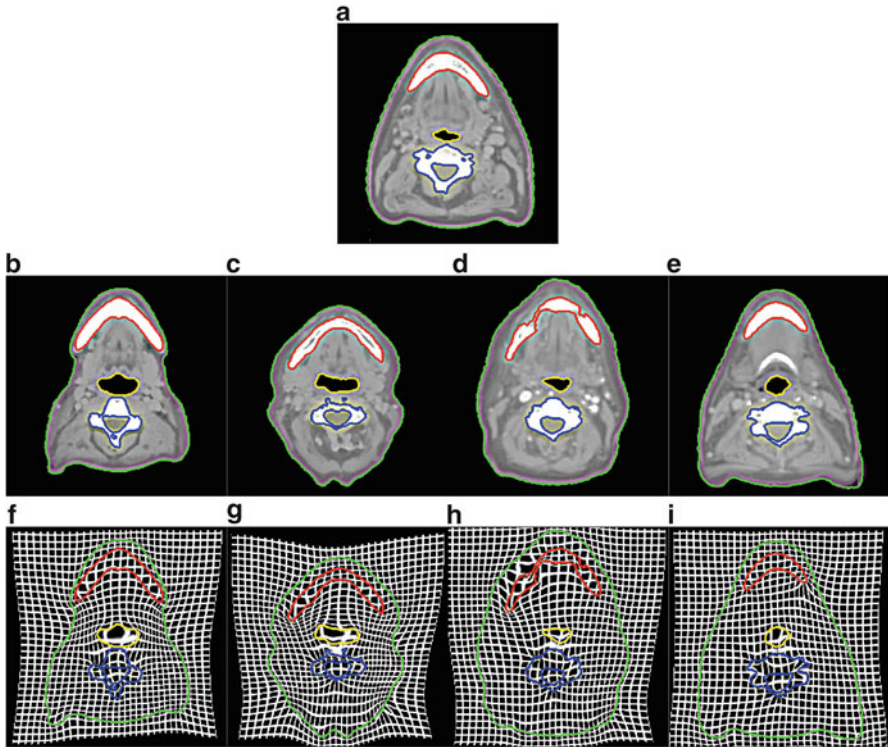


Fig. 6 Joint atlas-based registration and segmentation of Neck CT images. Row 1: Intensity Atlas with surrounded objects of interest (the external contour of the neck in green, the trachea in yellow, the jaw in red and the vertebra in blue). Row 2: Objects of interest delineated in the patients images. Rows 3: Corresponding deformation fields

task. To our knowledge, two different methods corresponding to this second approach were proposed. The first one consists to couple active contour segmentation to registration in a energy-based variational framework [1, 33, 70, 101, 114, 115]. The second one joints the labellisation of the tissue of the source and target image to the registration task by using Markov random fields [112, 113].

Figure 6 shows atlas-based segmentation results obtained on neck CT 2D images with the joint registration-segmentation model developed in our lab [34, 35, 37, 44, 45]. Our model follows the approach that consists to track the deformation of a level set function modeling selected atlas contours. For this case, the contours selected to drive the registration are enhanced by different colors: the external contour of the neck in green, the trachea in yellow, the jaw in red and the vertebra in blue. Columns 2 show the segmentations obtained on the patient images. The deformed grids of Row 3 help to visualize the dense deformation interpolated on the whole image domain. This example shows that the joint registration-segmentation methods permit to register only the objects that are consistent between the atlas and the

patient image. Here, we did not consider in the registration process the arteria, muscles and fat (structures inside the gray part of the neck) that do not correspond between these 2D images.

4.2 Hierarchical Atlas

Atlas-based segmentation methods proposed so far use two types of atlas: determinist atlases and statistical atlases (refer to Sect. 1.1 but the novel concept of atlas-based segmentation approaches combining registration and segmentation led to the concept of a new type of atlas, *the hierarchical atlas* [37, 52]).

The hierarchical atlas is an atlas composed of several layers. Each layer contains a subset of atlas objects and two successive layers are linked by a spatial relationship (the position of the objects defined in one layer is depending of the position of the objects defined in the previous layer). The design of such atlas requires to study the spatial dependance existing between anatomical structures. Intuitively, it seems logical that the position of the most rigid structures (the bones) determine the position of softer structures (the tissues or the fluids).

The registration process of a hierarchical atlas permits to perform the registration progressively by limiting the number of objects to register in the atlas. More visible and important objects (layer 1) are registered first. The resulting deformation field is then applied to get the initial condition for the registration of the objects of second layer, including also the contours of the first layer in order to keep a constraint on their registration. Indeed, such hierarchical approach aim at reducing the complexity of the registration problem by limiting the number of objects to register. Then, the solution of this simplified problem is used as initial condition for a more complex problem, i.e. with more objects to register. The process is repeated until the original image resolution is reached.

5 Conclusion

We reviewed here the existing approaches for atlas-based segmentation in medical imaging and we focus on those based on a volume registration method. Particularly, we presented the problem of using atlas information for pathological image analysis and we proposed our solution for atlas-based segmentation in MR image of the brain when large space-occupying lesions are present. We discussed also future research directions that aim at overcome current limitations of atlas-based segmentation approaches based on registration only by combining segmentation and registration strategies.

Acknowledgements Our acknowledgment goes to Prof. Reto Meuli from the Radiology Department of the Lausanne Hospital (CHUV) and to Dr. Simon Warfield from Harvard Medical School

for providing the patient images. Also, we thank Prof. Ron Kikinis who has provided us with the digitized atlas of the Harvard Medical School. This work has been supported by Center for Biomedical Imaging (CIBM) of the Geneva - Lausanne Universities, the EPFL, and the foundations Leenaards and Louis-Jeantet, as well as by the Swiss National Science Foundation under grant number 205320-101621.

References

1. J. An, Y. Chen, F. Huang, D. Wilson, and E. Geiser. A variational pde based level set method for a simultaneous segmentation and non-rigid registration. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 286–293, 2005.
2. E. Angelini, Y. Jin, and A. Laine. *Handbook of Biomedical Image Analysis*, chapter State of the Art of Level Set Methods in Segmentation and Registration of Medical Imaging Modalities, pages 47–101. Springer US, 2007.
3. M. Bach Cuadra. Atlas-based segmentation and classification of magnetic resonance brain images. *THSE NO 2875, École Polytechnique Fédérale De Lausanne*, 2003.
4. M. Bach Cuadra, L. Cammoun, T. Butz, O. Cuisenaire, and J. Thiran. Comparison and validation of tissue modelization and statistical classification methods in t1-weighted mr brain images. *IEEE Transactions on Medical Imaging*, 24(12):1548–1565, 2005.
5. M. Bach Cuadra, O. Cuisenaire, R. Meuli, and J.-P. Thiran. Automatic segmentation of internal structures of the brain in mri using a tandem of affine and non-rigid registration of an anatomical atlas. In *International Conference in Image Processing (ICIP)*, October 2001.
6. M. Bach Cuadra, M. De Craene, V. Duay, B. Macq, C. Pollo, and J. Thiran. Dense deformation field estimation for atlas-based segmentation of pathological mr brain images. *Methods and Programs in Biomedicine*, 84(2-3):66–75, 2006.
7. M. Bach Cuadra, C. Polio, A. Bardera, O. Cuisenaire, J.-G. Villemure, and J. Thiran. Atlas-based segmentation of pathological mr brain images using a model of lesion growth. *IEEE Trans. Med. Imag.*, 23(10):1301–1314, 2004.
8. C. Baillard, P. Hellier, and B. C. Cooperation between level set techniques and 3d registration for the segmentation of brain structures. In *International Conference on Pattern Recognition (ICPR)*, pages 991–994, 2000.
9. R. Bajcsy. Digital anatomy atlas and its registration to mri, fmri, pet: The past presents a future. In *Biomedical Image Registration, Second International Workshop (WBIR)*, pages 201–211, Philadelphia, USA, 2003.
10. R. Bajcsy and S. Kovacic. Multi resolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
11. R. Bajcsy, R. Lieberman, and M. Reivich. A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. *Journal of Computer Assisted Tomography.*, 7(4):618–625, 1983.
12. K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. In *IEEE International Symposium on Biomedical Imaging (ISBI): From Nano to Macro.*, pages 908–911, Arlington, USA, 2004.
13. P.-Y. Bondiau, G. Malandain, S. Chanalet, P. Marcy, J.-L. Habrand, F. Fauchon, P. Paquis, A. Courdi, O. Commowick, I. Rutten, and N. Ayache. Atlas-based automatic segmentation of mr images: validation study on the brainstem in radiotherapy context. *Int J Radiat Oncol Biol Phys.*, 61(1):289–298, 2005.
14. M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Visualization in Biomedical Computing (VBC '96)*, pages 267–276, 1996.
15. T. Brox, A. Bruhn, N. Papenberb, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *8th European Conf. Computer Vision, Part IV: Lecture Notes in Computer Science*, volume 3024, pages 25–36, 2004.

16. M. Cabezas, A. Oliver, X. Lladó, J. Freixenet, and M. Bach Cuadra. A review of atlas-based segmentation for magnetic resonance brain images. *Computer Methods and Programs in Biomedicine*, 104(3):e158–e177, 2011.
17. Y. Chen, F. Huang, R. Tagare, H. D. amd Murali, D. Wilson, and E. A. Geiser. Using prior shape and intensity profile in medical image segmentation. In *IEEE International Conference on Computer Vision*, pages 1117–1124, 2003.
18. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. 3d brain mapping using a deformable neuroanatomy. *Phys. Med. Biol.*, 39:609–618, 1994.
19. C. Ciofolo. Atlas-based segmentation using level sets and fuzzy labels. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 310–317, 2004.
20. A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3d multi-modality medical image registration using feature space clustering. In *Computer Vision, Virtual Reality, and Robotics in Medicine*, volume 905, pages 195–204, 1995.
21. D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998. <http://www.bic.mni.mcgill.ca/brainweb/>.
22. L. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans. Automatic 3-d model-based neuroanatomical segmentation. *Human Brain Mapping*, 3(3):190–208, 1995.
23. O. Commowick and G. Malandain. Evaluation of atlas construction strategies in the context of radiotherapy planning. In *Proceedings of the SA2PM Workshop (From Statistical Atlases to Personalized Models)*, Copenhagen, October 2006. Held in conjunction with MICCAI 2006.
24. O. Commowick, R. Stefanescu, P. Fillard, V. Arsigny, N. Ayache, X. Pennec, and G. Malandain. Incorporating statistical measures of anatomical variability in atlas-to-subject registration for conformal brain radiotherapy. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2, pages 927–934, 2005.
25. D. Cooper, C. Cootes, T.F. and Taylor, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 2(61):38–59, 1995.
26. T. Cootes, C. Beeston, and C. Edwards, G.J.and Taylor. A unified framework for atlas matching using active appearance models. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2:927–934, 2005.
27. O. Cuisenaire, J.-P. Thiran, B. Macq, C. Michel, A. De Volder, and F. Marques. Automatic registration of 3d mr images with a computerized brain atlas. In *SPIE Medical Imaging*, volume 1719, pages 438–449, 1996.
28. B. Dawant, S. Hartmann, and S. Gadamsetty. Brain Atlas Deformation in the Presence of Large Space-occupying Tumors. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 589–596, 1999.
29. B. Dawant, S. Hartmann, J.-P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel. Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations : Part I, methodology and validation on normal subjects. *IEEE Transactions on Medical Imaging*, 18(10):902–916, 1999.
30. B. M. Dawant, R. Li, E. Cetinkaya, C. Kao, J. M. Fitzpatrick, and P. E. Konrad. Computerized atlas-guided positioning of deep brain stimulators: A feasibility study. *WBIR*, pages 142–150, 2003.
31. M. De Craene, A. du Bois d’Aische, B. Macq, and S. K. Warfield. Multi-subject registration for unbiased statistical atlas construction. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 655–662, 2004.
32. P.-F. D’Haese. Automatic segmentation of brain structures for radiation therapy planning. In *SPIE Medical Image Processing*, pages 517–526, 2003.
33. M. Droske, W. Ring, and M. Rumpf. Mumford-shah based registration. *Computing and Visualization in Science (CVS)*, 2007. to appear in CVS.
34. V. Duay, M. Bach Cuadra, X. Bresson, and J.-P. Thiran. Dense deformation field estimation for atlas registration using the active contour framework. In *European Signal Processing Conference (EUSIPCO)*, 2006.

35. V. Duay, X. Bresson, N. Houhou, M. Bach Cuadra, and J.-P. Thiran. Registration of multiple regions derived from the optical flow model and the active contour framework. In *European Signal Processing Conference (EUSIPCO)*, 2007.
36. V. Duay, P. DHaese, R. Li, and B. Dawant. Non-rigid registration algorithm with spatially varying stiffness properties. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 408–411, 2004.
37. V. Duay, N. Houhou, and J.-P. Thiran. Atlas-based segmentation of medical images locally constrained by level sets. In *International Conference in Image Processing (ICIP)*, 2005.
38. M. Esiri and M. J. The neuropathology of dementia. *Cambridge University Press*, 2002.
39. A. Evans, D. Collins, P. Neelin, M. Kamber, and T. S. Marrett. Three-dimensional correlative imaging: applications in human brain mapping. *Functional Imaging: Technical Foundations*, pages 145–162, 1994.
40. M. Ferrant, A. Nabavi, B. Macq, P. M. Black, F. A. Jolesz, R. Kikinis, and S. K. Warfield. Serial registration of intraoperative mr images of the brain. *Medical Image Analysis*, 6(4): 337–359, 2002.
41. K. Friston, J. Ashburner, C. D. Frith, J.-B. Poline, J. Heather, and R. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995. <http://www.fil.ion.ucl.ac.uk/spm/>.
42. R. Galloway, R. Macuinas, W. Bass, and W. Carpin. Optical localization for interactive image-guided neurosurgery. *Medical Imaging*, 2164:137–145, 1994.
43. J. Gee, M. Reivich, and R. Bajcsy. Elastically deforming a three-dimensional atlas to match anatomical brain images. *J. Comput. Assist. Tomogr.*, 17:225–236, 1993.
44. S. Gorthi, V. Duay, X. Bresson, M. Bach Cuadra, F. J. Sánchez Castro, C. Pollo, A. S. Allal, and J. P. Thiran. Active deformation fields: dense deformation field estimation for atlas-based segmentation using the active contour framework. *Medical Image Analysis*, 15(6):787–800, 2011.
45. S. Gorthi, V. Duay, N. Houhou, M. Bach Cuadra, U. Schick, M. Becker, A. Allal, and J.-P. Thiran. Segmentation of head and neck lymph node regions for radiotherapy planning, using active contour based atlas registration. *IEEE Journal of selected topics in signal processing*, 3(1):135–147, 2009.
46. T. Greitz, C. Bohm, S. Holte, and L. Eriksson. A computerized brain atlas: construction, anatomical content and some applications. *Journal of Computer Assisted Tomography*, 15(1):26–38, 1991.
47. A. Guimond, J. Meunier, and J. Thirion. Average brain models: a convergence study. *Comput. Vis. Image Underst.*, 77(9):192–210, 2000.
48. P. Haese, V. Duay, R. Li, A. du Bois Aische, A. Cmelak, E. Donnelly, K. Niermann, T. Merchant, B. Macq, and B. Dawant. Automatic segmentation of brain structures for radiation therapy planning. *Medical Imaging Conference SPIE*, 2003.
49. J. Haller, A. Banerjee, G. Christensen, M. Gado, S. Joshi, M. Miller, Y. Sheline, M. Vannier, and J. Csernansky. 3d hippocampal morphometry by high dimensional transformation of a neuroanatomical atlas. *Radiology*, 202(2):504–510, 1997.
50. P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. Collins, A. Evans, G. Malandain, and N. Ayache. Retrospective evaluation of inter-subject brain registration. *IEEE Transactions on Medical Imaging*, 22(9):1120–1130, 2003.
51. K. Hohne, M. Bomans, M. Riemer, R. Schubert, U. Tiede, and W. Lierse. A volume based anatomical atlas. *IEEE Computer Graphics and Applications.*, 12(4):72–78, 1992.
52. N. Houhou, V. Duay, A. S. Allal, and J.-P. Thiran. Medical images registration with a hierarchical atlas. In *EUSIPCO*, 2005.
53. D. V. Iosifescu, M. E. Shenton, S. K. Warfield, R. Kikinis, J. Dengler, F. A. Jolesz, and R. W. Mccarley. An automated registration algorithm for measuring mri subcortical brain structures. *Neuroimage*, 6(1):13–25, July 1997.
54. S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage.*, 23(1):151–160, 2004.

55. T. Kapur, P. A. Beardsley, S. F. Gibson, W. E. L. Grimson, and W. M. Wells. Model based segmentation of clinical knee mri. In *Proc. IEEE Int'l Workshop on Model-Based 3D Image Analysis*, pages 97–106, 1998.
56. M. Kass, A. Witkin, and T. D. Snakes: active contour models. In *First international conference on computer vision*, pages 259–268, 1987.
57. M. Kaus, S. Warfield, A. Nabavi, E. Chatzidakis, P. Black, F. Jolesz, and R. Kikinis. Segmentation of meningiomas and low grade gliomas in mri. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 1–10, 1999.
58. R. Kikinis, M. Shenton, D. Iosifescu, R. McCarley, P. Saiviroonporn, H. Hokama, A. Robatino, D. Metcalf, C. Wible, C. Portas, R. Donnino, and F. Jolesz. A digital brain atlas for surgical planning, model driven segmentation and teaching. *IEEE Transactions on Visualization and Computer Graphics.*, 2(3):232–241, 1996.
59. S. Kyriacou and C. Davatzikos. Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model. *IEEE Trans. Med. Imaging*, 18(7):580–592, 1999.
60. Laboratory of Neuro Imaging, UCLA. International Consortium for Brain Mapping. <http://www.loni.ucla.edu/ICBM/>, 1993.
61. K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18(10): 897–908, 1999.
62. T. Liu, D. Shen, and C. Davatzikos. Deformable registration of tumor-diseased brain images. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 720–728, 2004.
63. P. Lorenzen, B. Davis, and S. Joshi. Unbiased atlas formation via large deformations metric mapping. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2, pages 411–418, Palm Springs, California, USA, 2005.
64. D. Louis Collins, G. Le Goualher, and A. Evans. Non-linear cerebral registration with sulcal constraints. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 974–984, 1998.
65. F. Maes and A. Collignon. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16, 1997.
66. C. R. Maurer and J. M. Fitzpatrick. *Interactive ImageGuided Neurosurgery*, chapter A review of medical image registration, pages 17–44. American Association of neurological surgeons, 1993.
67. J. C. Mazziotta, A. W. Toga, and R. S. J. Frackowiak. *Brain Mapping: The Disorders*. Academic Press, 2000.
68. McConnell Brain Imaging Center. BrainWeb: Simulated Brain Database. <http://www.bic.mni.mcgill.ca/brainweb/>, 1997.
69. M. Miga, T. Sinha, D. Cash, R. Galloway, and R. Weil. Cortical surface registration for image-guided neurosurgery using laser range scanning. *IEEE Transactions on Medical Imaging*, 22(8):973–985, 2003.
70. M. Moelich and T. Chan. Joint segmentation and registration using logic models. Technical Report 03-06, Mathematics Department, UCLA, 2003.
71. A. Mohamed and C. Davatzikos. Finite element modeling of brain tumor mass-effect from 3d medical images. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 400–408, 2005.
72. National Library of Medicine. The visible human project. <http://www.nlm.nih.gov/research/visible>, 1991.
73. A. Noe, S. Kovacic, and J. Gee. Segmentation of cerebral mri scans using a partial volume model, shading correction, and an anatomical prior. In *SPIE Medical Image Processing*, 2001.
74. W. L. Nowinski and D. Belov. Toward atlas-assisted automatic interpretation of mri morphological brain scans in the presence of tumor. *Academic Radiology*, 12(8):1049–1057, August 2005.

75. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging Vision and Graphics*, chapter Shape analysis towards model-based segmentation, pages 231–250. Springer Verlag, New York, 2003.
76. S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed - algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.
77. N. Pal and S. Pal. A review on image segmentation techniques. *Pattern Recognition*, 26(9):1277–1294, 1993.
78. N. Paragios. A variational approach for the segmentation of the left ventricle in mr cardiac images. In *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 153–160, 2001.
79. N. Paragios. A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Transactions on Medical Imaging*, 22:773–776, 2003.
80. H. Park, P. Bland, A. Hero, and C. Meyer. Least biased target selection in probabilistic atlas construction. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2, pages 419–426, 2005.
81. D. Perperidis, R. Chandrashekar, M. Lorenzo-Valdés, G. Sanchez-Ortiz, A. Rao, D. Rueckert, and R. Mohiaddin. Building a 4d atlas of the cardiac anatomy and motion using mr imaging. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 412–415, 2004.
82. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, August 2003.
83. C. Pollo, M. Bach Cuadra, O. Cuisenaire, J.-G. Villemure, and J.-P. Thiran. Segmentation of brain structures in presence of a space-occupying lesion. *Neuroimage*, 24(4):990–996, February 2005.
84. M. Prastawa, E. Bullitt, and N. Moon. Automatic brain tumor segmentation by subject specific modification of atlas priors. *Acad. Radiol.*, 10(12):1341–1348, 2003.
85. T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and C. R. Maurer, Jr. Quo vadis, atlas-based segmentation? In J. Suri, D. L. Wilson, and S. Laxminarayan, editors, *The Handbook of Medical Image Analysis – Volume III: Registration Models*, chapter 11, pages 435–486. Kluwer Academic / Plenum Publishers, 2005.
86. D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
87. F. Sanchez Castro, C. Pollo, J. G. Villemure, and T. J. P. Feature-segmentation-based registration for fast and accurate deep brain stimulation targeting. In *Proceedings of the 20th International Congress and Exhibition in Computer Assisted Radiology and Surgery*, 2006.
88. F. Sanchez Castro, C. Pollo, J. G. Villemure, and T. J. P. Validation of experts versus atlas-based and automatic registration methods for subthalamic nucleus targeting on mri. *International Journal of Computer Assisted Radiology and Surgery*, 1(1):5–12, 2006.
89. J. A. Schnabel, C. Tanner, A. Castellano Smith, M. Leach, R. Hose, D. Hill, and D. Hawkes. Validation of non-rigid registration using finite element methods. In *Lecture Notes in Computer Science*, Springer Verlag, Berlin, editor, *Information Processing in Medical Imaging (IPMI)*, pages 345–358, 2001.
90. D. Shattuck, S. Sandor-Leahy, K. Schaper, D. Rottenberg, and R. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13:856–876, 2001.
91. J. A. Stark and W. J. Fitzgerald. Model-based adaptive histogram equalization. *Signal Processing*, pages 193–200, 1994.
92. R. Stefanescu. *Parallel nonlinear registration of medical images with a priori information on anatomy and pathology*. Thèse de sciences, Université de Nice – Sophia-Antipolis, March 2005.

93. R. Stefanescu, O. Commowick, G. Malandain, P.-Y. Bondiau, N. Ayache, and X. Pennec. Non-rigid atlas to subject registration with pathologies for conformal brain radiotherapy. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 704–711, 2004.
94. C. Studholme, D. L. G. Hill, and D. J. Hawkes. Multiresolution voxel similarity measures for mr-pet registration. *Information Processing in Medical Imaging*, pages 287–298, 1995.
95. G. Subsol, J.-P. Thirion, and N. Ayache. A scheme for automatically building 3D morphometric anatomical atlases: application to a skull atlas. *Medical Image Analysis*, 2(1):37–60, 1998.
96. J. S. Suri, S. Singh, and L. Reden. Computer vision and pattern recognition techniques for 2-d and 3-d mr cerebral cortical segmentation (part i): A state-of-the-art review. *Pattern Analysis and Applications*, 5:46–76, 2002.
97. J. Talairach and P. Tournoux. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system - an approach to cerebral imaging*. Thieme Medical Publishers, 1998.
98. J. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.
99. L. Thurfjell, C. Bohm, T. Greitz, and L. Eriksson. Transformations and algorithms in a computerized brain atlas. *IEEE Transactions on Nuclear Sciences*, 40:1187–1191, 1993.
100. A. W. Toga. *Brain Warping*. Academic Press, 1999.
101. G. Unal and G. Slabaugh. Coupled pdes for non-rigid registration and segmentation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 168–175, 2005.
102. K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18: 885–896, 1999.
103. K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18:897–908, 1999.
104. B. C. Vemuri and Y. Chen. *Geometric Level Set Methods in Imaging, Vision and Graphics*, chapter Joint image registration and segmentation, pages 251–269. Springer Verlag, New York, 2003.
105. B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard. Image registration via level-set motion: Applications to atlas-based segmentation. *IEEE Transaction on Medical Image Analysis*, 7(1):1–20, 2003.
106. P. Viola and W. Wells. Alignment by maximization of mutual information. *Fifth Int. Conf. on Computer Vision*, pages 16–23, 1995.
107. S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis. Adaptive, template moderated, spatially varying statistical classification. *Medical Image Analysis*, 4(1):43–55, March 2000.
108. S. K. Warfield, J. Rexilius, P. Huppi, T. Inder, E. Miller, W. Wells, G. Zientara, F. Jolesz, and R. Kikinis. A binary entropy measure to assess nonrigid registration algorithms. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 266–274, 2001.
109. W. Wells, R. Kikinis, W. Grimson, and F. Jolesz. Adaptive segmentation of mri data. *IEEE Transactions on Medical Imaging*, 15:429–442, 1996.
110. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer Jr, R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, J. Maintz, M. Viergever, G. Malandain, and R. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21(4):554–566, 1997.
111. R. Woods, M. Dapretto, N. Sicotte, A. Toga, and J. Mazziotta. Creation and use of a talairach-compatible atlas for accurate, automated, nonlinear intersubject registration, and analysis of functional imaging data. *Human Brain Mapping*, 8(2-3):73–79, 1999.
112. P. Wyatt and J. A. Noble. Map mrf joint segmentation and registration of medical images. *Medical Image Analysis*, 7(4):539–552, 2003.

113. C. Xiaohua, M. Brady, and D. Rueckert. Simultaneous segmentation and registration for medical image. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 663–670, 2004.
114. A. Yezzi, L. Zollei, and T. Kapur. A variational framework for joint segmentation and registration. In *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (CVPR-MMBIA)*, pages 44–49, 2001.
115. Y.-N. Young and D. Levy. Registration-based morphing of active contours for segmentation of ct scans. *Mathematical Biosciences and Engineering*, 2(1):79–96, 2005.
116. E. Zacharaki, D. Shen, A. Mohamed, and C. Davatzikos. Registration of brain images with tumors: Towards the construction of statistical atlases for therapy planning. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2006.
117. Y. Zhan, D. Shen, J. Zeng, L. Sun, G. Fichtinger, J. Moul, and C. Davatzikos. Targeted prostate biopsy using statistical image analysis. *IEEE Trans Med Imaging*, 26(6):779–88, 2007.
118. L. Zollei, E. Learned Miller, W. Grimson, and W. Wells, III. Efficient population registration of 3d data. In *Computer Vision for Biomedical Image Applications.*, pages 291–301, 2005.

Integration of Topological Constraints in Medical Image Segmentation

F. Ségonne and B. Fischl*

Abstract Topology is a strong global constraint that can be useful in generating geometrically accurate segmentations of anatomical structures. Conversely, topological “defects” or departures from the true topology of a structure due to segmentation errors can greatly reduce the utility of anatomical models. In this chapter we cover methods for integrating topological constraints into segmentation procedures in order to generate geometrically accurate and topologically correct models, which is critical for many clinical and research applications.

1 Introduction

1.1 Description of the Problem

In medical imaging, the overall shape of a region of interest is typically prescribed by medical knowledge; it is usually known *a priori*. Excluding pathological cases, the shape of most macroscopic brain structures can be continuously deformed into

* Support for this research was provided in part by the National Center for Research Resources (P41-RR14075, R01 RR16594-01A1 and the NCRR BIRN Morphometric Project BIRN002, U24 RR021382), the National Institute for Biomedical Imaging and Bioengineering (R01 EB001550, R01EB006758), the National Institute for Neurological Disorders and Stroke (R01 NS052585-01) as well as the Mental Illness and Neuroscience Discovery (MIND) Institute, and is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149.

F. Ségonne (✉) • B. Fischl
Department of Radiology, MGH/Harvard Medical School, Building 149,
13th Street, Charlestown, MA 02129, USA
e-mail: fischl@nmr.mgh.harvard.edu; florent.segonne_medical@m4x.org

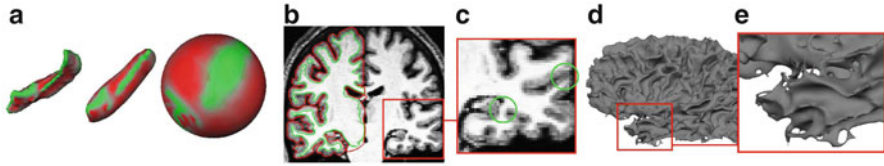


Fig. 1 a) Subcortical structures have a spherical topology. For instance, the shape of the hippocampus can be continuously deformed onto a sphere. b) The human cerebral cortex is a highly folded ribbon of gray matter that lies inside the cerebrospinal fluid (the red interface) and outside the white matter of the brain (the green interface). When the midline connections between the left and right hemisphere are artificially closed, these two surfaces have the topology of a sphere. c) Due to the partial volume effect, subject motion, etc . . . , it becomes hard to distinguish opposite banks of a the gray matter. d) Segmentation algorithms that do not constrain the topology often produce cortical segmentations with several topological defects (i.e. handles, cavities, disconnected components). e) A close-up of a topologically-incorrect cortical surface representation

a sphere. In mathematical terms, these structures have the topology of a sphere. Particularly, this implies that most brain structures consist of *one single connected* object that does not possess any *handles* (i.e. holes) or *cavities*. This is the case for noncortical structures (such as left and right ventricle, putamen, pallidum, amygdala, hippocampus, thalamus, and caudate nucleus - see Fig. 1-a), but it also holds for the cortex under some specific conditions.¹ In addition to their individual topological properties, anatomical structures occur in a characteristic spatial pattern relative to one another (e.g. in the human brain, the amygdala is anterior and superior to the hippocampus). The set of individual topological properties and specific relationships between anatomical structures determine the global topology of a region of interest.

Although many clinical and research applications require accurate segmentations that respect the true anatomy of the targeted structures, only a few techniques have been proposed to achieve accurate and topologically-correct segmentations.

1.2 Motivation

Many neurodegenerative disorders, psychiatric disorders, and healthy aging are frequently associated with structural changes in the brain. These changes, which can cause alterations in the imaging properties of the brain tissue, as well as in morphometric properties of brain structures, can be captured and detected by sophisticated

¹The human cerebral cortex is a highly folded ribbon of gray matter (GM) that lies inside the cerebrospinal fluid (CSF) and outside the white matter (WM) of the brain. Locally, its intrinsic “unfolded” structure is that of a 2D sheet, several millimeters thick. In the absence of pathology and assuming that the midline hemispheric connections are artificially closed, each cortical hemisphere can be considered as a simply-connected 2D sheet of neurons that carries the simple topology of a sphere - see Fig. 1-b

segmentation techniques. Certain clinical and research applications depend crucially on the accuracy and correctness of the representations (visualization [11, 12, 56], spherical coordinate system and surface-based atlases [12, 16–18, 22, 50, 56], shape analysis [16, 20, 31, 43, 53, 54], surface-based processing of functional data [12], and inter-subject registration [22, 51, 55], among others). Small geometric errors in a segmentation can easily change the apparent connectivity of the segmented structure, posing a problem to studies that aim at analyzing the connectedness of different regions (e.g. dramatic underestimation of true geodesic distances). The accuracy and correctness of the representations can be critical factors in the success of studies investigating the subtle early effects of various disease processes.

1.3 Challenges

Accurate segmentation under anatomical consistency is challenging. Segmentation algorithms, which operate on the intensity or texture variations of the image, are sensitive to the artifacts produced by the image acquisition process. These limitations include image noise, low frequency image intensity variations or non-uniformity due to radio frequency (RF) inhomogeneities, geometric distortions in the images, partial volume averaging effect, and subject motion. Segmentation techniques that do not integrate any topological constraints generate segmentations that often contain deviations from the true topology of the structures of interest. These deviations are called topological defects and can be of three types: cavities, disconnected components, or handles (which are topologically equivalent to holes) that incorrectly connect parts of the volumes.

The integration of topological constraints significantly increases the complexity of the task. Topology is both a global and a local concept; small and local modifications of a geometric shape can change its global connectivity. Furthermore, topology is intrinsically a continuous concept (Sect. 2.1) and topological notions are difficult to adapt into a discrete framework (Sect. 2.2). Due to these difficulties, the set of techniques applicable to the segmentation of images is quite limited (Sect. 3).

In this chapter, we present background material of central importance for the segmentation of medical images under topological constraints. We introduce some elementary notions of topology (Sect. 2.1) and show how these notions can be adapted into a discrete framework and applied to the segmentation problem (Sect. 2.2). Next, we describe the current state-of-the-art segmentation of images under topological constraints (Sect. 3) and detail the limitations of existing approaches (Sect. 3). Section 4 concludes.

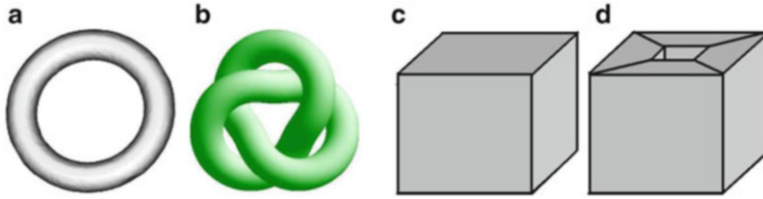


Fig. 2 a-b) Two tori that are homeomorphically equivalent. They share the same intrinsic topology. However, they do not share the same homotopy type as one cannot be continuously transformed into the other. c) A geometric object with a spherical topology; its Euler-characteristic is $\chi = v - e + f = 8 - 12 + 6 = 2$. d) A geometric object with a toroidal topology and an Euler-characteristic of $\chi = v - e + f = 16 - 32 + 16 = 0$

2 Topology in Medical Imaging

In medical image segmentation, one is interested in locating (accurately and also under topological constraints) specific regions of interest that are formed by one or several anatomical structures. Those can be represented equivalently by their volume or their surface and these two equivalent representations correspond to the two most common data structures used in medical imaging: 3D voxel grids and surfaces (such as triangulations or levelsets).

2.1 General Notions of Topology

Topology is a branch of mathematics that studies the properties of geometric figures by abstracting their inherent connectivity while ignoring their detailed form. The exact geometry of the objects, their location and the details of their shape are irrelevant to the study of their topological properties. Schematically, this amounts to characterizing a geometric object (i.e. a surface or a volume) by its number of disconnected components, holes and cavities, but not by their position. For instance, the surface of a coffee mug with a handle has the same topology as the surface of a doughnut (this type of surface is called a one-handled torus).

A - Homeomorphism, Genus, and Euler-Characteristic

In medical imaging, the geometric entities under consideration are anatomical structures, which can frequently be advantageously represented by their surfaces. The Gauss-Bonnet theorem in differential geometry links the geometry of surfaces with their topology. Any compact connected orientable surface is *homeomorphic* to a sphere with some number of handles. This number of handles is a *topological invariant* called the *genus*. For example, a sphere is of genus 0 and a torus is of genus 1. The genus g is directly related to another topological invariant called

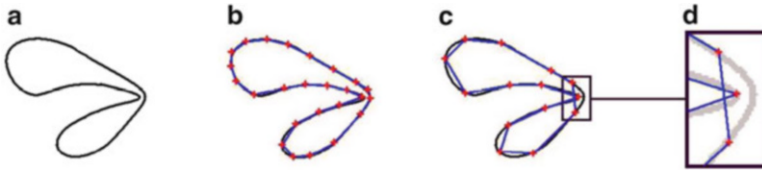


Fig. 3 a) a simple closed curve with the topology of a circle. b) One example of a polyhedral decomposition of the curve using 25 vertices and edges. The corresponding Euler-characteristic $\chi = v - e = 0$ is that of a circle. c) Another discretization of the same curve using 14 edges and vertices. Note that the Euler-characteristic is still that of a circle $\chi = 0$, even though the discrete representation of the curve self-intersects in the 2D embedding space. d) Close-up

the *Euler-characteristic* χ by the formula $\chi = 2(1 - g)$.² The Euler-characteristic is of great practical interest because it can be calculated from any polyhedral decomposition (e.g. triangulation) of the surface by the simple formula $\chi = v - e + f$, where v , e and f denote respectively the number of vertices, edges and faces of the polyhedron. The Euler-characteristic of a sphere is $\chi = 2$. This implies that any surface with $\chi = 2$ is topologically equivalent (i.e. homeomorphic) to a sphere and therefore does not contain any handles. Surfaces with an Euler-characteristic $\chi < 2$ have a topology that is different from that of a sphere. Note, however, that the Euler-characteristic does not provide any information about the localization of the handles.

B - Intrinsic Topology and Homotopy Type

Homeomorphisms are used to define the *intrinsic* topology of an object, independently of the embedding space. The topological invariance of the Euler characteristic implies that the way a surface is decomposed (i.e. tessellated) does not influence its (intrinsic) topology. Any polyhedral decomposition of a surface will encode the same intrinsic topology. For example, a knotted solid torus has the same genus (and the same Euler-characteristic $\chi = 0$) as a simple torus. In order to topologically differentiate surfaces, one needs a theory that considers the embedding space. *Homotopy*, which defines two surfaces to be homotopic if one can be continuously transformed into the other, is such a theory that provides a measure of an object's topology (see [8] for an excellent course in algebraic topology). We stress the fact that the Euler-characteristic does not define the homotopy type of a surface, since the embedding space is being ignored. In particular, this implies that a discrete representation of a surface using a polygonal decomposition with the desired Euler-characteristic might be self-intersecting in the 3D embedding space (Fig 3).

C - Topological Defects, Duality Foreground/Background

In this chapter, we call topological defect any deviation from spherical topology: cavities, disconnected components, or handles. We note that for each defect present in a geometric entity, referred to as the foreground object, there exists a correspond-

²In the case of multiple surfaces involving K connected components, the total genus is related to the total Euler-characteristic by the formula: $\chi = 2(K - g)$.

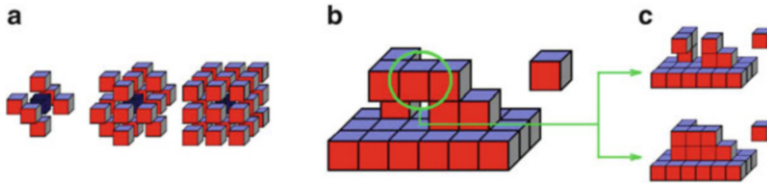


Fig. 4 a) 6-, 18- and 26-connectivity. b) The circled voxel is a non-simple point c) Several dual topological corrections are possible: either cutting the handle (top) or filling the hole (bottom)

ing defect in the background (i.e. the embedding space): a disconnected foreground component can be interpreted as a background cavity; a foreground cavity is a disconnected background component; and a handle in a foreground component defines another handle in the background component. This foreground/background duality is of crucial importance for all retrospective topology correction techniques, as it provides a dual methodology to correct a topological defect. For instance, the presence of a handle in an object could be corrected by either cutting the handle in the foreground object, or cutting the corresponding handle in the background object. Cutting the background handle can be interpreted as filling the corresponding foreground hole (Fig. 4-b,c).

2.2 Topology and Discrete Imaging

In order to apply topological concepts to a discrete framework and define the topology type (i.e. homotopy type) of digital segmentations, the notion of continuity must be adapted to discrete spaces and objects, such as 3D image grids and triangulations. This is obtained by replacing the notion of continuity with the weaker notion of connectivity. We describe how topological notions can be adapted to the two most common data structures used in medical imaging: 3D data structures and surfaces.

A - Digital Topology

Digital topology provides an elegant framework, which translates the continuous concepts of topology to discrete images. In this theory, binary images inherit a precise topological meaning. In particular, the concept of homotopic deformation, which is required to assign a topological type to a digital object, is clearly defined through the notion of *simple point*. An extensive discussion of these concepts can be found in [36]. In this section, some basic notions of digital topology are presented. All definitions are from the work of G. Bertrand, which we refer to for more details [7].

In the digital topology framework, a 3D image I is interpreted as a graph. The vertices of the graph are the digital points (i.e. the voxels) and the edges are defined through neighborhood relations between points (i.e. the connectivity).

A 3D binary digital image I is composed of a foreground object X and its inverse, the complement \bar{X} . We first need to define the concept of *connectivity*, which specifies the condition of adjacency that two points must fulfill to be regarded as connected. Three types of connectivity are commonly used in 3D: 6-, 18- and 26-connectivity. Two voxels are 6-adjacent if they share a face, 18-adjacent if they share at least an edge and 26-adjacent if they share at least a corner (Fig. 4-a). In order to avoid topological paradoxes, different connectivities, n and \bar{n} , must be used for one digital object X and its complement \bar{X} . This leaves us with four pairs of compatible connectivities: (6,26), (6,18), (26,6) and (18,6). One important consequence of the previous requirement is that digital topology does not provide a consistent framework for multi-label images. No compatible connectivities can be chosen for neighboring components of the same object. Therefore, digital topology is strictly limited to binary images.

We now come to the definition of a simple point. This concept is central to most digital segmentation methods that integrate topological constraints [3, 26, 32, 36, 42, 44, 49].

Definition 1.1 Simple point. *A point of a binary object is simple if it can be added or removed without changing the topology of both the object and the background, that is, without changing the number of connected components, cavities and handles of both X and \bar{X} (Fig. 4-b,c).*

A simple point is easily characterized by two topological numbers with respect to the digital object (X, \bar{X}) and a consistent connectivity pair (n, \bar{n}) . These numbers, denoted $T_n(\mathbf{x}, X)$ and $T_{\bar{n}}(\mathbf{x}, \bar{X})$, were introduced by G. Bertrand in [1] as an elegant way to classify the topology type of a given voxel. The values of $T_n(\mathbf{x}, X)$ and $T_{\bar{n}}(\mathbf{x}, \bar{X})$ characterize isolated, interior and border points as well as different kinds of junctions. In particular, a point is simple if and only if $T_n(\mathbf{x}, X) = T_{\bar{n}}(\mathbf{x}, \bar{X}) = 1$. Their efficient computation, which only involves the 26-neighborhood, is described in [6].

The definition of a discrete homotopy follows from the concept of simple point.

Definition 1.2 Homotopic deformation. *We define a homotopic deformation of an object X as a sequence of deletions or additions of simple points.*

Finally, two objects X and Y share the same homotopy type if there exists a sequence of transformations $X_0 \dots X_k$ and a sequence of points $\mathbf{x}_1 \dots \mathbf{x}_k$, such that $X_0 = X$ and $X_{i-1} = X_i \cup \{\mathbf{x}_i\}$ or $X_{i-1} = X_i \setminus \{\mathbf{x}_i\}$ and the point \mathbf{x}_i is simple relative to X_i for $i = 1, \dots, k$.

To be complete, we also mention some recent research in digital topology, such as the concept of multisimple point in [45, 44] to characterize the modification of the genus, a novel characterization of homeomorphic deformation in [4], as well as some octree grid topology concept in [1].

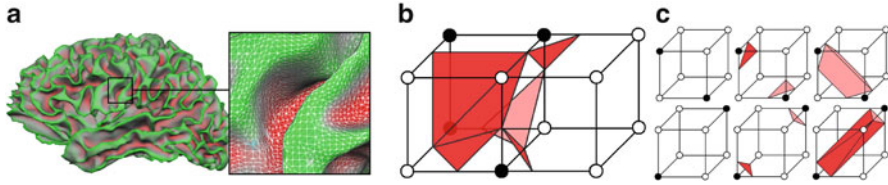


Fig. 5 a) A triangulated cortical surface b) Tiling inconsistency in isosurface extraction c) Different tessellations corresponding to different topology types

B - Surfaces in Discrete Imaging

We now turn to the translation of continuous topological concepts to discrete surface representations. There are essentially two ways of representing a surface in discrete imaging. Surfaces can be either represented explicitly, by using a parameterized polygonal decomposition, or implicitly as the level set of some function defined in the 3D embedding space. Each type of representation has advantages and disadvantages, and has been extensively used for the purpose of medical image segmentation [11, 13, 21, 23, 35, 37, 57–59].

B.1 - Explicit Representations

An explicit representation models a surface by a set of vertices, edges, and faces, associated with a chosen parameterization of each face. The set of vertices, edges, and faces composes the polyhedral representation of the surface. The parameterization of the faces determines the exact geometry of the surface. For instance, tessellations correspond to linear parameterizations of each face, while splines use higher-order approximations. Triangulations are a special kind of tessellation, in which each face is a triangle (Fig. 5-a).

The topological invariance of the Euler-characteristic implies that explicit models unambiguously encode the intrinsic topology of the surfaces. However, there is no guarantee that the surface representation is not self-intersecting. As previously mentioned, the topological equivalence defining the intrinsic topology of a geometric entity ignores the embedding space. Consequently, additional precautions must be taken in order to ensure that the discretization of a surface does not generate self-intersections. The self-intersection problem is important when the surfaces are iteratively deformed in order to match a targeted structure (i.e. theory of active contours).

Finally, we note that explicit representations can approximate surfaces at any level of precision, by using more refined meshes. Contrary to the theory of digital topology that constitutes a discrete approximation of the continuous space, and is therefore limited by the resolution of the 3D digital images, explicit representations can accurately approximate any surface by using high-resolution meshes.

B.2 - Implicit Representations

Implicit models encode the surface of interest \mathbf{C} as the level set of a higher-dimensional function ϕ defined in the embedding space \mathbb{R}^3 . The function ϕ , defined

on a 3D voxel grid, is usually the signed distance function of the surface with the contour being the zero level set of ϕ : $C = \phi^{-1}(0)$.

This type of representation has several advantages. First, no explicit representation and no parameterization are required. In the theory of active contours, this has proven to be a huge advantage as implicit representations can naturally change topology during the deformation of the model. Self-intersections, which are costly to prevent in parametric deformable models, are avoided and topological changes are automated. In addition, many fundamental properties of the surface C , such as its normal or its curvature, are easily computed from the level set function ϕ .

However, these models can only represent manifolds of codimension one without borders, i.e. closed surfaces in \mathbb{R}^3 . For the purpose of segmenting anatomical structures, the use of such representations is not a limitation. Another - more subtle - drawback of implicit representations is that, even though level sets achieve sub-voxel accuracy, the exact location of the contour depends on the image resolution. For instance, in the case of two adjacent banks of a sulcus that are closer than the resolution of the underlying voxel grid (or physically touching), the finite image resolution and the topological constraint necessitate some voxels to be labeled as outside voxels (ideally, these voxels should be the ones belonging to CSF), thus imposing a constraint on the location and accuracy of the surface model (some recent methods to alleviate this limitation have been proposed in [2, 27]).

So far, we have not specified how implicit representations can ensure that the topology of the encoded surface is the correct one. Since implicit representations make use of the underlying 3D voxel grid (through a signed distance function ϕ) to encode the contour of interest, digital topology (Sect. 2.2-A) can be used to specify the topology of the contour [28]. The foreground object X is simply defined as the set of negative grid points (i.e. $X = \{\mathbf{x} \in \mathbb{R}^3 \mid \phi(\mathbf{x}) \leq 0\}$), and the background object \bar{X} as the set of strictly positive grid points (i.e. $\bar{X} = \{\mathbf{x} \in \mathbb{R}^3 \mid \phi(\mathbf{x}) > 0\}$). Then, given a choice of compatible connectivities, the topology of the contour is determined unambiguously.

C - From Images to Surfaces: Isocontour Extraction

In the previous section, we described the manner in which topology can be adapted to the two most common data structures used in medical imaging. The ability to go from one representation to the next arises as a difficulty. Although it is possible to generate triangulations from 3D binary digital segmentations, such that the resulting topology of the surfaces is consistent with the choice of digital topology, it is not always possible to produce a digital binary representation, whose topology is similar to that of a given triangulation: digital topology constitutes a discrete approximation of the continuous space at a finite resolution, while triangulations approximate continuous surfaces at any level of precision.

The marching cubes (MC) algorithm was first introduced by Lorensen and Cline in 1987 [34] as a way to generate a polygonal decomposition (e.g. a triangulation) from a scalar field sampled on a rectilinear grid (e.g. an implicit representation). Given an isovalue, the MC algorithm quickly extracts a representation of the isosurface of the scalar field. The MC algorithm first partitions the data into a set

of cubic (or rectilinear) cells, the cell vertices being the grid points. Based on the relative polarity of their scalar value (above or below the isovalue), each vertex is assigned a binary label, which indicates whether the grid point is inside or outside the isosurface. Then, each cubic cell is processed sequentially. Patches (i.e. sets of triangles) that approximate the isosurface (based on tri-linear interpolation) are produced within each cube, and the polygon patches are naturally joined together to form the final isosurface representation.

Unfortunately, the standard marching squares or marching cubes algorithm does not generate topologically consistent tessellations, since the resulting tessellations may contain tiling and topological inconsistencies (Fig. 5-b). In order to alleviate this problem, Han et al. [28] have designed a modified connectivity-consistent marching contour algorithm, by building a specialized case table for each type of digital topology (Fig. 5-c). Extensive discussion of isocontour extraction algorithms can be found in the thesis of Han [25]. Note also some new research directions such as [1].

3 State of the Art in Segmentation under Topological Constraints

As noted previously, methods for producing topologically-correct segmentations can be broadly divided into two categories. A first set of approaches directly incorporates topological constraints into the segmentation process, while another set aims at correcting retrospectively the spherical topology of an already segmented image.

3.1 *Topologically-Constrained Segmentations*

The topology-enforcing techniques proceed by iteratively deforming a model of known topology onto a targeted structure, while preserving its topology. Several techniques have been used for the segmentation of anatomical structures, with the topological constraint taking different forms depending on the chosen method.

A - Active Contours

Depending on the representation, two different implementations are usually encountered. One encodes the manifold of interest with an explicit representation using a Lagrangian formulation [52], while another implicitly represents the contour as the level set of a function defined on higher-dimensional manifold in an Eulerian formulation [9, 40].

- Parameterized models maintain an explicit representation of the contour and preserve its initial *intrinsic* topology. Any level of accuracy can be achieved by using more refined meshes. However, the preservation of the whole topology also

requires the prevention of self-intersections, which proves to be computationally intensive and requires elaborate methods to detect and prevent surface intersection during the evolution. Also, note that the preservation of the initial topology is often a strong limitation to most explicit models, since a fully automatic and efficient handling of topology changes using explicit models remains an open issue [14, 15, 33, 41, 38, 39].

- The ability to automatically handle topological changes is a long acknowledged advantage of the level set method over explicit deformable models, but may not be desirable in some applications where some prior knowledge of the target topology is available. This is typically the case in biomedical image segmentation, where the topology of the organs and their mutual topological relations is prescribed by anatomical knowledge (in the absence of gross pathology). In order to overcome this problem, a topology-preserving variant of the level set method has been proposed in [28]. This method is based on the theory of digital topology and uses the underlying embedding space to constrain the topology of the interior of the level set. However, the strict topology preservation necessitates an initialization of the active contour that is close to its final configuration in order to avoid topological barriers that can easily generate large geometrical errors. In the case of complex structures, like the cortical surface, such initialization proves to be extremely difficult [26].

Recently, another variant was proposed to exert a more subtle topological control on a level set evolution by allowing connected components to merge, split or vanish without changing the genus of the active contours [44]. While the original level set model does not provide any topological control, topology-preserving level sets enforce a strong constraint that is often too restrictive. This framework establishes a trade-off between the two models. It offers a more fine-grained topological control that alleviates many problems of methods that enforce a strong topological constraint (e.g. sensitivity to initialization and noise, simultaneous evolution of multiple components and speed of convergence).

B - Digital Homotopic Deformations

Similar to active contour models, digital approaches [3, 5, 8, 36, 42] deform an initial region with a known given topology (typically a single voxel carrying a spherical topology), by addition/deletion of points, minimizing a global energy functional while preserving the correct digital topology. Regions are grown or shrunk by adding points that will not change the region topology. Most of these methods are based on the theory of digital topology and the notion of simple points.

C - Segmentation by Registration

Some approaches have been proposed to match a template with a given topology onto a specified MRI image [3, 10, 30]. These methods have the strong advantage of being able to enforce complex topology in the segmentation process, and to encode the spatial relationships that exist in between structures [4, 5]. Nevertheless, the design of elaborate templates that include several structures with the correct topology is challenging.

D - Limitations of Topologically-Constrained Segmentations

All these methods have the advantage of allowing the user to specify the proper topology and not allowing it to change. In the case of segmentation by registration, full brain models containing several structures can be matched onto a targeted image. Unfortunately, these methods are highly sensitive with regard to their initialization, and accurate final configurations most often require an initialization of the models that is close to its final configuration. One of the main reasons is that the energy functionals driving the deformation are typically highly non-convex and the evolution is therefore easily trapped in local minima.

A significant drawback of topologically constrained evolution is that it can lead to large geometric errors, due to the topological constraint and the presence of topological barriers (constituted by sets of non-simple points in the case of digital segmentations, self-touching and frozen surface regions in active contours segmentations). This is the case for methods that aim at segmenting the cortex starting from one single object located deep inside the cortical surface. Large topological barriers are often generated during the template deformation leading to inaccurate final segmentations. This is mostly a result of the presence of noise in the image and of the fact that topologically constrained segmentation prevents the formation of cavities (easy to detect and suppress) as well as the formation of handles.

Finally, we note that digital methods, as well as implicit representations that use the 3D embedding space to encode the surface of interest, are constrained by the finite resolution of the 3D grid and may not be able to represent deep folds in the target structure. To solve this problem, Han et al. [27] have implemented a moving grid algorithm, which aims at optimally deforming the underlying 3D grid for accurate implicit representations. Let us also mention some recent octree-based topology-preserving geometric deformable model [2].

3.2 Retrospective Topology Correction

Recently, new approaches have been developed to retrospectively correct the topology of an already-segmented image. These techniques, which do not impose any topological constraints on the segmentation process, can focus on attaining more accurate models with few topological inconsistencies to be identified and corrected post-hoc. These methods can be divided into two main classes: volume-based methods that work directly on the volume lattice and correct the topology by addition/deletion of voxels [26, 32, 46, 49], and surface-based methods that aim at modifying the tessellation by locating and cutting handles [19, 24, 47, 48].

A - Volume-Based Approaches

Most volume-based approaches have been specifically designed to enforce the spherical topology of cortical surface. These methods identify the location of the topological defects present in the segmentation by building a graph encoding the

connectivity of the segmentation; the topological defects present in the volume are then corrected by introducing some cuts in the connectivity graph (e.g. modifying the binary labels of some key voxels in the volume).

One of the most inspirational approaches in this domain is certainly the pioneering work of Shattuck and Leahy [49]. One drawback of their approach is that the “cuts”, which are necessary to correct the topological defects, can only be oriented along the Cartesian axes and give rise to “unnatural” topological corrections. Their method is based on the theory of digital topology but is limited to 6-connectivity and has not been generalized for any other connectivity rule.

Han et al. developed an algorithm to correct the topology of a binary object under any digital connectivity [26]. They detect handles by graph analysis, using successive foreground and background morphological openings to iteratively break the potential topological defects at the smallest scales. In contrast to the approach of Shattuck and Leahy, “cuts” are not forced to be oriented along cardinal axes. However, topological corrections at a specific scale depend on the choice of filter, either foreground or background morphological filter, which fails to evaluate simultaneously the effect of two complementary dual solutions (i.e. cutting the handle or filling the corresponding hole) on the corrected segmentation.

Kriegeskorte and Goeble proposed a region growing method prioritized by the distance-to-surface of the voxels in order to force the cuts to be located at the thinnest part of each topological defect [32]. The same process is applied to the inverse object, offering an alternative solution to each cut. An empirical cost is then assigned to each solution and the final decision is the one minimizing the global cost function.

While these methods can be effective, they cannot be used to correct the topology of arbitrary segmentations, as they make assumptions regarding the topology of the initial input image. Most frequently, fully-connected volumes are assumed and cavities are supposed to be removed as a preprocessing step. In addition, they do not integrate any statistical or geometric information into the topology correction process. To alleviate these limitations, Ségonne et al. [46, 45] propose a topology correction approach that is phrased within the theory of Bayesian parameter estimation and integrates statistical information into the topology correction process. In addition, no assumption is made about the topology of the initial input images.

B - Surface-Based Approaches

Approaches of the other type operate directly on the triangulated surface mesh. Topological defects are located either as intersections of wavefronts propagating on the tessellation [24, 29] or as non-homeomorphic regions between the initial triangulation and a sphere [19, 47, 48].

In [24, 29], a randomly selected vertex is used to initialize a region growing algorithm, which detects loops (i.e. topological defects) in the triangulation where wavefronts meet. Topological corrections are obtained through the use of opening operators on the triangle mesh, resulting in a fast method that depends on the initially selected vertex. In a similar work, Jaume [29] identifies minimal loops in the volume by wavefront propagation. This method assumes that the initial

triangulation was generated through the use of a topologically-consistent algorithm. The minimal loops are then used to identify non-simple voxels in the volume, which are subsequently deleted. Again, this approach orients the “cuts” along the Cartesian axes and generates “unnatural” topological corrections. In addition, both methods ignore all additional information, such as the underlying intensity profile or the expected local curvature, and the resulting topological corrections might not be accurate.

Fischl et al. [19] proposed an automated procedure to locate topological defects by homeomorphically mapping the initial triangulation onto a sphere. Topological defects are identified as regions in which the homeomorphic mapping is broken and a greedy algorithm is used to retessellate incorrect patches, constraining the topology on the sphere S while preserving geometric accuracy by a maximum likelihood optimization. Unfortunately, this method relies on a greedy algorithm and the reconstructed final surface might be inaccurate. In addition, even though the final intrinsic topology will be the correct one, the proposed method cannot guarantee that the final surface will not self-intersect.

Recently, Ségonne et al. proposed a methodology [47, 48] that alleviates most limitations of previous approaches, and is able to generate accurate topological corrections by integrating statistical and geometric information into the topology correction process while guaranteeing that the final surface will not self-intersect. Non-separating loops locate handles present in the volume, and produce topologically-corrected candidate solutions by discarding the faces that form the loops and by sealing the open mesh. The accuracy of each candidate solution is then maximized by active contour optimization. Finally, randomly-generated candidate solutions are selected based on their goodness of fit in a Bayesian framework.

C - Limitations of Retrospective Topology Correction Algorithms

Most of these methods assume that the topological defects in the segmentation are located at the thinnest parts of the segmented volume and aim to correct the topology by minimally modifying the volume or tessellation [24, 26, 29, 32, 49]. Although this will often lead to accurate results, due to the accuracy of initial segmentations, topological corrections may not be optimal: additional information, such as the expected local curvature or the local intensity distribution, may lead to different corrections (i.e. hopefully comparable to the ones a trained operator would make). In addition, digital methods often suffer from the finite resolution of the digital grids, posing a problem to the accurate location of the potential cuts.

Few methods have been proposed to integrate into the segmentation process some additional information, such as intensity or curvature [46–48]. Although the method introduced in [46] has the advantage of correcting the topology of any initial segmentation without making any assumption on the initial connectedness of the segmentation, it suffers from the same limitations as most digital methods (i.e. the location of the potential cuts). Also, as most retrospective topology correction methods it only evaluates a small number of potential topological corrections per defect (i.e. only two dual corrections), consequently failing often to produce optimal solutions. To our knowledge, only the approach developed in [48] has been proposed

thus far to generate multiple solutions and explore the full space of potential solutions in order to select the best correction of a topological defect.

4 Conclusion

In this chapter, we covered methods for integrating topological constraints into segmentation procedures in order to generate geometrically accurate and topologically correct models. We introduced some elementary but essential notions of topology (Sect. 2.1), such as the concepts of homeomorphism and homotopy that are necessary to characterize the topological type of a geometric object. We have clearly distinguished the *intrinsic* topology of an object from its homotopy type (Sect. 2.1). Also, we have emphasized the connections linking topology and differential geometry, such as the crucial notion of the Euler-characteristic of a surface.

The adaptation of the continuous concepts of topology into a discrete framework that is practical to the segmentation of medical images proves to be challenging. However, we have shown that topologically-consistent frameworks can be constructed by replacing the notion of continuity by the weaker notion of connectivity (Sect. 2.2), using concepts from the theory of digital topology is (Sect. 2.2-A). In particular, we have introduced the important concepts of simple point and topological numbers, and defined the discrete equivalent of homotopic deformations based on the notion of simple point. We have also presented isocontour extraction techniques (Sect. 2.2-C).

Finally, we described current state-of-the-art segmentation of images under topological constraints and detailed the limitations of existing approaches (Sect. 3). Among state-of-art techniques, retrospective methods (Sect. 3.2) achieve overall better results than topologically constrained segmentation methods (Sect. 3.1). In addition, techniques that integrate additional information, such as intensity or curvature, often lead to more accurate segmentations. However, segmentation under topological constraints remains a challenge with several promising research directions [2, 4, 5, 44, 48].

References

1. Y. Bai, X. Han, and J. Prince. Octree-based topology-preserving iso-surface simplification. In *MMBIA*, 2006.
2. Y. Bai, X. Han, and J. Prince. Octree grid topology preserving geometric deformable models for 3d medical image segmentation. In *International Conference on Information Processing in Medical Imaging*, 2007.
3. P. Bazin and P. D.L. Topology preserving tissue classification with fast marching and topology templates. *International Conference on Information Processing in Medical Imaging*, pages 234–245, 2005.

4. P. Bazin, L. Ellingsen, and D. Pham. Digital homeomorphisms in deformable registration. In *International Conference on Information Processing in Medical Imaging*, 2007.
5. P. Bazin and D. Pham. Statistical and topological atlas based brain image segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, 2007.
6. G. Bertrand. A boolean characterization of three-dimensional simple points. *Pattern Recognition Letters*, 17:115–124, 1996.
7. G. Bertrand and G. Malandain. A new characterization of three-dimensional simple points. *Pattern Recognition Letters*, 2(15):169–175, 1994.
8. S. Bischoff and L. Kobbelt. Isosurface reconstruction with topology control. *Pacific Graphics Proceedings*, pages 246–255, 2002.
9. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *The International Journal of Computer*, 22(1):61–79, 1997.
10. G. Christense, R. Rabbitt, and M. Miller. 3d brain mapping using a deformable neuroanatomy. *Phys. Med. Biol.*, 39:609–618, 1994.
11. A. Dale, B. Fischl, and S. M.I. Cortical surface-based analysis i: Segmentation and surface reconstruction. *NeuroImage*, 9:179–194, 1999.
12. A. M. Dale and M. I. Sereno. Improved localization of cortical activity by combining eeg and meg with mri cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, 5(2):162–176, 1993.
13. C. Davatzikos and R. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. *IEEE TMI*, 15:758–795, 1996.
14. H. Delingette. General object reconstruction based on simplex meshes. *The International Journal of Computer Vision*, 32(2):111–146, 1999.
15. H. Delingette and J. Montagnat. Shape and topology constraints on parametric active contours. *Computer Vision and Image Understanding*, 83(2):140–171, 2001.
16. R. Desikan, F. Ségonne, and et al. A computer generated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Human Brain Mapping*, 2005.
17. H. Drury, D. Van Essen, C. Anderson, C. Lee, T. Coogan, and J. Lewis. Computerized mappings of the cerebral cortex: A multiresolution flattening method and a surface-based coordinate system. *J. Cogn. Neurosci.*, 8(1):1–28, 1996.
18. B. Fischl and A. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97:11044–11049, 2000.
19. B. Fischl, A. Liu, and A. Dale. Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE TMI*, 20:70–80, 2001.
20. B. Fischl, D. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van der Kouwe, R. Killinay, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. Dale. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33:341–355, 2002.
21. B. Fischl, M. Sereno, and A. Dale. Cortical surface-based analysis ii: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9:195–207, 1999.
22. B. Fischl, M. Sereno, R. Tootell, and A. Dale. High-resolution inter-subject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8:272–284, 1999.
23. R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Cortex segmentation: A fast variational geometric approach. *IEEE TMI*, 21(2):1544–1551, 2002.
24. I. Guskov and Z. Wood. Topological noise removal. *Graphics I proceedings*, pages 19–26, 2001.
25. X. Han. *Anatomically Consistent Segmentation of Medical Imagery Using a Level Set Method and Digital Topology*. PhD thesis, Baltimore, Maryland, October 2003.
26. X. Han, C. Xu, U. Braga-Neto, and J. Prince. Topology correction in brain cortex segmentation using a multiscale, graph-based approach. *IEEE TMI*, 21(2):109–121, 2001.
27. X. Han, C. Xu, and J. Prince. A 2d moving grid geometric deformable model. *IEEE Conf. on Comp. Vis. Patt. Recog.*, pages 153–160, 2003.

28. X. Han, C. Xu, and J. Prince. A topology preserving level set method for geometric deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):755–768, 2003.
29. S. Jaume. *Topology Simplification Algorithm for the Segmentation of Medical Images*. PhD thesis, University of Louvain (Belgium), Feb 2004.
30. B. Karaç and C. Davatzikos. Topology preservation and regularity in estimated deformation fields. *International Conference on Information Processing in Medical Imaging*, pages 426–437, 2003.
31. R. Kikinis and etal. Temporal love sulco-gyral pattern anomalies in schizophrenia: An in vivo mr three-dimensional surface rendering study. *Neuroscience Letters*, 182:7–12, 1994.
32. N. Kriegeskorte and R. Goeble. An efficient algorithm for topologically segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage*, 14:329–346, 2001.
33. J.-O. Lachaud and A. Montanvert. Deformable meshes with automated topology changes for coarse-to-fine 3D surface extraction. *Medical Image Analysis*, 3(2):187–207, 1999.
34. W. Lorensen and H. Cline. Marching cubes: A high-resolution 3D surface reconstruction algorithm. *ACM Computer Graphics*, 21(4):163–170, 1987.
35. D. MacDonald, N. Kabani, D. Avis, and A. Evens. Automated 3d extraction of inner and outer surfaces of cerebral cortex from mri. *NeuroImage*, 12:340–356, 2000.
36. J.-F. Mangin, V. Frouin, I. Bloch, J. Regis, and J. Lopez-Krahe. From 3d magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical Imaging and Vision*, 5:297–318, 1995.
37. T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108, 1996.
38. T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey, 1999 update. *Handbook of Medical Image Processing*, 1999.
39. T. McInerney and D. Terzopoulos. T-snakes: Topology adaptive snakes. *Medical Image Analysis*, 4:73–91, 2000.
40. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.
41. J.-P. Pons and J.-D. Boissonnat. Delaunay deformable models: Topology-adaptive meshes based on the restricted delaunay triangulation. In *Conference on Computer Vision and Pattern Recognition*, 2007.
42. F. Poupon, J.-F. Mangin, D. Hasboun, C. Poupon, I. Magnin, and V. Frouin. Multi-object deformable templates dedicated to the segmentation of brain deep structures. *LNCS*, 1496:1134–1143, 1998.
43. D. Salat, R. Buckner, A. Snyder, D. Greve, R. Desikan, E. Busa, J. Morris, A. Dale, and B. Fischl. Thinning of the cerebral cortex in aging. *Cerebral Cortex*, 14(7):721–730, July 2004.
44. F. Ségonne. Active contours under topology control genus preserving level sets. *The International Journal of Computer*, 2007.
45. F. Ségonne. *Segmentation of Medical Images under Topological Constraints*. PhD thesis, Massachusetts Institute of Technology, December 2005.
46. F. Ségonne, E. Grimson, and B. Fischl. Topological correction of subcortical segmentation. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, volume 2879-2, pages 695–702, 2003.
47. F. Ségonne, E. Grimson, and B. Fischl. A genetic algorithm for the topology correction of cortical surfaces. In *Proceedings of Information Processing in Medical Imaging*, LNCS, volume 3565, pages 393–405, 2005.
48. F. Ségonne, J. Pacheco, and B. Fischl. A geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *TMI*, 26(4):518–529, 2007.
49. D. Shattuck and R. Leahy. Automated graph based analysis and correction of cortical volume topology. *IEEE TMI*, 20(11):1167–1177, 2001.
50. J. Tanabe, D. Amend, N. Schuff, V. DiScialfani, F. Ezekiel, D. Norman, G. Fein, and M. Wiener. Tissue segmentation of the brain in alzheimer’s disease. *J. Neuroradiol.*, 18:115–123, 1997.

51. X. Tao, X. Han, M. Rettmann, J. Prince, and C. Davatzikos. Statistical study on cortical sulci of human brains. *Proceedings of Inf. Proc. in Med. Imag.*, pages 37–49, 2001.
52. D. Terzopoulos, A. Witkin, and M. Kass. Constraints on Deformable Models: Recovering 3D shape and Nonrigid Motion. *Artificial Intelligence*, 36(1):91–123, 1988.
53. P. Thompson, D. MacDonald, M. Mega, C. Holmes, A. Evans, and A. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *J. Comput. Assist. Tomogr.*, 21(4):567–581, 1998.
54. P. Thompson, J. Moussai, S. Zohoori, A. Goldkorn, A. Khan, M. Mega, G. Small, J. Cummings, and A. Toga. Cortical variability and asymmetry in normal aging and alzheimer’s disease. *Cerebral Cortex*, 8(6):492–509, 1998.
55. M. Vaillant and C. Davatzikos. Hierarchical matching of cortical features for deformable brain image registration. *Proceedings of Inf. Proc. in Med. Imag.*, pages 182–195, 1999.
56. D. Van Essen and H. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Journal of Neuroscience*, 17(18):7079–7102, 1997.
57. C. Xu, D. Pham, and J. Prince. Medical image segmentation using deformable models. *Handbook of Medical Imaging - Medical Image Processing and Analysis*, 2:129–174, 2000.
58. C. Xu, D. Pham, M. Rettmann, D. Yu, and J. Prince. Reconstruction of the human cerebral cortex from magnetic resonance images. *IEEE TMI*, 18:467–480, 1999.
59. X. Zeng, L. Staib, R. Schultz, and J. Duncan. Segmentation and measurement of the cortex from 3d mr images using coupled surfaces propagation. *IEEE TMI*, 18:100–111, 1999.

Monte Carlo Sampling for the Segmentation of Tubular Structures

C. Florin, N. Paragios, and J. Williams

Abstract In this paper, we present a multiple hypotheses testing for the segmentation of tubular structures in medical imaging that addresses appearance (scanner artifacts, pathologies, ...) and geometric (bifurcations) non-linearities. Our method represents vessels/tubular structures as sequences of state vectors (vessel cuts/cross-sections), which are described by the position of the corresponding plane, the center of the vessel in this plane and its radius. Thus, 3D segmentation consists in finding the optimal sequence of 2D planes normal to the vessel's centerline. This sequence of planes is modeled by a probability density function (pdf for short) which is maximized with respect to the parameters of the state vector. Such a pdf is approximated in a non-parametric way, the Particle Filter approach, that is able to express multiple hypotheses (branches). Validation using ground truth from clinical experts and very promising experimental results for the segmentation of the coronaries demonstrates the potential of the proposed approach.

C. Florin (✉)

Corporate Technology, Siemens Corporation, 755 College Rd E, Princeton, NJ 08540, United States

e-mail: charles.florin@siemens.com

N. Paragios

Center for Visual Computing, Department of Applied Mathematics, Ecole Centrale Paris, Paris, France

e-mail: nikos.paragios@ecp.fr

J. Williams

CEO at Siemens Healthcare Molecular Imaging, Siemens Healthcare, Nürnberg, DE

e-mail: james.williams@siemens.com

1 Introduction

1.1 *Presentation of the problem*

Cardio-vascular diseases are the leading cause of death in the western world; there is a constant demand for improvement of diagnostic tools to detect and measure anomalies in the coronary tree. Coronary arteries are narrow vessels (between 3 to 5 mms next to the aorta, between 1.5 to 2.5 mms after two branchings). Their role is to feed the heart muscle with oxygenated blood, and their segmentation provides a valuable tool for clinicians to diagnose pathologies such as calcifications and stenoses. Nevertheless, their segmentation is a difficult task because of the low contrast conditions, bifurcations, intensity distortions produced by pathologies and scanner artifacts, and the coronaries' proximity to the heart chambers [26].

1.2 *Previous work in Tubular Structures Segmentation*

Tubular structure segmentation techniques consist of model-free and model-based methods. Skeleton-based techniques are the most fundamental among the model-free [31] and aim at detecting the vessel centerlines, from which the whole vessel tree is reconstructed. Region growing methods [37] progressively segment the vessels from a seed point, based on intensity similarity between adjacent pixels. These methods are successful for homogeneous regions, but pathological vessels are more challenging, and may leak into other structures of similar intensity. Morphological operators [11] can be applied to correct a segmentation, smooth its edges or eventually fill holes in the structure of interest, but fail to account for prior knowledge. Tracking approaches [18, 32] are based on the application of local operators to track the vessel. Given a starting condition, such methods recover the vessel centerline through processing vessel cross section information [17]. Various forms of edge-driven techniques, similarity/matching terms between the vessel profile in successive planes, as well as their combination, have been considered to perform tracking. In particular a method that relies on the average outward flux of the gradient vector field of the Euclidean distance from the vessel boundary to recover skeleton points has been developed in [3].

Deformable models may either be parametric or geometric. Parametric deformable models [27] can be viewed as elastic surfaces (often called snakes), and can handle topological changes with some difficulties. Geometric deformable models [4, 28], on the contrary, can change their topology during the process and may eventually leak into neighboring structures or vasculature. Like snakes, deformable models aim at minimizing the energy computed along the model. Level sets [24] are a way to apply deformable models to non-linear problems, such as vessel segmentation [21]. One may use the fast marching algorithm and its variant for vessel segmentation using the minimal path principle [2, 7, 35] to determine the

path of minimal length between two points, backtracking from one point toward the other crossing the isosurfaces perpendicularly. To discourage leaking, a local shape term that constrains the diameter of the vessel was introduced in [22]. One should also mention the method introduced in [20], where the optimization of a co-dimension two active contour was presented to segment brain vessels.

Model-based techniques, on the other hand, use prior knowledge and features to match a model with the input image and extract the vessels. The knowledge may concern the whole structure, or consist of a local region of the vessel. Along this direction, vessel template matching techniques (deformable template matcher) [25] have been investigated. The template model consists of a series of connected nodes that is deformed to best match the input image. Generalized Cylindrical models are modified in Extruded Generalized Cylinders in [23] to recover vessels in angiograms. For highly curved vessels, the local basis used for classical generalized cylinders may be twisted, and a non-orthogonality issue may occur. This problem is solved by keeping the vessel cross section orthogonal to the centerline and the two normal vectors always on the same side of the tangent vector spine as the algorithm moves along the vessel. In [19], the vessel is modeled by a tubular structure, and segmented by filtering the image with a multiscale structural term derived from the image intensity Hessian matrix [14, 36]. More recently, [30] a vessel likelihood is obtained from a classifier trained over a bag of multiscale filters. After the likelihood is computed over the whole image, a tracing algorithm tracks the vessel from a seed point.

1.3 Overview of our method

Existing approaches suffer from certain limitations. Techniques such as local operators, region growing, morphological filters and geometric contours are prone to be sensitive to local minima and fail to take into account prior knowledge on the form of the vessel. Alternatively, cylindrical models, parametric active contours and template matching techniques may not be well suited to account for the eventual distortions of vessel appearance produced by pathologies or scanner artifacts, and require special handling of bifurcations. Tracking methods, on the other hand, may often fail in the presence of missing and corrupted data, or sudden changes. Level sets are time-consuming when they are implemented in the most general way. On the other hand, their efficient implementation using the fast marching method [29, 33] reduces computational burden at the cost of losing the local implicit function properties. To improve the segmentation results, one must account for bifurcations, sudden changes of pixels intensity and missing data.

The segmentation problem is replaced by a tracking problem (Sect. 2): the course of the vessel is followed by obtaining the set of 2d planar (normal to the centerline) segmentation (see Fig. 1). On a particular plane, the vessel is represented by a model (see Sect. 2.1), whose parameters are optimized to fit the image data according to a shape and an appearance measure. However, to follow the centerline by always selecting the maximum likelihood in the parameter space is not sufficient. That is

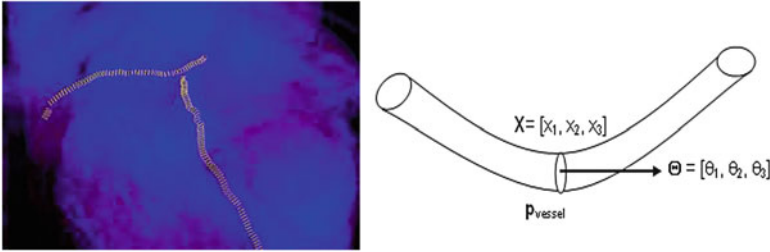


Fig. 1 The feature space is defined by the cross-section center position $\mathbf{x} = (x_1, x_2, x_3)$, the cross-section tangential direction $\Theta = (\theta_1, \theta_2, \theta_3)$ and the lumen pixel intensity distribution $\mathbf{p}_{\text{vessel}}$

the reason why the authors are driven toward a method that would handle multiple hypotheses, and keep only the few most probable following [12, 13]. At each step, a scheme based on particle filtering [8] is used to sample the parameters probability density function (pdf). To these samples is assign a probability measure (Sect. 2.2) that is updated at every step with the prior value and the model's fitness to the new image features. In Sect. 3, the experimental method is explained and the results are presented in Sect. 3.3. Finally, Sect. 4 concludes this chapter with a discussion.

2 Segmentation Model & Theoretical Foundations

2.1 Vessel Model & Particle Filters

To explain our method at a concept level, let us assume that a segment of the vessel has been detected: a 2D shape on a 3D plane. Similar to region growing and front propagation techniques, our method aims to segment the vessel in adjacent planes. To this end, one can consider the hypotheses ω of the vessel being at a certain location (\mathbf{x}), having certain orientation (Θ), and referring to certain shape - an elliptic model is a common choice (ϵ) - with certain appearance characteristics ($\mathbf{p}_{\text{vessel}}$).

$$\underbrace{\mathbf{x} = (x_1, x_2, x_3)}_{\text{position}}, \underbrace{\Theta = (\theta_1, \theta_2, \theta_3)}_{\text{orientation}}, \underbrace{\epsilon = (\alpha, \beta, \phi)}_{\text{shape}}, \underbrace{\mathbf{p}_{\text{vessel}}}_{\text{appearance}} \quad (1)$$

Then, segmentation consists in finding the optimal parameters of ω given the observed 3D volume. Let us consider a probabilistic interpretation of the problem with $\pi(\omega)$ being the posterior distribution that measures the fitness of the vector ω with the observation. Under the assumption that such a law is present, segmentation consists in finding at each step the set of parameters ω that maximizes $\pi(\omega)$. However, since such a model is unknown, one can assume an autoregressive

mechanism that, given prior knowledge, predicts the actual position of the vessel and a sequential estimate of its corresponding states. To this end, we define:

- a state vector ω composed of \mathbf{x} , Θ , ϵ and \mathbf{p}_{vessel} (Eq. (1))
- an iterative process to predict the next state and update the density function, that can be done using a Bayes sequential estimator and is based on the computation of the present state ω_t pdf of a system, based on observations from time 1 to time t $z_{1:t} : \pi(\omega_t | z_{1:t})$. Assuming that one has access to the prior pdf $\pi(\omega_{t-1} | z_{1:t-1})$, the posterior pdf $\pi(\omega_t | z_{1:t})$ is computed according to the Bayes rule:

$$\pi(\omega_t | z_{1:t}) = \frac{\pi(z_t | \omega_t) \pi(\omega_t | z_{1:t-1})}{\pi(z_t | z_{1:t-1})}$$

- a distance between prediction and actual observation, based on the observation.

Simple parametric models will be susceptible to fail with vessels' irregularities (pathologies, prosthesis, ...). Therefore instead of optimizing a single state vector, multiple hypotheses are generated and weighted according to actual observation. Nevertheless, in practical cases, it is impossible to compute exactly the posterior pdf $\pi(\omega_t | z_{1:t})$. An elegant approach to implement such a technique refers to the use of particle filters where each given hypothesis is a state in the feature space (or particle), and the collection of hypothesis is a sampling of the feature space.

Particle Filters [1, 8] are sequential Monte-Carlo techniques that are used to estimate the Bayesian posterior probability density functions (pdf) [16, 34]. In terms of a mathematical formulation, such a method approximates the posterior pdf by M random measures $\{\omega_t^m, m = 1..M\}$ associated to M weights $\{\lambda_t^m, m = 1..M\}$, such that

$$\pi(\omega_t | z_{1:t}) \approx \sum_{m=1}^M \lambda_t^m \delta(\omega_t - \omega_t^m), \tag{2}$$

where each weight λ_t^m reflects the importance of the sample ω_t^m in the pdf.

The samples ω_t^m are drawn using the principle of *Importance Density* [9], of pdf $q(\omega_t^m | x_{1:t}^m, z_t)$ and it is shown that their weights λ_t^m are updated according to

$$\lambda_t^m \propto \lambda_{t-1}^m \frac{\pi(z_t | \omega_t^m) \pi(\omega_t^m | \omega_{t-1}^m)}{q(\omega_t^m | \omega_{t-1}^m, z_t)}. \tag{3}$$

Once a set of samples has been drawn, $\pi \left(\omega_t^m | \omega_{t-1}^m, z_t \right)$ can be computed out of the observation z_t for each sample, and the estimation of the posteriori pdf can be sequentially updated.

2.2 Prediction & Observation: Distance

This theory is now applied to vessel tracking. Each one of the particles ω_t^m represents a hypothetic state of the vessel; a probability measure $p \left(z_t | \omega_t^m \right)$ is used to quantify how the image data z_t fits the vessel model ω_t^m . To this end, we are using the image terms, and in particular the intensities that do correspond to the vessel in the current cross-section. The vessel's cross-section is defined by the hypothetic state vector (see Eq. (1)) with a 3D location, a 3D orientation, a lumen's diameter and a pixel intensity distribution model (the multi-Gaussian). The observed distribution of this set is approximated using a Gaussian mixture model according to the Expectancy-Maximization principle. Each hypothesis is composed by the features given in Eq. (1), therefore, the probability measure is essentially the likelihood of the observation z , given the appearance A model. The following measures (loosely called probabilities) are normalized so that their sum over all particles is equal to one. Assuming statistical independence between shape S and appearance model A , $p(z_t | \omega_t) = p(z_t | S)p(z_t | A)$.

- Probability measure for shape based on contrast

Given the vessel model (see Eq. (1)), whose parameters are specified by the particle ω_t , a measure of contrast, that we call the *ribbon measure* R , is computed:

$$\begin{cases} R = -\infty, & \mu_{int} \leq \mu_{ext} \\ R = \frac{\mu_{int} - \mu_{ext}}{\mu_{int} + \mu_{ext}}, & otherwise \end{cases} \quad (4)$$

The probability of the observation given the shape model is then computed:

$$p(z | S) = e^{-\frac{|R|}{R_0}} \quad (5)$$

where R_0 is a normalizing constant (the average value of R from ground truth), μ_{int} is the mean intensity value for the voxels in the vessel, and μ_{ext} is the intensities mean value for the voxels in a band outside the vessel, such that the band and the vessel's lumen have the same area. This measure is normalized

to be equivalent to model a probability measure. Since the coronary arteries are brighter than the background, the best match maximizes R .

- Probability measure for appearance

For the vessel lumen pixels distribution \mathbf{p}_{vessel} Eq. (1), the probability is measured as the distance between the hypothesized distribution and the distribution actually observed.

The distance we use is the symmetrized Kullback-Leibler distance $D(p, q)$ between the model $p(x) = \mathbf{p}_{vessel}$ and the observation $q(x)$:

$$D(p, q) = \int p(x) \log \left(\frac{p(x)}{q(x)} \right) + q(x) \log \left(\frac{q(x)}{p(x)} \right) dx, \quad (6)$$

$$p(z | A) = e^{-\frac{|D(p,q)|}{D_0}}, \quad (7)$$

where D_0 is a normalizing constant, equal to the average value of D from prior knowledge. The appearance model p is a bi-modal Gaussian, and the observation q is the histogram of the 2D cross-section. In practice, q is modeled by a bi-modal Gaussian with an Expectancy-Minimization algorithm since such an approximation is used for the model p . Equation (6) is then computed summing the divergence over all pixels on the 2D cross-section. This method does not provide the exact solution for Eq. (6) but is fast and precise enough for our case, with 2D cross-sections of size 31×31 . Once again, this measure $p(z|A)$ is normalized to be equivalent to a probability measure.

2.3 Branching Detection

When a branching occurs, the particles naturally split up in the two daughter branches (the case of trifurcation is not studied here), and then track them separately (see Fig. 2). As branchings are never perfectly balanced, one of them attracts the majority of the particles after few resampling steps. To avoid the collapse of one of the modes, two techniques are available: either to increase the number of particles in the weakest branch, or to treat the two branches separately. The second approach is preferred in this paper, for the particular context of vessel segmentation; therefore, after branchings are detected, each mode is treated as an entirely separate new particle filter. To this end, a simple K-means [10] clustering in the joint space (position + orientation) of the particles is considered at each iteration. When the two clusters are well separated (when the distance between the clusters center is above a certain threshold), the number of particles is doubled and they are equally dispatched in the two branches. The segmentation goes on, according to Eq. (3), by treating the two modes as entirely distinct particle filters.

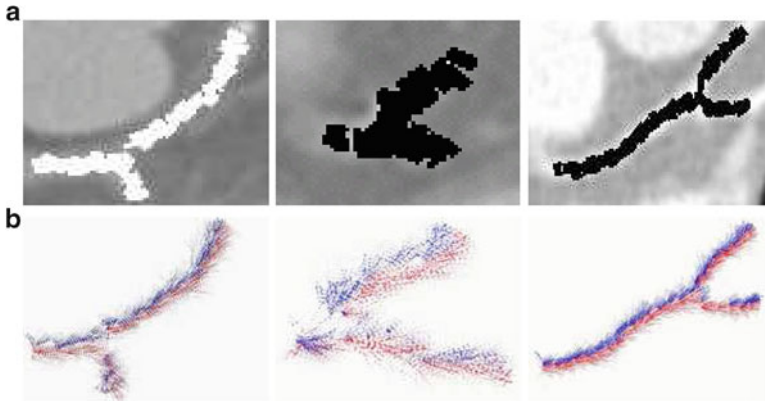


Fig. 2 (a) branching points between LCX and LAD for three patients with the particles' mean state overlaid, (b) the particles are clustered at each time step. The branching is detected when the distance between the two clusters center is above a certain threshold

Table 1 Intensity range for different organs coded on 12 bits

	myocardium	vessel lumen / ventricles	calcification	lungs
intensity	900-1100	1100-1300	1400-2000	0-200

3 Experimental Validation

3.1 Image Modality and Algorithm Specifications

The algorithm was tested on 34 CT images from different SOMATON scanners and patients who presented different or no pathologies. A typical voxel resolution is $0.3 \text{ mm} \times 0.3 \text{ mm} \times 1 \text{ mm}$. Contrast agent was used for all images, with different concentration and different products. Table 1 summarizes the typical intensity range for different tissues, as they are found in a CT angiography volume, with pixels' value coded on 12 bits. No preprocessing is applied before the segmentation procedure described in this article.

Regarding the initial configuration, the use of approximately 1, 000 particles gave sufficient results for our experiments. We performed a systematic resampling according to the SIR every time the effective sampling size $N_{eff} = \sum_i 1/w_i^2$ (where w_i is the weight of the i th particle) falls below half the number of particles. The preference for SIR is motivated by the robustness of the segmentation. The tracking stops when the sum of image measures at a given iteration falls below a given threshold.

Table 2 Results table showing the number of cases for which branches are incorrectly segmented, over a dataset of 34 patients, using Particle Filters (PF) and Front Propagation (FP), with respect to expert ground truth

vessel name	RCA	Acute Marg.	LAD	First Septal	LCX	Obtuse Marg.
# missed, PF	none	5	none	2	none	2
# missed, FP	12	28	16	23	21	26

3.2 Comparison with Front Propagation

Our method is compared with Front Propagation, implemented using the Fast Marching algorithm [5], based on a curvilinear structures detection [36]. The Hessian analysis is used to detect tubular structures; this measure (called “vesselness” in [15]) is integrated into a potential map on which the Fast Marching algorithm is run such as in [6]. In few words, Front Propagation computes isosurfaces in a Riemannian space, whose metric is based on the image: the vesselness measure in our case. The front propagates faster along the vessel than in other non-tubular structures. However, in the case of intensity inhomogeneities, this measure drops and the front either stops or leaks into neighboring structures.

In the synthetic case, the error measure Δ is defined as the symmetric difference between ground truth G and segmentation S :

$$\Delta = 1 - \frac{2|G \cap S|}{|G| + |S|}.$$

Since ground truth is not available for the real case studies, an expert visually validates the number of branches correctly segmented and missed.

3.3 Results

The algorithm has been evaluated on 34 patients, and has successfully recovered all the main arteries (RCA, LAD, LCX) for each patient as shown in Table 2, while a small portion of visual results are also presented in Fig. 3. The results in Table 2 corresponds to the number of branches segmented by Particle Filters and identified by a human expert. For comparison purposes, the same test is performed using Front Propagation based on the image Hessian matrix [36]. These results were achieved with a one-click initialization. All patients presented some kind of artery pathologies in one, at least, of their coronary vessels (12 cases with calcification, 8 stenosis, 4 stent, 2 bypasses), and many present intensity artifacts (7 stepping, 5 hardening effects). Our approach has successfully segmented both healthy and unhealthy coronaries without leaking into neighboring structures (over-segmentation). The method seems to outperform regarding the detection of the main branchings, while in some cases branching of lower clinical importance at the distal part of the tree have been missed.

Table 3 Results table showing the symmetric difference between maximum likelihood segmentation result (particle with maximum probability at each step), and ground truth on synthetic data

Radius of curvature	3.5	7	10.5	13.3	14
symmetric difference	7.80 %	9.29 %	9.24 %	12.11 %	11.07 %
Gap width	4	6	8	12	16
symmetric difference	0	2.22 %	7.11 %	11.56 %	24.44 %

A second test is performed on a 2D synthetic dataset with Gaussian pixel intensity distributions (vessel: mean = 1200, standard deviation = 50; background: mean = 1000, standard deviation = 100) noisier and with less contrast than real cases. Using real cases radiuses (3-4 pixels lumen diameters), and using the segmentation provided by the particle of maximum weight, the symmetric difference between segmentation and ground truth is 11.07 % for a tubular structure with a 14 pixels radius of curvature, and 2.22 % for a 6 pixels wide simulated lumen obstruction (Table 3). The algorithm is stopped when the sum of the particles' weight falls below a threshold.

4 Conclusion

In this chapter, we have shown that Monte-Carlo sampling and multiple hypotheses testing can be used for the segmentation of tubular structures. In the context of vascular segmentation, Particle Filters sequentially estimate the pdf of segmentations in a particular feature space. The case of coronary arteries was considered to validate such an approach where the ability to handle discontinuities on the structural (branching) as well as appearance space (calcifications, pathological cases, etc.) was demonstrated. The main advantage of such methods lies in their capability to handle intensity inhomogeneities from pathologies and bifurcations. Experiments were conducted on several healthy and diseased patients CTA data sets, segmenting the *Left Main Coronary Artery* and the *Right Coronary Artery* Fig. 3.

As a final remark, it may be underlined that particle filters require heavy computational time in a general context (around an hour for a CT $512 \times 512 \times 300$ volumetric image), compared to front propagation. However, in the case of non linear problems, with missing or corrupt data, particle filters provide a better segmentation than deterministic methods. In the particular case of coronary arteries segmentation, due to pathologies, contrast agent heterogeneities and branchings, the use of non deterministic methods have been proved successful. Therefore, when time is a constraint, a compromise is to be found to apply deterministic methods in linear cases and statistical modelization, such as Particle Filtering, in all other cases.

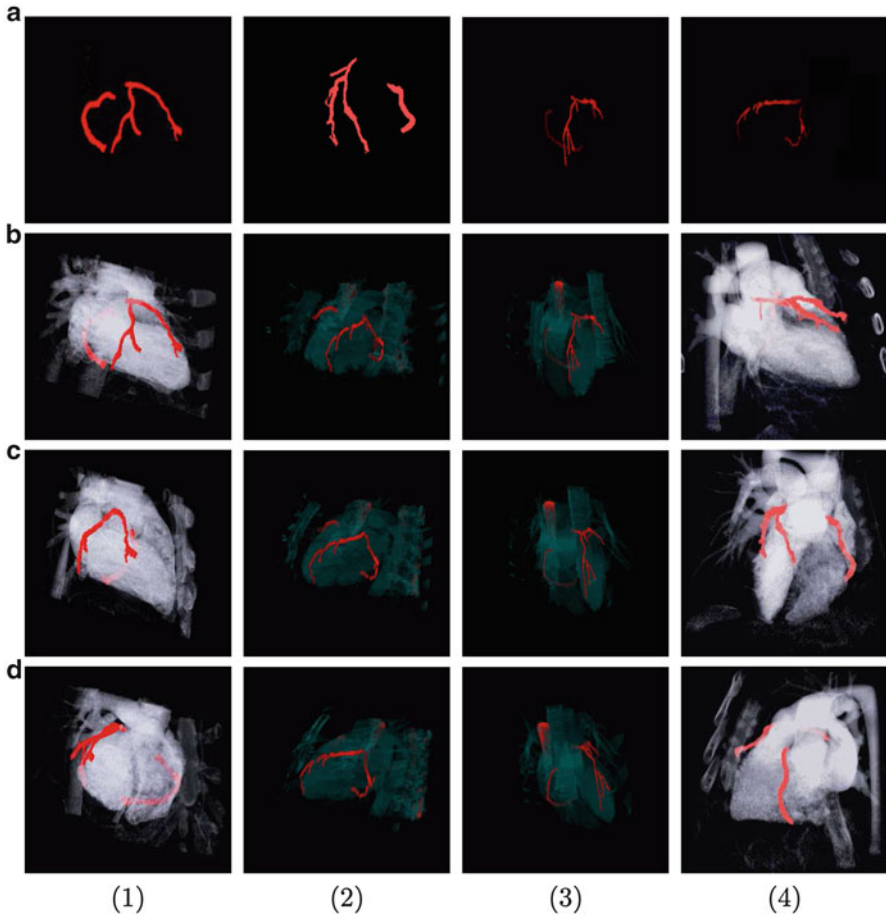


Fig. 3 Segmentation of the Left anterior descending coronary artery and Right coronary artery in CTA (in red) for four patients ; (a) coronary tree, (b,c,d) Different 3D views super-imposed to the cardiac volume are presented

References

1. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking. *IEEE Trans. on Signal Process.*, 50:174–188, 2002.
2. B. Avants and J. Williams. An adaptive minimal path generation technique for vessel tracking in CTA/CE-MRA volume images. In *Med. Image Comput. Comput. Assist. Interv. Int. Conf.*, volume 3749, pages 707–716. Springer, 2000.
3. S. Bouix, K. Siddiqi, and A. R. Tannenbaum. Flux driven automatic centerline extraction. In *Med. Image Anal.*, volume 9, pages 209–221(3), 2005.
4. V. Caselles, F. Catté, B. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.

5. T. Deschamps. *Curve and Shape Extraction with Minimal Path and Level-Sets techniques-Applications to 3D Medical Imaging*. PhD thesis, Université Paris-IX Dauphine, Place du maréchal de Lattre de Tassigny, 75775 Paris Cedex, Dec. 2001.
6. T. Deschamps and L. Cohen. Fast extraction of tubular and tree 3d surfaces with front propagation methods. In *IARP International Conference on Pattern Recognition*, volume 1, pages 731–734. IEEE Computer Society, 2002.
7. T. Deschamps and L. D. Cohen. Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. *Med. Image Anal.*, 5(4):281–299, Dec. 2001.
8. A. Doucet, J. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
9. A. Doucet, N. Gordon, and C. Andrieu. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10(3):197–208, 2000.
10. R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.
11. M. Figueiredo and J. Leitao. A nonsmoothing approach to the estimation of vessel contours in angiograms. *IEEE Trans. Med. Imaging*, 14:162–172, 1995.
12. C. Florin, N. Paragios, and J. Williams. Particle filters, a Quasi-Monte Carlo solution for segmentation of coronaries. In *Med. Image Comput. Comput. Assist. Interv. Int. Conf.*, pages 246–253, 2005.
13. C. Florin, N. Paragios, and J. Williams. Globally optimal active contours, sequential monte carlo and on-line learning for vessel segmentation. In *European Conference on Computer Vision*, volume 3953, pages 476–489, 2006.
14. A. Frangi, W. Niessen, P. Nederkoorn, O. Elgersma, and M. Viergever. Three-dimensional model-based stenosis quantification of the carotid arteries from contrast-enhanced MR angiography. In *IEEE Mathematical Methods in Biomedical Image Analysis*, pages 110–118, 2000.
15. A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. *Lecture Notes in Computer Science*, 1496, 1998.
16. N. Gordon. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEEE Proceedings*, 140:107–113, 1993.
17. M. Hart and L. Holley. A method of Automated Coronary Artery Tracking in Unsubtracted Angiograms. *IEEE Comput. in Cardiol.*, pages 93–96, 1993.
18. M. Isard and A. Blake. Contour Tracking by Stochastic Propagation of Conditional Density. In *European Conference on Computer Vision*, volume I, pages 343–356, 1996.
19. K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trussset. Model based detection of tubular structures in 3d images. *Computer Vision and Image Understanding*, 80:130–171, 2000.
20. L. Lorigo, O. Faugeras, E. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C. Westin. Codimension-Two Geodesic Active Contours for the Segmentation of Tubular Structures. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages I:444–451, 2000.
21. R. Malladi and J. Sethian. A Real-Time Algorithm for Medical Shape Recovery. In *IEEE International Conference in Computer Vision*, pages 304–310, 1998.
22. D. Nain, A. Yezzi, and G. Turk. Vessel Segmentation Using a Shape Driven Flow. In *Med. Image Comput. Comput. Assist. Interv. Int. Conf.*, 1, pages 51–59. Springer, 2004.
23. T. O’Donnell, T. Boult, X. Fang, and A. Gupta. The Extruded Generalized Cylinder: A Deformable Model for Object Recovery. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 174–181, 1994.
24. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
25. R. Petrocelli, K. Manbeck, and J. Elion. Three Dimensional Structure Recognition in Digital Angiograms using Gauss-Markov Models. In *Comput. in Radiol.*, pages 101–104. IEEE, 1993.
26. F. L. Ruberg. *Computed Tomography of the Coronary Arteries*. London, UK: Taylor & Francis, 2005.
27. D. Rueckert, P. Burger, S. Forbat, R. Mohiadin, and G. Yang. Automatic Tracking of the Aorta in Cardiovascular MR images using Deformable Models. *IEEE Trans. Med. Imaging*, 16: 581–590, 1997.

28. J. Sethian. A Review of the Theory, Algorithms, and Applications of Level Set Methods for Propagating Interfaces. *Cambridge University Press*, pages 487–499, 1995.
29. J. Sethian. *Level Set Methods*. Cambridge University Press, 1996.
30. M. Sofka and C. V. Stewart. Retinal vessel extraction using multiscale matched filters, confidence and edge measures. *IEEE Trans. Med. Imaging*, 25(12):1531–1546, 2006.
31. E. Sorantin, C. Halmai, B. Erbohelyi, K. Palagyi, K. Nyul, K. Olle, B. Geiger, F. Lindbichler, G. Friedrich, and K. Kiesler. Spiral-CT-based assessment of Tracheal Stenoses using 3D Skeletonization. *IEEE Trans. Med. Imaging*, 21:263–273, 2002.
32. K. Toyama and A. Blake. Probabilistic Tracking in a Metric Space. In *IEEE International Conference in Computer Vision*, pages 50–59, 2001.
33. J. Tsitsiklis. Efficient Algorithms for Globally Optimal Trajectories. *IEEE Transactions on Automatic Control*, 40:1528–1538, 1995.
34. W. West. Modeling with mixtures. In J. Bernardo, J. Berger, A. Dawid, and A. Smith, editors, *Bayesian Statistics 4*. Clarendon Press, 1993.
35. O. Wink, W. J. Niessen, and M. A. Viergever. Multiscale vessel tracking. *IEEE Trans. Med. Imaging*, 23(1):130–133, 2004.
36. S. Y., N. S., S. N., A. H., Y. S., K. T., G. G., and K. R. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med. Image Anal.*, 2:143–168(26), 1998.
37. P. Yim, P. Choyke, and R. Summers. Grayscale Skeletonization of Small Vessels in Magnetic Resonance Angiography. *IEEE Trans. Med. Imaging*, 19:568–576, 2000.

Non-rigid registration using free-form deformations

D. Rueckert and P. Aljabar

Abstract Free-form deformations are a powerful geometric modeling technique which can be used to represent complex 3D deformations. In recent years, free-form deformations have gained significant popularity in algorithms for the non-rigid registration of medical images. In this chapter we show how free-form deformations can be used in non-rigid registration to model complex local deformations of 3D organs. In particular, we discuss diffeomorphic and non-diffeomorphic representations of 3D deformation fields using free-form deformations as well as different penalty functions that can be used to constrain the deformation fields during the registration. We also show how free-form deformations can be used in combination with mutual information-based similarity metrics for the registration of mono-modal and multi-modal medical images. Finally, we discuss applications of registration techniques based on free-form deformations for the analysis of images of the breast, heart and brain as well as for segmentation and shape modelling.

1 Introduction

The analysis of medical images plays an increasingly important role in various clinical applications. Image registration is a key component in many image analysis applications. The goal of image registration is to find corresponding anatomical locations in two images. Image registration can be applied to images from the

D. Rueckert (✉)

Department of Computing, Imperial College London, 180 Queen's Gate,
London SW7 2AZ, UK

e-mail: d.rueckert@imperial.ac.uk

P. Aljabar

Department of Biomedical Engineering, Division of Imaging Sciences, King's College
London, Lambeth Wing, St Thomas' Hospital, London SE1 7EH, UK

e-mail: Paul.Aljabar@kcl.ac.uk

same subject acquired by different imaging modalities or at different time points as well as to images acquired from different subjects. To bring images into registration it is usually necessary to estimate a geometric transformation which aligns the images. Most non-rigid registration techniques use either elastic [3, 29], fluid [4, 8, 14, 15] or other deformation models [20, 33, 54, 58, 61] to represent this geometric transformation. In this chapter we focus on registration algorithms which use a particular deformation model, namely free-form deformations based on B-splines.

In general, finding the optimal geometric transformation is achieved by minimization of a cost function which measures the degree of (mis-)alignment of the images as a function of the geometric transformation. Most registration algorithms use a cost function based on image intensity information to directly measure the degree of (mis-)alignment of the images. These methods are called voxel-based registration techniques and are especially successful since they do not require any feature extraction or segmentation of the images. Comprehensive reviews of image registration techniques can be found in [31, 44, 66].

2 Image registration using free-form deformations

The goal of image registration is to relate any point in the reference or *target* image to the *source* image, i.e. to find the optimal transformation $\mathbf{T} : \mathbf{p} \mapsto \mathbf{p}'$ which maps any point in the target image I_A into its corresponding point in the source image I_B . The transformation \mathbf{T} can be separated into two components: A global component (e.g. a rigid or affine transformation) and a local component. Thus, the transformation \mathbf{T} can be written as:

$$\mathbf{T}(\mathbf{p}) = \mathbf{T}_{global}(\mathbf{p}) + \mathbf{T}_{local}(\mathbf{p}) \quad (1)$$

The global transformation typically accounts for variations in the position, orientation and scaling between the two images. However, the global transformation cannot account for any local deformations.

In the late eighties a number of techniques for modeling deformations emerged in the computer graphics community. In particular, Sederberg and Parry developed free-form deformations (FFD) [57] as a powerful modelling tool for 3D deformable objects. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3D object and produces a smooth and continuous transformation. In the original paper by Sederberg and Parry [57] trivariate Bernstein polynomials were used to interpolate the deformation between control points. A more popular choice is to use trivariate B-spline tensor products as the deformation function [37, 38]. The use of FFDs based on B-splines for image registration was first proposed by Rueckert et al. [53, 54]. Over the last decade the use of FFDs for image registration has attracted significant interest [36, 45, 50, 51].

2.1 Free-form deformations

To define a spline-based FFD we denote the domain of the image volume as $\Omega = \{\mathbf{p} = (x, y, z) \mid 0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$. Let Φ denote a $n_x \times n_y \times n_z$ mesh of control points $\phi_{i,j,k}$ with uniform control point spacing δ . Then, the FFD can be written as the 3D tensor product of the familiar 1D cubic B-splines:

$$\mathbf{T}_{local}(\mathbf{p}) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n} \quad (2)$$

where $i = \lfloor \frac{x}{\delta} \rfloor - 1, j = \lfloor \frac{y}{\delta} \rfloor - 1, k = \lfloor \frac{z}{\delta} \rfloor - 1, u = \frac{x}{\delta} - \lfloor \frac{x}{\delta} \rfloor, v = \frac{y}{\delta} - \lfloor \frac{y}{\delta} \rfloor, w = \frac{z}{\delta} - \lfloor \frac{z}{\delta} \rfloor$ and where B_l represents the l -th basis function of the B-spline [37,38]:

$$\begin{aligned} B_0(u) &= (1 - u)^3 / 6 \\ B_1(u) &= (3u^3 - 6u^2 + 4) / 6 \\ B_2(u) &= (-3u^3 + 3u^2 + 3u + 1) / 6 \\ B_3(u) &= u^3 / 6 \end{aligned}$$

In contrast to thin-plate splines [6] or elastic-body splines [23], B-splines are locally controlled which makes them computationally efficient even for a large number of control points. In particular, the basis functions of cubic B-splines have a limited support, i.e. changing control point $\phi_{i,j,k}$ affects the transformation only in the local neighbourhood of that control point.

The derivative of a coordinate transformation is the matrix of its partial derivatives. In the case of 3D coordinate systems this is always a 3×3 matrix called the Jacobian matrix and the determinant of this matrix is called the Jacobian determinant of the transformation, or simply the Jacobian. This determinant measures how infinitesimal volumes change under the transformation. For this reason, the Jacobian determinant is the multiplicative factor needed to adjust the differential volume form when applying the coordinate transformation.

An advantage of B-spline FFDs is the fact the derivatives of the transformation can be computed analytically. The derivatives of the transformation are often used in the optimization of the registration and for regularization as well as in the subsequent analysis of the resulting transformation. The Jacobian matrix of transformation is defined as:

$$\mathbf{J}(\mathbf{p}) = \begin{pmatrix} \frac{\partial T_x(\mathbf{p})}{\partial x} & \frac{\partial T_x(\mathbf{p})}{\partial y} & \frac{\partial T_x(\mathbf{p})}{\partial z} \\ \frac{\partial T_y(\mathbf{p})}{\partial x} & \frac{\partial T_y(\mathbf{p})}{\partial y} & \frac{\partial T_y(\mathbf{p})}{\partial z} \\ \frac{\partial T_z(\mathbf{p})}{\partial x} & \frac{\partial T_z(\mathbf{p})}{\partial y} & \frac{\partial T_z(\mathbf{p})}{\partial z} \end{pmatrix} \quad (3)$$

The determinant of this matrix measures how infinitesimal volumes change under the transformation and can be used to analyze the local behavior of the transformation. A positive value of the determinant of the Jacobian matrix can be interpreted as follows:

$$J(\mathbf{p}) = \det |\mathbf{J}(p)| = \begin{cases} > 1 & \text{volume expansion} \\ = 1 & \text{no volume change} \\ < 1 & \text{volume contraction} \end{cases} \quad (4)$$

If the value of the determinant changes from positive to negative the transformation is folding and is no longer a one-to-one transformation. Since the FFD is the tensor product of independent 1D B-splines, the derivative of the local transformation \mathbf{T}_{local} with respect to x is computed as follows:

$$\frac{\partial \mathbf{T}_{local}(\mathbf{p})}{\partial x} = \frac{1}{\delta_x} \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \frac{dB_l(u)}{du} B_m(v) B_n(w) \phi_{i+l, j+m, k+n} \quad (5)$$

The remaining derivatives have an analogous form. The computation of the derivatives of the B-spline basis functions B_i itself is straightforward.

The control points Φ act as parameters of the B-spline FFD and the degree of non-rigid deformation which can be modelled depends essentially on the resolution of the mesh of control points Φ . A large spacing of control points allows modelling of global non-rigid deformations while a small spacing of control points allows modelling of highly local non-rigid deformations. At the same time, the resolution of the control point mesh defines the number of degrees of freedom and consequently the computational complexity. A common approach uses a multi-resolution approach for FFDs in which the resolution of the control point mesh is increased in a coarse to fine fashion (see Fig. 1). An arbitrary FFD based on B-splines can be refined to an identical deformation with half the control point spacing along each dimension. In the 1D case, the control point positions ϕ' of the refined grid can be computed from the coarse control points ϕ [26]:

$$\phi'_{2i+1} = \frac{1}{2}(\phi_i + \phi_{i+1}) \quad \text{and} \quad \phi'_{2i} = \frac{1}{8}(\phi_{i-1} + \phi_{i+1} + 6\phi_i) \quad (6)$$

This equation can be easily generalized to 3D by applying the tensor product. Another possibility is the use of non-uniform FFDs [55]. This can be achieved by introducing a control point status associated with each control point in the mesh, marking it either active or passive. Active control points can be modified during the registration process, whereas passive control points remain fixed. An alternative approach for FFDs based on non-uniform rational B-splines (NURBS) has been proposed by Wang and Jiang [64].

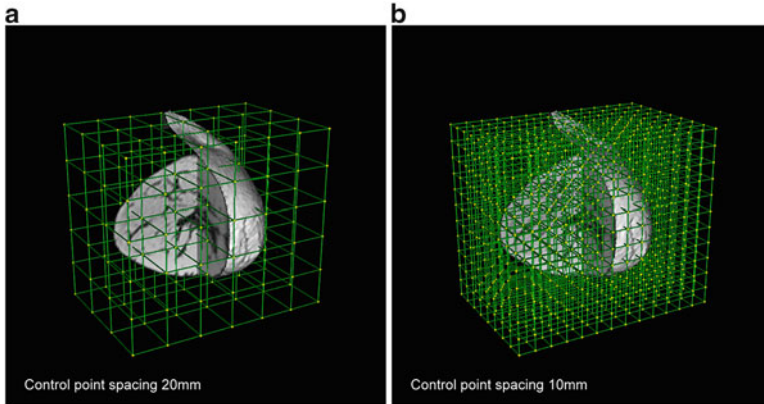


Fig. 1 A free-form deformation control point mesh (a) before subdivision (control point spacing 20 mm) and (b) after subdivision (control point spacing 10 mm)

In general FFDs are defined on a Cartesian coordinate system, however it is possible to use other coordinate systems. For example, Chandrashekara proposed the use of a FFD model defined in a cylindrical coordinate system for the registration of cardiac MR images [11]. In a similar fashion, Lin et al. [40] proposed the use of extended free-form deformations (EFFD) [19] for the registration of cardiac MR images. Another, more generic, approach has been recently proposed by Chandrashekara [10]: In this approach free-form deformations are defined on lattices with arbitrary topology [42]. An advantage of this approach is the fact that the control point mesh can be closely adapted to the geometry of the anatomy studied, e.g. epi- and endocardial surfaces of the left ventricle.

2.2 *Voxel-based similarity measures for free-form deformations*

To relate a point in the target image to the source image, one must define a similarity criterion (or cost function) which measures the degree of alignment between both images. A popular choice for this are voxel-based similarity measures which use the image intensities directly and do not require the extraction of any features such as a landmarks, curves or surfaces. Commonly used voxel-based similarity measures include the sum of squared differences (SSD) or cross-correlation (CC). However, these measures make rather strong assumptions about the relationship of the image intensities in both images which is not suitable for multi-modality registration. Even in the case of mono-modality registration this assumption is often violated, e.g. in contrast-enhanced imaging. An alternative voxel-based similarity measure is mutual information (MI) which was independently proposed by Collignon [16] and

Viola [62]. Mutual information is based on the concept of information theory and expresses the amount of information in one image A that explains a second image B ,

$$\mathcal{E}_{\text{similarity}}(A, B) = H(A) + H(B) - H(A, B) \quad (7)$$

where $H(A)$, $H(B)$ denote the marginal entropies of A , B and $H(A, B)$ denotes their joint entropy. These entropies can be estimated from the joint histogram of A and B or using kernel density estimators like Parzen windowing. If both images are aligned the mutual information is maximised. It has been shown by Studholme [60] that mutual information itself is not independent of the overlap between two images. To avoid any dependency on the amount of image overlap, Studholme suggested the use of normalised mutual information (NMI) as a measure of image alignment:

$$\mathcal{E}_{\text{similarity}}(A, B) = \frac{H(A) + H(B)}{H(A, B)} \quad (8)$$

Similar forms of normalised mutual information have been proposed by Maes et al. [43]. A more complete review of the use of mutual information for image registration can be found in [48].

2.3 Optimization of free-form deformations

Like many other problems in computer vision, registration can be formulated as an optimisation problem whose goal is to minimise an associated energy or cost function. The most general form of such a cost function in registration is

$$\mathcal{C} = -\mathcal{E}_{\text{similarity}} + \lambda \mathcal{E}_{\text{penalty}} \quad (9)$$

This type of cost function comprises two competing goals: The first term represents the cost associated with the image similarity $\mathcal{E}_{\text{similarity}}$ in Eqs. (7) or (8) while the second term penalizes certain transformations and thus constrains the behavior of the transformation (different penalty functions will be discussed in the next section). The parameter λ is a weighting parameter which defines the trade-off between the alignment of the two images and the penalty function of the transformation. From a probabilistic point of view, the cost function in Eq. (9) can be explained in a Bayesian context: The similarity measure can be viewed as a likelihood term which expresses the probability of a match between source and target image while the penalty function represents a prior which encodes a-priori knowledge about the expected transformation.

In the original paper by Rueckert et al. [54] the optimization of the FFD is carried out in a multi-resolution fashion via a steepest gradient optimization algorithm. More recently, Klein et al. has compared different optimization strategies for FFDs [35].

2.4 Penalty functions for free-form deformations

Typically, non-rigid image registration is an ill-posed problem. Thus, it is necessary to add some constraints to render the problem well-posed. A common approach is to enforce the smoothness of the deformation [54]. Free-form deformations based on B-splines are intrinsically smooth (at least relative to the control point spacing), however additional smoothness can be enforced by adding a penalty term which regularizes the transformation. The general form of such a smoothness penalty term has been described by Wahba [63]. In 3D, the penalty term takes the following form

$$\begin{aligned} \mathcal{E}_{smooth} = & \int \left(\frac{\partial^2 \mathbf{T}}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial y^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial z^2} \right)^2 \\ & + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial yz} \right)^2 d\mathbf{p} \end{aligned} \quad (10)$$

This quantity is the 3D counterpart of the 2D bending energy of a thin-plate of metal and defines a cost function which is associated with the smoothness of the transformation. Note that this regularization term is zero for an affine transformation and therefore penalises only non-affine transformations [63].

Another class of penalty functions aim to generate biomechanically plausible deformations. For example, Rohlfing et al. suggested a constraint which preserves volume [51]:

$$\mathcal{E}_{volume} = \int |\log(J(\mathbf{p}))| d\mathbf{p} \quad (11)$$

Here $J(\mathbf{p})$ is the determinant of the Jacobian matrix \mathbf{J} of the free-form deformation. As mentioned previously the Jacobian measures how infinitesimal volumes change under the transformation. This function therefore penalizes the compression or expansion of tissues or organs during the registration. It should be noted that the penalty term above penalizes volume changes over the entire domain, however due to the integration there may be small regions in the image which show a large volume change while the majority of regions show no volume change. Other authors have proposed a rigidity constraint which forces the deformation in certain regions to be nearly rigid [41], e.g.

$$\mathcal{E}_{rigidity} = \int \|\mathbf{J}(\mathbf{p})\mathbf{J}(\mathbf{p})^T - \mathbf{1}\| d\mathbf{p} \quad (12)$$

The penalty functions above do not guarantee that the resulting deformation field is diffeomorphic (smooth and invertible). In order to ensure that the FFD is diffeomorphic it is possible to add a penalty function which penalizes non-diffeomorphic

transformations, e.g. transformations which introduce folding. One suitable penalty function for this has the following form:

$$\mathcal{E}_{\text{folding}} = \int \mathcal{P}(\mathbf{p}) d\mathbf{p} \quad (13)$$

where

$$\mathcal{P}(\mathbf{p}) = \begin{cases} \frac{\gamma^2}{|J(\mathbf{p})|^2} - 2 & \text{if } |J(\mathbf{p})| \leq \gamma \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

A similar penalty function was first proposed by Edwards et al. [25] and effectively penalises any transformations for which the determinant of the Jacobian falls below a threshold γ . By penalising Jacobians that approach zero, one can prevent the transformation from collapsing and ensure diffeomorphisms. Note that simply using a smoothness penalty function would not be sufficient to guarantee a diffeomorphic transformation, since it is possible for a transformation to be smooth but non-diffeomorphic.

2.5 Diffeomorphic free-form deformations

In general, most registration algorithms make the assumption that similar structures are present in both images. Therefore it is desirable that the deformation field be smooth and invertible (so that every point in one image has a corresponding point in the other). Such smooth, invertible transformations are called diffeomorphisms. Choi and Lee [13] have derived sufficient conditions for the injectivity of FFDs which are represented in terms of control point displacements. These sufficient conditions can be easily tested and can be used to guarantee a diffeomorphic FFD. Without loss of generality we will assume in the following that the control points are arranged on a lattice with unit spacing. Let $\Delta \mathbf{c}_{i,j,k} = (\Delta x_{i,j,k}, \Delta y_{i,j,k}, \Delta z_{i,j,k})$ be the displacement of control point $\mathbf{c}_{i,j,k}$. Let $\delta_x = \max |\Delta x_{i,j,k}|$, $\delta_y = \max |\Delta y_{i,j,k}|$, $\delta_z = \max |\Delta z_{i,j,k}|$.

Theorem 1. *A FFD based on cubic B-splines is locally injective over all the domain if $\delta_x < \frac{1}{K}$, $\delta_y < \frac{1}{K}$ and $\delta_z < \frac{1}{K}$.*

Choi and Lee [13] have determined a value of $K \approx 2.48$ so that the maximum displacement of control points given by the bound $\frac{1}{K}$ is approximately 0.40. This means that the maximum displacement of control points is determined by the spacing of control points in the lattice. For example, for a lattice with 20mm control point spacing the maximum control point displacement is 8mm while for a lattice with 2.5mm control point spacing the maximum control point displacement is 1mm. In practice the bounds on the displacements are too small to model any realistic deformations. To model large deformations one can use a composition of FFDs as

proposed in [30]. For each FFD in this composition, the maximum control point displacement is limited by theorem 1. This is fundamentally different to the multi-level FFDs mentioned earlier since the FFDs are concatenated,

$$\mathbf{T}(\mathbf{p}) = \mathbf{T}_n \circ \mathbf{T}_{n-1} \circ \cdots \circ \mathbf{T}_2 \circ \mathbf{T}_1(\mathbf{p}) \quad (15)$$

so that the final deformation is a composition of FFDs. Since the composition of two diffeomorphisms produces a diffeomorphism one can construct a diffeomorphic deformation by ensuring that each individual FFD is diffeomorphic.

3 Applications

In this section we discuss some of the applications of non-rigid registration based on FFDs. In particular, we will concentrate on the use of non-rigid registration for the alignment of images of the breast, heart and brain. In addition we will discuss how registration techniques based on FFDs can be used for generic image analysis tasks, i.e. image segmentation or shape modelling.

3.1 *Non-rigid registration in mammography*

A common registration problem is the matching of images of the same subject acquired at different times. The need for non-rigid registration arises from the fact that most tissues are far from rigid and can deform considerably. This tissue deformation may be caused by patient motion, cardiac motion or respiratory motion. An example application is given the use of non-rigid registration of contrast-enhanced MRI of the breast. Here, the difference between the rate of uptake of contrast agent in healthy and cancerous tissue can be used to identify cancerous lesions. The rate of uptake of contrast agent is estimated using the difference between pre- and post-contrast image, and any motion between the images complicates its estimation. Due to the highly deformable nature of the breast tissue, non-rigid registration techniques are required to correct for this motion. One of the first applications of non-rigid registration based on FFDs was the alignment of pre- and post-contrast MRI of the breast [54]. The performance of the non-rigid registration algorithm has been evaluated in detail in a clinical study [24] as well as in a validation study in which biomechanically plausible deformations have been simulated using FEM techniques [56]. More recently it has been observed that non-rigid registration techniques based on FFDs can lead to biomechanically unrealistic deformations, e.g. deformations which include local tissue expansion or compressions even though normal tissue is nearly incompressible. Thus, the use of penalty functions which prevent local volume change has been proposed (see Sect. 2.4) [51].

3.2 *Non-rigid registration in cardiovascular image analysis*

Recent advances in non-invasive imaging modalities allow for high-resolution imaging of the cardiovascular system. Among these imaging modalities, MRI is playing an increasingly important role. The advantages of MRI include the flexibility of choosing tomographic planes at any desired position and orientation with high spatial and temporal resolution. Because of the high soft tissue contrast, different anatomical structures can be clearly delineated, enabling their accurate measurement and assessment. In addition to imaging of anatomical structures, MRI may also be used to acquire functional information (such as the perfusion or contractility of myocardial tissue) about the cardiac system. The combination of both types of measurement provides valuable information for the study of the cardiovascular system. Non-rigid registration is an important tool for the analysis of cardiovascular MRI as it allows the modelling cardiac and respiratory motion as well as the fusion of MRI with other imaging modalities such as ultrasound.

In cardiac MRI, non-rigid registration can be used to model the motion of the myocardium. This enables the extraction of quantitative information on the dynamics of the myocardium such as wall thickening or radial and circumferential strain. In order to extract myocardial motion MR tagging is often used where a pattern of stripes in the myocardial tissue is created by selective excitation to invert or saturate a series of planes through the heart at specific points in time [65] or by spatial modulation of magnetisation (SPAMM) [2]. The tracking of these tags throughout the cardiac cycle is a challenging task since it is difficult to locate the tags and fading of the tags towards the end of the cardiac cycle due to T1 relaxation. In addition, to fully reconstruct the 3D motion of the myocardium, it is necessary to acquire multiple-slice short-axis (SA) and long-axis (LA) images of the heart and to track the tags in both SA and LA images simultaneously. The problem of motion tracking in tagged MR images can be formulated as a registration problem [12]. The advantage of such a registration-based approach is that tag localisation and motion modelling can be achieved simultaneously as part of the non-rigid registration process. To track the myocardium in a sequence of short-axis and long-axis image volumes V_0, \dots, V_N one can use a multi-level FFD in which the transformation $\mathbf{T}(\mathbf{p}, t)$ is represented as the sum of FFDs [55]:

$$\mathbf{T}(\mathbf{p}, t) = \sum_{i=1}^t \mathbf{T}_{local}^i(\mathbf{p})$$

After registering the volume V_1 to V_0 we obtain a single FFD representing the motion of the myocardium at time $t = 1$. To register volume V_2 to V_0 a second FFD is added to the sequence of FFDs to yield the transformation at time $t = 2$. This process continues until all volumes in the sequence are registered, allowing us to relate any point in the myocardium at time $t = 0$ to its corresponding point throughout the sequence. An example of cardiac motion tracking using non-rigid registration is shown in Fig. 2. In this figure a virtual tag grid is aligned with the

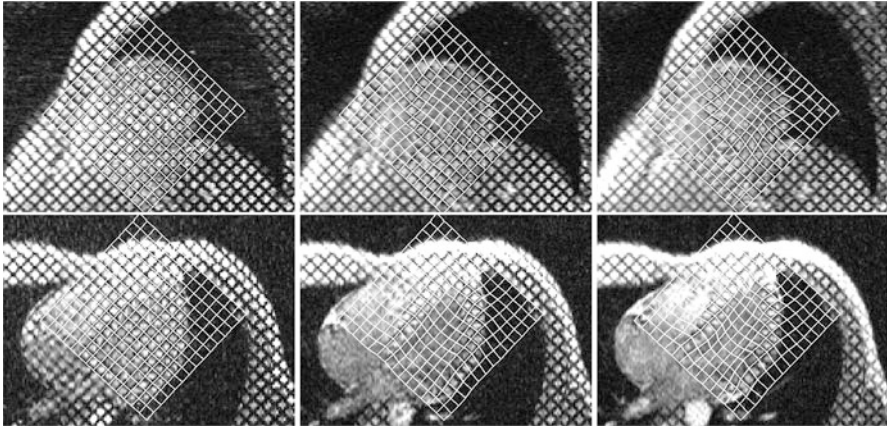


Fig. 2 The short-axis images (top) and the long-axis images (bottom) taken at different times are registered to their corresponding images taken at time $t = 0$ (left top and bottom) to recover the deformation within the myocardium. The short-axis and long-axis images show a virtual tag grid which has been aligned with the tag pattern at time $t = 0$. As time progresses, the virtual tag grid is deformed by the free-form deformation and follows the underlying tag pattern in the images

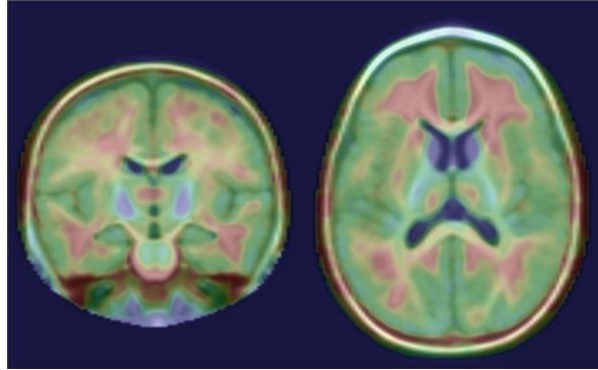
tag pattern at time $t = 0$. As time progresses the virtual tag grid is deformed by the free-form deformation according to the cardiac motion. If the cardiac motion has been recovered the virtual tag grid will follow the underlying tag pattern in the images.

A different application of non-rigid registration is the correction of respiratory motion in contrast-enhanced dynamic MR. Here the motion induced by free breathing during the acquisition degrades the images. It has been shown that respiratory motion induces significant deformations of the heart [46]. Non-rigid registration based on free-form deformations has been used successfully to correct for this respiratory motion [47].

3.3 *Non-rigid registration for neurological image analysis*

The non-rigid registration of brain images has led to a variety of applications in the context of neurological studies in which the transformations between image pairs are the focus of attention. In longitudinal studies, changes in the brain over time can be modeled using non-rigid transformations between serially acquired images. These longitudinal changes can range from the dramatic and complex growth of the early years to the subtle changes due to neurodegeneration later in life. In cross-sectional studies, non-rigid registrations between images of different subjects can be used to characterised inter-subject variability or the differences between an individual and a reference image. Approaches that focus on the transformations in

Fig. 3 An atlas illustrating patterns of volume change between one and two years of age for a group of 25 subjects - hotter colours indicate expansion and colder colours indicate contraction. The Jacobian determinant maps derived from multiple longitudinal FFD registrations were aligned in order to generate this atlas [1]



this way have been referred to as deformation- or tensor-based morphometry (DBM or TBM). In a simple cross-sectional study, for example, where two clinical groups are to be separated, it is possible to non-rigidly register all the images to a reference brain image. This implies that differences between the subjects are now encoded in the resulting transformations and the properties of these transformations can be used to identify group differences. An example application of this approach was the identification of volumetric changes in the brain due to pre-term birth carried out by Boardman et al. [5]. This work focussed on the Jacobian determinant maps of the subject-reference FFD transformations.

In a longitudinal study, a transformation between baseline and follow-up images of the brain can be used to identify the patterns of change occurring between scans. The analytic form of FFD transformations makes properties such as the Jacobian tensor and its determinant, the curl, divergence etc. readily calculated. The Jacobian determinant in particular can be used to identify regions of expansion and contraction over time due tissue growth or loss. TBM studies based on FFDs have been used to identify degenerative changes in the brain associated with alcoholism [9, 52] and with dementia [7, 59]. FFDs have also been used to characterise patterns of growth in the brain in children between the ages of one and two years [1] where Jacobian maps derived from multiple intra-subject registrations were spatially normalised and averaged to create an atlas of growth for the group (see Fig. 3).

3.4 Atlas-based segmentation using non-rigid registration

A natural consequence of the development of non-invasive imaging techniques has been the development of methods for delineating or segmenting anatomical structures within the images. Segmented images can then be used to provide clinically useful volumetric or morphometric information for example by identifying growth for particular anatomical structures or identifying blood flow or functional activ-

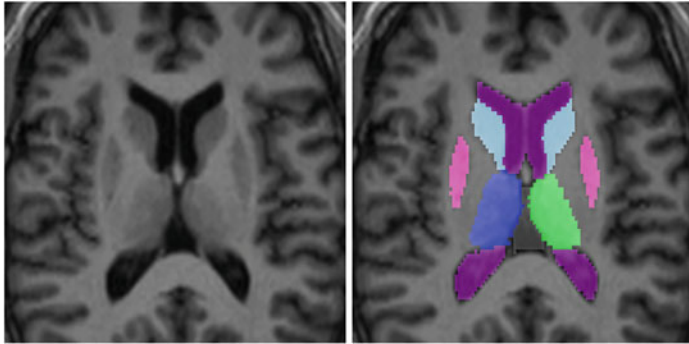


Fig. 4 Left: An MRI image of the anatomy of a subject. Right: Overlay of the segmentation for a group of sub-cortical structures obtained by using atlas-based segmentation and classifier fusion

ity within specific regions. However, accurate labelling of complex anatomical structures - such as those in the sub-cortical region of the brain - is a difficult task which is generally best achieved by trained and experienced human raters. The correspondence obtained by non-rigid registration can be used to overcome this problem and to create a segmentation by the propagation of a labelling; from an image for which an accurate segmentation is available to an unseen or query image. The image for which labels are available is often termed an ‘atlas’ and this approach to segmentation is known as atlas-based segmentation. The main prerequisites for accuracy in atlas-based segmentation are accurate registration of the atlas and query images, accurate labelling in the original atlas and that the labels are anatomically representative of the query. Many of the errors due to limitations in accuracy can, however, be overcome by repeating the process: registering multiple atlases to the query and transforming all of their labels. The propagated labels are then treated as classifiers and fused on a per-voxel basis in order to obtain the final segmentation estimate.

This classifier fusion approach can use any of a number of methods at the fusion stage but a simple majority vote rule for the final label assignment has been shown to perform very well against other atlas-based methods by Rohlfing et al. [49] where FFD registrations were used to align the atlases with the query image. FFD registrations were also used by Heckemann et al. [32] to carry out classifier fusion based on a pool of manually labelled human brain images. These images were used as query images in a comprehensive set leave-one-out experiments. The manual labels for each subject were used to assess the quality of segmentations resulting from classifier fusion and the results obtained showed the approach to be robust and accurate, reaching levels of accuracy comparable with expert raters.

3.5 *Statistical shape modeling using non-rigid registration*

Statistical models of shape variability have been successfully applied to perform various image analysis tasks in 2D and 3D images. In particular, their application for image segmentation in the context of active shape models (ASM) has been very successful [18]. In building such statistical models, a set of segmentations of the shape of interest is required as well as a set of landmarks that can be unambiguously defined in each sample shape. An extension of ASMs are the so-called Active Appearance Models (AAM) [17] which have been used for atlas matching. AAMs incorporate not only information about the spatial distribution of landmarks and the intensity information at the landmarks, but also about their underlying texture distribution. However, a fundamental problem when building these models is the fact that they require the determination of point correspondences between the different shapes. The manual identification of such correspondences is a time consuming and tedious task. This is particularly true in 3D where the number of landmarks required to describe the shape accurately increases dramatically compared with 2D applications.

In recent years several approaches have been proposed for the automatic identification point correspondences between shapes. For example, Leventon et al. [39] proposed an approach in which the statistical analysis is carried out directly on the signed distance maps of a set of aligned shapes. This approach effectively assumes that the corresponding points of two shapes are the closest points. Another solution to the point correspondence problem is to construct correspondences implicitly via the shape parameterisation, e.g. by using spherical harmonics [34]. Davies et al. [21, 22] proposed a method for establishing correspondences between shapes by optimising the quality of the statistical model based on the resulting minimum description length.

An alternative approach is to use non-rigid registration to establish correspondences across shapes: The approach by Frangi et al. [27, 28] assumes that each shape is represented as a labeled image. During the first step of the shape model construction an atlas of all shapes is constructed by rigid or affine registration of all labeled images into a common coordinate system. After this alignment, the distance transformations of the labeled images are averaged to produce a shape atlas. This atlas is then landmarked by extracting pseudo-landmarks along the surfaces of the shape, e.g. by applying a marching cubes algorithm followed by surface decimation. Non-rigid registrations based on FFDs [54] between the shapes and the atlas are then used to propagate the landmarks to each shape in the dataset.

4 Discussion

In this chapter we have discussed non-rigid registration techniques based on B-spline FFDs and mutual information. We have shown how B-spline FFDs can be constrained to achieve biomechanically plausible deformations as well as how

B-spline FFDs can be used for diffeomorphic registration. In our experience B-spline FFDs are a powerful and flexible tool for estimating local deformations. These local deformations can be either the result of motion, deformation, growth or disease processes. It should be pointed out that non-rigid registration is very much an area of on-going research and most algorithms are still in the stage of development and evaluation. The lack of a generic gold standard for assessing and evaluating the success of non-rigid registration algorithms is one of their most significant drawbacks. In the absence of any such gold standard, it is necessary to compare any non-rigid registration algorithm to other established techniques.

References

1. P. Aljabar, K. Bhatia, M. Murgasova, J. Hajnal, J. Boardman, L. Srinivasan, M. Rutherford, L. Dyet, A. Edwards, and D. Rueckert. Assessment of brain growth in early childhood using deformation based morphometry. *NeuroImage*, 39(1):348–358, 2008.
2. L. Axel and L. Dougherty. Heart wall motion: Improved method of spatial modulation of magnetization for MR imaging. *Radiology*, 172(2):349–350, 1989.
3. R. Bajcsy and S. Kovačič. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
4. M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.
5. J. P. Boardman, K. Bhatia, S. Counsell, J. Allsop, O. Kapellou, M. A. Rutherford, A. D. Edwards, J. V. Hajnal, and D. Rueckert. An evaluation of deformation-based morphometry in the developing human brain and detection of volumetric changes associated with preterm birth. In *Sixth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '03)*, Lecture Notes in Computer Science, 2003.
6. F. L. Bookstein. Principal Warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
7. R. Boyes, D. Rueckert, P. Aljabar, J. Whitwell, J. Schott, D. Hill, and N. Fox. Cerebral atrophy measurements using jacobian integration: Comparison with the boundary shift integral. *NeuroImage*, 32(1):159–169, 2006.
8. M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Proc. 4th International Conference Visualization in Biomedical Computing (VBC'96)*, pages 267–276, 1996.
9. V. Cardenas, C. Studholme, S. Gazdzinski, T. Durazzo, and D. Meyerhoff. Deformation-based morphometry of brain changes in alcohol dependence and abstinence. *NeuroImage*, 34(3): 879–887, 2006.
10. R. Chandrashekar, R. Mohiaddin, R. Razavi, and D. Rueckert. Nonrigid image registration with subdivision lattices: Application to cardiac MR image analysis. In *Tenth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '07)*, pages 335–342, 2007.
11. R. Chandrashekar, R. Mohiaddin, and D. Rueckert. Analysis of myocardial motion and strain patterns using a cylindrical B-spline transformation model. In *Surgery Simulation and Soft Tissue Modelling (IS4TM 03)*, pages 88–99, 2003.
12. R. Chandrashekar, R. H. Mohiaddin, and D. Rueckert. Analysis of myocardial motion in tagged MR images using non-rigid image registration. In *Proc. SPIE Medical Imaging 2002: Image Processing*, pages 1168–1179, San Diego, CA, Feb. 2002.

13. Y. Choi and S. Lee. Injectivity conditions of 2D and 3D uniform cubic B-spline functions. *Graphical Models*, 62(6):411–427, 2000.
14. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996.
15. G. E. Christensen, R. D. Rabbitt, M. I. Miller, S. C. Joshi, U. Grenander, T. A. Coogan, and D. C. van Essen. Topological properties of smooth anatomic maps. In *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, pages 101–112, 1995.
16. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Seutens, and G. Marchal. Automated multimodality image registration using information theory. In *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, pages 263–274, 1995.
17. T. F. Cootes, C. Beeston, G. J. Edwards, and C. J. Taylor. A unified framework for atlas matching using active appearance models. In *Information Processing in Medical Imaging: Proc. 16th International Conference (IPMI'99)*, pages 322–333, 1999.
18. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active Shape Models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
19. S. Coquillart. Extended free-form deformation: A sculpturing tool for 3D geometric modelling. *Computer Graphics*, 24(4):187–196, 1986.
20. C. Davatzikos. Spatial transformation and registration of brain images using elastically deformable models. *Computer Vision and Image Understanding*, 66(2):207–222, 1997.
21. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. 3D statistical shape models using direct optimization of description length. In *Proc. 7th European Conference on Computer Vision (ECCV'02)*, pages 3–20, 2002.
22. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5):525–537, 2002.
23. M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms. A physics-based coordinate transformation for 3-D image matching. *IEEE Transactions on Medical Imaging*, 16(3):317–328, 1997.
24. E. R. E. Denton, L. I. Sonoda, D. Rueckert, S. C. Rankin, C. Hayes, M. Leach, D. L. G. Hill, and D. J. Hawkes. Comparison and evaluation of rigid and non-rigid registration of breast MR images. *Journal of Computer Assisted Tomography*, 23:800–805, 1999.
25. P. J. Edwards, D. L. G. Hill, J. A. Little, and D. J. Hawkes. A three-component deformation model for image-guided surgery. *Medical Image Analysis*, 2(4):355–367, 1998.
26. D. R. Forsey and R. H. Bartels. Hierarchical B-spline refinement. *ACM Transactions on Computer Graphics*, 22(4):205–212, 1988.
27. A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen. Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration. In *Information Processing in Medical Imaging: Proc. 17th International Conference (IPMI'01)*, Lecture Notes in Computer Science, pages 78–91, Davis, CA, July 2001. Springer-Verlag.
28. A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. *IEEE Transactions on Medical Imaging*, 21(9):1151–1166, 2002.
29. J. C. Gee. On matching brain volumes. *Pattern Recognition*, 32(1):99–111, 1999.
30. M. Hagenlocker and K. Fujimura. CFFD: a tool for designing flexible shapes. *The Visual Computer*, 14(5/6):271–287, 1998.
31. J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, editors. *Medical Image Registration*. CRC Press, 2001.
32. R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain mri segmentation combining label propagation and decision fusion. *Neuroimage*, 33(1):115–126, 2006.
33. P. Hellier, C. Barillot, É. Mémin, and P. Perex. Hierarchical estimation of a dense deformation field for 3D robust registration. *IEEE Transactions on Medical Imaging*, 20(5):388–402, 2001.
34. A. Kelemen, G. Székely, and G. Gerig. Elastic model-based segmentation of 3-D neurological data sets. *IEEE Transactions on Medical Imaging*, 18(10):828–839, 1999.

35. S. Klein, M. Staring, and J. Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *IEEE Transactions on Image Processing*, 16(12):2879–2890, December 2007.
36. J. Kybic and M. Unser. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11):1427–1442, 2003.
37. S. Lee, G. Wolberg, K.-Y. Chwa, and S. Y. Shin. Image metamorphosis with scattered feature constraints. *IEEE Transactions on Visualization and Computer Graphics*, 2(4):337–354, 1996.
38. S. Lee, G. Wolberg, and S. Y. Shin. Scattered data interpolation with multilevel B-splines. *IEEE Transactions on Visualization and Computer Graphics*, 3(3):228–244, 1997.
39. M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR'00)*, pages 316–323, 2000.
40. N. Lin and J. S. Duncan. Generalized robust point matching using an extended free-form deformation model: Application to cardiac images. In *IEEE International Symposium on Biomedical Imaging*, 2004.
41. D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens. Nonrigid image registration using free-form deformations with a local rigidity constraint. In *Seventh Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '04)*, pages 639–646, 2004.
42. R. MacCracken and K. I. Joy. Free-form deformations with lattices of arbitrary topology. In *SIGGRAPH*, pages 181–188, 1996.
43. F. Maes, A. Collignon, D. Vandermeulen, G. Marechal, and R. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
44. J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
45. D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003.
46. K. McLeish, D. L. G. Hill, D. Atkinson, J. M. Blackall, and R. Razavi. A study of the motion and deformation of the heart due to respiration. *IEEE Transactions on Medical Imaging*, 21(9):1142–1150, 2002.
47. H. Ólafsdóttir, M. B. Stegmann, B. K. Ersbøll, and H. B. Larsson. A comparison of FFD-based nonrigid registration and AAMs applied to myocardial perfusion MRI. In *International Symposium on Medical Imaging 2006, San Diego, CA*, volume 6144, 2006.
48. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.
49. T. Rohlfing, R. Brandt, R. Menzel, and C. M. Jr. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4):1428–1442, 2004.
50. T. Rohlfing and J. C. R. Maurer. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. *IEEE Transactions on Information Technology in Biomedicine*, 7(1):16–25, 2003.
51. T. Rohlfing, J. C. R. Maurer, D. A. Bluemke, and M. A. Jacobs. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741, 2003.
52. T. Rohlfing, E. Sullivan, and A. Pfefferbaum. Deformation-based brain morphometry to track the course of alcoholism: Differences between intra-subject and inter-subject analysis. *Psychiatry Research: Neuroimaging*, 146(2):157–170, 2006.
53. D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. J. Hawkes. Non-rigid registration of breast MR images using mutual information. In *First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '98)*, Lecture Notes in Computer Science, pages 1144–1152, Cambridge, MA, 1998. Springer-Verlag.

54. D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
55. J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. C. Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. G. Hill, and D. J. Hawkes. A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. In *Fourth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '01)*, Lecture Notes in Computer Science, pages 573–581, Utrecht, NL, Oct. 2001. Springer-Verlag.
56. J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. G. Hill, and D. J. Hawkes. Validation of non-rigid image registration using finite element methods: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 22(2):238–247, 2003.
57. T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. *SIGGRAPH*, 20(4):151–160, 1986.
58. D. Shen and C. Davatzikos. Hammer: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, 21(11):1421–1439, 2002.
59. C. Studholme, V. Cardenas, R. Blumenfeld, N. Schuff, H. Rosen, B. Miller, and M. Weiner. A deformation tensor morphometry study of semantic dementia with quantitative validation. *NeuroImage*, 21(4):1387–1398, 2004.
60. C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86, 1998.
61. J.-P. Thirion. Image matching as a diffusion process: An analogy with Maxwell’s demons. *Medical Image Analysis*, 2(3):243–260, 1998.
62. P. Viola. *Alignment By Maximization of Mutual Information*. PhD thesis, Massachusetts Institute of Technology. A.I. Technical Report No. 1548, 1995.
63. G. Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, 1990.
64. J. Wang and T. Jiang. Nonrigid registration of brain MRI using NURBS. *Pattern Recognition Letters*, 28(2):214–223, 2007.
65. E. A. Zerhouni, D. M. Parish, W. J. Rogers, A. Yang, and E. P. Shapiro. Human heart: Tagging with MR imaging – a method for non-invasive assessment of myocardial motion. *Radiology*, 169:59–63, 1988.
66. B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

Image registration using mutual information

F. Maes, D. Loeckx, D. Vandermeulen, and P. Suetens

Abstract Different imaging modalities, such as CT, MRI and PET, are based on different physical principles and capture different and often complementary information. Many applications in clinical practice benefit from an integrated visualization and combined analysis of such multimodal images. In many applications it is also necessary to compare images acquired at a different time points, such as in the analysis of dynamic image sequences or of follow-up studies. Analysis of a single scene from multiple images assumes that the geometrical correspondence or registration between these images is known, such that anatomically identical points can be precisely identified and compared in each of the images. But reliable automated retrospective fusion or registration of multimodality images based on intrinsic image features is complicated by their different photometric properties, by the complexity of the scene and by the large variety of clinical applications. Maximization of mutual information of corresponding voxel intensities allows for fully automated registration of multimodality images without need for segmentation or user intervention, which makes it well suited for routine clinical use in a variety of applications.

1 Introduction

A fundamental problem in medical image analysis is the integration of information from multiple images of the same subject, acquired using the same or different imaging modalities and possibly at different time points. One essential aspect thereof is image registration, i.e. establishing the geometric relationship between corresponding points in multiple images of the same scene. While various more or

F. Maes (✉) • D. Loeckx • D. Vandermeulen • P. Suetens
ESAT/PSI-Medical Image Computing & iMinds, KU Leuven, Kasteelpark
Arenberg 10 - bus 2441, Leuven B-3001, Belgium
e-mail: frederik.maes@esat.kuleuven.be

less automated approaches for image registration have been proposed, one strategy in particular, namely maximization of mutual information (MMI) of corresponding voxel intensities, has been very successful in the field of medical image analysis. Mutual information (MI) is a basic concept from information theory, that is applied in the context of image registration to measure the amount of information that one image contains about the other. The MMI registration criterion postulates that MI is maximal when the images are correctly aligned. MMI has been demonstrated to be a very general and powerful criterion, that can be applied automatically and very reliably, without prior segmentation or pre-processing, on a large variety of applications. This makes the method highly suited for routine use in clinical practice. In this text, which is a largely based on [21], we focus on the application of MMI for global affine registration of three-dimensional (3-D) medical image volumes. MMI has also been applied for other registration problems, such as non-rigid image matching, 2D/3D image registration and registration of models to images, or registration in other contexts, such as microscopy, histology, remote sensing or computer vision.

2 Voxel-based multimodality image registration

Images acquired by different scanners or at different time points are usually acquired independently. Unless specific provisions were made prior to acquisition (e.g. the use of a stereotactic reference frame), their relative position is generally unknown and a retrospective registration procedure is required that recovers the registration transformation from the image content itself. Even when the images are acquired in the same session without the patient leaving the scanner and can be assumed to be registered by acquisition, explicit image registration may be required to correct for inter-scan patient or organ motion. Image registration may be performed manually by letting a human expert interactively displace one image relative to another, based on geometric clues provided by anatomical landmarks visible in each of the images and assisted by visualization software that supplies visual feedback of image alignment. While such subjective approach may be sufficient to support clinical decisions in some applications, a more formal and objective registration measure is required to provide a reliable registration solution in case registration clues are uncertain or inconsistent. Moreover, for the registration tool to be useful and successful in clinical practice, automating the registration process as much as possible is important to minimize the effort and time required by the user.

In many applications, local non-rigid tissue deformations are negligible or irrelevant and the geometric relationship between the images to be registered can be modeled by a rigid or affine linear transformation, composed of a 3-D translation and rotation and possibly also 3-D scaling and skew. The registration problem then consists of determining the 6 or 12 parameters of the rigid or affine geometrical transformation that correctly aligns both images. In other applications it may be needed to correct for local non-rigid image distortions, for instance to

compensate for breathing induced deformations in the thorax, or to quantify local morphological differences from similar images of different subjects, for instance for building geometrical models of the brain. In these cases a more general non-rigid deformation model, with typically much more degrees of freedom, needs to be used.

Registration procedures can be classified according to the image features and the correspondence criterion used to compute the registration [22]. Intrinsic registration methods can be classified as either point based, surface based or voxel based. *Point based* and *surface based* methods compute the registration solution by minimizing the distances between corresponding anatomical landmarks or surfaces respectively extracted from each of the images, while *voxel based* registration methods optimize a functional measuring the similarity of corresponding voxel pairs for some feature. The main advantage of voxel based methods is that feature calculation is straightforward or even absent when only grey-values are used, such that accuracy is not dependent on accurate segmentation as in point or surface based methods, which is difficult to automate.

While for intra-modality registration it can often be assumed that the image intensities in both images are identical or linearly related, this assumption is not valid for inter-modality registration and measures such as sum of squared differences or cross-correlation are therefore in general not applicable. Woods *et al.* [47] and Hill *et al.* [8, 9] observed that misregistration induces dispersion in the joint histogram of corresponding voxel intensities and proposed histogram 'peakedness' as a measure of registration. Collignon *et al.* [3] recognized the information theoretic nature of the image registration problem and proposed the much more general notion of joint image intensity entropy as a new matching criterion. However, this measure is sensitive to partial overlap of the images, as it does not account for the fact that the entropy or information complexity of each of the images separately within their region of overlap may vary during registration. Finally, two different groups, Collignon and Maes *et al.* [2, 19] at KU Leuven, Belgium, and Viola and Wells *et al.* [42, 43] at MIT, USA, almost simultaneously but independently of each other, introduced maximization of mutual information (MMI) of image intensities as a new registration criterion. Mutual information (MI), or relative entropy, is a basic concept from information theory, which can be considered a non-linear generalization of cross-correlation. MI measures the statistical dependence between two random variables or the amount of information that one variable contains about the other [4]. The MMI registration criterion postulates that the MI of the image intensity values of corresponding voxel pairs is maximal if the images are geometrically aligned. Related early work in this area includes the work by Studholme *et al.* [35], who demonstrated the superior performance of MMI over other voxel-based registration measures, by Pluim [25] and by Meyer *et al.* [13], who was the first to apply MMI for non-rigid matching.

3 The mutual information registration criterion

Multimodal medical images represent measurements of some physical property of the anatomical objects in the imaged scene. Hence, although the image intensity values at corresponding points in different images of the same scene may be very different between different modalities, in general they are not independent observations but statistically related measurements of the same underlying object or tissue. This is intuitively clear for modalities that capture information which is linked directly to anatomy, such as X-ray attenuation in CT or water content in MR, but also holds for modalities that represent different kinds of information, such as anatomy in MR and function in PET. Let \mathcal{A} and \mathcal{B} be two images that are geometrically related by the registration transformation $\mathbf{T}\boldsymbol{\alpha}$ with parameters $\boldsymbol{\alpha}$ such that voxels \mathbf{p} in \mathcal{A} with intensity a correspond to voxels $\mathbf{q} = \mathbf{T}\boldsymbol{\alpha}(\mathbf{p})$ in \mathcal{B} with intensity b . Taking random samples \mathbf{p} in \mathcal{A} , a and b can be considered as discrete random variables A and B with joint and marginal distributions $p_{AB}(a, b)$, $p_A(a)$ and $p_B(b)$ respectively. The mutual information $I(A, B)$ of A and B measures the degree of dependence between A and B as the distance between the joint distribution $p_{AB}(a, b)$ and the distribution associated to the case of complete independence $p_A(a) \cdot p_B(b)$, by means of the Kullback-Leibler measure [4], i.e.

$$I(A, B) = \sum_{a,b} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_A(a) \cdot p_B(b)} \quad (1)$$

The relationship $p_{AB}(a, b)$ between a and b and hence their mutual information $I(A, B)$ depends on $\mathbf{T}\boldsymbol{\alpha}$, i.e. on the registration of the images. The mutual information registration criterion postulates that the images are geometrically aligned by the transformation $\mathbf{T}\boldsymbol{\alpha}^*$ for which $I(A, B)$ is maximal:

$$\boldsymbol{\alpha}^* = \arg \max_{\boldsymbol{\alpha}} I(A, B)$$

Mutual information is related to the information theoretic notion of entropy by the equations $I(A, B) = H(A) + H(B) - H(A, B) = H(A) - H(A|B) = H(B) - H(B|A)$ with $H(A)$ and $H(B)$ being the entropy of A and B respectively, $H(A, B)$ their joint entropy and $H(A|B)$ and $H(B|A)$ the conditional entropy of A given B and of B given A respectively. The entropy $H(A)$ is a measure of the amount of uncertainty about the random variable A , while $H(A|B)$ is the amount of uncertainty left in A when knowing B . Hence, $I(A, B)$ measures the amount of information that one image contains about the other, which should be maximal at registration. If both marginal distributions $p_A(a)$ and $p_B(b)$ can be considered to be independent of the registration parameters $\boldsymbol{\alpha}$, the MI criterion reduces to minimizing the joint entropy $H_{AB}(A, B)$. If either $p_A(a)$ or $p_B(b)$ is independent of $\boldsymbol{\alpha}$, which is the case if one of the images is always completely contained in the

other, the MI criterion reduces to minimizing the conditional entropy $H(A|B)$ or $H(B|A)$. However, if both images only partially overlap, which is very likely during optimization, the volume of overlap will change when α is varied and both marginal distributions $p_A(a)$ and $p_B(b)$ and therefore also their entropies $H(A)$ and $H(B)$ will in general depend on α . Hence, maximizing mutual information will tend to find as much as possible of the complexity that is in the separate data sets so that at the same time they explain each other well.

Other information-theoretic registration measures can be derived from the MI criterion presented above, such as the entropy correlation coefficient $ECC(A, B) = 2 \cdot \frac{I(A, B)}{H(A) + H(B)}$ [19] or the normalized mutual information $NMI = \frac{H(A) + H(B)}{H(A, B)}$ [38] with $0 \leq ECC \leq 1$, $1 \leq NMI \leq 2$ and $ECC = 2 \cdot (1 - 1/NMI)$. Maximization of NMI or ECC may be superior to MMI itself in case the region of overlap of both images is relatively small at the correct registration solution, as MMI may be biased towards registration solutions with larger total amount of information $H(A) + H(B)$ within the region of overlap [38]. MI is only one example of the more general f – *information* measures of dependence, several of which were investigated for image registration in [31].

4 Implementation

The MMI registration criterion does not require any preprocessing or segmentation of the images. With each of the images is associated a 3-D coordinate frame in millimeter units, that takes the pixel size, inter-slice distance and the orientation of the image axes relative to the patient into account. One of the images to be registered is selected to be the *floating* image \mathcal{F} from which samples $s \in S$ are taken and transformed by the geometric transformation $\mathbf{T}\alpha$ with parameters α into the *reference* image \mathcal{R} . S may include all voxels in \mathcal{F} or a subset thereof to increase speed performance.

The MMI method requires the estimation of the joint probability density $p(f, r)$ of corresponding voxel intensities f and r in the floating and reference image respectively. This can be obtained from the joint intensity histogram of the region of overlap of the images. The joint image intensity histogram $h_\alpha(f, r)$ of the volume of overlap $s \in S_\alpha \subset S$ of \mathcal{F} and \mathcal{R} can be constructed by simple binning of the image intensity pairs $(f(s), r(\mathbf{T}\alpha(s)))$ for all $s \in S_\alpha$. In order to do this efficiently, the floating and the reference image intensities are first linearly rescaled to the range $[0, n_{\mathcal{F}} - 1]$ and $[0, n_{\mathcal{R}} - 1]$ respectively, with $n_{\mathcal{F}}$ and $n_{\mathcal{R}}$ the number of bins assigned to the floating and reference image respectively and $n_{\mathcal{F}} \times n_{\mathcal{R}}$ being the total number

of bins in the joint histogram. Estimations for the marginal and joint image intensity distributions $p_{\mathcal{F}}\alpha(f)$, $p_{\mathcal{R}}\alpha(r)$ and $p_{\mathcal{F}\mathcal{R}}\alpha(f, r)$ are obtained by normalization of $h_{\alpha}(f, r)$

$$p_{\mathcal{F}\mathcal{R}}\alpha(f, r) = \frac{h_{\alpha}(f, r)}{\sum_{f,r} h_{\alpha}(f, r)} \quad (2)$$

$$p_{\mathcal{F}}\alpha(f) = \sum_r p_{\mathcal{F}\mathcal{R}}\alpha(f, r) \quad (3)$$

$$p_{\mathcal{R}}\alpha(r) = \sum_f p_{\mathcal{F}\mathcal{R}}\alpha(f, r) \quad (4)$$

and the MI registration criterion is evaluated using

$$I(\alpha) = \sum_{f,r} p_{\mathcal{F}\mathcal{R}}\alpha(f, r) \log_2 \frac{p_{\mathcal{F}\mathcal{R}}\alpha(f, r)}{p_{\mathcal{F}}\alpha(f) p_{\mathcal{R}}\alpha(r)}. \quad (5)$$

Typically, $n_{\mathcal{F}}$ and $n_{\mathcal{R}}$ need to be chosen much smaller than the number of different values in the original images in order to assure a sufficient number of counts in each bin. If not, the joint histogram $h_{\mathcal{F}\mathcal{R}}\alpha$ would be rather sparse with many zero entries and entries that contain only one or a few counts, such that a small change in the registration parameters α would lead to many discontinuous changes in the joint histogram, with non-zero entries becoming zero and vice versa, that propagate into $p_{\mathcal{F}\mathcal{R}}\alpha$. Such abrupt changes in $p_{\mathcal{F}\mathcal{R}}\alpha$ induce discontinuities and many local maxima in $I(\alpha)$, which deteriorates optimization robustness of the MI measure. Appropriate values for $n_{\mathcal{F}}$ and $n_{\mathcal{R}}$ can only be determined by experimentation. Moreover, $\mathbf{T}_{\alpha}(s)$ will in general not coincide with a grid point of \mathcal{R} , such that interpolation of the reference image is needed to obtain the image intensity value $r(\mathbf{T}_{\alpha}(s))$. Zeroth order or nearest neighbor interpolation of \mathcal{R} is most efficient, but is insensitive to translations up to 1 voxel and therefore insufficient to guarantee subvoxel accuracy. But even when higher order interpolation methods are used, such as linear, cubic or B-spline interpolation, simple binning of the interpolated intensity pairs $(f(s), r(T_{\alpha}s))$ leads to discontinuous changes in the joint intensity probability $p_{\mathcal{F}\mathcal{R}}\alpha(f, r)$ and in the marginal probability $p_{\mathcal{R}}\alpha(r)$ for small variations of α when the interpolated values $r(T_{\alpha}s)$ fall in a different bin. Note that post-processing the histogram obtained by binning by convolution with a Gaussian or other smoothing kernel is not sufficient to eliminate these discontinuities.

To avoid the problem of discontinuities induced by intensity binning, two different solutions have been proposed. The first one involves the use of the Parzen windowing technique (PW) to estimate the joint probability density $p(f, r)$ as a sum of continuous and differentiable kernel functions $k(f - f_i, r - r_i)$ centered around each interpolated voxel sample pair (f_i, r_i) and satisfying the partitioning of unity constraint $\sum_{f,r} k(f - f_i, r - r_i) = 1, \forall f_i, r_i$. The function k distributes the contribution of each sample i over multiple adjacent histogram bins (f, r) , hence smoothing the histogram and making it a continuous function of the registration

parameters α . Different kernel functions can be used for this purpose, such as linear [18], Gaussian [42] or B-spline [39] functions. An alternative approach to update the joint histogram in a continuous way for each voxel pair $(s, T_\alpha s)$ was proposed in [1, 2, 19]. Instead of interpolating new intensity values in \mathcal{R} , This method, termed partial volume distribution interpolation (PV), distributes the contribution to the joint histogram of the sample s with intensity $f(s)$ in \mathcal{F} over the intensity values r of the nearest neighbors of $T_\alpha s$ on the 3-D grid of \mathcal{R} , using the same weights as for trilinear [19] or higher order [1] interpolation. Each entry in the joint histogram is then the sum of smoothly varying fractions of 1, such that the histogram changes smoothly as α is varied. PV interpolation results in a continuous and a.e. differentiable registration criterion with typically a large basin of attraction around the correct optimum [19]. However, in case the images to be registered have identical voxel grids or if their voxel sizes are multiples of each other in one or more dimensions, PV interpolation may introduce local optima in the MI measure at grid aligning registration positions [18, 28, 40]. These may deteriorate registration accuracy in case the true registration solution differs little from a grid-aligning position, such as when using image registration to recover subvoxel small displacements for motion correction in dynamic image sequences.

The optimal registration parameters α^* are found by maximization of $I(\alpha)$ for which different local optimization schemes can be applied, including heuristic search [36], Powell's method [19], simplex search [20, 24], gradient descent [39] or other gradient based optimization methods [20], as well as global optimization methods [11]. The gradient of MI w.r.t. the registration parameters can be evaluated numerically using a finite difference scheme as in [33], while analytical expressions for the gradient of MI w.r.t. the registration parameters have been derived for PW interpolation in [23, 39] and for PV interpolation in [20]. For high resolution images subsampling of the floating image can be applied without deteriorating optimization robustness of the MMI registration criterion [20, 29]. Important speed-ups can thus be realized by using a multiresolution optimization strategy, starting with a coarsely sampled image for efficiency and increasing the resolution as the optimization proceeds for accuracy [36, 39]. Multiple multiresolution strategies, involving different subsampling factors and number of resolution levels and using various optimization methods, were compared in [20] for affine and in [15] for non-rigid registration. A stochastic iterative gradient-based optimization method was used in [14], estimating the gradient of MI at each iteration from only a small subset of voxel pairs randomly sampled from the images.

5 Evaluation and applications

The robustness of the MMI registration algorithm was established in various studies with respect to implementation issues such as sampling [19, 29, 36], interpolation [19, 26, 28, 39], initial positioning of the images [35, 36] and optimization

strategy [20, 39], as well as with respect to partial overlap of the images [36] and image degradations such as noise, intensity inhomogeneity or geometric distortion [19].

The accuracy of the MMI registration algorithm has been validated for registration of CT, MR and PET brain images within the framework of the Retrospective Registration Evaluation Project (RREP) conducted by Fitzpatrick *et al.* at Vanderbilt University [12] and reported in West *et al.* [44, 45, 46], using as gold standard a prospective, marker-based registration method. To ensure blindness of the study, the frame and the fiducial markers were removed from the images by manual editing prior to retrospective image registration. The transformation differences between the reference and the submitted transformations were evaluated at different sites within the brain and the median and maximal error over all sites and all patients were recorded for each method that participated in the study. The RREP evaluation showed that the MMI approach achieves subvoxel registration accuracy for both CT/MR as well as PET/MR registration and performs better than the other methods in the study, although the number of registration experiments performed was too small to draw statistically significant conclusions. This was confirmed in other studies, such as [34].

Because of its reliability and generality and because of its full automation, image registration by MMI has large potential for routine use in clinical practice in a variety of applications, involving various organs and imaging modalities (see Fig. 1). For an extensive survey of MMI registration applications we refer to [30].

6 Discussion

Mutual information does not rely on the intensity values directly to measure correspondence between different images, but on their relative occurrence in each of the images separately and co-occurrence in both images combined. As such it is insensitive to intensity permutations or one-to-one intensity transformations and is capable of handling positive and negative intensity correlations simultaneously. Unlike other voxel-based registration criteria, the MI criterion does not make limiting assumptions about the nature of the relationship between the image intensities of corresponding voxels in the different modalities, which is highly data dependent, and does not impose constraints on the image content of the modalities involved. This explains the success of MMI for multimodal image registration in a wide range of applications involving various modality combinations, while also for unimodal registration applications MMI is often preferred [10].

Nevertheless, there are cases in which MMI fails as a registration criterion. Such failures occur due to insufficient mutual information in the images, ambiguity about the intensity relationship between both images if this is not spatially invariant, or inability to reliably estimate MI if the number of image samples is small. The fundamental assumption of MMI based registration that image intensities in both images are related to corresponding objects that should be aligned by registration,

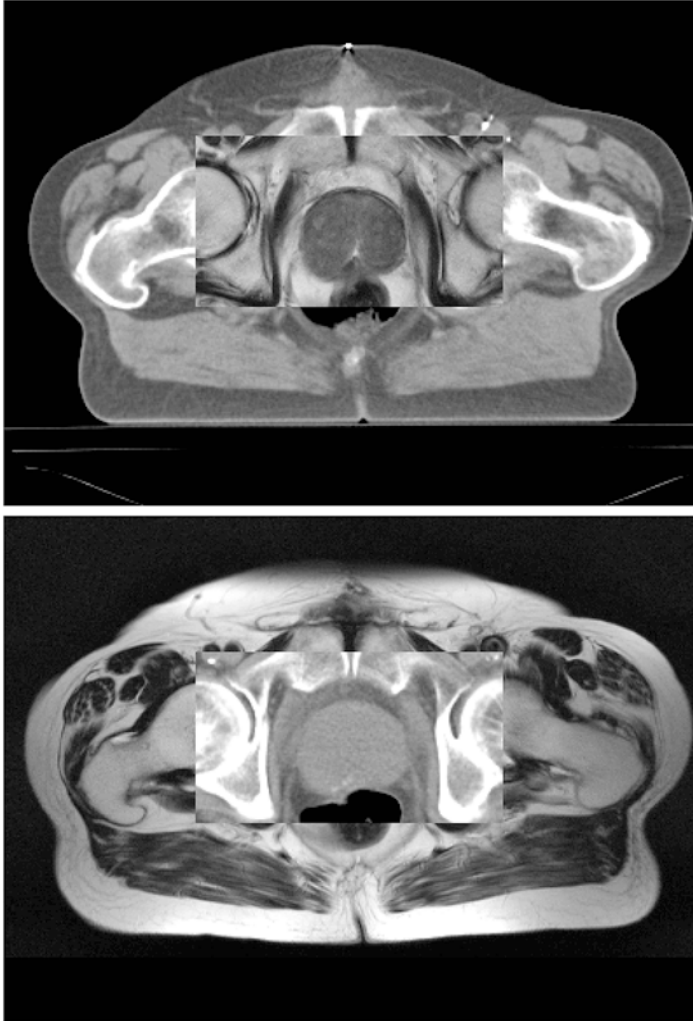


Fig. 1 Global affine registration using MMI of CT (top) and MR images (bottom) of the prostate used for radiotherapy planning. The central part of either image has been shown in overlay over the other for visual inspection. The CT image is needed for estimating the dose distribution, while the target volume and organs at risk can be more accurately delineated in the corresponding MR image. Registration of both images allows to transfer the MR contours into the CT volume, such that the complementary information of both scans can be combined during planning. Despite the presence of large local non-rigid deformations of soft tissues such as skin, fat and muscles, the MMI criterion succeeds at aligning corresponding rigid bony structures in the two images. The registration can be locally refined if needed by defining a suitable region of interest around the prostate

may be invalid if the information in both images is very different, such as anatomical information from CT with functional information from PET in non-brain regions such as the thorax [41]. Because MI is computed from the joint intensity probability of both images that is estimated by pooling contributions from everywhere in the image domain, the MMI criterion implicitly assumes that the statistical relationship between corresponding voxel intensities is identical over the whole area of overlap. However, the photometric relationship between two multimodal images of the same scene may not be spatially invariant, for instance if one of the images suffers from severe intensity inhomogeneity. Also, the MMI registration criterion assumes that the joint probability distribution of corresponding voxel intensities can be estimated reliably around the registration solution. In practice, this requires the volume of overlap at registration to contain a sufficiently large number of voxels. For low resolution images or if the region of overlap is small, the statistical relationship between both images needs to be derived from a small number of samples, which is not robust. In these cases, the computed MI may show multiple local optima around the correct registration solution or the registered position may not coincide with a local maximum of MI [29, 38].

Several adaptations of the MI measure have been proposed in order to increase registration robustness in cases where the joint intensity histogram by itself is insufficient, such as including gradient information in the registration measure [27], the use of higher-order mutual information to explicitly take the dependence of neighboring voxel intensities into account [32], or incorporating additional information channels with region labeling information in the mutual information criterion [6, 37].

While 3D affine image registration using MMI is well established and already used in routine clinical practice, extension of the MMI criterion to non-rigid image registration is still subject of research. Non-rigid registration involves finding a 3-D deformation field that maps each point in one image onto the corresponding point in the other image, displacing each voxel individually to correct for local distortions between both images up to voxel scale. Regularization of the deformation field is required to constrain the registration solution space to include only deformation fields that are physically acceptable and to smoothly extrapolate the registration solution from sites with salient registration features (e.g. object boundaries) towards regions where registration clues are absent or ambiguous (e.g. regions with homogeneous intensity). Various approaches for multimodal non-rigid image registration using MMI have been proposed that differ in their regularization of the deformation field, e.g. using thin plate splines [24] or B-splines [14, 17, 33], elastic [7] or viscous fluid [5] deformation models. Another distinction can be made regarding the way the variation of MI with changes in the deformation parameters is computed, namely globally over the entire image domain [14, 33] or locally within subregions only [16, 17].

7 Conclusion

Image registration by maximization of mutual information considers all voxels in the images to be registered to estimate the statistical dependence between corresponding voxel intensities, which is assumed to be maximal when the images are correctly aligned. The MMI criterion is histogram based rather than intensity based and does not impose limiting assumptions on the specific nature of the relationship between corresponding voxel intensities. Since its introduction in the field in 1995 by Collignon *et al.* [2] and by Viola and Wells [42], MMI has become the standard in the field and the method of choice for multimodal image registration in a wide range of applications. The success of MMI for multimodal image registration can be explained by the fact that it got rid of the need for image segmentation or preprocessing as required with previous registration algorithms and that it allows for completely automated registration without need for user interaction, making the method very well suited for application in clinical practice. Several approaches have been presented for extension of the MMI criterion to non-rigid image matching in the context of image rectification, shape normalization, motion estimation or tissue deformation correction, which is still an active area of research.

References

1. H. Chen and P. Varshney. Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation. *IEEE Transactions on Medical Imaging*, 22(9):1111–1119, 2003.
2. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multimodality medical image registration using information theory. In Y. Bizais, C. Barillot, and R. D. Paola, editors, *Proceedings of the XIV'th Int'l Conf. Information Processing in Medical Imaging (IPMI'95)*, volume 3 of *Computational Imaging and Vision*, pages 263–274, Ile de Berder, France, June 1995. Kluwer Academic Publishers.
3. A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3-D multi-modality medical image registration using feature space clustering. In N. Ayache, editor, *Proc. First Int'l Conf. Computer Vision, Virtual Reality and Robotics in Medicine (CVR'95)*, volume 905 of *Lecture Notes in Computer Science*, pages 195–204, Nice, France, April 1995. Springer.
4. T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, N.Y., USA, 1991.
5. E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens. A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*, 7(4):565–575, 2003.
6. E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens. An information theoretic approach for non-rigid image registration using voxel class probabilities. *Medical Image Analysis*, 10(3):413–431, 2006.
7. G. Hermosillo, C. Chef d'Hotel, and O. Faugeras. Variational methods for multimodal image matching. *International Journal of Computer Vision*, 50(3):329–343, 2002.

8. D. Hill, D. Hawkes, N. Harrison, and C. Ruff. A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics. In H. Barrett and A. Gmitro, editors, *Proc. XIIIth Int'l Conf. Information Processing in Medical Imaging (IPMI'93)*, volume 687 of *Lecture Notes in Computer Science*, pages 182–196, Flagstaff, Arizona, USA, June 1993. Springer-Verlag.
9. D. Hill, C. Studholme, and D. Hawkes. Voxel similarity measures for automated image registration. In *Visualization in Biomedical Computing (VBC'94)*, volume 2359 of *Proc. SPIE*, pages 205–216, 1994.
10. M. Holden, D. Hill, E. Denton, J. Jarosz, T. Cox, T. Rohlfing, J. Goodey, and D. Hawkes. Voxel similarity measures for 3-D serial MR brain image registration. *IEEE Transactions on Medical Imaging*, 19(7):94–102, 2000.
11. M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5:143–156, 2001.
12. J.M. Fitzpatrick, Principal Investigator. Retrospective Image Registration Evaluation, National Institutes of Health, Project Number 1 R01 CA89323, Vanderbilt University, Nashville, TN, 1994. See: <http://www.vuse.vanderbilt.edu/~image/registration/>.
13. B. Kim, J. Boes, and C. Meyer. Mutual information for automated multimodal image warping. *NeuroImage*, 3(3):158, June 1996. Second International Conference on Functional Mapping of the Human Brain.
14. S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. elastix: a toolbox for intensity based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.
15. S. Klein, M. Staring, and J. Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. *IEEE Transactions on Image Processing*, 16(12):2879–2890, 2007.
16. B. Likar and F. Pernuš. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19(1–2):33–44, 2001.
17. D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens. Nonrigid image registration using conditional mutual information. *IEEE Transactions on Medical Imaging*, 29(1):19–29, 2010.
18. F. Maes. *Segmentation and Registration of Multimodal Medical Images: from Theory, Implementation and Validation to a Useful Tool in Clinical Practice*. PhD thesis, KU Leuven, Dept. Electrical Engineering (ESAT/PSI), Leuven, Belgium, May 1998.
19. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multi-modality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
20. F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.
21. F. Maes, D. Vandermeulen, and P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10):1699–1722, 2003.
22. J. Maintz. *Retrospective Registration of Tomographic Brain Images*. PhD thesis, Universiteit Utrecht, Utrecht, The Netherlands, 1996.
23. D. Mattes, D. Haynor, H. Vesselle, T. Lewellen, and W. Eubank. Pet-ct image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003.
24. C. Meyer, J. Boes, B. Kim, P. Bland, R. Wahl, K. Zasadny, P. Kison, K. Koral, and K. Frey. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations. *Medical Image Analysis*, 1(3):195–206, 1997.
25. J. Pluim. Multi-modality matching using mutual information. Master's thesis, University of Groningen, Department of Computing Science, Groningen, The Netherlands, November 1996.

26. J. Pluim. *Mutual information based registration of medical images*. PhD thesis, Utrecht University, Utrecht, The Netherlands, 2001.
27. J. Pluim, J. Maintz, and M. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, 2000.
28. J. Pluim, J. Maintz, and M. Viergever. Interpolation artefacts in mutual information-based image registration. *Computer Vision and Image Understanding*, 77(2):211–232, 2000.
29. J. Pluim, J. Maintz, and M. Viergever. Mutual information matching in multiresolution contexts. *Image and Vision Computing*, 19(1-2):45–52, 2001.
30. J. Pluim, J. Maintz, and M. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.
31. J. Pluim, J. Maintz, and M. Viergever. f-information measures in medical image registration. *IEEE Transactions on Medical Imaging*, 23(12):1508–1516, 2004.
32. D. Rueckert, M. Clarkson, D. Hill, and D. Hawkes. Non-rigid registration using higher order mutual information. In K.M.Hanson, editor, *Medical Imaging: Image Processing*, volume 3979 of *Proc. SPIE*, pages 438–447, San Diego, CA, USA, February 2000. SPIE Press, Bellingham, WA.
33. D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
34. D. Skerl, B. Likar, and J. Fitzpatrick. Comparative evaluation of similarity measures for the rigid registration of multi-modal head images. *Physics in Medicine and Biology*, 52(18):5587–5601, 2007.
35. C. Studholme, D. Hill, and D. Hawkes. Multiresolution voxel similarity measures for MR-PET registration. In Y. Bizais, C. Barillot, and R. D. Paola, editors, *Proceedings of the XIVth Int’l Conf. Information Processing in Medical Imaging (IPMI’95)*, volume 3 of *Computational Imaging and Vision*, pages 287–298, Ile de Berder, France, June 1995. Kluwer Academic Publishers.
36. C. Studholme, D. Hill, and D. Hawkes. Automated 3-D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.
37. C. Studholme, D. Hill, and D. Hawkes. Incorporating connected region labelling into automated image registration using mutual information. In *Proc. 2nd IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 23–31, San Francisco, CA, USA, June 1996. IEEE Computer Society Press.
38. C. Studholme, D. Hill, and D. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86, 1999.
39. P. Thévenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.
40. J. Tsao. Interpolation artifacts in multimodality image registration based on maximization of mutual information. *IEEE Transactions on Medical Imaging*, 22(7):854–864, 2003.
41. J. Vansteenkiste, S. Stroobants, P. Dupont, P. De Leyn, W. De Wever, E. Verbeken, J. Nuyts, F. Maes, J. Bogaert, and the Leuven Lung Cancer Group. FDG-PET scan in potentially operable non-small cell lung cancer : do anatometabolic PET-CT fusion images improve the localisation of regional lymph node metastases? *European Journal of Nuclear Medicine*, 25(11):1495–1501, 1998.
42. P. Viola and W. Wells, III. Alignment by maximization of mutual information. In *Proc. of the Fifth International Conference on Computer Vision*, pages 16–23, Cambridge, MA, USA, June 1995.
43. W. Wells, III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.
44. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, Jr., R. Kessler, and R. Maciunas. Retrospective intermodality registration techniques for images of the head: surface-based versus volume-based. *IEEE Transactions on Medical Imaging*, 18(2):144–150, 1999.

45. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, Jr., R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. van den Elsen, P. Hemler, S. Napel, T. Sumanaweera, B. Harkness, D. Hill, C. Studholme, G. Malandain, X. Pennec, M. Noz, G. Maguire, Jr., M. Pollack, C. Pellizari, R. Robb, D. Hanson, and R. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. In *Medical Imaging: Image Processing*, volume 2710 of *Proc. SPIE*, pages 332–347, Newport Beach, California, USA, February 1996.
46. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, Jr., R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, J. Maintz, M. Viergever, G. Malandain, X. Pennec, M. Noz, G. Maguire, Jr., M. Pollack, C. Pellizari, R. Robb, D. Hanson, and R. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.
47. R. Woods, J. Mazziotta, and S. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17(4):536–546, 1993.

Physical Model Based Recovery of Displacement and Deformations from 3D Medical Images

P. Yang, C. Delorenzo, X. Papademetris, and J.S. Duncan

Abstract Estimating tissue displacement and deformation from time-varying medical images is a common problem in biomedical image analysis. For example, in order to better manage patients with ischemic heart disease, it would be useful to know their current extent of injury. This can be assessed by accurately tracking the motion of the left ventricle of the beating heart. Another example of this type of application is estimating the displacement of brain tissue during neurosurgery. The latter application is necessary because the presurgical planning for these delicate surgeries is based on images that may not accurately reflect the intraoperative brain (due to the action of gravity and other forces). In both examples, the tissue deformation cannot be measured directly. Instead, a sparse set of (potentially noisy) displacement estimates are extracted from acquired images. In this chapter, we explain how to use the physical properties of underlying organs or structures to guide such estimations of deformation, using neurosurgery and cardiac motion as example cases.

1 Introduction

Estimating displacements and deformations of real world objects from time-varying medical images is a common problem in biomedical image analysis. In such problems, the physical properties of the underlying organs or structures can be used to guide this information recovery process. For example, knowledge of the fiber architecture of the heart can be used as a constraint in the estimation of

P. Yang (✉) • C. Delorenzo • J.S. Duncan
Center for Understanding Biology using Imaging Technology (CUBIT), Stony Brook University,
Stony Brook, NY 11794, USA
e-mail: Christine.DeLorenzo@stonybrookmedicine.edu

X. Papademetris
Department of Diagnostic Radiology and Biomedical Engineering, Yale University,
300 cedar St, TAC N119, New Haven, CT 06519, USA
e-mail: xenophon.papademetris@yale.edu

N. Paragios et al. (eds.), *Handbook of Biomedical Imaging: Methodologies and Clinical Research*, DOI 10.1007/978-0-387-09749-7_17,
© Springer Science+Business Media New York 2015

cardiac deformation [41]. This type of real displacement estimation stands in contrast to the more common domain of non-rigid image registration methods where the goal is to estimate displacement fields that map, for example, brains from two different subjects to the same coordinate system. Confusingly, in the case of image registration, physical models are also often used to regularize the displacement estimation process (e.g. see [12, 21]) as they often possess desirable mathematical properties. However, in the cases of the problems described in this chapter, we employ physical models precisely to take advantage of research in the biomechanics community (e.g. [25]) about the material properties of the underlying tissues as an aid to the displacement estimation process.

Two of the most common areas of biomedical image analysis in which physical models have been employed are (i) the estimation of brain shift (and generally deformation) during neurosurgery and (ii) the estimation of left ventricular deformation over time. While these are, on the surface, very different problems we note that there exist a number of common threads between them. Fundamentally, in both cases, the goal is to estimate a dense and smooth displacement field over the whole space covered by the structure of interest (e.g. the brain or the left ventricle of the heart). However, usually what is available directly from the images are sparse, noise-corrupted and often partial (i.e. only some of the components are measurable) measurements of the actual displacement, often clustered on the outer surface of the organ.

The goal of much of this work is to use these measured displacements as an input and to leverage physical and geometrical models to essentially approximate/interpolate between these measurements to generate a smooth, dense and complete displacement field. To accomplish this we use *patient specific* geometric models derived from the image data to build either a finite element or a boundary element model of the underlying organ together with *generic* biomechanical models with parameters set based on the biomechanics literature. We emphasize that these models do not aim to predict the deformation in the style of forward models, but rather should be thought of as smart and optimal interpolation/approximation techniques.

We use two example studies to illustrate the use of this underlying mathematical framework. In Sect. 2 we describe methodology to compensate for brain shift in image guided neurosurgery using finite element methodology and in Sect. 3 we describe algorithms to estimate the deformation of the left ventricle of the heart using a boundary element based approach.

2 A Finite Element Based Approach: Brain Shift Compensation for Image Guided Neurosurgery

2.1 Background

Prior to brain surgery, neurosurgeons may acquire images of different modalities to delineate pathologic regions. One of the major challenges during neurosurgery is the localization of these pathologic targets within the brain anatomy. Surgical

navigation systems can aid intraoperative navigation by rigidly registering the patient to preoperative images and then displaying the current position of surgical instruments relative to these images. In these cases, the patient data is usually registered only once [10, 24, 43], assuming that the brain and other intracranial structures are rigid and will stay fixed relative to the skull. However, during surgery, the brain can deform up to five centimeters or more due to gravity, loss of blood and cerebrospinal fluid, swelling, surgical manipulation and the action of certain medications [4, 14, 20, 22, 23, 27, 45]. This nonrigid brain shift is the primary cause of intraoperative localization inaccuracies [16].

Due to the large number of contributing factors, it is nearly impossible to predict the exact pattern of brain deformation preoperatively [19, 44]. Therefore, intraoperative information is often necessary to compute the magnitude of brain shift. This information can be either surface or volume based. Though volumetric images, acquired by intraoperative magnetic resonance imaging (iMRI), computed tomography (iCT) or ultrasound (iUS), provide visualization of deformation throughout the entire brain, these methods are either invasive or costly [27, 39]. Additionally, the necessary image acquisition time can substantially lengthen surgery [44]. Using surface based information, acquired by laser range scanners (LRS) [3, 9, 18, 39, 46] or stereo cameras [47, 51], avoids these problems. However, while sometimes capable of achieving high accuracy, both LRS and stereo can be affected by resolution issues, which can compromise surface deformation estimations.

Since the LRS and stereo methods only acquire intraoperative information at the cortical surface, these methods are typically used in conjunction with a biomechanical model to infer volumetric deformation. Biomechanical models are also often used with volumetric imaging data because preoperative to intraoperative volumetric image matching can be difficult.

2.2 The Finite Element Method (FEM)

The finite element method is a numerical analysis technique for obtaining approximate solutions to a wide variety of engineering problems [29]. The key to this method is that the domain of problem is divided into small areas or volumes called *elements*. The problem is then discretized on an element by element basis and the resulting equations *assembled* to form the global solution.

An Example Problem: In this section, we describe an example problem and outline a possible solution using FEM in an energy minimization framework. The goal is to estimate a displacement field, $u(x, y, z)$, that is the optimal trade off between an internal energy function, $W(u)$,¹ and an approximating noisy displacement field, $u^m(x, y, z)$.

¹Note that although W is defined as function of the strain, e , as e is a function of the displacement, u , W can also be written as a function of the displacement field, u .

We define the optimal solution displacement field, u , as the one that minimizes functional $P(u)$ in a weighted least squares sense:

$$P(u) = \int_{vol} (W(u) + V(u, u^m)) d(vol)$$

$$W(u) = e(u)^t C e(u)$$

$$V(u, u^m) = \alpha (u^m - u)^2$$

$W(u)$ can be defined using a strain energy function as $W = e^t C e$, where e is local tissue strain and C , the material stiffness matrix, is a function of the displacement field (u), spatial position (m), and the tissue's material properties-Young's modulus (E) and Poisson's Ratio (ν). $V(u, u^m)$ is the external energy term, based on u^m , an original displacement estimate, and α , the confidence in the match. We will focus here primarily on the first term, $W(u)$.

Outline of the Solution Procedure:

Step 1: Divide Volume into elements (tetrahedra or hexahedra) to provide the basis functions for the discretization.

Step 2: Discretize the problem by approximating the displacement field in each element as a linear combination of displacements at the nodes of each element. For a hexahedral element, this discretization can be expressed as: $u \approx \sum_{i=1}^8 N_i u_i$, where N_i is the interpolation shape function for node i and u_i is the displacement at node i of the element.

Step 3: Write the internal energy equation as the sum of the internal energy for each element, v_{el} :

$$W(u) = \sum_{all\ elements} \left[\int_{v_{el}} e^t C e d(v_{el}) \right] \quad (1)$$

We can approximate the derivatives of u in an element with respect to components of the global coordinate system, x , as follows (note that the u_i are constant in this expression):

$$\frac{\partial u}{\partial x_k} = \sum_{i=1}^8 \frac{\partial (N_i u_i)}{\partial x_k} = \sum_{i=1}^8 \frac{\partial N_i}{\partial x_k} u_i$$

The derivatives of the displacement field, u , (i.e. $\frac{\partial u}{\partial x_k}$) are linear functions of the nodal displacements, u_i . Since the infinitesimal strain tensor consists of only sums and differences of partial derivatives, this tensor can also be expressed as a linear function of the nodal displacements by $e = B u$. (See Bathe [7] for nonlinear extensions of the finite strain deformation case.) Substituting this into equation (1) yields:

$$W(u) = \sum_{\text{all elements}} U^{et} \left[\int_{v_{el}} B^t C B d(v_{el}) \right] U^e = \sum_{\text{all elements}} U^{et} [K^e] U^e$$

where K^e is the element stiffness matrix, and U^e is a vector obtained by concatenating all the displacements of nodes in the element, i.e., $U^e = [u_{1,x}, u_{1,y}, u_{1,z}, \dots, u_{8,x}, u_{8,y}, u_{8,z}]$ where $u_i = (u_{i,x}, u_{i,y}, u_{i,z})$ is the displacement of node i .

Step 4: Rewrite the internal energy function in matrix form. First, define the global displacement vector, U , as $U = [u_{1,x}, u_{1,y}, u_{1,z}, \dots, u_{n,x}, u_{n,y}, u_{n,z}]^t$, where n is the total number of nodes in the solid. Then, define the global stiffness matrix, K , as the assembly of all the local element stiffness matrices, K^e :

$$K = \sum_{\text{all elements}} \mathcal{I}(K^e) \quad (2)$$

where \mathcal{I} is the re-indexing function, which takes element K_{ij}^e and adds it to the element K_{kl} , where k and l are the global node numbers of local nodes i and j .²

The internal energy can now be written as $W(U) = U^t K U$.

Applying External Data Constraints: The rest of the process involves the formulation of the external energy term, which is problem specific. Let's assume for now that this term can be posed in quadratic form as $V(U) = (U^m - U)^t A (U^m - U)$. Once this is in place, the sum of $W(U) + V(U)$ can be minimized, by differentiation with respect to U , and set to zero. This results in the final equation:

$$K U = A (A U^m - U) \quad (3)$$

Equation (3) can be solved for U using sparse matrix methods.³ Since U represents the values of u at each node, we can compute the resulting values of the displacement field u anywhere in the volume by means of the finite element approximation ($u \approx \sum_{i=1}^8 N_i u_i$).

²Within an element, the nodes are always numbered from 1 to 8. However this is a local index (short-hand) to the global node numbers. When the global matrix is assembled, the local indices (1 to 8) need to be converted back to the global indices (e.g. 1 to n). K^e has dimensions 24×24 and K has dimensions $3n \times 3n$. K_{14}^e , which is the stiffness between the x -directions of local nodes 1 and 2 would be part of K_{kl} where $k = 3(a - 1) + 1$ and a is the global index of local node 1 and $l = 3(b - 1) + 1$, where b is the global index of local node 2. Since nodes appear in more than one element the final value of K_{kl} is likely to be the sum of a number of local K_{ij}^e 's.

³In the case of finite deformations, we end up with an expression of the form $K(U) = A(U^m - U)$ which is solved iteratively.

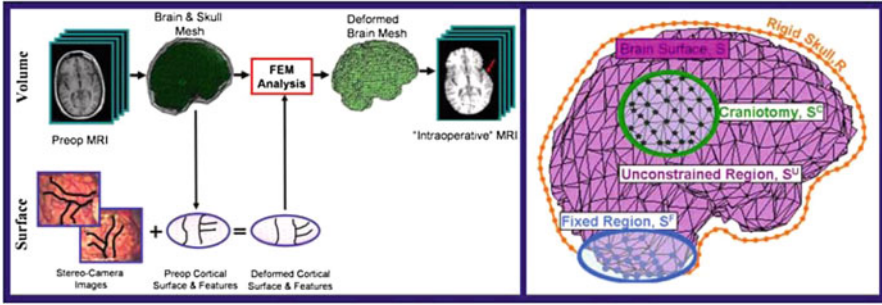


Fig. 1 Left: The brain and skull are extracted from the preoperative MRI, serving as inputs, along with the deformed cortical surface, to the FEM volume calculation. The FEM result is the deformed position of every brain mesh node. This mesh is then resampled into the original image space to yield a simulated “intraoperative” MRI. The red arrow indicates the region of greatest deformation. **Right:** Sample brain mesh with surface node labels. (See Sect. 2.3)

2.3 System Description

Our approach to brain shift compensation, outlined on the left side of Fig. 1, employs a 3D biomechanical brain model guided by sparse intraoperative data. The intraoperative data is acquired by stereo camera images and consists of cortical surface intensities and features (sulci). We use a game theoretic framework to overcome the resolution issues associated with stereo imaging and predict the cortical deformation. The application of this method requires discretizing the preoperative brain into small elements (right side of Fig. 1) and applying the finite element method (Sect. 2.2) to determine each element’s displacement.

Image-based Displacement Estimates: In order to obtain accurate quantitative information from stereo cameras, as with any imaging system, calibration is usually necessary. However, in many real world situations, accurate calibrations are not possible [34]. This is especially true in the operating room, where extreme time and space constraints limit the calibration procedure possibilities. (See [52] for calibration methods and the use of calibration parameters to project 3D points onto images.) The resulting inaccurate camera calibrations decrease image resolution and compound the difficulty of image-derived deformation estimation.

Therefore, in order to track the deforming cortical surface, a framework with the ability to solve for competing variables (surface displacement field/accurate camera calibration parameters) is needed. Game theory, the study of multiperson decision making in a competitive environment [6, 36], can provide this framework. In a game theoretic formulation, the players are computational modules. In the context of intraoperative neurosurgical guidance, the players would be 1) \underline{U}_{dns} , the dense displacement field applied to the preoperative cortical surface and 2) $\underline{A} = [\underline{A}_0, \underline{A}_1]$, the camera calibration parameters for the left (0) and right (1)

stereo cameras. This analysis therefore updates surface displacement and calibration parameter estimates at every iteration.

The model for determining these variables, expressed in a game theoretic formulation is:

$$C_1(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}}) = \underbrace{T_U(\underline{\mathbf{U}}_{dns})}_{\text{smoothness constraint}} + \alpha \times \underbrace{T_F(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}})}_{\text{feature matching}} + \underbrace{T_I(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}})}_{\text{intensity correlation}} \quad (4)$$

$$C_2(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}}) = \underbrace{T_A(\underline{\mathbf{A}})}_{\text{fiducial matching}} + \underbrace{\beta \times T_C(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}})}_{\text{reconstructed sulci matching}} \quad (5)$$

where C_1 , C_2 are the cost functions for the dense displacement field and camera calibration parameters, respectively. These cost functions can be iteratively minimized until they reach equilibrium, representing the algorithm solution [11]. The constants in these functions, α and β , are chosen using game theoretic constructs for noncooperative games, in which collusion between players is prevented and the players can pursue their own interests, which may conflict [6, 11]. The other terms are presented below.

Displacement Field Determination: The intensity correlation term matches the stereo image intensities that are backprojected onto the exposed brain surface. A backprojected image (Fig. 2D) is defined as $B_i^S = I_i(P(A_i, x)) \forall x \in S$, where S is the deformed surface, I is an intraoperative stereo image, P is a standard camera projection function from the surface to the images and i represents the camera number (0 or 1). This term can be written as $T_I(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}}) = \eta NCC(B_0^S, B_1^S)$ where NCC is the normalized cross correlation and η is a normalizing constant, ensuring that all terms have similar orders of magnitude.

A feature matching term measures the distance between 3D sulci on the cortical surface, $\underline{\mathbf{C}}$, which are projected into the stereo images, and the intraoperative sulci outlined in those images, $\underline{\mathbf{K}} = [K_0, K_1]$. The imaged sulci are manually extracted by an expert user and stored as 2D curves. This term can be expressed by:

$$T_F(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}}) = - \left(\int d \left[K_0, P \left(A_0, \left(\underline{\mathbf{C}} + \underline{\mathbf{U}}_{dns}^C \right) \right) \right] dS + \int d \left[K_1, P \left(A_1, \left(\underline{\mathbf{C}} + \underline{\mathbf{U}}_{dns}^C \right) \right) \right] dS \right),$$

where $\underline{\mathbf{U}}_{dns}^C$ is the dense displacement field restricted to the sulci and d is a mean Euclidean distance metric.

A prior term, $T_U(\underline{\mathbf{U}}_{dns}) = \rho e - \int \|\underline{\mathbf{U}}_{dns}''\| dS$, where ρ is a normalizing constant and $\underline{\mathbf{U}}_{dns}''$ is the second derivative of the dense displacement field, ensures deformation smoothness.

Camera Calibration Optimization: The reconstructed sulci matching term in the camera calibration optimization also takes advantage of the cortical surface features. Using standard stereo camera geometry, a reconstruction function, Φ , based on the triangulation of imaged points, can be written [52]. Sulci outlined from stereo images can therefore be reconstructed in 3D space and compared to their positions on the deformed cortical surface using the function $T_C(\underline{\mathbf{U}}_{dns}, \underline{\mathbf{A}}) = \int d[\Phi(K_0, K_1, \underline{\mathbf{A}}), (\underline{\mathbf{C}} + \underline{\mathbf{U}}_{dns}^C)]dS$.

The initial camera calibration is performed by locating fiducial points on the cortical surface and stereo images and finding the best fit transformation between the two fiducial point sets [52]. The fiducial matching term, acting as a prior, ensures that as the camera parameters are updated, the projection of the n fiducial points, $L_{0\dots n}$, onto the stereo camera images will be close to imaged fiducial points from cameras 0 and 1, $m_{0i\dots n}$ and $m_{1i\dots n}$, respectively. This yields: $T_A(\underline{\mathbf{A}}) = \sum_{i=1}^n \| [P(A_0, Li) - m_{0i}] \| + \| [P(A_1, Li) - m_{1i}] \|$.

Biomechanical Model: Because brain shift is small relative to brain size, linear elastic models (LEMs) can provide a good approximation of the tissue movement [2, 14, 19, 33, 47, 49]. An LEM is used in this work that (1) incorporates the main tissue characteristics (elasticity/near-incompressibility) without including more model parameters that are difficult to determine, (2) is static, since intraoperative brain deformation is a relatively slow process with negligible dynamic components, (3) is isotropic (Although brain tissues are not isotropic, especially the fibrous white matter, since the fiber directions are not always known, we assume that brain tissues are isotropic materials.) and (4) is guided by intraoperative inputs.

Based on the properties of LEMs, the deformation throughout the whole brain, B , can be calculated using an energy minimization framework (Sect. 2.2):

$$\hat{u} = \arg \min_u \left(\int_B W(u, m, E, v) dB \right) \quad (6)$$

where B is the entire brain, consisting of all nodes in the brain volume (V) and on the surface (S) and $W(u, m, E, v)$ is defined in Sect. 2.2.

Equation (6) is subject to three constraints: (1) The first constraint forces the displacements of the intraoperatively exposed cortical nodes, $u(S_i^c)$, in the region of the craniotomy, S^c , to exactly equal the game theoretic results at those nodes, $\underline{\mathbf{U}}_{dns}(s_i^c)$. It can be written as $u(s_i^c) = \underline{\mathbf{U}}_{dns}(s_i^c), \forall s_i^c \in S^c$. Due to this exact matching, the external energy term of equation (1) is not used. (2) The second constraint $\| (s_k^u + u(s_k^u)), -r_j \| > \delta, \forall s_k^u \in S^u, r_j \in R$, ensures that all nodes on the brain surface remain some small distance away from the skull (except in the craniotomy region). Here, s_k^u are brain surface nodes in the region not exposed during surgery, S^u , r_j are nodes on the rigid skull surface, R , and δ is an arbitrarily small constant. Due to the surface connectivity, this constraint also prevents any nodes from crossing the skull boundary, as this would force other surface nodes to violate this condition. Rather than forcing some surface nodes to be fixed, this

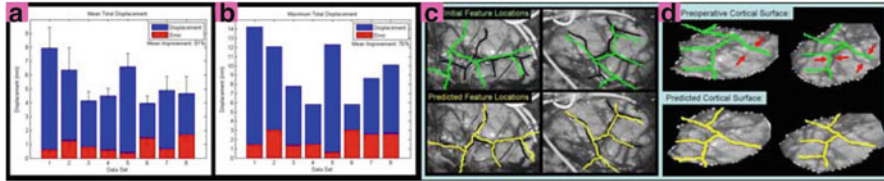


Fig. 2 Mean (A) and maximum (B) displacement and algorithm error ($\alpha = 4$, $\beta = 0.83$, $\rho = 0.1$, $\eta = 25$). In (C), intraoperative images of Data Set #2 show the misalignment of projected preoperative sulci (green) with intraoperative sulci positions (black) due to brain shift and camera calibration errors. Predicted sulci positions, projected with the updated calibration parameters (yellow), show better alignment. Backprojected intensities (D) are found by projecting each point on the surface to the left (left column) or right (right column) image and assigning the associated image intensity value to that point. Red arrows indicate the misalignment between the sulci on the preoperative surface (green) and those seen in the backprojected image. Sulci on the algorithm-predicted surface (yellow) are better aligned with the image intensities

model more realistically constrains the anatomy. (3) Finally, the third constraint, $u(s_l^f) = 0, \forall s_l^f \in S^f$, states that the model deformation of the fixed nodes, s_l^f , within the fixed region, S^f , must be zero. The fixed surface region is at the base of the brain in the inferior occipital lobes. Due to the tough tentorium cerebelli, as well as their distance from the craniotomy site, these nodes will not move during surgery. (See right side of Fig. 1.)

The constrained minimization was performed using the finite element analysis software package ABAQUS (ABAQUS, Inc., Providence, RI). The model inputs were a tetrahedral brain mesh (created using the automated algorithm suggested in Stokking [50]), a manually-outlined surface representing the surrounding skull (with craniotomy), Young's modulus (66.7 kPa) and Poisson's ratio (0.48) of brain [30]. The finite element analysis output is the displacement of all brain mesh nodes, which can be resampled into image space using trilinear interpolation.

2.4 Experimental Results

Game theoretic cortical surface tracking was used in five separate surgeries for a total of eight data sets. The algorithm results for all cases are shown in Figs. 2A & B. In Fig. 2A, the blue bars represent the calculated mean average displacement of the cortical surface as predicted by the game theoretic algorithm. Mean residual error of the algorithm (red bars) is calculated by averaging the closest distances between the predicted surface and sparse cortical surface points touched with a 3D locator intraoperatively.

Five out of eight of the cases (62.5 %) resulted in a mean algorithm error of less than 1.0 mm, and the mean error never exceeded 1.70 mm, representing an 81 % improvement over uncompensated error. Also, for half the cases, the maximum error was under 1.6 mm (Fig. 2B), representing a 76 % decrease in the maximum errors

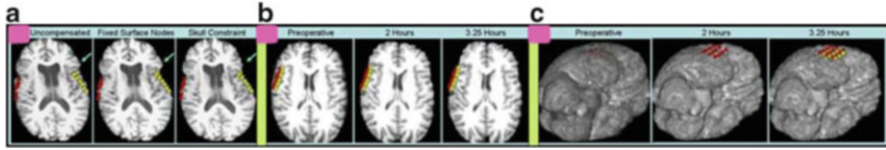


Fig. 3 **A)** A slice of the preoperative MRI (left), deformed using either a fixed surface (middle) or a skull constraint (right). Red and yellow spheres indicate intraoperatively acquired surface points. The aqua arrow is in the same location on each image. **B)** One slice of the preoperative (right) and predicted intraoperative initial (middle) and final (left) MR image. The spheres were acquired by the neurosurgeon either 2 (red) or 3.25 (yellow) hours into surgery. **C)** Volume Renderings of the images in (B)

using the model guidance. Thus, for all eight cases, every part of the surface was found more accurately using the game theoretic algorithm than by relying on the preoperative surface. Figures 2C & D illustrate these results using images from a typical sample case, Data Set #2.

These surface results were then used in conjunction with the biomechanical model to obtain volumetric deformation. As mentioned above, rather than artificially fixing the non-exposed surface nodes, the rigid skull was used to constrain the surface deformation (constraint #2). Figure 3A illustrates this effect on a case in which a bilateral craniotomy was performed. For this patient, the surface deformation for each side of the brain was calculated separately, using data sets 2 (left side of brain) and 3 (right side of brain) from Fig. 2, and the modeled skull contained two craniotomy sections. With a fixed surface, the nodes near the craniotomy cannot move, even when the deformation becomes large. This effect is most obvious in the region indicated by the aqua arrow, which is located in the same relative position on all three images. The deformation decreases sharply to zero outside the craniotomy region when the surface nodes are fixed. However, when constrained by the skull, the region indicated by the aqua arrow is allowed to deform inward as well, resulting in a more natural deformation.

To validate model consistency, volumetric deformation was also calculated for two data sets (4 & 5) from the same surgery. Figures 3B & C show the calculated downward surface shift, which is propagated through the volume.

3 A Boundary Element Based Approach: Estimation of 3D Left Ventricular Deformation

3.1 Background

Acute coronary artery occlusion results in myocardial injury, which will progress from the endocardium to the epicardium of the heart wall in a wavefront fashion. A primary goal in the treatment of patients presenting with acute myocardial

infarction is to reestablish coronary flow, and to interrupt the progression of injury, thereby salvaging myocardium. A number of laboratories have shown that a comprehensive quantitative analysis of myocardial strain can more accurately identify ischemic injury than simple analysis of endocardial wall motion or radial thickening [5]. Furthermore, the characterization of segmental strain components has shown great promise for defining the mechanical mechanisms of tethering or remodeling [31, 35]. At present, quantitative non-invasive measurement of 3D strain properties from images has been limited to special forms of magnetic resonance (MR) acquisitions, specifically MR tagging and restricted to mostly research settings.

In general, there are three different approaches to estimating displacement data from MR tagging. The first approach involves tagging in multiple intersecting planes at the same time, and using the tag intersections as tokens for tracking (e.g. [1, 54]). The second approach involves tagging in multiple intersecting planes, one set of parallel planes at a time. Then, each tagging plane is used separately to estimate the normal direction of motion perpendicular to the plane. This generates a set of partial displacements (i.e. the component parallel to the tag lines is missing) to be combined later (e.g. [26, 17]). The final approach uses a lower resolution modulation technique and attempts to model the tag fading over time using the Bloch equations. The displacements are then extracted using a variable brightness optical flow technique.

As an alternative to MR tagging, several investigators have employed changes in phase due to motion of tissue within a fixed voxel or volume of interest to assist in estimating instantaneous, localized velocities, and ultimately cardiac motion and deformation [55, 38]. This technique basically relies on the fact that a uniform motion of tissue in the presence of a magnetic field gradient produces a change in the MR signal phase that is proportional to velocity.

The use of computer vision-based techniques to estimate displacement is also possible. One approach to establishing correspondence is to track shape-related features on the LV over time as reported by Cohen [15], Papademetris [42], Lin [32]. This is the basis for much of our own work and is expanded later. We note that such methods were applied to modalities other than magnetic resonance such as X-ray CT [42] and ultrasound [42, 32, 53].

Finally, some investigators have used the intensity of the images directly to track local LV regions. Song and Leahy [48] used the intensity in ultrafast CT images to calculate the displacement fields for a beating heart. In addition, other investigators have used local image intensity or intensity-based image texture from echocardiographic image sequences to track local positions over 2D image sequences [37]. These efforts, along with some related MR tagging approaches roughly fall into the category of optical flow-based methods. With the exception of methods based on magnetic resonance tagging and to a lesser extent MR phase contrast velocities, none of the other methods is capable of estimating complete three-dimensional deformation maps of the left ventricle.

In this work we propose a method for estimating non-rigid cardiac motion which combines the shape based Generalized Robust Point Matching (GRPM) framework and an biomechanical model constructed by BEM.

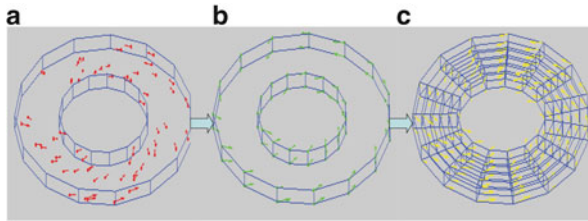


Fig. 4 The procedure of a BEM-based regularization model: (a) Displacements associated with scattered points (red); (b) Model interpolated displacements on the boundary nodes (green); (c) Model interpolated dense displacement fields (yellow). *Figure reprinted from [53], ©2007 by permission from Elsevier*

3.2 The Boundary Element Method (BEM)

The BEM is an alternative technique to FEM for solving Biomechanical problem. The BEM reformulates a dense field problem to the integral equation over the boundary of the field [8]. The procedure for employing the BEM-based regularization model is described in Fig. 4. The red points and arrows in Fig. 4a are the feature points and estimated displacements derived from image sequences. First, we map these displacements associated with feature points to the boundary grid nodes as shown in Fig. 4b. Then we interpolate the dense displacements on the dense grids as shown in Fig. 4c.

Let $Q = \{q_j, j = 1, 2, \dots, J\}$ be the nodes at regular positions of the boundary. Consider a set of scattered points $A = \{a_i, i = 1, 2, \dots, I\}$ that are extracted from the image. $G = \{g_l, l = 1, 2, \dots, L\}$ are the dense points at regular positions. With known displacements associated with the scattered points in A , our goal is finding the optimal displacements for all the points in G . The points in A , Q and G are red, green and yellow points as shown in Fig. 4, respectively. In the BEM, the displacements $U(q_j)$ and tractions $P(q_j)$ of the boundary node q_j can be written as:

$$HU(q_j) = GP(q_j) \quad (7)$$

The H and G are the fundamental matrices for the boundary nodes and they are defined by the elastic material property and shape information. The detail definition of the H and G can be found in [8]. The displacement at any point can be derived from known boundary displacements $U(q_j)$ and tractions $P(q_j)$:

$$u(a_i) = \sum_{j=1}^J \widehat{G} P(q_j) - \widehat{H} U(q_j) \quad (8)$$

$u(a_i)$ is the displacement at the point a_i . \widehat{G} and \widehat{H} are the fundamental matrix of the point a_i . To determine the unknown displacements at boundary nodes, we

first consider one data point a_i in A. Here we denote the measured displacement associating with the point a_i as $\tilde{u}(a_i)$. For the displacement at the point a_i to take on the value $\tilde{u}(a_i)$, the displacements $\tilde{U}_i(q_j)$ at the boundary nodes must satisfy:

$$\tilde{u}(a_i) = \sum_{j=1}^J C_{ij} \tilde{U}_i(q_j) \quad (9)$$

where $C = \widehat{G}G^{-1}H - \widehat{H}$. There are many values of $\widehat{U}_i(q_j)$ that are solutions to Eq. (9). Here we choose the solution in the least-squared sense such that:

$$\tilde{U}_i(q) = \arg \min \left(\sum_{j=1}^J \tilde{U}_i^2(q_j) \right) \quad (10)$$

Simple linear algebra using the pseudoinverse is used to derive the solution [28]:

$$\tilde{U}_i(q) = (C_{ij} \tilde{u}(a_i)) / \left(\sum_{j=1}^J C_{ij}^2 \right) \quad (11)$$

Now we consider all the data points in A. Here we choose the displacements $U(q)$ of boundary nodes to minimize the sum of squared differences between the displacements $u(a_i)$ and $\tilde{u}(a_i)$ of all the points in the point-set A, such as:

$$U(q) = \arg \min \left(\sum_{i=1}^I \lambda_i \| u(a_i) - \tilde{u}(a_i) \|^2 \right) \quad (12)$$

where λ_i is the weight based on the confidence in the $\tilde{u}(a_i)$ Then we get:

$$U(q) = \left(\sum_{i=1}^I \lambda_i C_{ij}^2 \tilde{U}_i(q) \right) / \left(\sum_{i=1}^I \lambda_i C_{ij}^2 \right) \quad (13)$$

Once we know the displacements of all the boundary nodes, the displacements of any point g_l in the dense point-set G can be calculated by:

$$U(g_l) = \sum_{j=1}^J C_{lj} U(q_j) \quad (14)$$

3.3 Generalized Robust Point Matching

The GRPM extends the Robust Point Matching (RPM) framework [13] to use a more general metric form that includes curvature information [32]. In particular, the use of shape information is embedded to guide more precise motion recovery and points that are mistaken as features or unmatched real feature points are automatically treated as outliers by the GRPM algorithm during the optimization annealing process. The objective function of GRPM is:

$$\begin{aligned}
 E(M) = & \sum_{i=1}^I \sum_{k=1}^K m_{ik} [\| f(a_i) - b_k \|^2 + \lambda_{\mathcal{L}} g(\| \kappa_A(a_i) - \kappa_B(b_k) \|^2)] \\
 & + \lambda \| \mathcal{L} f \|^2 + T \sum_{i=1}^I \sum_{k=1}^K m_{ik} \log m_{ik} \\
 & + T_0 \sum_{i=1}^I m_{i,K+1} \log m_{i,K+1} + T_0 \sum_{k=1}^K m_{I+1,k} \log m_{I+1,k} \quad (15)
 \end{aligned}$$

where M is the correspondence matrix. A mesh is generated on the boundary of I_1 . The T is annealing temperature gradually decreasing to zero as the matching iteration begins. The T_0 is annealing temperature for outliers. The f is the non-rigid transformation function and L is an operating function on f . The $g(\cdot)$ is a strictly increasing function and $\lambda_{\mathcal{L}}$ balances the significance between distance and new information. In our case, $\kappa_A(a_i)$ and $\kappa_B(b_k)$ are two principal curvatures at a_i and b_k , respectively.

3.4 Shape-based BEM-GRPM Tracking System

Figure 5 shows the design strategy: 1. Segmentation at the first frame; 2. Extracting the feature points; 3. Calculating the curvature associated with each feature point; 4. Calculating the surface mesh at the next time frame using the BEM-GRPM; 5. If the last time frame is reached, stop. Otherwise, repeat the step 4; 6. Calculating the dense displacement fields; 7. Calculating the strains.

Following steps summarize the scheme of BEM-GRPM framework:

Step 1: Estimate the correspondence matrix M between the points a_j and b_k The correspondence matrix M is calculated:

$$m_{ik} = \frac{1}{\sqrt{2\pi T^2}} e^{-\frac{\|f(a_i) - b_k\|^2 + \lambda_A g(\|\kappa_A(a_i) - \kappa_B(b_k)\|^2)}{T}}, \sum_{i=1}^{I+1} m_{ik} = 1, \text{ and } \sum_{k=1}^{K+1} m_{ik} = 1 \quad (16)$$

We denote the outlier clusters as a_{N+1} and b_{K+1} .

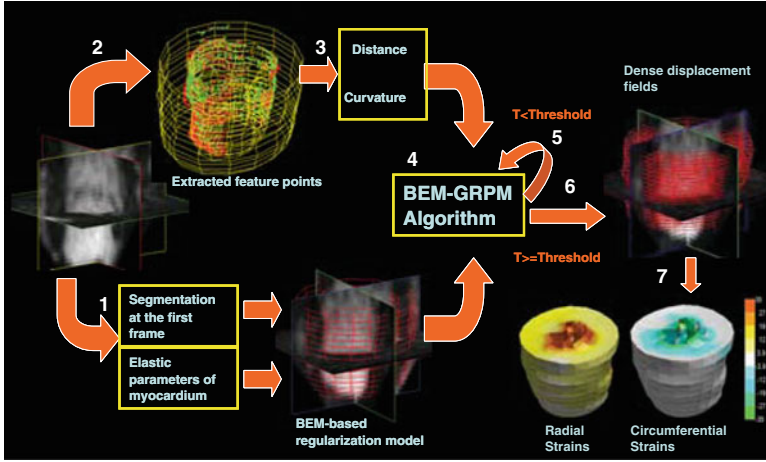


Fig. 5 The overview of the shapebased motion tracking system

Step 2: Calculate the corresponding points and confident parameters: Let \hat{b}_i be the corresponding point to a_i and λ_i the confidence in the match:

$$\hat{b}_i = \sum_{k=1}^K m_{ik} b_k / \sum_{k=1}^K m_{ik}, \lambda_i = \sum_{k=1}^K m_{ik} \quad (17)$$

Step 3: Update the non-rigid transformation function: the non-rigid transformation function f is defined by:

$$f = \arg \min_{f_i=1} \sum_{f_i=1}^I \lambda_i \left\| f(a_i) - \hat{b}_i \right\|^2 \quad (18)$$

here $f(a_i) = a_i + u(a_i)$, $\hat{b}_i = a_i + \tilde{u}(a_i)$. After we get the smoothed displacements of all the feature points, we update $f(a_i) : f(a_i) = a_i + u(a_i)$.

Step 4: Annealing process: If T is bigger than a threshold, $T = T \times \delta$, and go back to step 2. δ is the annealing rate of T .

Step 5: Dense displacement fields and strains: If annealing process is finished, we derive the dense displacement fields and the Lagrangian strains.

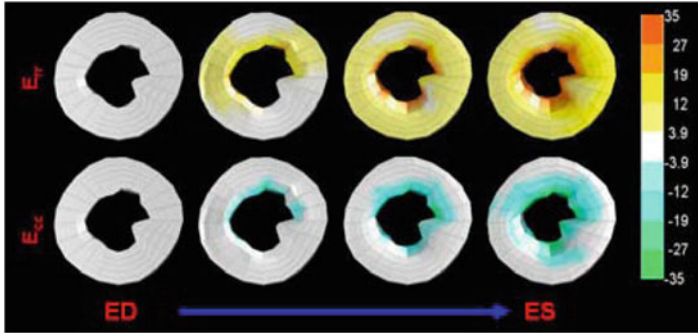


Fig. 6 A 2D slice of 3D radial strains E_{rr} (top) and circumferential strains E_{cc} (bottom) derived from an in-vivo cardiac MRI data (ED \rightarrow ES). Positive values represent thickening and negative values for shortening. *Figure reprinted from [53], ©2007 by permission from Elsevier*

3.5 Experimental Results

We first present the experimental results using the shape-based tracking system. The key practical application includes the estimation of left ventricular deformation from $3D+t$ MR and echocardiographic images. The results from MR images were compared to displacements found using implanted markers.

Experiments on 3D Canine MR images We test the *BEM-GRPM* algorithm on 3D cardiac MRI data from the baseline MR images (without post-occlusion) [40]. The node number of the BEM mesh is set to be 20×15 (θ, z). We track the myocardium frame-by-frame along the image sequence (from end-diastole (ED) to end-systole (ES)). The feature points are extracted by using thresholded curvature in this experiment because MRI has good image quality. The coordinates and the curvature value of the feature points are normalized in the beginning. We set the starting temperature T as 4 pixels. In the annealing process, we gradually reduce it by an annealing rate of 0.9. For each dataset, strains are calculated between end-diastole (ED) and end-systole (ES). A 2D slice of the 3D-derived radial strains (E_{rr}) and circumferential strains (E_{cc}) of one dataset are illustrated in Fig. 6. Note the normal behavior in the LV, showing positive radial strain (thickening) and negative circumferential strain (shortening) as we move from ED to ES.

To quantitatively validate the resulting motion trajectories, we use four canine MRI datasets with implanted markers for point-by-point displacement comparison (see [42] for more details on the marker implantation and localization). The mean displacement errors of BEM-GRPM for the four datasets are calculated and compared to the errors of EFFD-GRPM [32] which uses Extended Free Form Deformation (EFFD) as the regularization model with the GRPM. The displacement errors of BEM-GRPM are less than those of EFFD-GRPM as shown in Fig. 7.

Fig. 7 Absolute displacement error vs. implanted markers. The errors are estimated between ED to ES for 4 canine MR image datasets. **Blue**: the error of BEM-GRPM; **Red**: the error of EFFD-GRPM. *Figure reprinted from [53], ©2007 by permission from Elsevier*

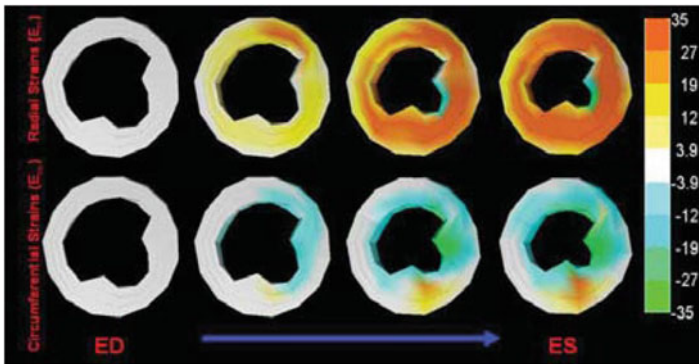
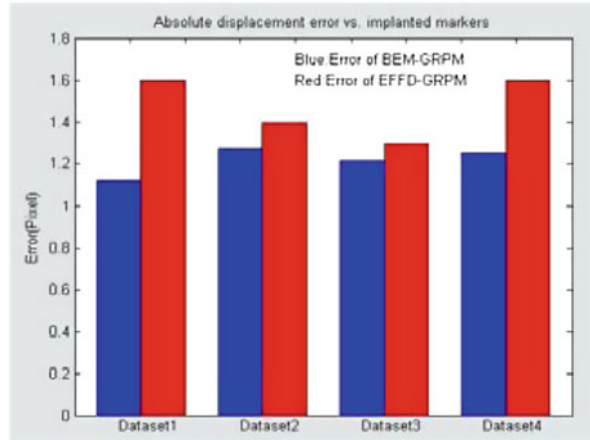


Fig. 8 Radial strains E_{rr} (top) and circumferential strains E_{cc} (bottom) derived from a contrast-enhanced echocardiographic images (ED \rightarrow ES). Positive values represent thickening and negative values for shortening. *Figure reprinted from [53], ©2007 by permission from Elsevier*

Experiments on 3D Echocardiographic B-mode images The experiments were carried on contrast-enhanced full volume 3D B-mode canine data. A 2D slice of the 3D-derived radial (E_{rr}) and circumferential strains (E_{cc}) are illustrated in Fig. 8. Although we can't evaluate our results from echocardiographic images quantitatively, the derived LV motion patterns are visually correct.

4 Concluding Remarks

We have presented two distinct application areas where biomechanical models are used to recover information from image sequences ...

Edit this some more

We have described the development of an automatic, robust and efficient based on a physical model. We have presented a sequence of results: First, dense displacements and strains were derived from MR image sequences using the shape-based BEM-GRPM tracking system. The results were compared to the displacements found using implanted markers, taken as the ground truth. Second, we applied the shape-based BEM-GRPM tracking system on a 3D echocardiographic image sequence and got the reasonable results.

References

1. A.A. Amini, Y. Chen, R.W. Curwen, V. Mani, and J. Sun. Coupled B-snake grids and constrained thin-plate splines for analysis of 2-D tissue deformations from tagged MRI. *IEEE Trans. Med. Imag.*, 17(3):344–356, 1998.
2. Neculai Archip, Andriy Fedorov, Bryn Lloyd, Nikos Chrisochoides, Alexandra Golby, Peter M. Black, and Simon K. Warfield. Integration of patient specific modeling and advanced image processing techniques for image-guided neurosurgery. In *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display, Proceedings of the SPIE*, volume 6141, pages 422–429, San Diego, CA, February 12-14 2006.
3. Michel A. Audette, Kaleem Siddiqi, Frank P. Ferrie, and Terry M. Peters. An integrated range-sensing, segmentation and registration framework for the characterization of intra-surgical brain deformations in image-guided surgery. *Computer Vision and Image Understanding*, 89(2-3):226–251, February - March 2003.
4. Michel A. Audette, Kaleem Siddiqi, and Terry M. Peters. Level-set surface segmentation and fast cortical range image tracking for computing intrasurgical deformations. In *Medical Image Computing and Computer-Assisted Intervention (MIC-CAI)*, volume 1679, pages 788–797, Cambridge, UK, September 19-22 1999.
5. H. Azhari, J. Weiss, W. Rogers, C. Siu, and E. Shapiro. A noninvasive comparative study of myocardial strains in ischemic canine hearts using tagged MRI in 3D. *American Journal of Physiology*, 268, 1995.
6. Tamer Başar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory, 2nd Ed.* Academic Press, New York, 1995.
7. K. Bathe. *Finite Element Procedures in Engineering Analysis.* Prentice-Hall, New Jersey, 1982.
8. C.A. Brebbia and J. Dominguez. *Boundary Elements An Introductory Course.* Computational Mechanics Publications, 1998.
9. Alize Cao, Prashanth Dumpuri, and Michael I. Miga. Tracking cortical surface deformations based on vessel structure using a laser range scanner. In *International Symposium on Biomedical Imaging (ISBI)*, pages 522–525, Washington, DC, USA, April 6-9 2006.
10. Alexandra Chabrierie, Fatma Ozlen, Shin Nakajima, Michael Leventon, Hideki Atsumi¹, Eric Grimson, Erwin Keeve, Sandra Helmers, James Rivello Jr., Gregory Holmes, Frank Duffy, Ferenc Jolesz, Ron Kikinis, and Peter Black. Three-dimensional reconstruction and surgical navigation in pediatric epilepsy surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1496, pages 74–83, Cambridge, MA, October 11-13 1998.
11. Amit Chakraborty and James S. Duncan. Game-Theoretic integration for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(1):12–30, January 1999.
12. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. 3D brain mapping using deformable neuroanatomy. *Physics in Medicine and Biology*, 39:609–618, 1994.
13. H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume 2, pages 44–51, 2000.

14. Olivier Clatz, Hervé Delingette, Ion-Florin Talos, Alexandra J. Golby, Ron Kikinis, Ferenc A. Jolesz, Nicholas Ayache, and Simon K. Warfield. Robust nonrigid registration to capture brain shift from intraoperative MRI. *IEEE Transactions on Medical Imaging*, 24(11):1417–1427, November 2005.
15. L. Cohen, N. Ayache, and P. Sulger. Tracking points on deformable objects using curvature information. *Lecture Notes in Computer Science (ECCV92)*, pages 458–466, 1992.
16. Hervé Delingette, Xavier Pennec, Luc Soler, Jacques Marescaux, and Nicholas Ayache. Computational models for image-guided robot-assisted and simulated medical interventions. *Proceedings of the IEEE*, 94(9):1678–1688, September 2006.
17. T.S. Denney Jr and J.L. Prince. Reconstruction of 3d left ventricular motion from planar tagged cardiac mr images: An estimation theoretic approach. *IEEE Transactions on Medical Imaging*, 14(4):625–635, 1995.
18. Valerie Duay, Tuhin K. Sinha, Pierre-François D’Haese, Michael I. Miga, and Benoit M. Dawant. Non-rigid registration of serial intra-operative images for automatic brain shift estimation. In *Workshop on Biomedical Image Registration (WBIR)*, volume 2717, pages 61–70, Philadelphia, PA, June 23-24 2003.
19. Prashanth Dumpuri, Reid C. Thompson, Tuhin K. Sinha, and Michael I. Miga. Automated brain shift correction using a pre-computed deformation atlas. In *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display, Proceedings of the SPIE*, volume 6141, pages 430–437, San Diego, CA, February 12-14 2006.
20. Matthieu Ferrant, Arya Nabavi, Benoît Macq, P. M. Black, Ferenc A. Jolesz, Ron Kikinis, and Simon K. Warfield. Serial registration of intraoperative MR images of the brain. *Medical Image Analysis*, 6(4):337–359, December 2002.
21. J. C. Gee, D. R. Haynor, L. Le Briquer, and R. K. Bajcsy. Advances in elastic matching theory and its implementation. In *CVRMed-MRCAS*, Grenoble, France, March 1997.
22. David T. Gering, Arya Nabavi, Ron Kikinis, W. Eric L. Grimson, Nobuhiko Hata, Peter Everett, Ferenc A. Jolesz, and William M. Wells III. An integrated visualization system for surgical planning and guidance using image fusion and interventional imaging. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1679, pages 809–819, Cambridge, UK, September 19-22 1999.
23. David G. Gobbi, Roch M. Comeau, and Terry M. Peters. Ultrasound/MRI overlay with image warping for neurosurgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1935, pages 106–114, Pittsburgh, PA, October 11-14 2000.
24. W. E. L. Grimson, G. J. Ettinger, S. J. White, T. Lozano-Pérez, W. M. Wells, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE Transactions on Medical Imaging*, 15(2):129–140, April 1996.
25. J. M. Guccione and A. D. McCulloch. Finite element modeling of ventricular mechanics. In P. J. Hunter, A. McCulloch, and P. Nielsen, editors, *Theory of Heart*, pages 122–144. Springer-Verlag, Berlin, 1991.
26. E. Haber, D.N. Metaxas, and L. Axel. Motion analysis of the right ventricle from mri images. *Medical Image Computing and Computer Aided Intervention*, pages 177–188, 1998.
27. Peter Hastreiter, Christof Rezk-Salama, Grzegorz Soza, Michael Bauer, Günther Greiner, Rudolf Fahlbusch, Oliver Ganslandt, and Christopher Nimsky. Strategies for brain shift evaluation. *Medical Image Analysis*, 8(4):447–464, December 2004.
28. W. Hsu, J. Huaahes, and H. Kaufman. Direct manipulation of free-form deformations. *Computer Graphics*, 26(2):177–184, 1992.
29. K. H. Huebner, E. A. Thornton, and T. G. Byrom. *The Finite Element Method For Engineers*. John Wiley & Sons, New York, 1995.
30. Albert I. King, King H. Yang, and Tom Khalil. WSU Brain Injury Model, <http://ttb.eng.wayne.edu/brain/>.
31. C. Kramer, W. Rogers, T. Theobald, T. Power, S. Petruolo, and N. Reichek. Remote noninfarcted regional dysfunction soon after first anterior myocardial infarction: A magnetic resonance tagging study. *Circulation*, 94(10):660–666, 1996.

32. N. Lin and J.S. Duncan. Generalized robust point matching using an extended free-form deformation model: Application to cardiac images. In *International Symposium on Biomedical Imaging*, 2004.
33. Karen E. Lunn, Keith D. Paulsen, Daniel R. Lynch, David W. Roberts, Francis E. Kennedy, and Alex Hartov. Assimilating intraoperative data with brain shift modeling using the adjoint equations. *Medical Image Analysis*, 9(3):281–293, June 2005.
34. M. Machacek, M. Sauter, and T. Rösgen. Two-step calibration of a stereo camera system for measurements in large volumes. *Measurements in Science and Technology*, 14:1631–1639, 2003.
35. J. Marcus, M. Gotte, A. V. Rossum, J. Kuijter, R. Heethaar, L. Axel, and C. Visser. Myocardial function in infarcted and remote regions early after infarction in man: Assessment by magnetic resonance tagging and strain analysis. *Magnetic Resonance in Medicine*, 38:803–810, 1997.
36. Elliott Mendelson. *Introducing Game Theory and Its Applications*. Chapman & Hall/CRC, Boca Raton, 2004.
37. J. Meunier. Tissue motion assessment from 3d echographic speckle tracking. *PHYSICS IN MEDICINE AND BIOLOGY*, 43(5):1241–1254, 1998.
38. F.G. Meyer, R.T. Constable, A.G. Sinusas, and J.S. Duncan. Tracking myocardial deformation using spatially constrained velocities. *Information Processing in Medical Imaging*, pages 26–30, 1995.
39. Michael I. Miga, Tuhin K. Sinha, David M. Cash, Robert L. Galloway, and Robert J. Weil. Cortical surface registration for image-guided neurosurgery using laser-range scanning. *IEEE Transactions on Medical Imaging*, 22(8):973–985, August 2003.
40. X Papademetris. *Estimation of 3D Left Ventricular Deformation from Medical Images Using Biomechanical Models*. PhD thesis, Department of Electrical Engineering, Yale University, 2000.
41. X. Papademetris, A. J. Sinusas, D. P. Dione, R. T. Constable, and J. S. Duncan. Estimation of 3D left ventricular deformation from medical images using biomechanical models. *IEEE Trans. Med. Imag.*, 21(7), July 2002.
42. X. Papademetris, A.J. Sinusas, D.P. Dione, R.T. Constable, and J.S. Duncan. Estimation of 3-D left ventricular deformation from medical images using biomechanical models. *IEEE TMI*, 21(7):786–800, 2002.
43. Terry Peters, Bruce Davey, Patrice Munger, Roch Comeau, Alan Evans, and Andr e Olivier. Three-Dimensional multimodal image-guidance for neurosurgery. *IEEE Transactions on Biomedical Engineering*, 15(2):121–128, April 1996.
44. Ingerid Reinertsen, Maxime Descoteaux, Simon Drouin, Kaleem Siddiqi, and D. Louis Collins. Vessel driven correction of brain shift. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3217, pages 208–216, Saint-Malo, France, September 26-29 2004.
45. David W. Roberts, Alexander Hartov, Francis E. Kennedy, Michael I. Miga, and Keith D. Paulsen. Intraoperative brain shift and deformation: A quantitative analysis of cortical displacement in 28 cases. *Neurosurgery*, 43(4):749–758, October 1998.
46. Tuhin K. Sinha, Benoit D. Dawant, Valerie Duay, David M. Cash, Robert J. Weil, Reid C. Thompson, Kyle D. Weaver, and Michael I. Miga. A method to track *Medical Imaging*, 24(6):767–781, June 2005.
47. Oskar Škrinjar, Arya Nabavi, and James S. Duncan. Model-driven brain shift compensation. *Medical Image Analysis*, 6(4):361–373, December 2002.
48. S. Song and R. Leahy. Computation of 3D velocity fields from 3D cine CT images. *IEEE TMI*, 10:295–306, 1991.
49. Grzegorz Soza, Roberto Grosso, Christopher Nimsky, Guenther Greiner, and Peter Hastreiter. Estimating mechanical brain tissue properties with simulation and registration. In *Medical Image Computing and Computer-Assisted Intervention (MIC-CAI)*, volume 3217, pages 276–283, Saint-Malo, France, September 26-29 2004.
50. Rik Stokking. *Integrated Visualizatoin of Functional and Anatomical Brain Images*. PhD thesis, University Utrecht, February 1998.

51. Hai Sun, Karen E. Lunn, Hany Farid, Ziji Wu, David W. Roberts, Alex Hartov, and Keith D. Paulsen. Stereopsis-guided brain shift compensation. *IEEE Transactions on Medical Imaging*, 24(8):1039–1052, 2005.
52. Emanuele Trucco and Alessandro Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1998.
53. P. Yan, A.J. Sinusas, and J.S. Duncan. Boundary element method-based regularization for recovering of lv deformation. *Medical Image Analysis*, 11(6):540–554, 2007.
54. A. A. Young, D. L. Kraitchman, L. Dougherty, and L. Axel. Tracking and finite element analysis of stripe deformation in magnetic resonance tagging. *IEEE Transactions on Medical Imaging*, 14(3):413–421, 1995.
55. Y. Zhu, M. Drangova, and N.J. Pelc. Estimation of deformation gradient and strain from cine-pc velocity data. *IEEE Transactions on Medical Imaging*, 16(6):840–851, 1997.

Graph-based Deformable Image Registration

A. Sotiras*, Y. Ou*, N. Paragios, and C. Davatzikos

Abstract Deformable image registration is a field that has received considerable attention in the medical image analysis community. As a consequence, there is an important body of works that aims to tackle deformable registration. In this chapter we review one class of these techniques that use discrete optimization, and more specifically Markov Random Field models. We begin the chapter by explaining how one can formulate the deformable registration problem as a minimal cost graph problem where the nodes of the graph corresponds to the deformation grid, the graph connectivity encodes regularization constraints, and the labels correspond to 3D displacements. We then explain the use of discrete models in intensity-based volumetric registration. In the third section, we detail the use of Gabor-based attribute vectors in the context of discrete deformable registration, demonstrating the versatility of the graph-based models. In the last section of the chapter, the case of landmark-based registration is discussed. We first explain the discrete graphical

*The first two authors contributed equally to this work.

A. Sotiras (✉) • C. Davatzikos
Section of Biomedical Image Analysis, Center for Biomedical Image Computing and Analytics,
University of Pennsylvania, Philadelphia, USA
e-mail: aristeidis.sotiras@uphs.upenn.edu; christos.davatzikos@uphs.upenn.edu

Y. Ou
Athinoula A. Martinos Center for Medical Imaging, Massachusetts General Hospital,
Harvard Medical School, Boston, USA
e-mail: yangming.ou@mgh.harvard.edu

N. Paragios
Center for Visual Computing, Department of Applied Mathematics,
Ecole Centrale Paris, Paris, France
e-mail: nikos.paragios@ecp.fr

model behind establishing landmark correspondences, and then continue to show how one can integrate it with the intensity-based model towards creating enhanced models that combine the best of both worlds.

1 Introduction

Medical image analysis plays an increasingly important role in many clinical applications. The increased amount and complexity of medical image data, which often involve multiple 3D image modalities as well as multiple acquisitions in time, result in a challenging analysis setting. Image registration, as well as image segmentation, are the two principal tools that allow for automatic and timely data analysis.

Image registration consists of determining a spatial transformation that establishes meaningful anatomical, or functional, correspondences between different image acquisitions. The term deformable is used to specify that the transformation is allowed to spatially vary (in contrast to the case of linear or global registration). In general, registration can be performed between two or more images. Nonetheless, in this chapter, we will focus on registration methods that involve pairs of images. The pairs of images may consist of acquisitions that image either the same subject (intra-subject registration) or different subjects (inter-subject registration).

In intra-subject registration, the subject is typically imaged either under different protocols, or at different time points. In the first case, different imaging modalities are used to capture complementary anatomical or functional information, and image registration is used to fuse this information towards enhancing the analytical and diagnostic abilities of the clinicians. In the second case, one may study short- or long-term longitudinal processes that range from tumor perfusion properties to normal aging and development. Another application of image registration is surgical or treatment planning. The registration of pre-operative and interventional data allows the clinical experts to refine their planning and improve care-giving.

Inter-subject registration is the cornerstone of population studies. Mapping members of a population to a common domain allows the study of within-population variability and the quantitative analysis of the form of anatomical structures. On the other hand, when distinct populations are spatially aligned, it is possible to discover the focal differences that distinguish them by contrasting them in the common domain.

In general, an image registration algorithm involves three components (see Fig. 1 [75]): i) a transformation model; ii) a similarity criterion; and iii) an optimization method. Image registration has been studied extensively during the past decades, leading to a rich body of works. These works differ mainly in their choices with respect to these three components. While an extensive overview of these components is beyond the scope of this chapter, let us briefly discuss some of the most common choices and models. For a more comprehensive review, we refer the interested reader to the books [26, 54], the surveys [53, 78, 93] and [75] that provide thorough overviews of the advances of the past decades in deformable registration.

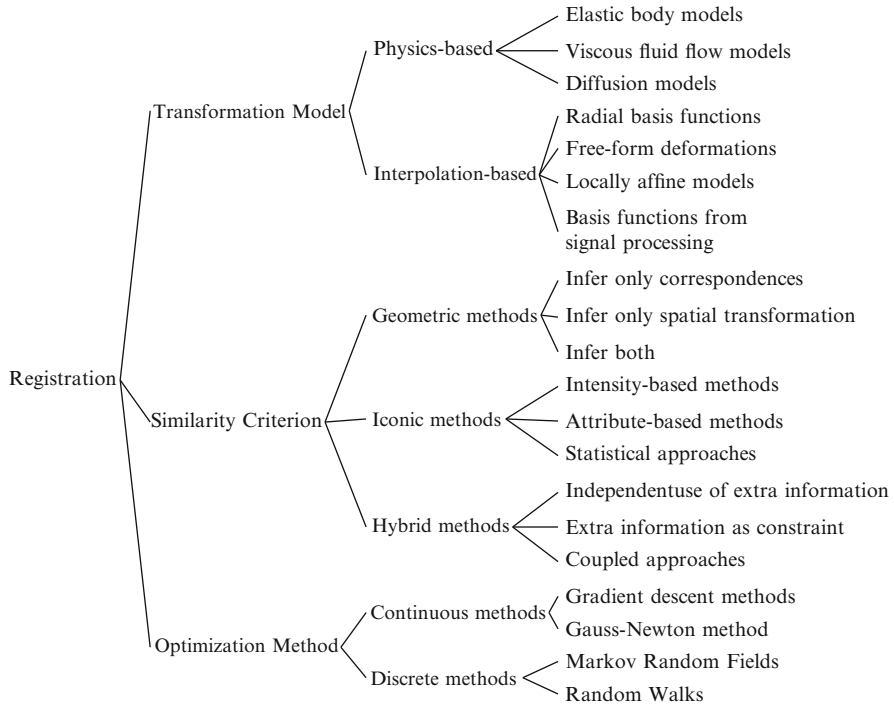


Fig. 1 Typical components of registration algorithms

The choice of the transformation model is usually dictated by the application at hand and is related to the nature of the deformation to be recovered. High-dimensional nonlinear models are necessary to cope with highly variable soft tissue, while low degrees of freedom models can represent the mapping between rigid bone structures. It is important to note that increasing the degrees of freedom of the model, and thus enriching its descriptive power, often comes at the cost of increased computational burden.

Several transformation models have been introduced in medical imaging for non-rigid alignment. These models can be coarsely classified into two categories (see Fig. 1 [30,75]): i) models derived from physical models, and ii) models derived from the interpolation theory or geometric models. Among the most prominent choices of the first class, one may cite elastic [19, 20], fluid [14, 18] or diffusion models [22, 80, 85]. Whereas, the second class comprises radial basis functions [10, 67], free-form deformations [68, 69], locally affine [55] and poly-affine models [2], or models parametrized by Fourier [1, 4] or Wavelet basis functions [87].

The similarity criterion quantifies the degree of alignment between the images. Registration methods can be classified into three categories (see Fig. 1) depending on the type of information that is utilized by the similarity criterion: i) geometric

registration (*a.k.a.* landmark/feature-based registration); ii) iconic registration (*a.k.a.* voxel-wise registration); and iii) hybrid registration.

Geometric registration aims to align meaningful anatomical locations or salient landmarks, which are either automatically extracted from the images [51] or provided by an expert. Geometric information is typically represented as point-sets and registration is tackled by first estimating the point correspondences [43, 82] and then employing an interpolation strategy (*e.g.* thin-plate splines [10]) to determine a dense deformation field that will align the images. Alternatively, geometric methods may infer directly the transformation that aligns the images without explicitly estimating point correspondences. This is possible by representing geometric information either as probability distributions [23, 83] or through the use of signed distance transformations [32]. Last, there exist methods that opt to simultaneously solve for both the correspondences and the transformation [15].

Iconic methods employ a similarity criterion that takes into account the intensity information of all image elements. The difficulty of choosing an appropriate similarity criterion varies greatly depending on the problem. In the mono-modal case, where both images are acquired using the same device and one can assume that the intensity profiles for the two images differ only by Gaussian noise, the use of sum of squared differences can be sufficient. Nonetheless, in the multi-modal case, where images from different modalities are involved, the criterion should be able to account for the different principles behind the acquisition protocols and capture the relation between the distinct intensity profiles. Towards this end, criteria based on statistics and information theory have been proposed. Examples include correlation ratio [65], mutual-information [52, 86] and Kullback-Leibler divergence [16]. Last, attribute-based methods that construct rich descriptions by summarizing intensity information over local regions have been proposed for both mono-modal and multi-modal registration [48, 62, 74].

Hybrid methods opt to exploit both iconic and geometric information in an effort to leverage their complementary nature towards more robust and accurate registration. Depending on how one combines the two types of information, three subclasses can be distinguished. In the first case, geometric information is used to initialize the alignment, while intensity-based volumetric registration refines the results [35, 64]. In the second case, geometric information can be used to provide additional constraints that are taken into account during iconic registration [27, 29]. In the third case, iconic and geometric information are integrated in a single objective function that allows for the simultaneous solution of both problems [11, 25, 76].

Once the transformation model and a suitable similarity criterion have been defined, an optimization method is used in order to infer the optimal set of parameters by maximizing the alignment of the two images. Solving for the optimal parameters is particularly challenging in the case of image registration. The reason behind this lies in the fact that image registration is, in general, an ill-posed problem and the associated objective functions are typically non-linear and non-convex. The optimization methods that are typically used in image registration fall under the umbrella of either continuous or discrete methods.

Typically, continuous optimization methods are constrained to problems where the variables take real values and the objective function is differentiable. This type of problems are common in image registration. As a consequence, these methods (typically gradient descent approaches) have been widely used in image registration [8, 69] because of the fact that they are rather intuitive and easy to implement. Moreover, they can handle a wide class of objective functions allowing for complex modeling assumptions regarding the transformation model. Nonetheless, they are often sensitive to the initial conditions, while being non-modular with respect to the similarity criterion and the transformation model. What is more, they are often computationally inefficient [24].

On the other hand, discrete optimization methods tackle problems where the variables take discrete values. Discrete optimization methods based on the Markov Random Field theory have been recently investigated in the context of image registration [24, 25]. Discrete optimization methods are constrained by limited precision due to the necessary quantization of the solution space. Moreover, they can not efficiently model complex variable interactions due to increased difficulty in inference. However, recent advances in higher-order inference methods have allowed the modeling of more sophisticated regularization priors [42]. More importantly, discrete optimization methods are versatile and can handle a wide range of similarity metrics (including non-differentiable ones). What is more, they are more robust to the initial conditions due to the global search they perform, while often converging faster than continuous methods.

In this chapter, we review the application of Markov Random Fields (MRFs) in deformable image registration. We explain in detail how one can map image registration from the continuous domain to discrete graph structures. We first present graph-based deformable registration in the case of iconic registration and show how one can encode intensity-based and statistical approaches. We then present discrete attribute-based registration methods and complete the presentation by describing MRF models for geometric and hybrid registration. Throughout this chapter, we discuss the underlying assumptions as well as implementation details. Experimental results that demonstrate the value of graph-based registration are given at the end of every section.

2 Graph-based Iconic Deformable Registration

In this chapter, we focus on pairwise deformable registration. The two images are usually termed as source (or moving) and target (or fixed) images, respectively. The source image is denoted by $S : \Omega_S \subset \mathbb{R}^d \mapsto \mathbb{R}$, while the target image by $T : \Omega_T \subset \mathbb{R}^d \mapsto \mathbb{R}$, $d = \{2, 3\}$. Ω_S and Ω_T denote the image domain for the source and target images, respectively. The source image undergoes a transformation $\mathcal{T} : \Omega_S \mapsto \Omega_T$.

Image registration aims to estimate the transformation \mathcal{T} such that the two images get aligned. This is typically achieved by means of an energy minimization problem:

$$\arg \min_{\boldsymbol{\theta}} \mathcal{M}(T, S \circ \mathcal{T}(\boldsymbol{\theta})) + \mathcal{R}(\mathcal{T}(\boldsymbol{\theta})). \quad (1)$$

Thus, the objective function comprises two terms. The first term, \mathcal{M} , quantifies the level of alignment between a target image T and a source image S under the influence of the transformation \mathcal{T} parametrized by $\boldsymbol{\theta}$. The second term, \mathcal{R} , regularizes the transformation and accounts for the ill-posedness of the problem. In general, the transformation at every position $\mathbf{x} \in \Omega$ (Ω depicting the image domain) is given as $\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$ where \mathbf{u} is the deformation field.

The previous minimization problem can be solved by adopting either continuous or discrete optimization methods. In this chapter, we focus on the application of discrete methods that exploit Markov Random Field theory.

2.1 Markov Random Fields

In discrete optimization settings, the variables take discrete values and the optimization is formulated as a discrete labeling problem where one searches to assign a label to each variable such that the objective function is minimized. Such problems can be elegantly expressed in the language of discrete Markov Random Field theory.

An MRF is a probabilistic model that can be represented by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set of vertices \mathcal{V} encodes the random variables, which take values from a discrete set \mathcal{L} . The interactions between the variables are encoded by the set of edges \mathcal{E} . The goal is to estimate the optimal label assignment by minimizing an energy of the form:

$$E_{MRF} = \sum_{p \in \mathcal{V}} \mathcal{U}_p(l_p) + \sum_{pq \in \mathcal{E}} \mathcal{P}_{pq}(l_p, l_q). \quad (2)$$

The MRF energy also comprises two terms. The first term is the sum of all unary potentials \mathcal{U}_p of the nodes $p \in \mathcal{V}$. This term typically corresponds to the data term since the unary terms are usually used to encode data likelihoods. The second term comprises the pairwise potentials \mathcal{P}_{pq} modeled by the edges connecting nodes p and q . The pairwise potentials usually act as regularizers penalizing disagreements in the label assignment of tightly related variables.

Many algorithms have been proposed in order to perform inference in the case of discrete MRFs. In the works that are presented in this chapter, the fast-PD¹ algorithm [39, 40] has been used to estimate the optimal labeling. The main

¹Fast-PD is available at <http://cvc-komodakis.centrale-ponts.fr/>.

motivation behind this choice is its great computational efficiency. Moreover, the fast-PD algorithm is appropriate since it can handle a wide-class of MRF models allowing us to use different smoothness penalty functions and has good optimality guarantees.

In the continuation of this section, we detail how deformable registration is formulated in terms of Markov Random Fields. First, however, the discrete formulation requires a decomposition of the continuous problem into discrete entities. This is described below.

2.2 Decomposition into Discrete Deformation Elements

Without loss of generality, let us consider a grid-based deformation model that combines low degrees of freedom with smooth local deformations. Let us consider a set of k control points distributed along the image domain using a uniform grid pattern. Furthermore, let k be much smaller than the number of image points. One can then deform the embedded image by manipulating the grid of control points. The dense displacement field is defined as a linear combination of the control point displacements $\mathbf{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_k\}$, with $\mathbf{d}_i \in \mathbb{R}^d$, as:

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^k \omega_i(\mathbf{x}) \mathbf{d}_i, \quad (3)$$

and the transformation \mathcal{T} becomes:

$$\mathcal{T}(\mathbf{x}) = \mathbf{x} + \sum_{i=1}^k \omega_i(\mathbf{x}) \mathbf{d}_i. \quad (4)$$

ω_i corresponds to an interpolation or weighting function which determines the influence of a control point i to the image point x – the closer the image point the higher the influence of the control point. The actual displacement of an image point is then computed via a weighted sum of control point displacements. A dense deformation of the image can thus be achieved by manipulating these few control points.

The free-form deformation is a typical choice for such a representation [71]. This model employs a weighting scheme that is based on cubic B-splines and has found many applications in medical image registration [69] due to its efficiency and the local support of the control points. We also employ this model. Nonetheless, let us note that the discrete deformable registration framework is modular with respect to the interpolation scheme and one may use this preferred strategy.

The parametrization of the deformation field leads naturally to the definition of a set of discrete deformation elements. Instead of seeking a displacement vector for every single image point, now, only the displacement vectors for the control

points need to be sought. If we take them into consideration, the matching term (see Eq. (1)) can be rewritten as:

$$\mathcal{M}(S \circ \mathcal{T}, T) = \frac{1}{k} \sum_{i=1}^k \int_{\Omega_S} \hat{\omega}_i(\mathbf{x}) \rho(S \circ \mathcal{T}(\mathbf{x}), T(\mathbf{x})) d\mathbf{x}, \quad (5)$$

where $\hat{\omega}_i$ are weighting functions similar to the ones in Eq. (4) and ρ denotes a similarity criterion.

Here, the weightings determine the influence or contribution of an image point x onto the (local) matching term of individual control points. Only image points in the vicinity of a control point are considered for the evaluation of the intensity-based similarity measure with respect to the displacement of this particular control point. This is in line with the local support that a control point has on the deformation. The previous is valid when point-wise similarity criteria are considered. When a criterion based on statistics or information theory is used, a different definition of $\hat{\omega}_i$ is adopted,

$$\hat{\omega}_i(\mathbf{x}) = \begin{cases} 1, & \text{if } \omega_i(\mathbf{x}) \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Thus, in both cases the criterion is evaluated on a patch. The only difference is that the patch is weighted in the first case. These local evaluations enhance the robustness of the algorithm to local intensity changes. Moreover, they allow for computationally efficient schemes.

The regularization term of the deformable registration energy (Eq. (1)) can also be expressed on the basis of the set of control points as:

$$\mathcal{R} = \frac{1}{k} \sum_{i=1}^k \int_{\Omega_S} \hat{\omega}_i(\mathbf{x}) \psi(\mathcal{T}(\mathbf{x})) d\mathbf{x}, \quad (7)$$

where ψ is a function that promotes desirable properties of the dense deformation field such as the smoothness and topology preservation.

2.3 Markov Random Field Registration Energy

Having identified the discrete deformation elements of our problem, we need to map them to MRF entities, *i.e.*, the graph vertices, the edges, the set of labels, and the potential functions.

Let \mathcal{G}_{ico} denote the graph that represents our problem. In this case, the random variables of interest are the control point displacement updates. Thus, the set of vertices \mathcal{V}_{ico} is used to encode them, *i.e.*, $|\mathcal{V}_{\text{ico}}| = |\Delta\mathbf{D}| = k$. Moreover, assigning

a label $l_p \in \mathcal{L}_{\text{ico}}$ to a node $p \in \mathcal{V}_{\text{ico}}$ is equivalent to displacing the corresponding control point p by an update $\Delta \mathbf{d}_p$, or $l_p \equiv \Delta \mathbf{d}_p$. In other words, the label set for this set of variable is a quantized version of the displacement space ($L_{\text{ico}} \subset \mathbb{R}^d$). The edge system \mathcal{E}_{ico} is constructed by following either a 6-connected neighborhood system in the 3D case, or a 4-connected system in the 2D case. The edge system follows the grid structure of the transformation model.

According to Eq. (5) we define the unary potentials as:

$$\mathcal{U}_{\text{ico},p}(l_p) = \int_{\Omega_S} \hat{\omega}_p(\mathbf{x}) \rho(S \circ \mathcal{F}_{\text{ico},l_p}(\mathbf{x}), T(\mathbf{x})) d\mathbf{x}, \quad (8)$$

where $\mathcal{F}_{\text{ico},l_p}$ denotes the transformation where a control point p has been updated by l_p . Region-based and statistical measures are again encoded in a similar way based on a local evaluation of the similarity measure.

Conditional independence is assumed between the random variables. As a consequence, the unary potential that constitutes the matching term can only be an approximation to the real matching energy. That is because the image deformation, and thus the local similarity measure, depends on more than one control point since their influence areas do overlap. Still, the above approximation yields very accurate registration as demonstrated by the experimental validation results that are reported in latter sections (Sect. 2.4, Sect. 3.3 and Sect. 4.3). Furthermore, it allows an extremely efficient approximation scheme which can be easily adapted for parallel architectures yielding extremely fast cost evaluations.

Actually, the previous approximation results in a weighted block matching strategy encoded on the unary potentials. The smoothness of the transformation derives from the explicit regularization constraints encoded by the pairwise potentials and the implicit smoothness stemming from the interpolation strategy.

The evaluation of the unary potentials for a label $l \in \mathcal{L}_{\text{ico}}$ corresponding to an update $\Delta \mathbf{d}$ can be efficiently performed as follows. First, a global translation according to the update $\Delta \mathbf{d}$ is applied to the whole image, and then the unary potentials for this label and for all control points are calculated simultaneously. This results in an one pass through the image to calculate the cost and distribute the local energies to the control points. The constrained transformation in the unary potentials is then simply defined as $\mathcal{F}_{\text{ico},l_p}(\mathbf{x}) = \mathcal{F}_{\text{ico}}(\mathbf{x}) + l_p$, where $\mathcal{F}_{\text{ico}}(\mathbf{x})$ is the current or initial estimate of the transformation.

The regularization term defined in Eq. (7) could be defined as well in the above manner. However, this is not very efficient since the penalties need to be computed on the dense field for every variable and every label. If we consider an elastic-like regularization, we can employ a very efficient discrete approximation of this term based on pairwise potentials as:

$$\mathcal{P}_{\text{ico, elastic},pq}(l_p, l_q) = \frac{\|(\mathbf{d}_p + \Delta \mathbf{d}_p) - (\mathbf{d}_q + \Delta \mathbf{d}_q)\|}{\|\mathbf{p} - \mathbf{q}\|}. \quad (9)$$

The pairwise potentials penalize deviations of displacements of neighboring control points $(p, q) \in \mathcal{E}_{ico}$ which is an approximation to penalizing the first derivatives of the transformation. Recall that $l_p \equiv \Delta \mathbf{d}_p$. Note, we can also remove the current displacements \mathbf{d}_p and \mathbf{d}_q from the above definition yielding a term that only penalizes the updates on the deformation. This would change the behavior of the energy from an elastic-like to a fluid-like regularization.

Let us detail how the label set \mathcal{L}_{ico} is constructed since that entails an important accuracy-efficiency trade-off. The smaller the set of labels, the more efficient is the inference. However, few labels result in a decrease of the accuracy of the registration. This is due to the fact that the registration accuracy is bounded by the range of deformations covered in the set of labels. As a consequence, it is reasonable to assume that the registration result is sub-optimal. In order to strike a satisfactory balance between accuracy and efficiency, we opt for an *iterative labeling* strategy combined with a *search space refinement* one. At each iteration, the optimal labeling is computed yielding an update on the transformation, *i.e.* $l_p \equiv \Delta \mathbf{d}_p$. This update is applied to the current estimate, and the subsequent iteration continues the registration based on the updated transformation and a refined label set. Thus, the error induced by the approximation stays small and incorrect matches can be corrected in the next iteration. Furthermore, the overall domain of possible deformations is rather bounded by the number of iterations and not by the set of finite labels.

The iterative labeling allows us to keep the label set quite small. The refinement strategy on the search space is rather intuitive. In the beginning we aim to recover large deformations and as we iterate, finer deformations will be added refining the solution. In each iteration, a sparse sampling with a fixed number of samples s is employed. The total number of labels in each iteration is then $|\mathcal{L}_{ico}| = g \cdot s + 1$ including the zero-displacement and g is the number of sampling directions. We uniformly sample displacements along certain directions up to a maximum displacement magnitude \mathbf{d}_{max} . Initially, the maximum displacement corresponds to our estimation of the larger deformation to be recovered. In the subsequent iterations, it is decreased by a user-specified factor $0 < f < 1$ limiting and refining the search space.

The number and orientation of the sampling directions g depend on the dimensionality of the registration. One possibility is to sample just along the main coordinate axes, *i.e.* in positive and negative direction of the x -, y -, and z -axis (in case of $3D$). Additionally, we can add samples for instance along diagonal axes. In $2D$ we commonly prefer a star-shape sampling, which turns out to be a good compromise between the number of samples and the sampling density. In our experiments we found that also very sparse samplings (*e.g.*, just along the main axes) gives very accurate registration results but might increase the total number of iterations that are needed until convergence. However, a single iteration is much faster to compute when the label set is small. In all our experiments we find that small label sets provide an excellent performance in terms of computational speed and registration accuracy.

The explicit control that one has over the creation of the label set \mathcal{L} enables us to impose desirable properties on the obtained solution without further modifying the discrete registration model. Two interesting properties that can be easily enforced by adapting appropriately the discrete solution space are *diffeomorphisms* and *symmetry*. Both properties are of particular interest in medical imaging and have been the focus of the work of many researchers.

Diffeomorphic transformations preserve topology and both they and their inverse are differentiable. These transformations are of interest in the field of computational neuroanatomy. Moreover, the resulting deformation fields are, in general, more physically plausible since foldings, which would disrupt topology, are avoided. As a consequence, many diffeomorphic registration algorithms have been proposed [3, 5, 8, 68, 85].

In this discrete setting, it is straightforward to guarantee a diffeomorphic result through the creation of the label set. By bounding the maximum sampled displacement by 0.4 times the deformation grid spacing, the resulting deformation is guaranteed to be diffeomorphic [68].

The majority of image registration algorithms are asymmetric. As a consequence, when interchanging the order of input images, the registration algorithm does not estimate the inverse transformation. This asymmetry introduces undesirable bias upon any statistical analysis that follows registration because the registration result depends on the choice of the target domain. Symmetric algorithms have been proposed in order to tackle this shortcoming [5, 13, 56, 79, 84].

Symmetry can also be introduced in graph-based deformable registration in a straightforward manner [77]. This is achieved by estimating two transformations, \mathcal{T}^f and \mathcal{T}^b , that deform both the source and the target images towards a common domain that is constrained to be equidistant from the two image domains. In order for this to be true, the transformations, or equivalently the two update deformation fields, should sum up to zero. If one assumes a transformation model that consists of two isomorphic deformation grids, this constraint translates to ensuring that the displacement updates of corresponding control points in the two grids sum to zero and can be simply mapped to discrete elements.

The satisfaction of the previous constraint can be easily guaranteed in a discrete setting by appropriately constructing the label set. More specifically, by letting the labels index pairs of displacement updates (one for each deformation field) that sum to zero, *i.e.* $l_p \equiv \{\Delta \mathbf{d}_p^f, -\Delta \mathbf{d}_p^b\}$. The extension of the unary terms is also straightforward, while the pairwise potentials and the graph construction are the same.

2.4 Experimental Validation

In this section, we present experimental results for the graph-based symmetric registration in 3D brain registration. The data set consists of 18 T1-weighted brain volumes that have been positionally normalized into the Talairach orientation

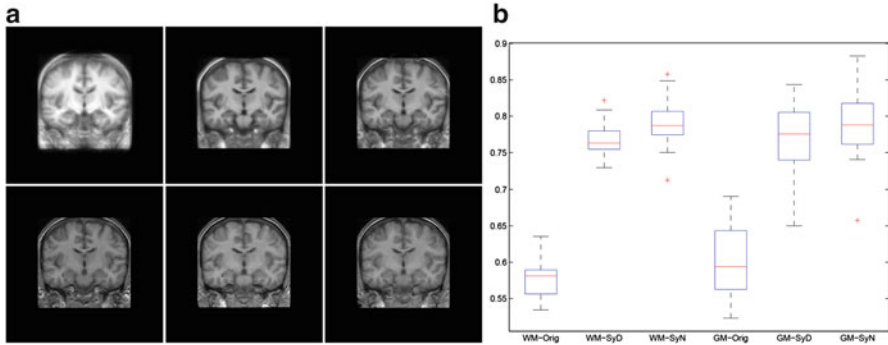


Fig. 2 a) In the first row, from left to right, the mean intensity image is depicted for the data set, after the graph-based symmetric registration method and after [5]. In the second row, from left to right, the target image is shown as well as a typical deformed image for the graph-based symmetric registration method and [5]. For all cases, the central slice is depicted. b) Boxplots for the DICE criterion initially, with the graph-based symmetric registration method and with [5]. On the left, the results for the WM. On the right, the results for the GM. The figure is reprinted from [77]

(rotation only). The MR brain data set along with manual segmentations was provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available online². The data set was rescaled and resampled so that all images have a size equal to $256 \times 256 \times 128$ and a physical resolution of approximately $0.9375 \times 0.9375 \times 1.5000$ mm.

This set of experiments is based on intensity-based similarity metrics (for results using attribute-based similarity metrics, we refer the reader to the next section of this chapter). The results are compared with a symmetric registration method based on continuous optimization [5] that is considered to be the state of the art in continuous deformable registration [38]. Both methods use Normalized Cross Correlation as the similarity criterion.

A multiresolution scheme was used in order to harness the computational burden. A three-level image pyramid was considered while a deformation grid of four different resolutions was employed. The two finest grid resolutions operated on the finest image resolution. The two coarsest operated on the respective coarse image representations. The initial grid spacing was set to 40 mm resulting in a deformation grid of size $7 \times 7 \times 6$. The size of the grid was doubled at each finer resolution. A number of 90 labels, 30 along each principal axis, were used. The maximum displacement indexed by a label was bounded to 0.4 times the grid spacing. The pairwise potentials were weighted by a factor of 0.1.

²<http://www.cma.mgh.harvard.edu/ibsr/data.html>

The qualitative results (sharp mean and deformed image) suggest that both methods successfully registered the images to the template domain. The results of [5] seem to have produced more aggressive deformation fields that have resulted to some unrealistic deformations in the top of the brain and can also be observed in the borders between white matter (WM) and gray matter (GM). This aggressive registration has also resulted in slightly more increased DICE coefficients for WM and GM. However, the results reported for the graph-based registration method were obtained in 10 min. On the contrary, 1 hour was necessary to register the images with [7] approximately. This important difference in the computational efficiency between the two methods can outweigh the slight difference in the quality of the solution in practice.

3 Graph-based Attribute-Based Deformable Registration

In the previous section, we studied the application of intensity-based deformable registration methods that involve voxel-wise and statistical similarity criteria. While these criteria are easy to compute and widely used, they suffer from certain shortcomings. First, they often have difficulties to reflect the underlying anatomy because pixels belonging to the same anatomical structure are often assigned different intensity values due to variabilities arising from scanners, imaging protocols, noise, partial volume effects, contrast differences, and image inhomogeneities. Moreover, single intensities are not informative enough to uniquely characterize image elements, and thus reliably guide image registration. For instance, hundreds of thousands of gray matter voxels in a brain image share similar intensities; but they belong to different anatomical structures. As a consequence, matching ambiguities arise in the matching between two images.

In order to reduce matching, one needs to characterize each voxel more distinctively. This may be achieved by creating richer high-dimensional descriptors of image elements that capture texture or geometric regional attributes. Therefore, attribute-based similarity criteria have been increasingly used in image registration. Typical examples include the use of geometric-moment-invariant (GMI) attributes coupled with tissue membership attributes and boundary/edge attributes [74], neighborhood intensity profile attributes [28], local frequency attributes [34, 49], local intensity histogram attributes [73, 88], geodesic intensity histogram attributes [44, 47] and scale-invariant attributes [81].

3.1 Gabor Attributes

The versatility of graph-based deformable registration models allows the seamless integration of any of the previous attribute-based similarity criteria. Nonetheless, the previous approaches involve features that are application-specific and fail to generalize to other applications, or require sophisticated pre-processing steps

(e.g. segmentation). As a consequence, it is important to appropriately choose the attribute-based description so that, when coupled with the highly modular discrete approaches, a general-purpose registration method is possible.

Gabor-attributes, which involve image convolution with Gaussian filters at multiple scales and orientations, present an interesting choice for general-purpose deformable registration. The reason is threefold. First, all anatomical images have texture information, at some some scale and orientation, reflecting the underlying geometric and anatomical characteristics. As a results, Gabor features that are able to capture this information can be, and have been, applied in a variety of studies. Second, Gabor filters are able to capture edge information that is relatively encoded by various image modalities, thus making them suitable for both mono- and multi-modal registration tasks. Third, their multi-scale and multi-orientation nature render image elements more distinctive and better identifiable for establishing correspondences. For example, the scale information helps differentiate voxels that are the center of a small and a bigger plate, respectively. The orientation information can help distinguish, for example, a voxel on a left-facing edge from a voxel on a right-facing edge. Moreover, it is also possible to automatically select a subset of Gabor attributes such that the information redundancy is reduced and the distinctiveness of the descriptor is increased.

The effect of characterizing voxels using Gabor attributes (with and without optimal Gabor attribute subset selection) is presented in Fig. 3 (reprinted from [62]). These effects are contrasted to the effect of using only intensities and using Gray-Level-Cooccurrence-Matrix (GLCM) texture attributes through the use of similarity maps between voxels from the source image (labeled under crosses) and all voxels in the target image. The similarity between two voxels, \mathbf{x} in the source image and \mathbf{y} in the target image, was defined as $\text{sim}(\mathbf{x}, \mathbf{y}) = \frac{1}{1 + \|A(\mathbf{x}) - A(\mathbf{y})\|^2}$, with $A(\cdot)$ being the attribute vector at each voxel. This similarity ranged from 0 (when the attributes between two voxels differed infinitely) and 1 (when the attributes between two voxels were identical). This figure shows that, as one replaced the intensity-based similarity to (optimal-)attribute-based similarities, even very ordinal voxels under the blue crosses were distinctively characterized or better localized in the space, therefore we only needed to search for their corresponding voxels within a much smaller range in the target image, largely removing matching ambiguities.

Let us detail in the next section how one can introduce Gabor-based attributes in the case of graph-based deformable registration [62]. More specifically, let us detail how the Markov Random Field energy (see Eq. (2)) changed in this regard.

3.2 Markov Random Field Registration Energy

The ease with which one can adopt attribute-based similarity criteria in the case of graph-based formulations for deformable registration is evidence of their high versatility and modularity. The key elements of the graphical model (*i.e.*, graph

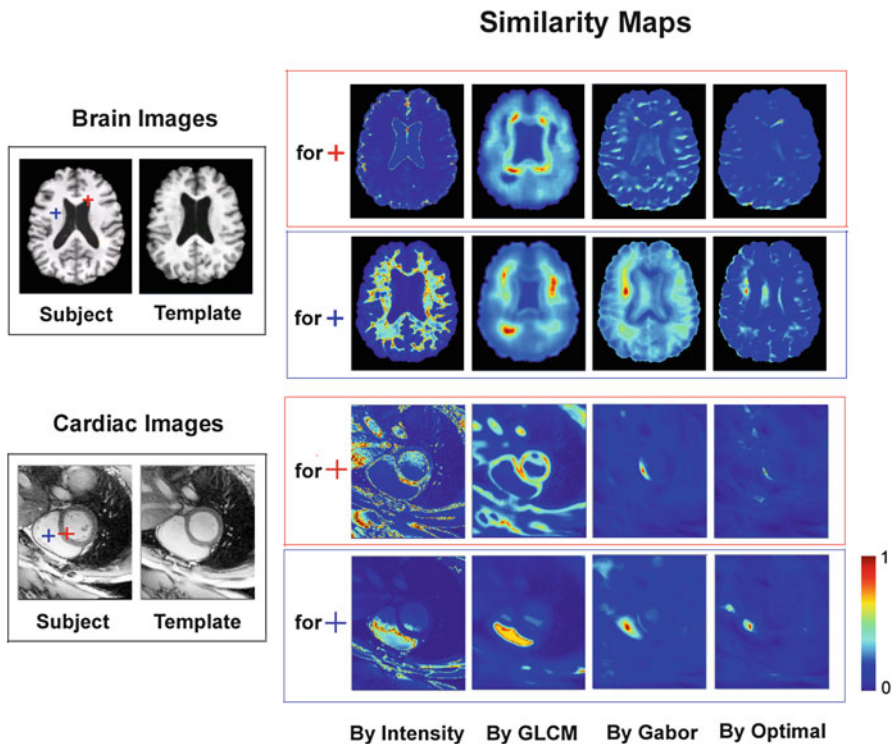


Fig. 3 The similarity maps between special/ordinary voxels (labeled by red/blue crosses) in the source (a.k.a, subject) images and all voxels in the target (a.k.a, template) images. As correspondences were sought based on voxel similarities (subject to spatial smoothness constraints), (optimal-)Gabor-attribute-based similarity maps returned a much smaller search range for correspondences. This figure is reprinted from [62]

construction, pairwise potentials, inference) need not change. One only needs to slightly change the definition of the unary potentials.

The unary potentials need only be modified in two regards: i) to evaluate the similarity criterion ρ over the attribute vectors $A(\cdot)$; and ii) to optionally, as suggested by [62], take into account a spatially-varying weighting parameter $ms(\mathbf{x})$, namely “mutual-saliency”, which automatically quantified the confidence of each voxel \mathbf{x} to establish reliable correspondences across images. Therefore, the modified unary potentials are defined as:

$$\mathcal{U}_{ico,p}(l_p) = \int_{\Omega_S} ms(\mathbf{x}) \cdot \hat{\omega}_p(\mathbf{x}) \cdot \rho(A_S \circ \mathcal{F}_{ico,l_p}(\mathbf{x}), A_T(\mathbf{x})) d\mathbf{x}. \quad (10)$$

	FFD + GradDes		FFD + DisOpt	
	AM w/ MS	AM w/o MS	AM w/ MS	AM w/o MS
Brain images ($256 \times 256 \times 171$)	534.67	268.82	115.78	36.49
Prostate images ($256 \times 256 \times 34$)	245.46	104.77	44.72	15.41
Cardiac images ($256 \times 256 \times 20$)	181.54	88.77	39.07	13.61

Fig. 4 The computational times (in minutes) when combining the MRF registration formulation with the discrete optimization strategy versus with the traditional gradient descent optimization strategy. The discrete optimization strategy on the MRF registration formulation helped significantly reduce the computational time. AM refers to attribute matching; MS refers to mutual-saliency weighting, which is a second component in DRAMMS but was not described in full detail in this section; basically it is an automatically computed weighting for adaptively utilizing voxels based on how much confidence we have for those voxels to find correspondences across images. FFD is the free form deformation transformation model as used in the MRF registration formulation. And DisOpt and GradDes are the discrete optimization and gradient descent optimization strategies. This figure is reprinted from [62]

3.3 Experimental Validation

In this section we present results obtained with an attributed-based discrete deformable registration termed DRAMMS (Deformable Registration via Attribute Matching and Mutual-Saliency) [62]. The presented results demonstrate the advantageous computational efficiency of graph-based registration method in comparison to the traditional gradient descent optimization strategy. Moreover, the results demonstrate the generality, accuracy and robustness of coupling attributed-based similarity criteria with graph-based formulations.

As far as the computational efficiency is concerned, Fig. 4 summarizes the computational time that is required to register brain, prostate, and cardiac images using a gradient descent optimization strategy and a discrete optimization strategy [39, 40], respectively. The discrete approach requires significantly reduced computational time.

In the second part of this section, we report results for DRAMMS in two different cases: i) skull-stripped brain MR images; and ii) brain MR images from the large-scale, multi-institutional, with-skull ADNI database.

In the first case, DRAMMS was compared to 11 other popular and publicly-available registration tools, all used with the optimized parameters as reported in [38] whenever applicable. In the public NIREP dataset containing T1-weighted MR images ($256 \times 300 \times 256$ voxels and $1.0 \times 1.0 \times 1.0 \text{ mm}^3/\text{voxel}$) of 16 healthy subjects, each registration method was applied to all the possible 210 pair-wise registrations, leading to 2,520 registrations in total. DRAMMS had been shown

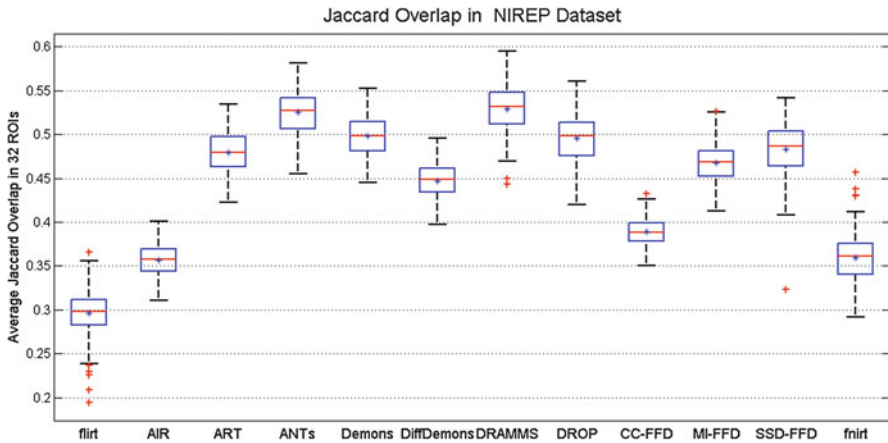


Fig. 5 The average Jaccard overlap among all ROIs in all possible pair-wise registrations within the NIREP database, for different registration tools. Reprinted from [57]

to yield the highest average Jaccard overlap among 32 regions-of-interest (ROIs) annotated by human experts, indicating the high accuracy (Fig. 5). Such a trend had also been observed in several other databases containing skull-stripped brain MR images from healthy subjects [57].

In the second case, DRAMMS was validated using brain MR images from the ADNI study. This study presents particular challenges because it contains data acquired at different imaging vendors/centers, and some of those data contain regions affected by pathologies. In Fig. 6 (re-printed from [57]) one can observe that DRAMMS can align largely variable ventricles, whereas other registration tools encountered great challenges. This is characteristic of the accuracy and robustness of the attribute-based discrete deformable registration method.

These results emphasize the generality, accuracy and robustness of the attribute-based discrete deformable registration. Because of these characteristics and its public availability³, DRAMMS has found applications in numerous translational studies including neuro-degenerative studies [17, 41, 72, 90], neuro-developmental ones [21, 33, 60, 70] as well as oncology studies [6, 59]. These applications underline the versatility of combining attribute-based similarity criteria with graph-based formulations.

³DRAMMS is available at <http://www.nitrc.org/projects/dramms/>.

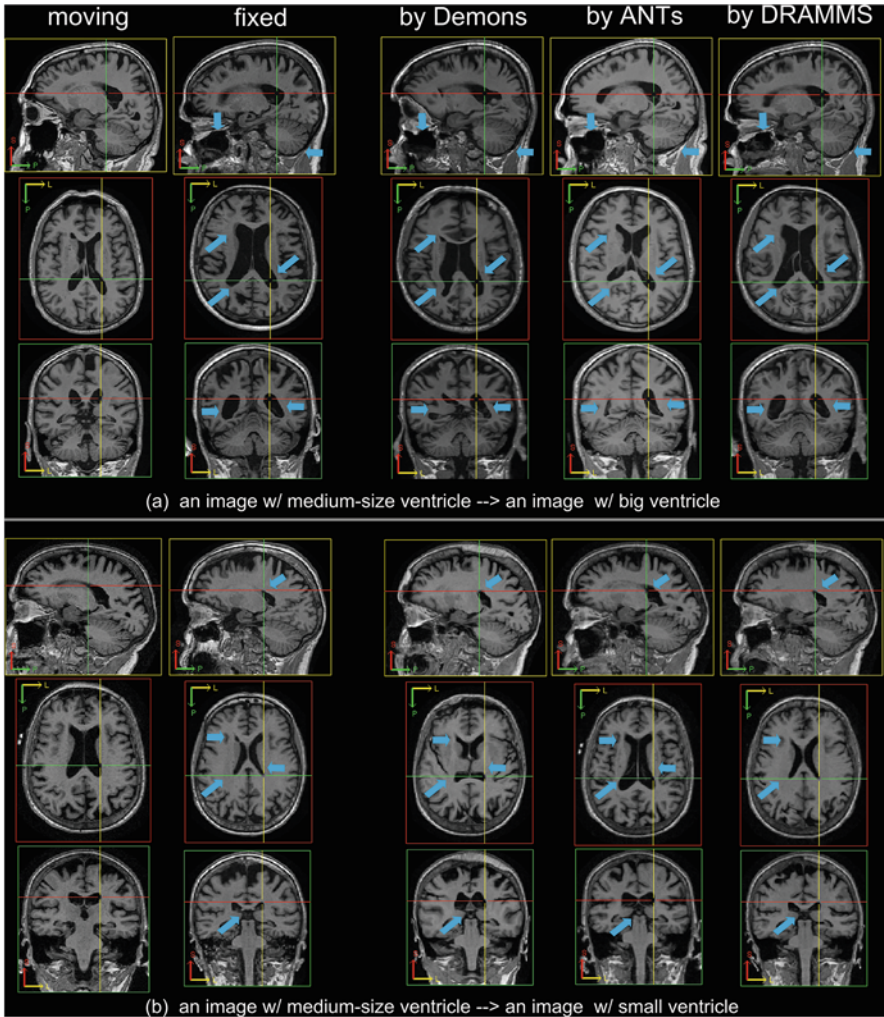


Fig. 6 Example registration results between subjects in the multi-site Alzheimer's Disease Neuroimaging Initiative (ADNI) database, by different registration methods. Blue arrows point out regions where the results from various registration methods differ. Reprinted from [57]

4 Graph-based Geometric and Hybrid Deformable Registration

The previous two sections presented MRF-based iconic (*a.k.a.* voxel-wise) registration using intensity- and attribute-based similarities. Typically, iconic approaches evaluate the similarity criteria over the whole image domain and have the potential to better quantify and represent the accuracy of the estimated dense deformation

field, albeit at an important computational cost. Nonetheless, iconic approaches do not explicitly take into account salient image points, failing to fully exploit image information. Moreover, the performance of iconic methods, especially methods based on continuous optimization, is greatly influenced by the initial conditions.

On the other hand, geometric methods utilize only a sparse subset of image elements that correspond to salient geometry or anatomy. Exploiting relevant information results in increased robustness. Nonetheless, the quality of the estimated deformation field is high only on the vicinity of the landmarks.

Hybrid registration methods exploit both types of information towards bridging the gap between the two basic classes of registration and enjoying the advantages of both worlds. Iconic and geometric information are integrated in a unified objective function and the solutions of the two problems satisfy each other. In this setting, iconic methods may profit from geometric information in the cases they encounter difficulties arising, for example, from large deformations (*e.g.*, the largely different ventricle size in Alzheimer's Disease population), or from missing correspondences such as the existence of pathologies. At the same time, geometric correspondences can be refined based on the iconic information that is available throughout the image domain.

In this section, we consecutively study the graph-based formulation of geometric and hybrid deformable registration. Similar to the previous sections, we first study the two problems in their continuous form and show how they can be decomposed in discrete entities. Then, we detail the graph-based formulation and present experimental results.

4.1 Decomposition into Discrete Deformation Elements

4.1.1 Geometric Registration

A prerequisite for geometric registration is the availability of landmarks that encode salient geometry or anatomy. Landmarks can be annotated by experts, or, to reduce intra-/inter-expert variability, by (semi-)automated methods. The latter is an open problem and an active topic of research.

Automatic approaches to detect landmarks include, but are not limited to, edge detection [31, 66], contour delineation [45], anatomical structure segmentation [9, 12], scale space analysis [36, 46, 63], and feature transformation (*e.g.*, SIFT [37, 50, 89], SURF [7, 92]). In [61, 91], for example, the authors used Laplacian operations to search for blob-like structures, and used the centers of the blobs as landmarks. In [58], the authors used regional centers or edges at various scales and orientations as landmarks, which were of strong response to Gabor filters. While a detailed survey of landmark detection is outside the scope of this section, we want to emphasize that the described graph-based formulation can seamlessly integrate landmark information coming from any algorithm or expert.

Given two sets of landmarks K ($\kappa \in K$) and Λ ($\lambda \in \Lambda$), one aims to estimate the transformation \mathcal{T}_{geo} that will bring them into correspondence by minimizing an objective function of the form of Eq. (1). More specifically, the goal is to bring every landmark belonging to the set K as close as possible to the landmark in the set Λ that is most similar to it. In other words, the matching term is expressed as:

$$\mathcal{M}_{\text{geo}}(K \circ \mathcal{T}_{\text{geo}}, \Lambda) = \frac{1}{n} \sum_{i=1}^n \delta(\mathcal{T}_{\text{geo}}(\kappa_i), \tilde{\lambda}_i) \quad (11)$$

where δ measures the Euclidean distance between two landmark positions, and

$$\tilde{\lambda}_i = \arg \min_{\lambda_j} \rho(\mathcal{T}_{\text{geo}}(\kappa_i), \lambda_j). \quad (12)$$

Note that the Euclidean position of the landmarks λ and κ is denoted in bold.

As far as the regularization term \mathcal{R}_{geo} is concerned, it aims to preserve the smoothness of the transformation. More specifically, it aims to locally preserve the geometric distance between pairs of landmarks:

$$\mathcal{R}_{\text{geo}}(\mathcal{T}_{\text{geo}}) = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \|(\mathcal{T}_{\text{geo}}(\kappa_i) - \mathcal{T}_{\text{geo}}(\kappa_j)) - (\kappa_i - \kappa_j)\|. \quad (13)$$

This implies the assumption that a linear registration step that has accounted for differences in scales has been applied prior to the deformable registration.

An equivalent way of formulating the geometric registration problem consists of first pairing landmarks $\kappa \in K$ with the most similar in appearance landmarks $\lambda \in \Lambda$ and then pruning the available pairs by keeping only those that are geometrically consistent as quantified by the regularization term (Eq. (13)). Let us note that, in both cases, the problem is inherently discrete.

4.1.2 Hybrid Registration

As discussed in the introduction, there are various ways of integrating geometric and iconic information. The most interesting, and potentially more accurate, is the one that allows both problems to be solved at the same time through the optimization of a universal energy that enforces the separate solutions to agree. This is possible by combining the previous energy terms for the iconic and geometric problem along with a hybrid term that acts upon the separate solutions:

$$\mathcal{H}(\mathcal{T}_{\text{ico}}, \mathcal{T}_{\text{geo}}) = \frac{1}{n} \sum_{i=1}^n \|\mathcal{T}_{\text{ico}}(\kappa_i) - \mathcal{T}_{\text{geo}}(\kappa_i)\|. \quad (14)$$

Note that we only need to enforce the agreement of the two solutions in the landmark positions. If we now also consider a connection between control point displacements \mathbf{D} and landmark displacements, the previous relation can be rewritten as:

$$\mathcal{H}(\mathcal{T}_{\text{ico}}, \mathcal{T}_{\text{geo}}) = \frac{1}{n} \sum_{i=1}^n \|\kappa_i + \mathbf{u}_{\text{geo}}(\kappa_i) - \kappa_i - \sum_{j=1}^k \omega_j(\kappa_i) \mathbf{d}_j\|, \quad (15)$$

where $\mathbf{u}_{\text{geo}}(\kappa_i) = \widetilde{\lambda}_i - \kappa_i$, *i.e.* the displacement for the correspondence of the two landmarks κ_i and $\widetilde{\lambda}_i$. As a principle, we would like this displacement to be ideally equal to the one that is given as a linear combination of the displacements of the control points at the position of a landmark. However, we can relax the previous requirement in order to increase the computational efficiency of the method. If we apply the triangular inequality and exploit the fact that the coefficients ω_j are positive, the coupling constraint is redefined as:

$$\mathcal{H}(T_{\text{ico}}, T_{\text{geo}}) \leq \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k \omega_j(\kappa_i) \|\mathbf{u}_{\text{geo}}(\kappa_i) - \mathbf{d}_j\|. \quad (16)$$

The previous constraint comprises only pairwise interactions between discrete elements.

4.2 Markov Random Field Registration Energy

Having identified the discrete elements for both geometric and hybrid registration, let us map them to MRF entities.

4.2.1 Geometric Registration

Let us now introduce a second graph $\mathcal{G}_{\text{geo}} = (\mathcal{V}_{\text{geo}}, \mathcal{E}_{\text{geo}})$ for the geometric entities K, Λ . We recall that they are two sets of landmarks having different cardinalities and we seek the transformation which will bring each landmark into correspondence with the best candidate. Equivalently, we may state that we are trying to solve for the correspondence of each landmark, which naturally results in a set of sparse displacements.

The second graph consists of a set of vertices \mathcal{V}_{geo} corresponding to the set of landmarks extracted in the source image, *i.e.* $|\mathcal{V}_{\text{geo}}| = |K|$. A label assignment $l_p \in \mathcal{L}_{\text{geo}} := \Lambda$ (where $p \in \mathcal{V}_{\text{geo}}$) is equivalent to matching the landmark $\kappa_p \in K$ to a candidate point $l_p \equiv \lambda \in \Lambda$. Assigning a label l_p implicitly defines a displacement $\mathbf{u}_{\text{geo}, l_p}(\kappa_p) = \lambda - \kappa_p$, since κ_p is mapped on the landmark l_p .

According to Eq. (12), the unary potentials are defined as:

$$\mathcal{U}_{\text{geo},p}(l_p) = \varrho(\kappa_p, l_p). \quad (17)$$

The two different though equivalent ways to see the label assignment problem are depicted in the previous equation. Assigning a label l_p can be interpreted as applying a transformation $\mathcal{T}_{\text{geo},l_p} = \kappa_p + \mathbf{u}_{\text{geo},l_p}(\kappa_p)$ or stating that the landmark κ_p corresponds to the l_p . Contrary to the iconic case, the set of transformations that can be applied is specified by the candidate landmarks and is sparse in its nature.

There is a number of ways to define the dissimilarity function ϱ . One approach would be to consider neighborhood information. That can be easily done by evaluating the criterion over a patch centered around the landmarks,

$$\mathcal{U}_{\text{geo},p}(l_p) = \int_{\Omega_{S,p}} \varrho(S \circ \mathcal{T}_{\text{geo},l_p}(\mathbf{x}), T(\mathbf{x})) d\mathbf{x}, \quad (18)$$

where $\Omega_{S,p}$ denotes a patch around the point κ_p . Another approach is to exploit attribute-based descriptors and mutual saliency [58] and define the potential as:

$$\mathcal{U}_{\text{geo},p}(l_p) = \exp\left(-\frac{ms(\kappa_p, l_p) \cdot sim(\kappa_p, l_p)}{2\sigma^2}\right). \quad (19)$$

where σ is a scaling factor, estimated as the standard deviation of the mutual saliency values of all the candidate pairs.

The regularization term defined in Eq. (13) can be encoded by the edge system \mathcal{E}_{geo} of the graph. In this setting, the regularization term can be expressed as:

$$\mathcal{E}_{\text{geo},pq}(l_p, l_q) = \|(\mathcal{T}_{\text{geo},l_p}(\kappa_p) - \mathcal{T}_{\text{geo},l_q}(\kappa_q)) - (\kappa_p - \kappa_q)\|. \quad (20)$$

The pairwise potential will enforce an isometric constraint. Moreover, by considering the vector differences flipping of the point positions is penalized.

Last, it is interesting to note that the same graph \mathcal{G}_{geo} is able to encode both ways of formulating the geometric registration problem that were presented in Sect. 4.1.1. This model was presented in [58] and [76].

4.2.2 Hybrid Registration

In this case, the graph-based formulation will consist of the discrete model for the iconic and geometric registration along with a coupling penalty (Eq. (16)). Therefore, the graph that represents the problem comprises \mathcal{G}_{ico} and \mathcal{G}_{geo} along with a third set of edges \mathcal{E}_{hyb} containing all possible connections between the iconic random variables and the geometric variables. The pairwise label assignment penalty on these coupling edges is then defined as:

$$\mathcal{P}_{\text{hyb},pq}(l_p, l_q) = \omega_q(\kappa_p) \|\mathbf{u}_{\text{geo},l_p}(\kappa_p) - (\mathbf{d}_q + l_q)\|, \quad (21)$$

Fig. 7 An example landmark pair (denoted by red and blue crosses) detected based on the Gabor response-based similarity metric and the mutual-saliency measure. (a) Source and (b) target images. This figure is re-printed from [58]

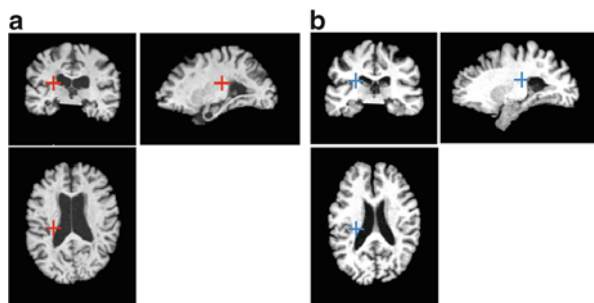
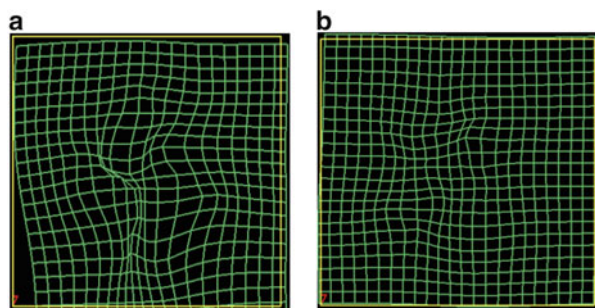


Fig. 8 The Dense deformation fields generated by (a) M1 – no MRF regularization and (b) M2 – with MRF regularization. This figure is re-printed from [58]



where $p \in \mathcal{V}_{\text{geo}}$ and $q \in \mathcal{V}_{\text{ico}}$, $l_p \in \mathcal{L}_{\text{geo}}$ and $l_q \in \mathcal{L}_{\text{ico}}$, and $(p, q) \in \mathcal{E}_{\text{hyb}}$. Such a pairwise term couples the displacements given by the two registration processes and imposes consistency. To conclude, the coupled registration objective function is represented by an MRF graph $\mathcal{G}_{\text{hyb}} = (\mathcal{V}_{\text{geo}} \cup \mathcal{V}_{\text{ico}}, \mathcal{E}_{\text{geo}} \cup \mathcal{E}_{\text{ico}} \cup \mathcal{E}_{\text{hyb}})$ with its associated unary and pairwise potential functions. This model was presented in [76].

4.3 Experimental Validation

4.3.1 Geometric Registration

Figure 7 shows a typical landmark pair detected by Gabor response and matched by the MRF formulation. Many such pairs found by the MRF formulation resulted in a deformation that was smoother with the MRF regularization rather than without, as can be seen in Fig. 8.

4.3.2 Hybrid Registration

In order to validate the coupled geometric registration method in a way that is invariant to landmark extraction, a multi-modal synthetic data set is used. In this setting, the ground truth deformation is known allowing for a quantitative analysis

Table 1 End point error (in millimeters) for the registration of the Synthetic MR Dataset. The grid spacing is denoted by h . This figure is reprinted from [25]

#	Iconic ($h = 60$ mm)		Hybrid ($h = 60$ mm)		Iconic ($h = 20$ mm)		Hybrid ($h = 20$ mm)	
	mean	std	mean	std	mean	std	mean	std
1	1.33	0.69	1.25	0.59	1.38	1.21	0.98	0.61
2	1.32	0.75	1.18	0.53	2.46	3.21	1.06	0.68
3	1.44	0.97	1.22	0.56	2.05	2.40	1.03	0.67
4	1.40	0.74	1.16	0.50	1.40	1.02	1.08	0.69
5	1.23	0.60	1.15	0.56	1.38	1.01	1.03	0.67
6	1.35	0.74	1.24	0.62	1.58	1.39	1.05	0.71
7	1.16	0.56	1.09	0.50	1.45	1.18	1.05	0.67
8	1.29	0.68	1.23	0.58	1.93	2.61	1.11	0.79
9	1.23	0.62	1.19	0.53	1.72	1.89	1.04	0.71
10	1.54	1.08	1.19	0.58	2.60	3.43	1.05	0.73
all	1.33	0.11	1.19	0.05	1.79	0.45	1.05	0.03

of the registration performance regarding both the dense deformation field accuracy and the quality of the established landmark correspondences.

The goal of this experiment is to demonstrate the added value from considering geometric information on top of standard iconic one. Thus, a comparison of the proposed framework with and without the geometric registration part takes place. Regarding the results, if we look at the registration accuracy in terms of end point error (Table 1), we see that the coupled iconic geometric registration method is able to further improve the results of the iconic one. This is evident, as the end point error has decreased by taking advantage of the geometric information.

As we expect the hybrid approach to be able to cope with large displacements better than the pure iconic one, we repeated the experiments by decreasing the initial control point spacing to 20 mm and thus limiting the maximum amount of deformation that can be handled. The results are also reported in Table 1. In this case, we can observe a more significant difference between the performance of the two proposed approaches. Therefore, we should conclude that the additional computational cost demanded by the coupled approach can be compensated by the better quality of the results.

5 Conclusion

This chapter presents a comprehensive overview of graph-based deformable registration. Discrete models for the cases of deformable registration involving point-wise intensity-based similarity criteria, statistical intensity-based criteria, attribute-based ones as well as for geometric and hybrid registration were presented. The increased computational efficiency, accuracy and robustness of graph-based formulations were also demonstrated.

Acknowledgements We would like to acknowledge Dr. Ben Glocker, from Imperial College London, whose work formed the basis of the subsequent works that are presented here.

References

1. Amit, Y.: A nonlinear variational problem for image matching. *SIAM Journal on Scientific Computing* **15**(1), 207–224 (1994)
2. Arsigny, V., Pennec, X., Ayache, N.: Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations – Application to the registration of histological slices. *Medical Image Analysis* **9**(6), 507–523 (2005)
3. Ashburner, J.: A fast diffeomorphic image registration algorithm. *NeuroImage* **38**(1), 95–113 (2007)
4. Ashburner, J., Friston, K.J.: Nonlinear spatial normalization using basis functions. *Human Brain Mapping* **7**(4), 254–266 (1999)
5. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis* **12**(1), 26–41 (2008)
6. Baumann, B.C., Teo, B.K., Pohl, K., Ou, Y., Doshi, J., Alonso-Basanta, M., Christodouleas, J., Davatzikos, C., Kao, G., Dorsey, J.: Multiparametric processing of serial mri during radiation therapy of brain tumors: ‘finishing with flair?’. *International Journal of Radiation Oncology* Biology* Physics* **81**(2), S794 (2011)
7. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). *Computer vision and image understanding* **110**(3), 346–359 (2008)
8. Beg, M.F., Miller, M.I., Trounev, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision* **61**(2), 139–157 (2005)
9. Betke, M., Hong, H., Thomas, D., Prince, C., Ko, J.P.: Landmark detection in the chest and registration of lung surfaces with an application to nodule registration. *Medical Image Analysis* **7**(3), 265–281 (2003)
10. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(6), 567–585 (1989)
11. Cachier, P., Mangin, J.F., Pennec, X., Rivière, D., Papadopoulos-Orfanos, D., Régis, J., Ayache, N.: Multisubject non-rigid registration of brain MRI using intensity and geometric features. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 734–742 (2001)
12. Can, A., Stewart, C.V., Roysam, B., Tanenbaum, H.L.: A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24**(3), 347–364 (2002)
13. Christensen, G.E., Johnson, H.J.: Consistent image registration. *IEEE transactions on medical imaging* **20**(7), 568–82 (2001)
14. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing* **5**(10), 1435–1447 (1996)
15. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* **89**(2-3), 114–141 (2003)
16. Chung, A.C., Wells III, W.M., Norbash, A., Grimson, W.E.L.: Multi-modal image registration by minimizing Kullback-Leibler distance. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 525–532 (2002)
17. Da, X., Toledo, J.B., Zee, J., Wolk, D.A., Xie, S.X., Ou, Y., Shacklett, A., Parnpi, P., Shaw, L., Trojanowski, J.Q., et al.: Integration and relative value of biomarkers for prediction of mci to ad progression: Spatial patterns of brain atrophy, cognitive scores, apoe genotype and csf biomarkers. *NeuroImage: Clinical* **4**, 164–173 (2014)

18. D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P.: A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis* **7**(4), 565–575 (2003)
19. Davatzikos, C.: Spatial transformation and registration of brain images using elastically deformable models. *Computer Vision and Image Understanding* **66**(2), 207–222 (1997)
20. Droske, M., Rumpf, M.: A variational approach to nonrigid morphological image registration. *SIAM Journal on Applied Mathematics* **64**(2), 668–687 (2004)
21. Erus, G., Battapady, H., Satterthwaite, T.D., Hakonarson, H., Gur, R.E., Davatzikos, C., Gur, R.C.: Imaging patterns of brain development and their relationship to cognition. *Cerebral Cortex* p. bht425 (2014)
22. Fischer, B., Modersitzki, J.: Fast diffusion registration. *AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging* **313**, 117–127 (2002)
23. Glaunès, J., Trounev, A., Younes, L.: Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In: *International Conference on Computer Vision and Pattern Recognition*, pp. 712–718 (2004)
24. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis* **12**(6), 731–741 (2008)
25. Glocker, B., Sotiras, A., Komodakis, N., Paragios, N.: Deformable medical image registration: setting the state of the art with discrete methods. *Annual Review of Biomedical Engineering* **13**, 219–244 (2011)
26. Hajnal, J.V., Hill, D.L., Hawkes, D.J. (eds.): *Medical image registration*. CRC Press, Boca Raton, FL (2001)
27. Hartkens, T., Hill, D.L.G., Castellano-Smith, A., Hawkes, D.J., Maurer, C.R., Martin, A., Hall, W., Liu, H., Truweit, C.: Using points and surfaces to improve voxel-based non-rigid registration. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 565–572 (2002)
28. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A.: Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis* **16**(7), 1423–1435 (2012)
29. Hellier, P., Barillot, C.: Coupling dense and landmark-based approaches for nonrigid registration. *IEEE Transactions on Medical Imaging* **22**(2), 217–227 (2003)
30. Holden, M.: A review of geometric transformations for nonrigid body registration. *IEEE Transactions on Medical Imaging* **27**(1), 111–128 (2008)
31. Hsieh, J.W., Liao, H.Y.M., Fan, K.C., Ko, M.T., Hung, Y.P.: Image registration using a new edge-based approach. *Computer Vision and Image Understanding* **67**(2), 112–130 (1997)
32. Huang, X., Paragios, N., Metaxas, D.N.: Shape registration in implicit spaces using information theory and free form deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(8), 1303–1318 (2006)
33. Ingalhalikar, M., Parker, D., Ghanbari, Y., Smith, A., Hua, K., Mori, S., Abel, T., Davatzikos, C., Verma, R.: Connectome and maturation profiles of the developing mouse brain using diffusion tensor imaging. *Cerebral Cortex* p. bhu068 (2014)
34. Jian, B., Vemuri, B., Marroquin, J.: Robust nonrigid multimodal image registration using local frequency maps. In: *Information Processing in Medical Imaging (IPMI)*, pp. 504–515 (2005)
35. Johnson, H.J., Christensen, G.E.: Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging* **21**(5), 450–461 (2002)
36. Kadir, T., Brady, M.: Saliency, scale and image description. *International Journal of Computer Vision* **45**(2), 83–105 (2001)
37. Ke, Y., Sukthankar, R.: Pca-sift: A more distinctive representation for local image descriptors. In: *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, pp. II–506. IEEE (2004)
38. Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage* **46**(3), 786–802 (2009)

39. Komodakis, N., Tziritas, G.: Approximate labeling via graph cuts based on linear programming. *IEEE transactions on pattern analysis and machine intelligence* **29**(8), 1436–53 (2007)
40. Komodakis, N., Tziritas, G., Paragios, N.: Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. *Computer Vision and Image Understanding* **112**(1), 14–29 (2008)
41. Koutsouleris, N., Davatzikos, C., Borgwardt, S., Gaser, C., Bottlender, R., Frodl, T., Falkai, P., Riecher-Rössler, A., Möller, H.J., Reiser, M., et al.: Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders. *Schizophrenia bulletin* p. sbt142 (2013)
42. Kwon, D., Lee, K., Yun, I., Lee, S.: Nonrigid image registration using dynamic higher-order mrf model. In: *European Conference on Computer Vision*, pp. 373–386 (2008)
43. Leordeanu, M., Hebert, M.: A spectral technique for correspondence problems using pairwise constraints. In: *International Conference on Computer Vision*, pp. 1482–1489 (2005)
44. Li, G., Guo, L., Liu, T.: Deformation invariant attribute vector for deformable registration of longitudinal brain MR images. *Computerized Medical Imaging and Graphics* **33**(5), 273–297 (2009)
45. Li, H., Manjunath, B., Mitra, S.K.: A contour-based approach to multisensor image registration. *Image Processing, IEEE Transactions on* **4**(3), 320–334 (1995)
46. Lindeberg, T.: Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *International Journal of Computer Vision* **11**(3), 283–318 (1993)
47. Ling, H., Jacobs, D.: Deformation invariant image matching. In: *The Tenth International Conference in Computer Vision (ICCV)*. Beijing, China. (2005)
48. Liu, J., Vemuri, B.C., Marroquin, J.L.: Local frequency representations for robust multimodal image registration. *IEEE Transactions on Medical Imaging* **21**(5), 462–469 (2002)
49. Liu, J., Vemuri, B.C., Marroquin, J.L.: Local frequency representations for robust multimodal image registration. *IEEE Transactions on Medical Imaging* **21**(5), 462–469 (2002)
50. Lowe, D.G.: Object recognition from local scale-invariant features. In: *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150–1157. Ieee (1999)
51. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60**(2), 91–110 (2004)
52. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging* **16**(2), 187–198 (1997)
53. Maintz, J.A., Viergever, M.A.: A survey of medical image registration. *Medical Image Analysis* **2**(1), 1–36 (1998)
54. Modersitzki, J.: *FAIR: Flexible algorithms for image registration*. SIAM, Philadelphia (2009)
55. Narayanan, R., Fessler, J.A., Park, H., Meyer, C.R.: Diffeomorphic nonlinear transformations: a local parametric approach for image registration. In: *International Conference on Information Processing in Medical Imaging*, pp. 174–185 (2005)
56. Noblet, V., Heinrich, C., Heitz, F., Arnschlag, J.P.: Symmetric nonrigid image registration: application to average brain templates construction. In: *Medical Image Computing and Computer-Assisted Intervention : MICCAI '08*, no. Pt 2 in LNCS, pp. 897–904 (2008)
57. Ou, Y., Akbari, H., Bilello, M., Da, X., Davatzikos, C.: Comparative evaluation of registration algorithms for different brain databases with varying difficulty: Results and Insights. *IEEE Transactions on Medical Imaging* (2014). doi:10.1109/TMI.2014.2330355
58. Ou, Y., Besbes, A., Bilello, M., Mansour, M., Davatzikos, C., Paragios, N.: Detecting mutually-salient landmark pairs with MRF regularization. In: *International Symposium on Biomedical Imaging*, pp. 400–403 (2010)
59. Ou, Y., Weinstein, S.P., Conant, E.F., Englander, S., Da, X., Gaonkar, B., Hsiao, M., Rosen, M., DeMichele, A., Davatzikos, C., Kontos, D.: Deformable registration for quantifying longitudinal tumor changes during neoadjuvant chemotherapy: In Press. *Magnetic Resonance in Medicine* (2014)

60. Ou, Y., Reynolds, N., Gollub, R., Pienaar, R., Wang, Y., Wang, T., Sack, D., Andriole, K., Pieper, S., Herrick, C., Murphy, S., Grant, P., Zollei, L.: Developmental brain adc atlas creation from clinical images. In: Organization for Human Brain Mapping (OHBM) (2014)
61. Ou, Y., Shen, D., Feldman, M., Tomaszewski, J., Davatzikos, C.: Non-rigid registration between histological and MR images of the prostate: A joint segmentation and registration framework. In: Computer Vision and Pattern Recognition workshop, 2009. CVPR 2009. IEEE Conference on, pp. 125–132 (2009)
62. Ou, Y., Sotiras, A., Paragios, N., Davatzikos, C.: DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Medical Image Analysis* **15**(4), 622–639 (2011)
63. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **12**(7), 629–639 (1990)
64. Postelnicu, G., Zollei, L., Fischl, B.: Combined volumetric and surface registration. *IEEE Transactions on Medical Imaging* **28**(4), 508–522 (2009)
65. Roche, A., Malandain, G., Pennec, X., Ayache, N.: The correlation ratio as a new similarity measure for multimodal image registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 1115–1124 (1998)
66. Rohr, K.: On 3d differential operators for detecting point landmarks. *Image and Vision Computing* **15**(3), 219–233 (1997)
67. Rohr, K., Stiehl, H.S., Sprengel, R., Buzug, T.M., Weese, J., Kuhn, M.: Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on Medical Imaging* **20**(6), 526–534 (2001)
68. Rueckert, D., Aljabar, P., Heckemann, R.A., Hajnal, J.V., Hammers, A.: Diffeomorphic registration using B-splines. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 702–709 (2006)
69. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging* **18**(8), 712–721 (1999)
70. Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughhead, J., Prabhakaran, K., Calkins, M.E., Hopson, R., Jackson, C., Keefe, J., Riley, M., et al.: Neuroimaging of the philadelphia neurodevelopmental cohort. *NeuroImage* **86**, 544–553 (2014)
71. Sederberg, T.W., Parry, S.R.: Free-form deformation of solid geometric models. *ACM Siggraph Computer Graphics* **20**(4), 151–160 (1986)
72. Serpa, M.H., Ou, Y., Schaufelberger, M.S., Doshi, J., Ferreira, L.K., Machado-Vieira, R., Menezes, P.R., Scazufca, M., Davatzikos, C., Busatto, G.F., et al.: Neuroanatomical classification in a population-based sample of psychotic major depression and bipolar i disorder with 1 year of diagnostic stability. *BioMed Research International* **2014** (2014)
73. Shen, D.: Image registration by local histogram matching. *Pattern Recognition* **40**(4), 1166–1172 (1997)
74. Shen, D., Davatzikos, C.: HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE transactions on Medical Imaging* **21**(11), 1421–39 (2002)
75. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: a survey. *IEEE Transactions on Medical Imaging* **32**(7), 1153–90 (2013)
76. Sotiras, A., Ou, Y., Glocker, B., Davatzikos, C., Paragios, N.: Simultaneous geometric–iconic registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 676–683 (2010)
77. Sotiras, A., Paragios, N.: Discrete symmetric image registration. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 342–345 (2012)
78. Szeliski, R.: Image alignment and stitching: A tutorial. *Foundations and Trends® in Computer Graphics and Vision* **2**(1), 1–104 (2006)
79. Tagare, H., Groisser, D., Skrinjar, O.: Symmetric non-rigid registration: A geometric theory and some numerical techniques. *Journal of Mathematical Imaging and Vision* **34**(1), 61–88 (2009)

80. Thirion, J.P.: Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis* **2**(3), 243–260 (1998)
81. Toews, M., Wells III, W.M.: Efficient and robust model-to-image alignment using 3d scale-invariant features. *Medical image analysis* **17**(3), 271–282 (2013)
82. Torresani, L., Kolmogorov, V., Rother, C.: Feature correspondence via graph matching: Models and global optimization. In: *European Conference on Computer Vision*, pp. 596–609 (2008)
83. Tsin, Y., Kanade, T.: A correlation-based approach to robust point set registration. In: *European Conference on Computer Vision*, pp. 558–569 (2004)
84. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric log-domain diffeomorphic Registration: a demons-based approach. In: *Medical Image Computing and Computer-Assisted Intervention : MICCAI'08*, no. Pt 1 in LNCS, pp. 754–61 (2008)
85. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic Demons: Efficient non-parametric image registration. *NeuroImage* **45**(1, Supplement 1), S61–S72 (2009)
86. Viola, P., Wells III, W.M.: Alignment by maximization of mutual information. *International Journal of Computer Vision* **24**(2), 137–154 (1997)
87. Wu, Y.T., Kanade, T., Li, C.C., Cohn, J.: Image registration using wavelet-based motion model. *International Journal of Computer Vision* **38**(2), 129–152 (2000)
88. Yang, J., Shen, D., Davatzikos, C.: Diffusion tensor image registration using tensor geometry and orientation features. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 905–913 (2008)
89. Yi, Z., Zhiguo, C., Yang, X.: Multi-spectral remote image registration based on sift. *Electronics Letters* **44**(2), 107–108 (2008)
90. Zanetti, M.V., Schaufelberger, M.S., Doshi, J., Ou, Y., Ferreira, L.K., Menezes, P.R., Scazufca, M., Davatzikos, C., Busatto, G.F.: Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **43**, 116–125 (2013)
91. Zhan, Y., Ou, Y., Feldman, M., Tomaszewski, J., Davatzikos, C., Shen, D.: Registering histologic and mr images of prostate for image-based cancer detection. *Academic Radiology* **14**(11), 1367–1381 (2007)
92. ZHANG, R.j., Zhang, J.q., Yang, C.: Image registration approach based on surf [j]. *Infrared and Laser Engineering* **1**, 041 (2009)
93. Zitova, B., Flusser, J.: Image registration methods: a survey. *Image and Vision Computing* **21**(11), 977–1000 (2003)

Part IV
Clinical Biomarkers

Cardiovascular Informatics

I.A. Kakadiaris, U. Kurkure, A. Bandekar, S. O'Malley, and M. Naghavi

Abstract As cardiac imaging technology advances, large amounts of imaging data are being produced which are not being mined sufficiently by current diagnostic tools for early detection and diagnosis of cardiovascular disease. We aim to develop a computational framework to mine cardiac imaging data and provide quantitative measures for developing a new risk assessment method. In this chapter, we present novel methods to quantify pericardial fat in non-contrast cardiac computed tomography images automatically, and to detect and quantify neovascularization in the coronary vessels using intra-vascular ultrasound imaging.

1 Introduction

Cardiovascular disease has long been the leading cause of death in developed countries, and it is rapidly becoming the number one killer in developing countries [3]. In 2006, it is estimated that more than 19 million people worldwide experienced a life-threatening heart attack. In the US alone, 1.4 million people suffer from a heart attack annually. Of the 140 million Americans between the ages of 35-44, 17.4 % of males (24.36 million) and 13.6 % of females (19.04 million) have coronary heart disease. Approximately 50 % of heart attack related deaths occur in people with no prior symptoms [6]. Hence, **sudden heart attacks remain the number one cause of death in the US**. Unpredicted heart attacks account for the majority of the \$280 billion burden of cardiovascular diseases.

I.A. Kakadiaris (✉) • U. Kurkure • A. Bandekar • S. O'Malley
CBL, University of Houston, Houston, TX

e-mail: ioannisk@uh.edu; udaykurkure@gmail.com; alokb@cbl.uh.edu; ffrog.geo@yahoo.com

M. Naghavi
AEHA, Houston, TX
e-mail: mn2@vp.org

Coronary artery disease occurs as a result of atherosclerosis, a condition in which a fatty substance called *plaque* builds up on the walls of the arteries. If these plaques rupture, blood clots form that obstruct the flow of blood to the heart and may cause a heart attack, which can often be fatal. Some plaques present a particularly high risk of rupture and a subsequent heart attack. All types of atherosclerotic plaques with a high likelihood of thrombotic complications and rapid progression are recognized as *vulnerable plaques*. The field of cardiology has witnessed a major paradigm shift in its determination of patients at risk of coronary artery disease. In the past, increasing fat deposition and gradual inward luminal narrowing of coronary arteries were thought to be the culprits in heart attacks. Today, cardiovascular specialists know that heart attacks are caused by inflammation of the coronary arteries and thrombotic complications of vulnerable plaques. As a result, the discovery of vulnerable plaque has recently evolved into the definition of “vulnerable patient.” *A vulnerable patient is defined as a person with more than 10 % likelihood of having a heart attack in the next 12 months.* Over 45 world leaders in cardiology have collectively introduced the field of vulnerable patient detection as the new era in preventive cardiology [7, 8].

Until a few years ago many cardiovascular specialists did not have many of the novel diagnostic tools that are becoming increasingly available today. These emerging diagnostic tests include: 1) novel genetic and serum biomarkers such as CRP and inflammatory markers; 2) noninvasive imaging tests such as Computed Tomography (CT), and CT Angiography (CTA), and ultrasound-based intima-media thickness measurement; and 3) new interventional catheter-based diagnostic tools (e.g., intravascular ultrasound (IVUS)). All of these provide unprecedented opportunities for early detection and risk stratification of subjects at risk of heart attack. A non-profit initiative pioneered by Association for Eradication of Heart Attacks (AEHA), Screening for Heart Attack Prevention and Education (SHAPE), presents a practice guideline for doctors to implement public screening of at-risk populations, calling for men 45 and older and women 55 and older to undergo a comprehensive vascular health assessment [6]. It comprises a pyramid of tests with serum tests and Framingham Risk Score assessment at the bottom of the pyramid, non-invasive imaging such as CT and CTA in the middle, and IVUS imaging at the top. If a patient is found to be at risk with the tests at one level, then s/he is referred to tests in the next level to allow better localization. Considering the large amounts of data that the SHAPE program will produce, there is an urgent need for computational tools to assist in screening for the conditions that underlie sudden cardiac events. Existing cardiovascular risk scoring methods do not take into account all the wealth of information that is available in imaging data. The reason is two-fold: 1) there is a lack of automatic techniques to mine the data for the required information, and 2) validation in large epidemiological studies is needed to determine which type of information will offer additive predictive value. Our long term vision is to contribute to the development of quantitative methods to assess cumulative risk of vulnerable patients by developing new techniques to mine additional information from their imaging data. In this chapter, we will focus on presenting methods for the analysis of CT and IVUS data.

On one hand, CT can provide information to assess abdominal and pericardial fat. Recent evidence indicates that pericardial fat may be a significant cardiovascular risk factor [16]. Although pericardial fat is routinely imaged during CT for coronary calcium scoring, it is currently ignored in the analysis of CT images. The primary reason for this is the absence of a tool capable of automatic quantification of pericardial fat. Recent studies on pericardial fat imaging were limited to manually outlined regions-of-interest and preset fat attenuation thresholds, which are subject to inter-observer and inter-scan variability. In general, there is an increased demand for automated, robust, and accurate quantitative measures of these tissues in the heart region. An accurate segmentation allows for quantitative and morphological analysis. Thus, the development of fully automated methods which will provide unbiased and consistent results is highly desirable.

On the other hand, IVUS, which is a catheter-based ultrasound technology, can provide real-time video of the interior of a blood vessel at video frame rates (10 to 30 frames/s). It is used routinely for the detailed assessment of vessel morphology and atherosclerotic plaque, often as a guide for interventional procedures (e.g., angioplasty) where alternative non-invasive imaging techniques such as X-ray angiography are insufficient. Recent years have witnessed a plethora of advancements [11, 12] in IVUS technology in an effort to extract additional diagnostic information from these sensors. A number of techniques have been developed to allow increased performance of IVUS contrast imaging, in which an echogenic solution of microbubbles is incorporated into the bloodstream as a tracer of perfusion. Unfortunately, the newest methods for contrast imaging remain experimental and tend to require non-standard IVUS hardware [4] (e.g., harmonic imaging catheters), non-standard contrast agents, or both. In this paper, we present a computer-aided technique [5, 17, 18] which allows IVUS contrast imaging to be accomplished with commercially-available IVUS systems and off-the-shelf microbubble contrast agents. Our contributions include a novel IVUS contrast imaging protocol, as well as the algorithms necessary to process the acquired imagery, that enabled quantification of vasa vasorum (VV) presence in vivo for the first time.

Specifically, in this chapter, we describe a unified knowledge-based medical image segmentation framework and a multi-class, multi-feature fuzzy connectedness-based tissue classification and segmentation method, which were used for automated detection and quantification of pericardial fat in non-contrast cardiac CT images. We also present our recently developed novel imaging protocol for contrast imaging in IVUS, employing stationary-catheter sequences, along with the computational tools necessary for frame-gating, contour-tracking and processing the resulting data, to image the extra-luminal perfusion due to blood flow through the VV.

The rest of the chapter is organized as follows. In Sect. 2, we present methods for CT data analysis. In Sect. 3, we present methods for IVUS data analysis. In Sect. 4, we present the validation results from our methods. In Sect. 5, we discuss the advantages and limitations of our work.

2 CT Data Analysis

In this section, we describe a unified knowledge-based medical image segmentation framework [2]. We present a knowledge-based atlas which includes spatial information from a probabilistic atlas, texture-based features relationships between organs in the atlas, and image analysis methods for segmentation of specific tissues or organs in the atlas. Our framework consists of a training phase and a deployment phase. In the training phase, we construct a probabilistic atlas of the region of interest and learn various parameters for each component in the framework. In the deployment phase, we use the probabilistic atlas for initialization of the region of interest and employ the learned parameters for further refinement of the region of interest using the image analysis methods predefined for the segmentation of a specific organ or tissue in the region of interest. Such a unified knowledge-based framework can be generalized for the segmentation of various human body regions such as the head, neck, pelvis, and other anatomical structures. We also present a multi-class multi-feature region classification and segmentation method based on fuzzy-connectedness formulation for fat tissue detection. By relaxing the connectedness criterion, this method can be used for fat tissue classification [1].

Knowledge-based atlas: We consider the data D to be a bounded subset of \mathbb{R}^3 . We assume that we have *a priori* knowledge of the $N \geq 1$ organs present in the region D . Our knowledge-based atlas \mathcal{K} used to describe the region of interest D consists of four elements, namely, a probabilistic atlas (\mathcal{P}), texture-based features (\mathcal{T}), relationships (\mathcal{R}), and specific segmentation methods (\mathcal{M}) for organs in the region D , $\mathcal{K}(D) = \{\mathcal{P}, \mathcal{T}, \mathcal{R}, \mathcal{M}\}$. We construct a probabilistic atlas \mathcal{P} for our data D , which is a map that assigns to each voxel a set of probabilities to belong to each of the organs, $\mathcal{P}(D) = (p_0, p_1, p_2, \dots, p_N)$, where p_i is the probability for the voxel to belong to organ i , ($1 \leq i \leq N$). All the organs other than those that are manually delineated are assigned the probability p_0 where $p_0 = 1 - \sum_{k=1}^N p_k$. The probabilistic atlas is constructed in the training phase of the framework. We initialize the organs present in the atlas during deployment using registration techniques described in [15]. This helps us to gain insight on the spatial information of different organs and variations with respect to one another, and also guides automated segmentation procedures.

Every organ will present different texture-based features depending on the image modality and the tissue type of the organ. We denote the set of texture-based features by, $\mathcal{T} = \{T_1, T_2, \dots, T_N\}$, where $T_i = [f_1^i, f_2^i, \dots, f_k^i]$ is set of optimal texture-based features selected in the training phase for discrimination for organ i , ($1 \leq i \leq N$) and k is the number of features for organ i . The optimal texture-based features are selected in the training phase of our framework. Texture-based features are used by the image analysis methods to segment the organs and refine the initialization by the probabilistic atlas.

Relationships are a set of rules which describe relationships between two organs in the region. Based on prior anatomical knowledge about the organs, we consider two types of relationships: hierarchical (e.g., child-parent) and spatial

(e.g., posterior-anterior). We denote the set of relationships by $\mathcal{R} = \{R_{r,t} | r, t \text{ organs}\}$, where $R_{r,t} \in \{\textit{posterior}, \textit{anterior}, \textit{right}, \textit{left}, \textit{child}, \textit{parent}\}$, r is the reference organ and t is the target organ. For example, $R_{i,j} = \textit{child}$ means that, the reference organ i (e.g., ventricle) is the child of the target organ j (e.g., heart); while $R_{i,j} = \textit{right}$ implies that the reference organ i (e.g., right lung) is to the right of the target organ j (e.g., heart).

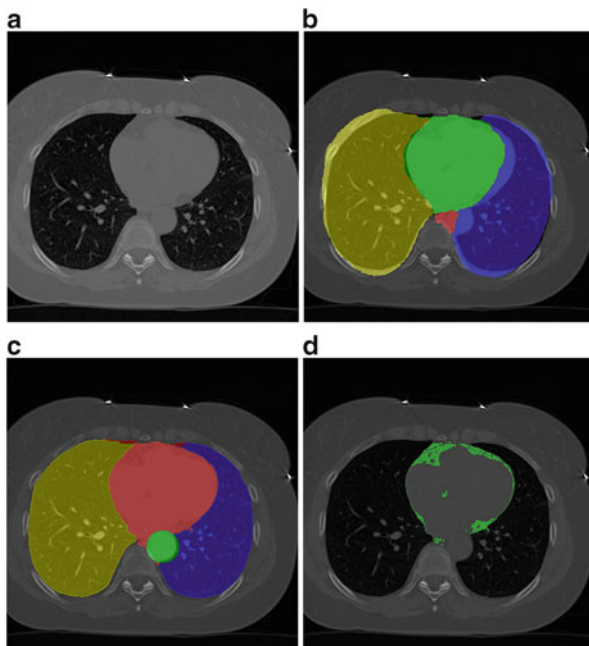
Image analysis methods are specific methods set during the training phase for a specific organ. We denote the set of all methods by, $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$. Due to high anatomic variations and the large amount of structural information found in medical images, global information based segmentation methods yield inadequate results in region extraction. We use the knowledge-based atlas to guide the automatic segmentation process and then use specific image analysis methods for segmentation of anatomical structures. We examine the performance of these methods in the training phase and accordingly select methods with higher accuracy and true positive rate as well as lower false positive rate. One of the segmentation methods used is the multi-class multi-feature fuzzy connectedness method, which is described in the following section.

Multi-class, multi-feature fuzzy connectedness: The anatomical objects in medical data are characterized by certain intensity level and intensity homogeneity features. Also, additional features can be computed to characterize certain properties of a tissue class. Such features can be used to distinguish between different types of tissue classes. Our multi-class, multi-feature fuzzy connectedness method is able to take advantage of multiple features to distinguish between multiple classes for segmentation and classification.

We define three kinds of fuzzy affinities: local fuzzy spel affinity, global object affinity, and global class affinity. The *local fuzzy spel affinity* (μ_κ) consists of three components: 1. the object feature intensity component (μ_ϕ), 2. the intensity homogeneity component (μ_ψ), and 3. the texture feature component (μ_φ). The similarity of the pixels' feature vectors is computed using the Mahalanobis metric: $m_{d(c \rightarrow d)}^2 = (X_{(c \rightarrow d)} - \bar{X}_{(c \rightarrow d)})^T S_{(c \rightarrow d)}^{-1} (X_{(c \rightarrow d)} - \bar{X}_{(c \rightarrow d)})$, where $X_{(c \rightarrow d)}$, $\bar{X}_{(c \rightarrow d)}$, $S_{(c \rightarrow d)}$ are the feature vector, the mean feature vector, and the covariance matrix in the direction from c to d , respectively. The bias in intensity in a specific direction is accounted for by allowing different levels and signs of intensity homogeneities in different directions of adjacency [13]. Thus, this formulation accounts for different levels of the increase or decrease in intensity values in the horizontal (left, right) or vertical (up, down) directions. The advantage of using the Mahalanobis metric is that it weighs the differences in various feature dimensions by the range of variability in the direction of the feature dimension. These distances are computed in units of standard deviation from the mean. This allows us to assign a statistical probability to the measurement. The local fuzzy spel affinity is computed as: $\mu_\kappa(c, d) = \frac{1}{1 + m_{d(c \rightarrow d)}}$ in order to ensure that $\mu_\kappa(c, d) \in Z^2 \rightarrow [0, 1]$ and it is reflexive and symmetric, where Z^2 is a set of all pixels of a two-dimensional Euclidean space.

Fuzzy connectedness captures the global hanging-togetherness of pixels by using the local affinity relation and by considering all possible paths between two, not

Fig. 1 Pericardial fat detection in CT. (a) Original CT image, (b) all labels overlaid after atlas initialization, (c) all labels after segmentation, and (d) pericardial fat overlaid



necessarily nearby, pixels in the image. It considers the strengths of all possible paths between given two pixels, where the strength of a particular path is the weakest affinity between the successive pairs of pixels along the path. Thus, the strongest connectedness path between the given two pixels specifies the degree of global hanging togetherness between the given two pixels. *Global object affinity* is the largest of the weakest affinities between the successive pairs of pixels along the path p_{cd} of all possible paths P_{cd} from c to d and is given by $\mu_{\kappa}(c,d) = \max_{p_{cd} \in P_{cd}} \{\min_{1 \leq i \leq m} [\mu_{\kappa}(c_{(i)}, c_{(i+1)})]\}$.

In our framework, the global object affinity and local pixel affinity are assigned only if the *global class affinity* (or discrepancy measure) of c and d belonging to the neighboring objects' classes is more (or less) than a predefined value, γ (note that the affinity value has an inverse relationship with the Mahalanobis distance metric in our formulation). The minimum discrepancy measure $J(c,d) = \min_{1 \leq i \leq b} m_d(c, d)$, where b is the number of neighboring classes of the target object, gives the maximum membership value of a pixel pair belonging to a certain class. If $J(c,d) < \gamma$, and the class to which the pixel pair belongs is not the target object class, then the local pixel affinity $\mu_{\kappa}(c,d)$ is set to zero, else its local pixel affinity is computed as described earlier.

Since pericardial fat tissue appears as disjoint sets of pixels distributed all over in the thoracic cavity, we relax the spatial connectedness constraint in our formulation to allow scattered tissue classification (Fig. 1). Our segmentation algorithm uses dynamic statistics of fat tissue for automatic segmentation of the fat tissue. We

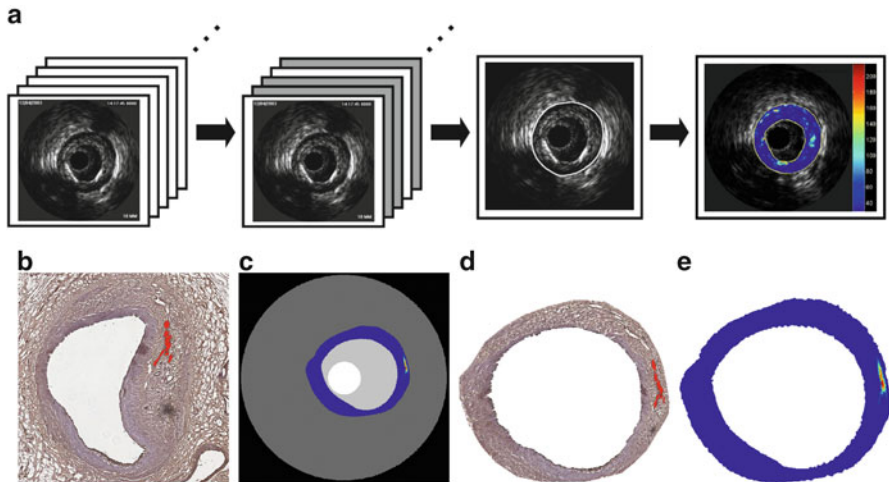


Fig. 2 Top: (a) Flowchart of an analysis of a contrast-enhanced IVUS sequence. From left to right - the original sequence, the sequence decimated by gating, the contour-tracking step, and difference imaging and overlay of results. Bottom: Demonstration of registration between histology and IVUS. (b) A stained histological image, (c) an IVUS image to which the histology image will be co-registered by manually defining corresponding landmarks, (d) deformed histology image based on the landmarks, and (e) the IVUS image highlighting the correspondence. Note the excellent agreement of our analysis with the histology

estimate the statistics of fat tissue by using a sample region around a seed point, hence the selection of the seed point is very critical. We obviate the need for manual seed selection by automatic seed initialization using relationships defined in the training phase. Thus, instead of applying a space-invariant global threshold value, our method adapts the threshold value locally in the feature space. In addition to fat segmentation in CT, this method has also been used for left ventricular segmentation in MRI [14].

3 IVUS Data Analysis

In this section, we introduce a number of techniques for IVUS image analysis [9, 10] which enable the use of difference imaging to detect those changes which occur in the IVUS imagery due to the perfusion of an intravascularly-injected contrast agent into the plaque and vessel wall. To determine if a particular physical region of the frame experiences an increase in echogenicity due to contrast perfusion into the wall, it is necessary to compare the appearance of that region under IVUS before and after the introduction of contrast. Since multiple sources of motion are present, we follow three steps to detect perfusion: motion compensation, image subtraction and deriving statistics from the resulting difference images (Fig. 2(a)).

The goal of two-step motion compensation (frame gating and contour-tracking) is to provide pixelwise correspondence between a region of interest (e.g., the plaque) in each frame. Unlike previous efforts utilizing ECG signals, we perform an appearance-based grouping of frames. By formulating the problem in terms of multidimensional scaling (MDS), a number of other useful operations may be performed. The MDS transform places points defined only by inter-point proximities into a metric space such that the proximities are preserved with minimal loss. In our context, this allows the creation of a frame-similarity space which may be employed as a concise visual and numerical summary of an entire frame sequence. Clustering this space allows sets of frames with various similarity properties to be extracted efficiently. We begin by taking our n -frame sequence and creating a square matrix D in which each entry $d_{i,j}$ represents the dissimilarity between frames i and j . As normalized cross-correlation returns values in the range $[-1, +1]$, we clamp these values to $[0, +1]$ and subtract them from one. This results in a matrix with zero along the diagonal and values on the range $[0, 1]$ everywhere else. Next, we let A be the matrix where $a_{i,j} = -\frac{1}{2}d_{i,j}^2$ and let $C_n = I_n - \frac{1}{n}1_n1_n^T$, where I_n is the $n \times n$ identity matrix and 1_n is the unit vector of length n . Let $B = C_nAC_n$. We let $\lambda_1 \geq \dots \geq \lambda_n$ and v_1, \dots, v_n be the eigenvalues and associated eigenvectors of B , and p be the number of positive eigenvalues. By forming the matrix $Y = (\sqrt{\lambda_1}v_1, \dots, \sqrt{\lambda_p}v_p)$, we obtain a p -dimensional point cloud in which each frame is represented in space by a single point. The distance between any two points i and j in Y approximates the dissimilarity between those frames as represented in D . We use randomly-initialized k -means with multiple runs to converge to a lowest-error clustering [9, 10].

To eliminate residual motion artifacts after frame gating, a more precise contour tracking operation is performed to provide pixel-wise correspondence. Two contours are drawn which define the inner and outer boundaries. Initialization for a single contour comes in the form of a manually-drawn contour on the first frame. Contours are found for all frames after the first by pairwise matching. Our method follows a two-step approach to complete the process: a rigid alignment step followed by an elastic refinement step. In the rigid step, starting with static-image contour x , the following rigid transformations are modeled to match the contour to the moving image: $\pm x$ translation, $\pm y$ translation, \pm rotation, and \pm dilation. Then, we proceed using a gradient ascent. In the elastic step, given the contour x' , which is itself a rigid transformation of the initial contour x , we deform x' elastically in order for it to better conform to the image features associated spatially with x . The output of this elastic matching step is a refined contour, x'' . We define a contour energy function for any contour, and by manipulating x' we seek to maximize the energy function to produce x'' . Lastly, if information about of the regions on the inside and outside of the contour is known, it is possible to use histogram statistics to influence the deforming contour.

Given tracked contour pairs for each image in our gated IVUS sequence, the region between the contours is resampled from image space into a rectangular region space. Difference imaging is accomplished by subtracting the pre-injection baseline from all region images in the gated sequence. As the same baseline is subtracted

from both the pre- and post-injection frames, it is simpler to determine which deviations from the baseline occur as a result of the contrast injection and which are the result of noise unrelated to the presence of contrast. To quantify enhancement in the difference-imaged regions of interest after they have been mapped back to the IVUS image space, we consider the set of pixels in the region of interest which are flagged as enhanced. Averaging the grey levels of these enhanced pixels, we obtain the average enhancement per enhanced pixel (AEPEP) statistic. Each frame in our gated sequence will have one associated AEPEP value; due to noise, pre-injection frames will have some positive value but, if enhancement occurs, the post-injection value will be greater.

4 Results

We applied our knowledge-based framework to detect and quantify the pericardial fat tissue in 300 non-contrast cardiac CT scans using the multi-class, multi-feature fuzzy connectedness method. We compared the results of our method (Automatic Fat Analysis in Computed Tomography - AFACT) to expert manual segmentations of pericardial fat. The manual delineation of the fat region in the CT images, performed by experienced physicians/radiologists, was used as the gold standard. We evaluated the results of our algorithm by computing accuracy, true positive rate and true negative rate. Specifically, we computed the false negatives (FN), false positives (FP), true negatives (TN), and true positives (TP) by computing the number of pixels that were classified as the background and the fat, both correctly and incorrectly. For 300 subjects, the mean accuracy for pericardial fat was $98.13\% \pm 2.2\%$. The mean true negative rate was $99.48\% \pm 2.2\%$. Finally, the mean true positive rate was $86.63\% \pm 9.32\%$ [2].

In vasa vasorum imaging, from 35 patients that were screened initially, 19 patients met the study criteria. The data from three patients were not analyzed due to poor image quality of the arterial layers, suboptimal infusion of the microbubbles, and displacement of the catheter. From the 16 patients, ten had unstable angina, four had ST elevation acute myocardial infarction, and two had non-ST elevation myocardial infarction. We performed three different types of analysis to evaluate the performance of our method. First, we examined the stability of the sequences before and after gating. "Stability" is measured as the mean cross-correlation between every frame pair in a sequence; this is performed for both the gated and ungated sequences. A significant increase was observed in mean cross-correlation for gated sequences (from 0.76 ± 0.08 to 0.84 ± 0.06). Second, for three normal sequences, the luminal border of 100 frames was manually segmented twice by the same observer. One of these sets of segmentations was used as a ground truth against which the other was compared; the root-mean-squared distance between corresponding points on these contours were then measured for each frame as a metric of error. All cases manifested the expected difference in error between our algorithm (2.73 ± 0.47 , 2.78 ± 0.91 , 4.46 ± 0.92 for cases 1, 2, and 3 respectively)

and normal intra-observer error (2.57 ± 0.74 , 3.35 ± 1.11 , 1.96 ± 0.97 for cases 1, 2, and 3 respectively). Note that these sequences were ungated, and hence in practice when we employ our gating algorithm as a pre-processing step, we can expect the error of the contour-tracking step to be lower. Last, contrast-enhanced IVUS studies were performed on the recordings from the 16 human patients. We observed an increase in the mean grey level of all examined regions as expressed by AEPEP from pre- to post-injection of microbubbles image. The percent increase of AEPEP in the region of interest after the injection of microbubbles was less than 10 % in 3 patients and 10–40 % in 13 patients [17, 18].

5 Discussion

We have presented a novel method that allows detection and quantification of pericardial fat in CT. Instead of using a fixed threshold, which is the standard method to detect fat in CT, our method is data-driven and adapts to the variations in individual scans. In addition to grey-level information, our method also uses additional derived information (e.g., inhomogeneity and texture features to detect pericardial fat). Our method is fully automatic except for the manual selection of top and bottom axial slices containing the heart. To our knowledge, our method is the first to use multiple features for adaptive data-driven classification of pericardial fat in CT. Clinical investigation of our fat quantification method is warranted to evaluate the role of pericardial fat in risk assessment.

We also presented a method which, for the first time, enabled (in vivo) imaging of extra-luminal perfusion under IVUS. Current evidence suggests that this perfusion is related to the presence of VV microvessels which, in turn, are a potential marker for plaque inflammation and consequent vulnerability. The significant changes in the IVUS signal after microbubble passage leave no doubt as to its ability to show contrast enhancement. The primary limitation of our method is the requirement that we image every area of interest twice: once before and once after the injection of contrast. However, this limitation is necessitated by our desire to restrict ourselves to currently-available IVUS hardware and standard contrast agents. As such, contrast-enhanced IVUS presents a promising imaging approach to the assessment of plaque vulnerability. We are currently in the process of validating our results with histological ground truth to obtain an estimate of the sensitivity of our approach in a clinical setting (Fig. 2(b-e)).

6 Conclusion

Our long term goal is to develop a new risk assessment scoring index for a patient's risk of a cardiovascular event. If these scores are to be used as markers of subclinical disease, they should provide the best representation of subclinical information.

This requires optimal use of all available data. Currently used tools do not use the wealth of information present in the images. In this chapter, we presented novel methods to mine information from CT and IVUS data to obtain pericardial fat measurements and to image extra-luminal perfusion due to blood flow through the vasa vasorum. The results we obtained are encouraging, and show the possibility of these methods being used in clinical settings.

Acknowledgment We would like to thank all members of the Ultimate IVUS team for their valuable assistance. This work was supported in part by NSF Grant IIS-0431144 and an NSF Graduate Research Fellowship (SMO). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

1. A. Bandekar, M. Naghavi, and I. Kakadiaris. Automated pericardial fat quantification in CT data. In *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pages 932–936, New York, NY, 2006.
2. A.N. Bandekar. *A Unified Knowledge-based Segmentation Framework for Medical Images*. PhD thesis, University of Houston, Dec. 2006.
3. R. Cooperand *et al.* Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States: Findings of the National Conference on Cardiovascular Disease Prevention. *Circulation*, 102(25):3137–3147, 2000.
4. D.E. Goertz *et al.* Subharmonic contrast intravascular ultrasound for vasa vasorum imaging. *Ultrasound Med Biol*, 33(12):1859–1872, December 2007.
5. I. Kakadiaris, S. O'Malley, M. Vavuranakis, S. Carlier, R. Metcalfe, C. Hartley, E. Falk, and M. Naghavi. Signal processing approaches to risk assessment in coronary artery disease. *IEEE Signal Processing Magazine*, 23(6):59–62, 2006.
6. M. Naghavi *et al.* From vulnerable plaque to vulnerable patient part III: Executive summary of the screening for heart attack prevention and education (SHAPE) task force report. *The American Journal of Cardiology*, 98(2):2–15, July 2006.
7. M. Naghavi *et al.* From vulnerable plaque to vulnerable patient: A call for new definitions and risk assessment strategies: Part I. *Circulation*, 108(14):1664–1672, October 2003. (In Press).
8. M. Naghavi *et al.* From vulnerable plaque to vulnerable patient: A call for new definitions and risk assessment strategies: Part II. *Circulation*, 108(15):1772–1778, October 2003.
9. S. O'Malley, S. Carlier, M. Naghavi, and I. Kakadiaris. Image-based frame gating of IVUS pullbacks: A surrogate for ecg. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, pages 433–436, Honolulu, Hawaii, 2007.
10. S. O'Malley, M. Naghavi, and I. Kakadiaris. Image-based frame gating for stationary-catheter IVUS sequences. In *Proc. Int. Workshop on Computer Vision for Intravascular and Intracardiac Imaging*, pages 14–21, Copenhagen, Denmark, 2006.
11. S. O'Malley, M. Naghavi, and I. Kakadiaris. One-class acoustic characterization applied to blood detection in ivus. In *Proc. Medical Image Computing and Computer-Assisted Intervention*, Brisbane, Australia, 2007.
12. S. O'Malley, M. Vavuranakis, M. Naghavi, and I. Kakadiaris. Intravascular ultrasound-based imaging of vasa vasorum for the detection of vulnerable atherosclerotic plaque. In *Proc. Int. Conf. on Medical Image Computing and Computer Assisted Intervention*, volume 1, pages 343–351, Palm Springs, CA, USA, 2005.

13. A. Pednekar and I.A. Kakadiaris. Image segmentation based on fuzzy connectedness using dynamic weights. *IEEE Trans. Image Processing*, 15(6):1555–1562, 2006.
14. A. Pednekar, U. Kurkure, I. A. Kakadiaris, R. Muthupillai, and S. Flamm. Left ventricular segmentation in MR using hierarchical multi-class multi-feature fuzzy connectedness. In *Proc. Medical Image Computing and Computer Assisted Intervention*, Rennes, Saint-Malo, France, 2004. Springer.
15. D. Rueckert *et al.* Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. on Medical Imaging*, 18:712–721, 1999.
16. R. Taguchi *et al.* Pericardial fat accumulation in men as a risk factor for coronary artery disease. *Atherosclerosis*, 157(1):203–9, 2001.
17. M. Vavuranakis, I. Kakadiaris, S. O’Malley, T. Papaioannou, E. Sanidas, S. Carlier, M. Naghavi, and C. Stefanadis. Contrast enhanced intravascular ultrasound for the detection of vulnerable plaques: a combined morphology and activity-based assessment of plaque vulnerability. *Expert Review of Cardiovascular Therapy*, 5:917–915, 2007.
18. M. Vavuranakis, I.A. Kakadiaris, S. M. O’Malley, T. G. Papaioannou, E. A. Sanidas, M. Naghavi, S. Carlier, D. Tousoulis, and C. Stefanadis. A new method for assessment of plaque vulnerability based on vasa vasorum imaging, by using contrast-enhanced intravascular ultrasound and differential image analysis. *Int. Journal of Cardiology*, 2008 (In Press).

Rheumatoid Arthritis Quantification using Appearance Models

G. Langs, P. Peloschek, H. Bischof, and F. Kainberger

Abstract Rheumatoid arthritis (RA) is a chronic disease that affects joints of the human skeleton. During therapy and during clinical trials, the accurate and precise measurement of the disease development is of crucial importance. Manual scoring frameworks exhibit high inter-reader variability and therefore constrain therapeutical monitoring or comparative evaluations during clinical trials.

In this chapter an automatic method for the quantification of rheumatoid arthritis is described. It is largely based on appearance models, and analyses a radiograph with regard to the two main indicators of RA progression: joint space width narrowing and erosions on the bones.

With the automatic approach a transition from global scoring methods that integrate over the entire anatomy, towards local measurements and the tracking of individual pathological changes becomes feasible. This is expected to improve both specificity and sensitivity of imaging biomarkers. It can improve therapy monitoring in particular if subtle changes occur, and can enhance the significance of clinical trials.

G. Langs (✉) • P. Peloschek • H. Bischof • F. Kainberger
Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16,
8010 Graz, Austria
e-mail: bischof@icg.tugraz.at

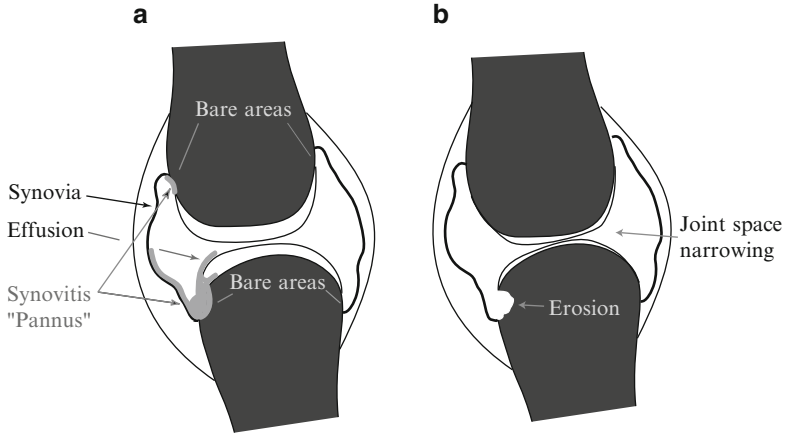


Fig. 1 A joint affected by synovitis and succeeding RA exhibiting joint space narrowing and erosive destructions in the bare area

1 Introduction

Rheumatoid arthritis (RA) is a chronic systemic inflammatory disease involving joints on e.g. the fingers, wrists and feet. It is chronic and progressive, and can result in severe pain. Recurring inflammation of the affected joints (i.e. arthritis) leads to a degradation of cartilage (*joint space narrowing*) and to erosive destructions (*erosions*) of the bone. If not treated properly RA can lead to the complete destruction of the joints. It affects physical function and mobility, and causes substantial short-term and long-term morbidity and increased mortality.

In Fig. 1a a joint affected by the disease is depicted. On the left side the course of the synovitis is shown. On its inner side the joint capsule is covered by the *synovia*. During the early phases of the disease the synovia becomes thicker and hypervascularized, also called synovitis (i.e. inflamed synovia, or *pannus*). The synovitis invades the bone and destroys the cartilage starting at the *bare areas* where the bone is not protected by cartilage. In Fig. 1b the two radiological markers for the disease progression are indicated: the thinned cartilage causes the joint space width to decrease since the two bones constituting the joint converge. The bare area lies between the cartilage and the osseous fixation of the joint capsule. It is not covered by cartilage, which has an osteo-protective effect against synovitis, thus the inflamed synovia starts here with the destruction of the bone (*erosions*).

1.1 *The evolving state of the art in RA assessment*

The radiographic surrogates associated with the progression of RA are *joint space narrowing (JSN)* which reflects an impairment of the cartilage and *erosions*, the destructions of the bony structure. The precise quantification of these markers is a decisive factor during treatment and during clinical multi-center trials. While ultrasound and magnetic resonance imaging are used to monitor synovitis, radiography is the standard modality for long-term RA progression monitoring [34]. Scoring of radiographic joint space width and erosions are accepted imaging biomarkers to assess the progression of RA during therapy.

Several manual scoring systems [21, 27] have been published and refined over the last 30 years. They use a discrete scale of categories to integrate the status of multiple joints. They are time consuming, require specialized training, and suffer from significant inter- and intra-reader variation [35]. This limits long-term assessment of disease progression, as well as the feasibility and discriminative power of multi-center studies.

Automated continuous local measurements can be expected to exceed traditional ordinal scoring scales in terms of sensitivity and precision.

The availability of digital image acquisition systems has prompted the development of new, increasingly automatic, quantitative measurement methods for radiographic features such as bone axes, bone density, or joint space width. In one study [1] an interactive joint space measurement method for rheumatoid arthritis of the finger joints gave highly reproducible results, thus increasing the study power while remaining consistent with traditional scoring.

Another interactive approach [25, 26], where points had to be placed manually in the metacarpophalangeal joint (MCP), achieved a considerably high reproducibility rating. In [8, 10] the diagnostic performance of a computer based scoring method was also found to surpass traditional scoring. These methods are restricted to angle- and distance measurements, include manual annotation on the image as part of the measurement procedure, and integrate only a limited degree of automation with user input. In [30] different methods for the measurement of the joint space width with various states of automation were compared. In [24] initial experiences with a fully automatic joint space width measurement method for RA were reported.

In a number of recent clinical studies [2, 3, 15, 22] erosions proved to provide more discriminative information on disease progression with respect to the treatment as opposed to the joint space width. The manual assessment of erosions in the traditional scoring frame work impedes the precise follow-up quantification and suffers from the afore-mentioned limitations.

2 Localized pathology quantification and tracking

A fully automatic local assessment of the pathological changes is highly desirable since it is a prerequisite for an accurate, reproducible and ultimately more sensitive measurement of the disease status. It is a development from discrete global scoring schemes towards continuous measurement methodologies [26, 28]. Model-based methods allow for a location consistent monitoring or *tracking* of subtle changes of anatomical structures. This can result in higher accuracy and becomes essential, due to advances in therapy that decelerate the disease development considerably [9]. Model-based strategies for an automated analysis of radiographs are applicable to osteoarthritis, spondylosis, osteoporosis, and axial malalignments in a straightforward manner.

Appearance models are used to solve three different problems during the disease assessment:

- Detection of the anatomical structures of interest in the radiography data.
- Consistent identification of individual positions in the anatomy over time for a single patient (pathology tracking), and localization of corresponding positions between patients.
- Analysis and measurement of the deviation from a healthy training population: the residual error after appearance model matching can be used to indicate pathologies if the training population is sufficiently representative.

In this chapter, we describe an automatic method to measure both joint space width and erosive destructions caused by RA. Parts of it have been proposed in [17, 19, 20, 24]. The basis of the approach are appearance models, and we will emphasize the particular uses one can make during the application of this family of methods.

3 Automatic Quantification of Rheumatoid Arthritis

Given a hand radiograph, the RA quantification is initialized by a local linear mapping which uses texture appearance, to locate the relevant joint positions. Subsequently active shape model (ASM) driven snakes identify the individual bone contours with a statistical shape and texture model. They account for deviations from the training population due to pathological changes. This allows for an accurate bone contour delineation even with difficult and ambiguous image data and results in a reproducible identification of landmark positions on the bones.

The joint space widths are measured between the detected contours and the location information covered by the model landmarks (i.e. the definition of the joint region on the bone). Erosions and their extent are determined by a boosted classifier ensemble that analyses texture features extracted from the bone contour region. It results in a label for each contour point describing the extension of erosive destructions. Furthermore erosions can be marked by the residual appearance model fitting error [20]. For both measures the model landmarks serve to establish correspondences across individuals, e.g. in order to define the joint space width measurement region and to track changes during follow-up assessment of one individual.

4 Locating the Joints: ShapeLLMs – Appearance Models and Local Linear Mappings

As a first step before segmenting the individual bone contours, a coarse estimate of the joint positions is necessary to provide the ASMs with a reliable initialization. The repetitive appearance and the variability of the hand positioning during acquisition is beyond the capture range of the ASMs that delineate the bone contours. Local linear mappings (LLMs) are trained to detect the positions of 12 joints (Fingertips, metacarpophalangeal joint metacarpophalangeal joints (MCP) and carpo metacarpal joints (CMC) of Fingers 2,3,4, and 5). They have been used for hand gesture recognition [23] and perform well on salient structures. In the case of radiographs they model dependencies between local image texture and landmark positions, and give reliable results for the initial detection of the hand, its orientation and the positions of the individual bones for the ASM initialization [17]. For each of the 12 joints a cascade of LLMs is trained on a series of increasingly small texture patches. The initial estimate is based on features extracted from a configuration of 12 overlapping windows on the radiograph. Subsequent iterations are based on features extracted locally at the current joint position estimate. The deformation of the configuration is constrained by a statistical shape model. In Fig. 2 an overview of this estimation is shown, a detailed explanation is given in [17, 19].

5 Delineating the Bone Contours: Active Shape Model Driven Snakes

Based on coarse joint position estimates active shape model (ASM) driven snakes [18, 19] delineate the bone contours and identify regions relevant to analysis. ASM driven snakes (ASMdS) segment the bones in a two-phase process: first they rely on a statistical shape model (ASM) to obtain a stable contour estimate, this is followed by an active contour approach, that controls the statistical model constraint in a gradual manner and adapts to deviations from the training population, while retaining the landmark identities.

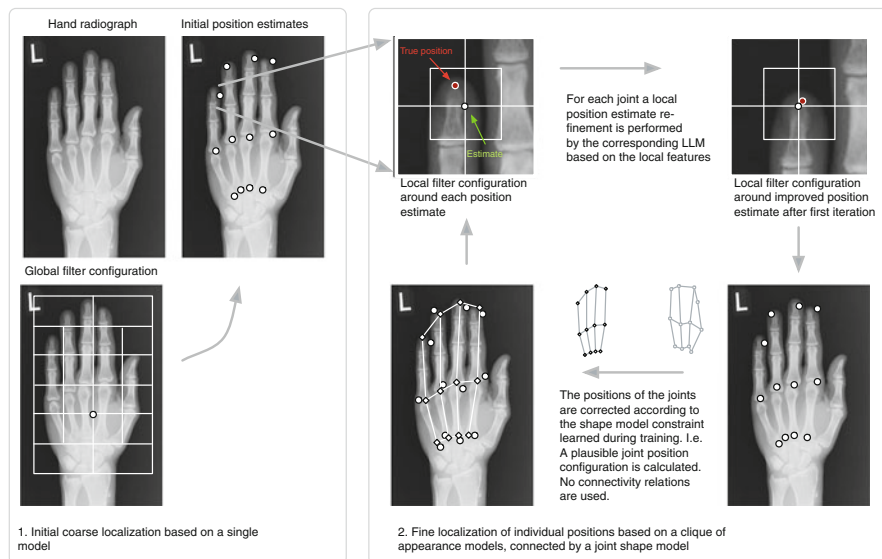
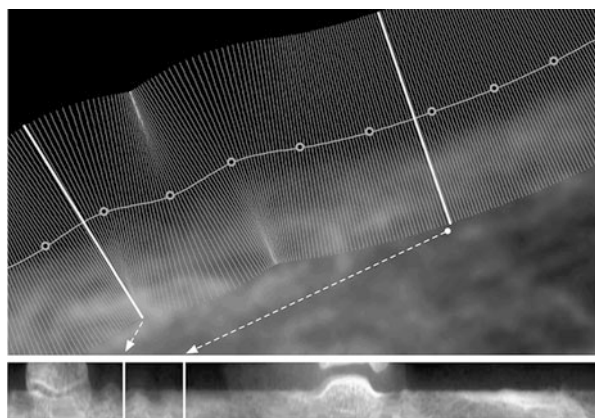


Fig. 2 Schematic overview of the joint localization by ShapeLLMs

Fig. 3 Extracting grey value profiles along the bone contour: a radiograph with ASM result landmarks and the profiles orthogonal to the interpolated contour; the resulting set of profiles extracted from the contour of a proximal phalanx, showing both adjoining bones, metacarpal and middle phalanx



5.1 ASM driven snakes

ASMs [6] are statistical landmark based models that represent the variation of shape and local gray values of a training set of examples in a compact manner. Although they provide very reliable estimates, the accuracy of ASMs with respect to the true image structure or contour is limited because of the shape model constraint that is based on a finite training population. This is a drawback if fine details that possibly deviate from the training population have to be analyzed, and the ASM training set stems from bones with no deviations from normal anatomy,

e.g. because the variability of local pathological changes cannot be modeled with a reasonable training set size and evades the representative power of a linear model. A similar issue was addressed in [7]. ASM driven snakes overcome this constraint by gradually decreasing the influence of the model constraint. They refine the contour estimate obtained by ASM and fit slight deviations from the normal training anatomy. However, they retain the location information captured by the model landmarks. The result is a dense delineation of the bone contour, and known positions of the landmarks on this contour.

5.2 ASM driven snakes search

The ASM search results in landmark estimates $x^{ASM} = (x_1, \dots, x_n)$. These landmarks are interpolated by a spline which is sampled at smaller intervals: $\mathbf{c} = (c_1, \dots, c_d)$. Gray level profiles p_i with length m are extracted from the image orthogonally to the interpolated contour. They build a matrix $\mathbf{P} = [p_1, \dots, p_d]$. Figure 3 shows the original positions of the grey level extraction points for a contour section of a proximal phalanx, and the set of extracted profiles $\mathbf{P}(x, y)$ for the entire bone contour. The spline interpolation in between the ASM landmarks is positioned at points $\mathbf{s}_1 = (x, y)_{i=1, \dots, d}: (x, y)_i = (i, m/2)$ in $\mathbf{P}(x, y)$, and can serve as initialization for a search with an active contour in \mathbf{P} . As in the standard active contour search, we minimize an energy functional

$$E = \int_0^1 \frac{1}{2} \left(\alpha |s'(s)|^2 + \beta |s''(s)|^2 \right) + E_{ext}(s(s)) ds \quad (1)$$

where the first term is the internal energy E_{int} capturing elasticity and rigidity properties of the curve, and the second term is the external energy E_{ext} . Their influence is determined by the factors α and β . The external energy E_{ext} is derived from the image and usually drags the contour towards high gradient edges in the data. In our case the local texture model of the ASM is taken advantage of by deriving E_{ext} from the image in the following way: the mean grey-level profiles corresponding to the landmarks in the ASM: $g_1^{ASM}, \dots, g_n^{ASM}$ are interpolated in order to derive an ordered set g_1, \dots, g_d . Thus, for every column in \mathbf{P} a corresponding grey level profile g_i is generated approximating the mean appearance in the training set from the model grey value profiles. The identification between g_i and p_i is defined by the landmarks on the fitted ASM. Then

$$E_{ext} = -\gamma |\nabla \mathbf{P}| - \mu \mathbf{P} \star \mathbf{G} \quad (2)$$

where $\mathbf{P} \star \mathbf{G}$ denotes the column wise convolution of p_i with the corresponding mean grey value profile g_i . Thus, with a high value for μ the lowest energy is obtained if the snake follows a path corresponding to a maximal similarity of the

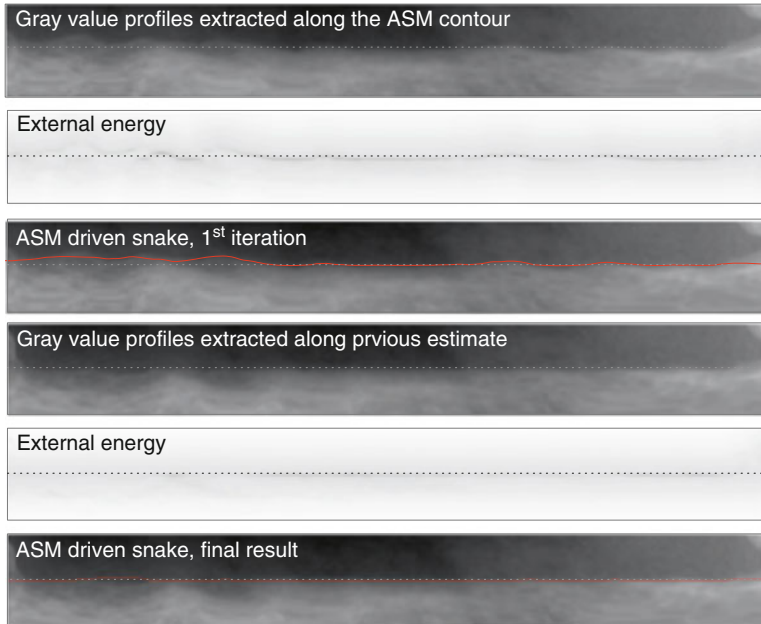


Fig. 4 ASM driven snakes; alternating: section of the grey value profiles extracted along the contour after ASM search, the resulting E_{ext} and the snake result

actual profile and the expected ASM profile. A high γ decreases the cost analogous to a traditional external force field depending on the local gradient. This can support the search if the convolution with the learned profiles does not result in a strong minimum, e.g. if the variability in the training set appearance at the contour position results in an undistinctive mean profile. In such cases the gradient in the search image still can provide information on the contour location. The lowest cost path passing $\mathbf{P}(x, y)$: \mathbf{s}_2 is a new estimate of the bone contour. Let $\mathbf{c}^2 = (c_1^2, \dots, c_d^2)$ be the corresponding curve in the image \mathbf{I} . Then $\mathbf{P}_2(x, y)$ is recalculated and the snake is fitted again. The resolution of the succeeding profile sets can be increased while the search region decreases.

In the case of bones exhibiting pathological changes caused by rheumatoid arthritis an outward pressure force can support convergence by preventing the snake from being captured on isolated local minima within the bone. In Fig. 4 a section of the profile sets extracted from a proximal phalanx, the corresponding cost-maps and the resulting snakes are depicted for the first and the fourth iteration of the process. In experiments with bones showing mild to moderate rheumatoid arthritis this number usually sufficed to give good results.

The final active contour result in $\mathbf{P}_{final}(x, y)$, \mathbf{s}_{final} is projected on the image resulting in the dense bone contour estimate \mathbf{c}_{final} . The corrected positions of the landmarks \hat{x}_i lie on this curve. The delineation of the bone contour thus retains the ASM landmarks, making a repeatable addressing of positions on the bone possible.

Fig. 5 The region for joint space width measurement on the MCP joint marked by model landmarks, and measurements for two MCP joints

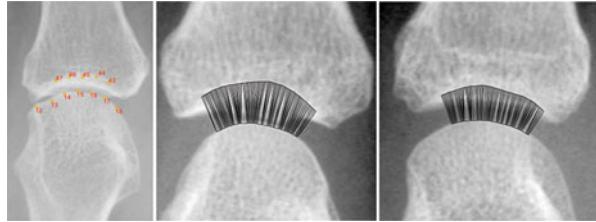
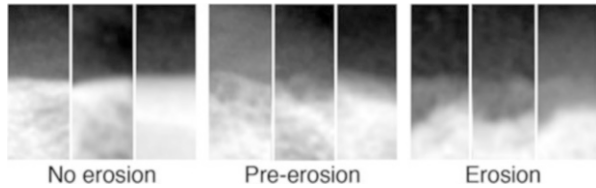


Fig. 6 Examples for patches extracted from the bone contour showing three different cases: no erosion, pre-erosion and erosion



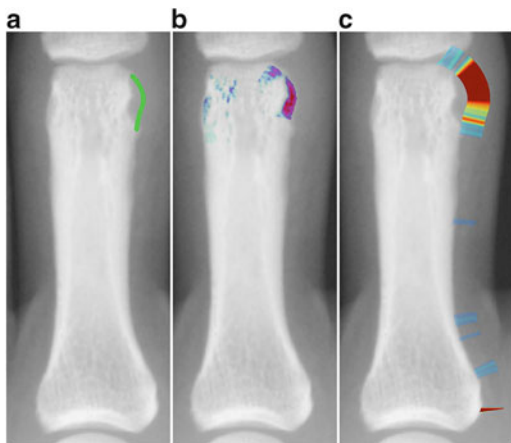
6 Measuring Joint Space Width

The joint space width can be measured straightforwardly after the bone segmentation. Since the development over time or *joint space narrowing (JSN)* is clinically relevant, the precision of the measurement is of prime importance. To measure the joint space width of the metacarpal phalangeal joint (MCP) ASM landmarks in the proximal region of the proximal phalanx (PP) and in the distal region of the metacarpal (MC) identify the measurement region. They are depicted in Fig. 5. For each point on an interpolation of the landmarks the minimum distance to the MC bone is measured, the mean value of these measurements within the region is defined as the JSW. In Fig. 5 an example of the resulting measurements for 2 MCP joints is depicted. The vertical lines indicate the closest neighbor on the MC for every point in the measurement region on the PP contour. In contrast to approaches that rely on measuring gradients along a line that passes through the joint region this method takes advantage of a model fitted to the entire bone, hence improving stability in case of possible ambiguous local texture or centre line orientation in the joint region.

7 Detecting and Visualizing Erosions

After bone contours have been delineated, the task of detecting erosions on the bone contour can be formulated as a classification of contour points into the classes healthy bone contour, i.e. *Non-Erosion* and contour affected by RA, i.e. *Erosion*. The result is a continuous value that captures the amount of affected bone contour. For this, texture features are extracted from a sequence of rectangular patches along

Fig. 7 Marked Erosions:
 a. ground truth, b. classifier
 result, and c. visualization
 with residual error



the bone contour. The dimensionality of the feature vector is reduced by feature selection [16] and the classification is accomplished by an AdaBoost classifier [11] for each point on the contour. In order to account for the high variability of erosions [32] two classes are referred to, *Erosion I* or *pre-erosions* and *Erosion II*, that differ in their appearance. Class II erosions are unequivocal erosions exhibiting all radiographic signs, while Class I erosions lack one or more of these features, and appear as an unsharpening of bone architecture (Fig. 6).

7.1 Detection

For each point $c_i \in \mathbf{c}_{final}$ on the bone contour, the bone texture is extracted in the form of a rectangular patch \mathbf{p}_i with borders parallel and orthogonal to the bone contour normal vector n_i at the position s_i . This results in a set of patches $(\mathbf{p}_i)_{i=1, \dots, n}$ for a bone, each corresponding to a single contour point (Fig. 6). Erosion detection performs a classification on these patches and assigns them one of the two classes *Erosion* or *Non Erosion*.

For the patches \mathbf{p}_i texture features comprising gradient, gradient orientation, and grey value deviation on a grid of 10 sub-patches are extracted. Before a classifier is trained, the number of features is decreased by a feature selection procedure that is based on an iterative re-weighting scheme similar to AdaBoost [16]. The point-wise classification of the bone contour is accomplished by an AdaBoost classifier working with linear discriminant analysis as a weak learner. Each weak learner is trained on a sub-set of example patches and features. Thereby the variability of erosion appearance can be accounted for. The classification of the contour points

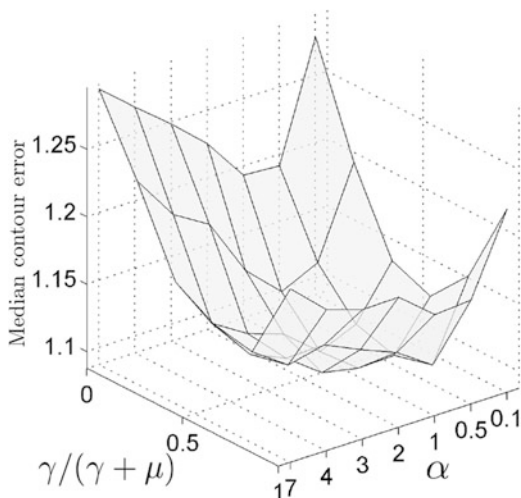
is achieved by a voting of 4 classifiers. For each of the erosion classes type 1 and type 2, two classifiers are trained: one on the entire body bone contour C_1^f and one in the joint region C_2^j . The combined classifier is defined by $C = C_2^j C_2^f + C_1^j C_1^f \hat{C}_1^f \hat{C}_2^f$, where \hat{C} is the indicator function of $C > t$. The inputs for the classifiers are the features extracted from a single patch, the output is the class label *No-Erosion* vs. *Erosion*. The voting results in a value indicating the presence of erosive destructions. For each contour point a classifier response is calculated, and thereby the extent and the location of the erosions are determined.

7.2 Visualization

In addition to the numerical value associated to each contour point, a more specific visualization of the deviation of erosion appearance from normal anatomy is necessary to allow the clinician to approve the results. This can be achieved by a generic appearance model in a straightforward fashion. To visualize the deviation of the observed local bone appearance from healthy bone texture a model is built from healthy training regions on bones and the best fit of the model is compared to the texture along the bone contour [20]. The appearance of healthy patches is modeled by *Gaussian mixture models (GMMs)* [4]. These models are generative models, and are able to synthesize data examples, which are plausible with regard to the training data. During the visualization the generative model of an intact bone contour is used to determine the deviation of erosion appearance from intact bone texture.

The GMM of patch appearance is a generative model. That is, it can simulate bone appearance that is plausible with regard to the training population of bone texture patches. By fitting a model to a patch, i.e., by *reconstructing* it, the algorithm determines the closest estimate of the patch appearance within the constraints imposed by the statistical model, i.e., the distribution of model parameters in the training set. The residual error, the difference between the actual appearance and the reconstruction by the model provides information about the local texture and its difference from the learned distribution. Showing this residual in the radiograph in addition to the labels C from the classifier, provides the musculoskeletal radiologist not only with location information but also with an estimate about the deviation of the bone contour texture from healthy bone. Note that the mere residual error could also be used for the detection of erosions. However, experiments indicate, that the discriminative power is not sufficient for classification, in particular in the case of pre-erosions. For the high variation of erosion appearances a classifier utilizing texture as described above is necessary.

Fig. 8 Median contour error for ASM driven snakes. The x-axis corresponds to the ratio $\gamma/(\gamma + \mu)$ (Eq. 2), the y-axis corresponds to the snake elasticity parameters α i.e. the higher α , the less elastic the snake is



8 Experimental Results and Discussion

In this section we report experimental results on several data sets of hand radiographs with mild to moderate rheumatoid arthritis. More detailed evaluations are reported in [19, 20, 24]. A comparative study of different approaches to semior fully-automatically measure the joint space width is reported in [30].

Contour delineation: The effect of ASM driven snakes is important, if small local deviations from the training population are to be expected. For meta-carpal and proximal phalangeal bones and a varying ratio between γ and μ , and elasticity parameter α , a minimum median error is reached by ASM driven snakes for intermediate values. In Fig. 8, the median contour error is plotted. For standard ASMs the mean/median error is 2.41/2.03px. The mean ASMdS error reaches its lowest value 1.73 for high α and highest influence γ of the model gray value model. The median exhibits a minimum of 1.09 within the parameter range. Results indicate a better fit of the contour estimate toward small local *rough* sections of the bone contour compared to standard ASMs. These occur particularly often on bones of patients suffering from chronic joint disease.

Joint space width measurement: For joint space narrowing i.e. the development over time the precision of the measurement method is of prime interest. For 10 repeated measurements on 160 MCP joints with a mean JSW of 1.75 mm, the mean absolute error of the automatic measurement results compared to a standard of reference annotated by experts was 0.19 to 0.4 mm. The standard deviation for 10 repeated measurements was between 0.04 to 0.13 mm corresponding to a coefficient of variation of 2% for non-overlapping bones and 7% if overlap did occur at the joint. The smallest detectable difference (SDD) for the joint space width is 0.08 mm. The joint space width measurement offers precision similar to existing semi-automatic approaches [5, 13, 14] but in a fully automatic fashion.

Erosions: When comparing automatic erosion detection (i.e. contour classification) with a manual expert standard of reference by a ROC analysis, the area under the ROC curve (AUC) is 0.89 (0.88 for class 1, and 0.92 for class 2). The AUC indicates the probability of a detected erosion coinciding with a manual standard of reference erosion [12]. In Fig. 7 an example bone with standard of reference annotation, detected, and marked erosion is depicted. There are no alternative automated erosion detection algorithms reported in previous literature. However for a group of expert readers overall scoring variability was reported in [29]. The intra-class correlation coefficient (ICC) [31] for this group ranged from 0.465 to 0.999, which indicates that absolute scores cannot be compared across readers or different studies, as mentioned in Sect. 1.1.

9 Conclusion

The method explained in this chapter automatically measures joint space width and erosion extent on hands affected by rheumatoid arthritis. We expect the automatic assessment of rheumatoid arthritis and other diseases to decrease inter-reader variability and artifacts introduced by individual readers as described in [29, 33]. Increasing reliability and sensitivity in detecting treatment effects would help to speed up the development of new and effective disease-controlling, anti-rheumatic therapies and to reduce the number of patients necessary for clinical trials. It is particularly relevant, since improved therapies make the changes caused by RA more subtle. This makes a specific and local tracking of changes as opposed to global and therefore less sensitive scores necessary. Future research in the field of automated quantification methods will have to focus on the improvement of their ability to adapt to different image acquisition protocols and to extend the area of application to more complex anatomical sites.

References

1. A. Angwin, A. Lloyd, G. Heald, G. Nepom, M. Binks, and M. James. Radiographic hand joint space width assessed by computer is a sensitive measure of change in early rheumatoid arthritis. *J Rheumatol*, 31:1062–1072, 2004.
2. B. B. Bresnihan, R. Newmark, S. Robbins, and H. Genant. Effects of anakinra monotherapy on joint damage in patients with ra. extension of a 24-week randomized, placebo-controlled trial. *Journal of Rheumatology*, 2004.
3. J. M. Bathon, R. W. Martin, R. M. Fleischmann, J. R. Tesser, M. H. Schiff, E. C. Keystone, M. C. Genovese, M. C. Wasko, L. W. Moreland, A. L. Weaver, J. Markenson, and B. K. Finck. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med*, 343(22):1586–1593, 2000.
4. C. Bishop. *Neural Networks for Pattern Recognition*. Oxford Univ. Press, 1995.

5. J. Buckland-Wright, D. Macfarlane, S. Williams, and R. Ward. Accuracy and precision of joint space width measurements in standard and macroradiographs of osteoarthritic knees. *Ann Rheum Dis*, 54:872–880, 1995.
6. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proceedings of BMVC'92*, pages 266–275, 1992.
7. T. F. Cootes and C. J. Taylor. Combining elastic and statistical models of appearance variation. In *ECCV (1)*, pages 149–163, 2000.
8. J. Duryea, Y. Jiang, M. Zakharevich, and H. Genant. Neural network based algorithm to quantify joint space width in joints of the hand for arthritis assessment. *Med. Phys.*, 27(5):1185–1194, 2000.
9. A. Finckh, H. Choi, and F. Wolfe. Progression of radiographic joint damage in different eras: trends towards milder disease in rheumatoid arthritis are attributable to improved treatment. *Ann Rheum Dis*, 65(6):1192–1197, 2006.
10. A. Finckh, P. de Pablo, J. N. Katz, G. Neumann, Y. Lu, F. Wolfe, and J. Duryea. Performance of an automated computer-based scoring method to assess joint space narrowing in rheumatoid arthritis, a longitudinal study. *Arthritis and Rheumatism*, 54(5):1444–1450, 2006.
11. Y. Freund and R. Shapire. A decision-theoretic generalization of online learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139, 1997.
12. J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982.
13. J. Higgs, D. Smith, K. D. Rosier, and R. C. Jr. Quantitative measurement of erosion growth and joint space loss in rheumatoid arthritis hand radiographs. *J Rheumatol*, 23:265–272, 1996.
14. M. James, G. Heald, J. Shorter, and R. Turner. Joint space measurement in hand radiographs using computerized image analysis. *Arthritis Rheum*, 38:891–901, 1995.
15. E. Keystone, A. Kavanaugh, J. Sharp, H. Tannenbaum, Y. Hua, L. Teoh, S. Fischkoff, and E. Chartash. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human antitumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: A randomized, placebo-controlled, 52-week trial. *Arthritis and Rheumatism*, 50(5):1400–1411, 2004.
16. P. Křížek, J. Kittler, and V. Hlaváč. Feature selection based on the training set manipulation. In *Proceedings of ICPR'06*, volume 2, pages 658–661, 2006.
17. G. Langs. *Autonomous Learning of Appearance Models in Medical Image Analysis*. PhD thesis, Graz University of Technology, Institute for Computer Graphics and Vision, May 2007.
18. G. Langs, P. Peloschek, and H. Bischof. ASM driven snakes in rheumatoid arthritis assessment. In *Proceedings of 13th Scandinavian Conference on Image Analysis, SCIA 2003, Goeteborg, Schweden*, pages 454–461. Springer, 2003.
19. G. Langs, P. Peloschek, H. Bischof, and F. Kainberger. Automatic quantification of joint space narrowing and erosions in rheumatoid arthritis. Submitted to *IEEE Transactions on Medical Imaging*, Under Review.
20. G. Langs, P. Peloschek, H. Bischof, and F. Kainberger. Model-based erosion spotting and visualization in rheumatoid arthritis. *Acad Radiol*, 14(10):1179–1188, 2007.
21. A. Larsen, K. Dale, and M. Eek. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiol Diagn*, 18:481–491, 1977.
22. P. E. Lipsky, D. van der Heijde, E. W. S. Clair, D. E. Furst, F. C. Breedveld, J. R. Kalden, J. S. Smolen, M. Weisman, P. Emery, M. Feldmann, G. R. Harriman, and R. N. Maini. Infliximab and methotrexate in the treatment of rheumatoid arthritis. *N Engl J Med*, 343(22):1594–1602, 2000.
23. C. Noelker and H. Ritter. GREFFIT: Visual recognition of hand postures. In *Gesture Workshop*, pages 61–72, 1999.
24. P. Peloschek, G. Langs, M. Weber, J. Sailer, M. Reissegger, H. Imhof, H. Bischof, and F. Kainberger. An automatic model-based system for joint space measurements on hand radiographs: Initial experience. *Radiology*, 245(3):855–862, 2007.
25. J. Sharp. Measurement of structural abnormalities in arthritis using radiographic images. *Radiol Clin North Am*, 42(1):109–119, 2004.

26. J. Sharp, J. Gardner, and E. Bennett. Computer-based methods for measuring joint space and estimating erosion volume in the finger and wrist joints of patients with rheumatoid arthritis. *Arthritis and Rheumatism*, 43(6):1378–1386, 2000.
27. J. Sharp, M. Lidsky, L. Collins, and J. Moreland. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. *Arthritis and Rheumatism*, 14:706–720, 1971.
28. J. Sharp, D. van der Heijde, J. Angwin, J. Duryea, H. Moens, J. Jacobs, J. Maillefert, and C. Strand. Measurement of joint space width and erosion size. *J Rheumatol*, 32(12):2456–2461, December 2005.
29. J. Sharp, F. Wolfe, M. Lassere, MaartenBoers, D. von der Heijde, A. Larsen, H. Paulus, R. Rau, and V. Strand. Variability of precision in scoring radiographic abnormalities in rheumatoid arthritis by experienced readers. *Journal of Rheumatology*, 31(6):1062–1072, 2004.
30. J. T. Sharp, J. Angwin, M. Boers, J. Duryea, G. von Ingersleben, J. R. Hall, J. A. Kauffman, R. Landewé, G. Langs, C. Lukas, J.-F. Maillefert, H. J. B. Moens, P. Peloschek, V. Strand, and D. van der Heijde. Computer based methods for measurement of joint space width: Update of an ongoing omeract project. *Journal of Rheumatology*, 34(4):874–83, 2007.
31. P. Shrout and J. Fleiss. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2):420–428, 1979.
32. O. Sommer, A. Kladosek, V. Weiler, H. Czembirek, M. Boeck, and M. Stiskal. Rheumatoid arthritis: a practical guide to state-of-the-art imaging, image interpretation, and clinical implications. *Radiographics*, 25:381–398, 2005.
33. H. Swinkels, R. Laan, M. van 't Hof, D. van der Heijde, N. de Vries, and P. van Riel. Modified sharp method: factors influencing reproducibility and variability. *Semin Arthritis Rheum*, 31(3):176–190, Dec 2001.
34. D. van der Heijde. Radiographic imaging: the 'gold standard' for assessment of disease progression in rheumatoid arthritis. *Rheumatology*, 39(Suppl 1):9–16, 2000.
35. D. v.d.Heijde, A. Boonen, M. Boers, P. Kostense, and S. van der Linden. Reading radiographs in chronological order, in pairs or as single films has important implications ofor the discriminative power of rheumatoid erthritis clinical trials. *Rheumatology*, 38:1213–1220, 1999.

Medical Image Processing for Analysis of Colon Motility

N. Navab, B. Glocker, O. Kutter, S.M. Kirchhoff,
and M. Reiser

Abstract A precise analysis and diagnosis of colon motility dysfunctions with current methods is almost unachievable. This makes it extremely difficult for the clinical experts to decide for the right intervention such as colon resection. The use of Cine MRI for visualizing the colon motility is a very promising technique. In addition, if image segmentation and qualitative motion analysis provide the necessary tools, it could provide the appropriate diagnostic solution. In this work we define necessary steps in the image processing chain to obtain clinical relevant measurements for a computer aided diagnosis of colon motility dysfunctions. For each step, we develop methods for an efficient handling of the MRI time sequences. There is need for compensating the breathing motion since no respiratory gating can be used during acquisition. We segment the colon using a graph-cuts approach in 2D over time for further analysis and visualization. The analysis of the large bowel motility is done by tracking the diameter of the colon during the propagation of the peristaltic wave. The main objective of this work is to automatize the assessment of clinical parameters which can be used to define a clinical index for motility pathologies.

N. Navab (✉) • O. Kutter
Institut fuer Informatik, Computer Aided Medical Procedures (CAMP), Technische Universitaet Muenchen, Boltzmannstr. 3, 85748 Garching, Germany
e-mail: navab@cs.tum.edu; oliver.kutter@stryker.com

S.M. Kirchhoff • M. Reiser
Institut fuer Klinische Radiologie, Ludwig-Maximilians-Universitaet Muenchen, Marchioninstr. 15, 81377 Muenchen, Germany
e-mail: sonja.kirchhoff@med.uni-muenchen.de; maximilian.reiser@med.uni-muenchen.de

B. Glocker
Department of Computing, BioMedIA Group, Imperial College London,
180 Queen's Gate, London SW7 2AZ, UK
e-mail: b.glocker@imperial.ac.uk

1 Introduction

Dysfunctions of large bowel motility presents a common problem in our society that may be attributable to a great variety of possible etiologies resulting most commonly in either constipation or diarrhea. Today's available diagnostic imaging techniques, such as bowel enema and the determination of the bowel transit time by the application of radiopaque markers, provide only snapshots of the dynamic large bowel motility and approximate clues concerning motility dysfunctions. These examination techniques cannot sufficiently demonstrate bowel movement in real-time.

To date, only the pancolononic manometry technique has been used to evaluate colonic motor function over either the entire length or limited segments of the human colon [1, 8, 9]. Manometry as well as the barostat technique present both scientifically established but very complicated methods with several drawbacks such as the invasive examination itself being very time intensive and inconvenient for the patients and thus, both methods are rarely used in clinical routine. Noninvasive, well-established and examiner-independent methods are Scintigraphic examinations. They allow for determining the colonic transit time. However, a considerable radiation exposure is associated with this method [2]. Thus, classical examination techniques are not suited to monitor large bowel motility. The ideal technique to visualize and quantify large bowel motility would allow for a fast and repeated imaging over larger time frames with a high temporal resolution reducing movement and respiratory artifacts. Additionally, such a technique should allow to include newly available technologies such as biofeedback, electrical stimulation of intestinal pacemakers, or the administration of specific stimulating drugs [6].

Magnetic resonance imaging (MRI) allows noninvasive ultrafast dynamic imaging with a high soft tissue contrast. The potential of functional *Cine MRI* for the visualization of the abdominal organs has been presented previously [12] and makes this technique a very promising tool in our application because it allows for the visualization of morphology and function of the large bowel at the same time, provided that fast image acquisition is used. However, without any stimulation, the activity of the large bowel exhibits a broad range of individual differences, such that an examination within an acceptable time frame of approximately 30 minutes is not feasible [7, 12]. To this end, different prokinetic agents were used in our previous studies [6] in order to achieve a predictable activity of the large bowel which lead to the achievement of the first real-time visualization of the large bowel movement and its peristaltic wave (see Fig. 1). A further and detailed analysis of large bowel motility and possible dysfunctions provides an immense progress in diagnosis of the lower gastrointestinal tract also concerning an individual and more suitable therapeutical approach from surgical as well as the internal medical side.

In order to integrate such a novel imaging technique into the daily clinical routine, it is mandatory to develop computer based analysis approaches. The huge amount of data which is acquired per patient needs automatic and semi-automatic methods in order to support the clinical expert in findings and diagnosis. Our main objective in

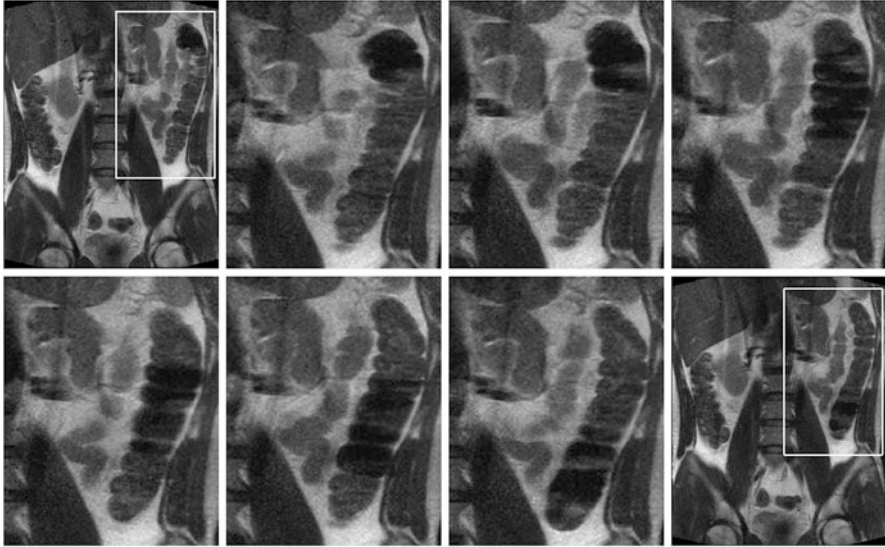


Fig. 1 Coronal Cine MRI slices visualizing the peristaltic wave in the descending colon

this work is to model and quantify the activity of the large bowel hopefully providing the means to define the clinical significance of a variety of motility disorders in a wide range of patients. Necessary steps in the image processing workflow are defined, and technical approaches towards a computer aided diagnosis tool are proposed.

In the following, we present our preliminary results on a set of experimental data. Volunteer data was acquired and processed using the imaging techniques and algorithms presented in the next Sections.

2 Image Acquisition using Cine MRI

The volunteers undergo functional Cine MRI. The standardized functional Cine MRI examination is performed at 6 AM, after a minimum starving phase of 8 hours, on 1.5-T system (Siemens Avanto). The volunteers are examined in supine position. Neither premedication nor contrast agent is applied. The dynamic part of the examination consists of 2 blocks of repeated measurements covering the entire abdomen, using a T2-weighted HASTE-sequence. Each block contains a stack of approximately 200 slices over time orientated in coronal plane adapted to the anatomic course of the descending colon (see Fig. 1). The image resolution is 256×320 . Between the 2 dynamic blocks of measurements a prokinetic agent is administered in order to stimulate the colon motility. For the further analysis of the



Fig. 2 Our image processing chain within a computer aided diagnosis tool for colon motility dysfunctions

peristaltic motion, the subsequence (usually about 20 slices) showing this motion is manually selected from the image blocks. The pre-scan without stimulation was up to now only used in our previous studies [6, 7].

In order to achieve the fastest possible frame rate for our MRI acquisition, no respiratory gating can be used during the scans. The time between two successive frames could be reduced to approximately 1.4 seconds which is fast enough to visualize the peristaltic wave. Still, the sequences suffer from breathing motion artifacts which makes the identification of corresponding points in the colon quite hard. This was not a problem in our previous studies where manually extracted diameters were measured over time and stimulated and non-stimulated sequences were compared, the identification of corresponding points could be achieved manually by the medical expert. In order to automatize such procedures, there is need for an accurate motion compensation in a post-processing step. This leads us to the first step in the image processing chain (see Fig 2).

3 Image Processing

In order to develop a computer aided diagnosis tool for colon motility dysfunctions, we first identify necessary steps within the image processing chain. We already mentioned the problem of breathing motion artifacts in the Cine MRI data sets. Once, these artifacts can be successfully removed, all further steps will benefit from the motion compensation. Our later analysis of the motility is based on the segmentation of the colon in all slices over time. We propose a semi-automatic approach based on interactive graph-cuts. This will be explained in more detail in Sect. 3.2. The actual analysis and our approach for extracting clinically valuable measurements is then described in Sect. 4. The full image processing chain is sketched in Fig. 2.

3.1 Motion Compensation

As already mentioned, no respiratory gating techniques are used during the image acquisition in order to achieve a high frame rate. The resulting breathing artifacts in the image sequences are visible in a vertical jumping of the abdominal organs effected by breathing motion such as liver, kidney, and of course the colon itself. For our further processing steps, a compensation of this motion is of great interest. In order to make the identification of corresponding points in the colon much

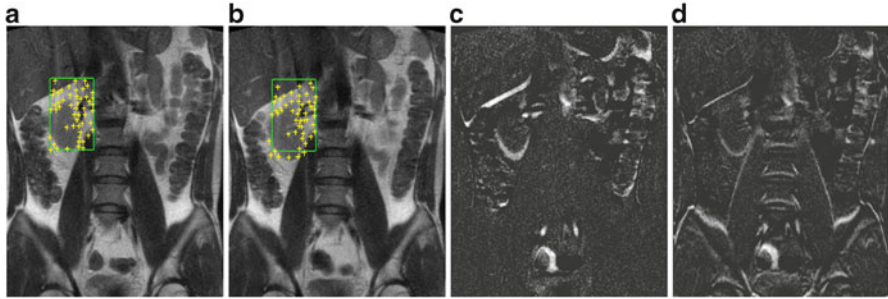


Fig. 3 (a) Selected region and extracted Harris feature points. (b) Tracked features in one of the following frames. (c) Difference image between reference frame and consecutive frame before compensation. (d) and the difference image after compensation. Clearly visible, the motion at the lower boundary of the liver has been reduced

easier we try to stabilize the image parts of interest. We propose a semi-automatic motion compensation method. Therefore, the user selects a subregion in one of the slices, which we call the reference frame. The selected region should represent the overall breathing motion but should not include any parts of the colon. So we can avoid to eventually compensate for colon motility. Within the selected region features are extracted which fulfill the Harris [10, 14] assumptions. The robust and fast implementation of the Lucas-Kanade optical flow method [3] is then used to compute the displacements of each single feature in every frame in respect to the reference frame. Afterwards we compute a mean vertical displacement for every frame which represents the overall breathing motion. By translating each frame by its corresponding mean displacement we can compensate for this motion. In practice, the image part containing the lower liver boundary and right kidney turned out to be a good region for tracking the breathing motion (see also Fig. 3). These parts show a very similar breathing motion such as the colon itself. The result of the motion compensation is a clear stabilization of all organs with a similar movement. Naturally, former stable parts (e.g. the spine, see also Fig. 3d) are consecutively jumping within the image series. However, this fact does not present a problem for the further processing of the colon.

3.2 Colon Segmentation

The segmentation of the colon is crucial for our further analysis. Shape and appearance have to be well preserved by the segmentation. The individual patient's anatomy is particularly reflected in varying orientation and form of the large bowel. A segmentation method should be highly flexible to handle these variations. To this end, we use the interactive graph-cuts approach proposed by Boykov and Jolly [4]. The segmentation is defined as an energy formulation

$$E(A) = R(A) + \lambda \cdot B(A) \quad (1)$$

where A indicates a segmentation of the pixels x of domain Ω of the image series I into two subsets \mathcal{O} (object pixels) and \mathcal{B} (background pixels) with

$$A_{x \in \Omega} = \begin{cases} \text{"obj"} & \text{if } x \in \mathcal{O} \\ \text{"bkg"} & \text{if } x \in \mathcal{B} \end{cases} \quad (2)$$

Here, the regional term R represents a priori knowledge given through the user interactions. Interactively, the user sets so-called seed brushes for the object \mathcal{S}_{obj} that is considered to be segmented (i.e. the colon) and additionally, seeds for the background \mathcal{S}_{bkg} (see Fig. 4a). The function R is then defined as

$$R(A) = \sum_{x \in \Omega} R_x(A_x) \quad (3)$$

where

$$R_x(A_x) = \begin{cases} \infty & \text{if } A_x = \text{"obj"} \wedge x \in \mathcal{S}_{bkg} \\ \infty & \text{if } A_x = \text{"bkg"} \wedge x \in \mathcal{S}_{obj} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Intuitively, the regional term forces the pixels belonging to seed brushes to keep their assignment to the object respectively background segmentation subset. The second part B of the segmentation energy is the so-called boundary term. Here, it represents the interaction energy for pairs of neighboring pixels $x, y \in \mathcal{N}$ to belong to the same segmentation subset and thus provides a certain smoothness on the segmentation result, or

$$B(A) = \sum_{x, y \in \mathcal{N}} B_{x, y} \cdot \delta(A_x, A_y) \quad (5)$$

with

$$\delta(A_x, A_y) = \begin{cases} 1 & \text{if } A_x \neq A_y \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

In our experiments, we use a simple boundary term based on intensity differences which is define in terms of a penalty function, or

$$B_{x, y} \propto \exp\left(-\frac{(I_x - I_y)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(x, y)} \quad (7)$$

This function penalizes discontinuities in the segmentation result for neighboring pixels of similar intensities. The weighting λ controls the influence of the boundary term. We set this value to 0.01 which was determined empirically for our kind of

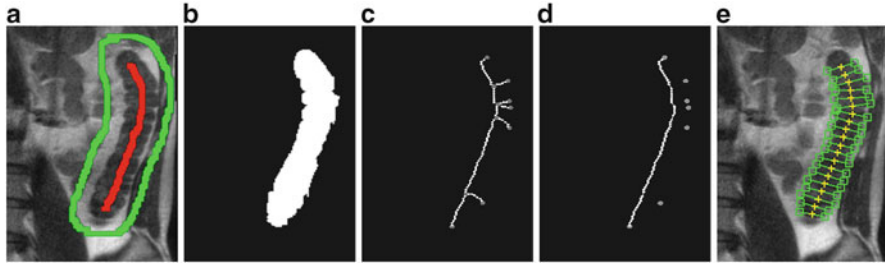


Fig. 4 (a) Seed brushes in one frame of the image sequence. (b) Resulting segmentation of the descending colon. (c) Skeletonization of the segmentation. (d) Extracted longest path which is used as centerline. (e) Diameter measurement at 20 sample points

image data. The exact global minimum of the energy formulation in (1) can be computed by using a max-flow algorithm (e.g [5]).

Since we are dealing with MRI time series of 2D slices, each frame is showing a similar image of the patient’s abdomen with slightly moved, or in case of motility, extended bowel diameter (see Fig. 1). Furthermore, after performing the motion compensation, we can assume that the only large bowel motion left in the images is due to motility. We can make use of this minimal changes within the segmentation method in order to minimize the user interaction. Actually, we can perform a full segmentation of the whole time series very efficiently. The user sets the seed brushes only in one frame. Object seeds have to be set roughly at the centerline of the colon and background seeds are placed around the colon part of interest (see Fig. 4a). Thanks to the motion compensation we can benefit from this strategy in two ways: on the one hand, these brushes can be set automatically in all other frames of the time series in a “copy & paste” fashion, on the other hand, the surrounding background seeds can be used as a restriction or bounding box for the computation of the graph-cuts algorithm. The segmentation of the whole series is then done in one single energy formulation. Thus, the boundary term B also acts as a smoothness constraint in the temporal domain. The computation time of one segmentation for a subsequence of about 20 frames showing the peristaltic wave is less than 10 seconds.

One important property of such a segmentation approach is the extreme flexibility. This method can be used for all parts of the large bowel and can deal with extreme shape variations which are likely to occur. This is crucial for our further analysis of the colon motility which is based on the segmentation result.

4 Analysis of Colon Motility

The aim of the colon motility analysis is to obtain as much information as possible about the peristaltic motion visible in the Cine MRI sequences. Our approach is based on the idea of measuring the bowel diameter over time which was previously

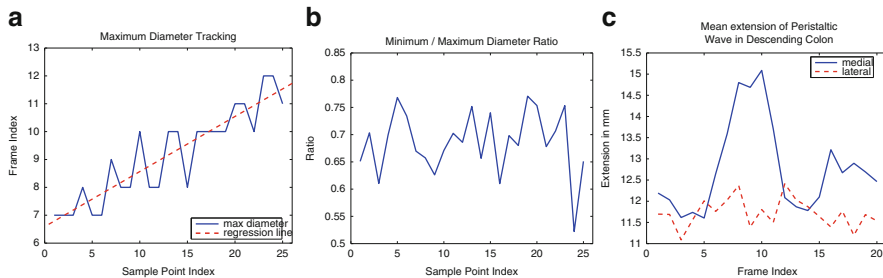


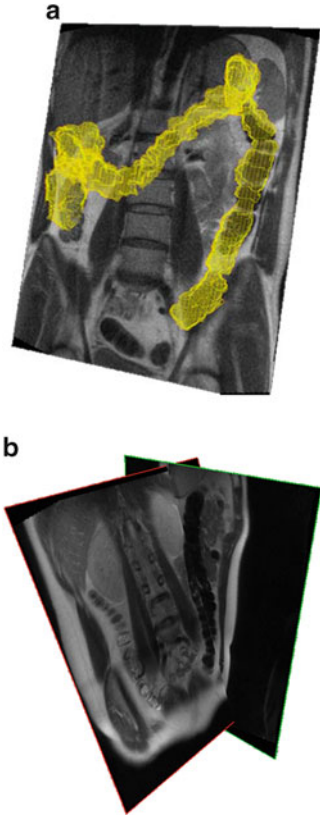
Fig. 5 (a) The maximum diameter tracked over time at each of the 25 sample points. Such a measure can be used to estimate the speed of the peristaltic wave. (b) Ratio of minimum and maximum diameter at each sample point is used to assess the activity of the colon. (c) The mean extension for the lateral and medial side of the colon during a propagating peristaltic wave over 20 frames

presented in [7]. In a first clinical trial, this fully manual approach was able to measure significant changes in the motility after administration of stimulating drugs [6]. However, this method was extremely time consuming and tedious in means of reproducibility. A limited number of 5 diameters were measured manually in one frame. Then, the corresponding points were identified manually in the successive frames and again the diameters had to be measured. In order to improve and automatize this approach and increase the number of measurements to a user selected bound, we are making use of the two proposed image processing steps so far, the motion compensation and the segmentation.

At first, we extract the skeleton of the segmented colon in one frame using a thinning algorithm proposed by Palagyi *et al.* [13]. From this skeletonization we construct a graph using a wave propagation approach presented by Zahlten *et al.* [15]. The result of these two steps is shown exemplary for one image sequence in Fig. 4c. We extract the longest path from the resulting graph in order to obtain a good approximation of the real centerline of the colon (see Fig. 4d). A B-Spline is fitted to the centerline which then can be subdivided into a user defined number of segments. At each segment we determine the diameter of the colon by measuring the extension of the segmentation perpendicular to the centerline (see Fig. 4e). Since the segmentation is already computed for all slices and thanks to the motion compensation, we can easily measure the colon diameter at all these specific positions over time. The user can change the number of measurement points without any recomputation on the segmentation or centerline and gets the new measurements within milliseconds.

In order to analyze the present colon motility, we extract several values with clinical relevance from our measurements. One significant parameter of interest is the propagation speed of the peristaltic wave. This can be measured by tracking the maximum diameter along the measurement points (see Fig. 5a). A value also important in assessing pathologies and dysfunctions of the colon motility is the ratio of the average maximum and minimum diameter (see Fig. 5b). A ratio close to 1.0 could indicate a local defect on the contraction ability. Besides these values

Fig. 6 (a) Fusion of dynamic 2D and static 3D. (b) Multi-plane MRI showing the colon motion in three dimensions



hopefully leading to a clinical index in near future, other phenomena reported in the medical literature could be measured for the first time. We could show that there is a significant difference in the contraction of the large bowel on the lateral and medial side (see Fig. 5c). This is what clinical experts actually expected. However, up to now there did not exist any image based method to proof these assumptions.

5 Conclusion

Our experiments focused on the descending part of the colon. This part is most relevant for clinical interventions as it is suitable for minimally invasive surgery. We tested our method on volunteer data and compared it to the manual approach of diameter calculation [7]. Our approach significantly reduces user-interaction and is fast in delivering quantitative results on colon motility assessment. We increased the number of clinical parameters by for instance the speed of propagation or the min-max ratio.

In future work, we consider fully-automatic methods, *e.g.* learning-based template matching approaches may provide a tool to remove user interactions from the motion compensation task. Deformable registration methods will be considered in order to compute dense motion fields. We also would like to propose advanced visualization techniques, *e.g.* fused 2D-3D data, which may help during planning and intervention (see Fig. 6a). New experiments using manometry are realized allowing a comparison of our non-invasive approach to an invasive but established technique. Our goal is to achieve a higher sensitivity while providing precise spatial information for defects. Recent advances in fast multi-plane imaging allow for capturing motion in three dimensions (see Fig. 6b) and thus may further improve the assessment of colon motility [11].

References

1. G. Bassotti and M. Crowell. Colon and rectum: normal function and clinical disorders. *Schuster Atlas of Gastrointestinal Motility in Health and Disease*, pages 241–252, 2002.
2. E. Bonapace, A. Maurer, S. Davidoff, and et al. Whole gut transit scintigraphy in the clinical evaluation of patients with upper and lower gastrointestinal symptoms. *Am J Gastroenterol.*, 95:2838–2847, 2000.
3. J.-Y. Bouguet. *Pyramidal Implementation of the Lucas-Kanade Feature Tracker*. OpenCV Documentation, Microprocessor Research Labs, Intel Corporation, 1999.
4. Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In *Proc. International Conference on Computer Vision (ICCV)*, volume I, pages 105–112, 2001.
5. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(9):1124–1137, Sept. 2004.
6. S. Buhmann, C. Kirchhoff, and et al. Assessment of large bowel motility by cine magnetic resonance imaging using two different prokinetic agents. *Investigative Radiology*, 40, 11: 689–694, 2005.
7. S. Buhmann, C. Kirchhoff, and et al. Visualization and quantification of large bowel motility with functional cine-mri. *Fortschritte auf dem Gebiet der Roentgenstrahlen und der bildgebenden Verfahren (RoeFo)*, 177:35–40, 2005.
8. R. Hagger, D. Kumar, M. Benson, and et al. Periodic colonic motor activity identified by 24-h pancolonic ambulatory manometry in humans. *Neurogastroenterol Motil.*, pages 271–278, 2002.
9. M. Hansen. Small intestinal manometry. *Physiol Res.*, 51:541–556, 2002.
10. C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conf.*, pages 147–151, 1988.
11. O. Kutter, S. Kirchhoff, M. Berkovic, M. Reiser, and N. Navab. Spatio-temporal registration in multiplane mri acquisitions for 3d colon motility analysis. In *SPIE Medical Imaging*, 2008.
12. A. Lienemann, D. Sprenger, H. Steitz, and et al. Detection and mapping of intraabdominal adhesions by using functional cine mr imaging: preliminary results. *Radiology*, 217:421–425, 2000.

13. K. Palágyi, E. Sorantin, E. Balogh, A. Kuba, C. Halmai, B. Erdöhelyi, and K. Hausegger. A sequential 3d thinning algorithm and its medical applications. In *Proc. Int'l Conf. Information Processing in Medical Imaging (IPMI)*. Springer, 2001.
14. J. Shi and C. Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593–600, 1994.
15. C. Zahlten, H. Jürgens, and H.-O. Peitgen. Reconstruction of branching blood vessels from ct-data. In *Eurographics Workshop of Visualization in Scientific Computing*, pages 161–168. Springer, 1994.

Segmentation of Diseased Livers: A 3D Refinement Approach

R. Beichel, C. Bauer, A. Bornik, E. Sorantin, and H. Bischof

Abstract Liver segmentation is the first data analysis step in computer-aided planning of liver tumor resections. For clinical applicability, the segmentation approach must be able to cope with the high variation in shape and gray-value appearance of the liver. In this article we present a novel segmentation scheme based on a true 3D segmentation refinement concept utilizing a hybrid desktop/virtual reality user interface. The method consists of two main stages. First, an initial segmentation is generated using graph cuts. Second, a segmentation refinement step allows to fix arbitrary segmentation errors. We demonstrate the robustness of our method on ten contrast enhanced liver CT scans and compare it to fifteen other methods. Our segmentation approach copes successfully with the high variation found in patient data sets and allows to produce a segmentation in a time-efficient manner.

R. Beichel (✉)

Department of Electrical and Computer Engineering and Department of Internal Medicine, The University of Iowa, 4016 Seamans Center, Iowa City, IA 52442, USA
e-mail: reinhard-beichel@uiowa.edu

C. Bauer

Department of Electrical and Computer Engineering, The University of Iowa, 4016 Seamans Center, Iowa City, IA 52442, USA
e-mail: christian-bauer@uiowa.edu

A. Bornik • H. Bischof

Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, 8010 Graz, Austria
e-mail: bischof@icg.tugraz.at; bornik@icg.tugraz.at

E. Sorantin

Research Unit for Digital Information and Image Processing, Department of Radiology, Medical University Graz, Auenbruggerplatz 4, 8010 Graz, Austria
e-mail: erich.sorantin@medunigraz.at

1 Introduction

Liver cancer is one of the four most common deadly malignant neoplasms in the world. Approximately 618,000 deaths due to liver cancer were reported in 2002¹. Tomographic imaging modalities like X-ray computed tomography (CT) play an important role in diagnosis and treatment of liver diseases like hepatocellular carcinoma (HCC). Liver resection has evolved as the treatment of choice for various benign and malignant hepatic tumors. Deriving a digital geometric model of hepatic (patho)anatomy from preoperative image data facilitates the planning procedure of liver tumor resections [13]. Thus, methods for liver segmentation in volume data are needed which are applicable in clinical routine. In this context, several problems have to be addressed: (a) **high shape variation** due to natural anatomical variation, disease (e.g., cirrhosis), or previous surgical interventions (e.g., liver segment resection), (b) **inhomogeneous gray-value appearance** caused by tumors or metastasis, and (c) **low contrast** to neighboring structures/organs like colon or stomach. For clinical application, liver segmentation must be able to handle all possible cases in a time-efficient manner. Especially, livers with large or multiple tumors are of interest, since treatment selection is crucial in these cases.

A large number of approaches to liver segmentation have been developed using methods like live wire, level sets, deformable models, or active shape models [8, 11, 12, 15]. Since the performance of the algorithms heavily depends on patient selection, image quality, imaging protocol, and severeness of the disease, different approaches can hardly be compared. Therefore, an international competition was organized by Heimann et al. [7], addressing the problem of liver segmentation in CT data. This competition was held in the form of a workshop at the 2007 Medical Image Computing and Computer Assisted Intervention conference. The proceedings give an overview of current liver segmentation methods and their performance [7]. To this competition, nine fully automated methods, five semi-automated methods, and two interactive methods were submitted, including our approach [1].

In summary, despite the progress made in the development of liver segmentation methods, segmentation errors or failure are still common. Tumors or previously performed resection of liver segments can lead to large variations in the gray value appearance or in the shape of the liver. No pure bottom-up approach or model-based approach is able to successfully cope with all kinds of possible variations that occur in clinical practice [7]. In order to make liver segmentation applicable for clinical routine a method for efficient refinement of segmentation errors is essential.

One of the rare examples of segmentation refinement are reported in [9] and [2], where Rational Gaussian (RaG) Surfaces are used to represent segmented objects. Segmentation errors can be corrected by manipulation of control points using a 2D desktop setup. Another system allowing to alter the boundary of segmented objects by morphological operations and surface dragging was reported in [10]. A tool for

¹ <http://www.who.int/whr/2004/en>

data driven editing of presegmented images/volumes based on graph cuts or alternatively random walker algorithms was proposed in [6]. All approaches mentioned so far are based on 2D interaction and monoscopic desktop-based visualization techniques, despite the fact that 3D objects are targeted. In general, 2D interaction methods are not sufficient for refinement of 3D models extracted from volumetric data sets. This is for two reasons: Interpreting the complex 3D anatomy when only viewing 2D cross-sections is a complicated task; provided a 3D model of the liver helps in understanding its complex shape [13]. Second, interaction with a 3D model using a 2D input device is more time-consuming and complex in terms of interaction patterns than directly using the 3D information provided by a VR setup [4].

To tackle these problems, we propose a novel refinement approach to 3D liver segmentation. Based on an initial highly automated graph cut segmentation, refinement tools allow to manipulate the segmentation result in 3D, and thus, to correct possible errors. Segmentation refinement is facilitated by a hybrid user interface, combining a conventional desktop setup with a virtual reality (VR) system.

After presenting details of the method, the proposed approach is evaluated on ten liver CT data sets from an international competition [7]. Achieved results and benchmarks against competing methods are presented. At the end of this work the clinical applicability of the system and possible extensions are discussed.

2 Methodology

The proposed approach to liver segmentation consists of two main stages: initial segmentation and interactive segmentation refinement. As input for the first stage, a CT volume and one or more start regions, marking liver tissue, are used. The segmentation is then generated using a graph cut approach². In addition, a partitioning of the segmentation and the background into volume chunks is derived from edge/surface features calculated from the CT volume. These two types of output are passed on to the second stage which allows for the correction/refinement of segmentation errors remaining after the first stage. Refinement takes place in two steps. First, volume chunks can be added or removed. This step is usually very fast, and the majority of segmentation errors occurring in practice can be fixed or at least significantly reduced. Second, after conversion of the binary segmentation to a simplex mesh, arbitrary errors can be addressed by deforming the mesh using various tools. Each of the refinement steps is facilitated using interactive VR-enabled tools for true 3D segmentation inspection and refinement, allowing for stereoscopic viewing and true 3D interaction. Since the last stage of the refinement procedure is mesh-based, a voxelization method is used to generate a labeled volume [14].

²Note that graph cut segmentation is not used interactively, as proposed by Boykov et al. in [5], since the behavior of graph cuts is not always intuitive.

2.1 Graph-Cut-based initial segmentation

An initial segmentation is generated using a graph cut [5] approach. From image data, a graph $G = (V, E)$ is built, where nodes are denoted by V and undirected edges by E . Nodes V of the graph are formed by data elements (voxels), and two additional terminal nodes, a source node s and sink node t . Edge weights allow to model different relations between nodes (see [5] for details). Let P denote the set of voxels from the input volume data set \mathbf{V} —to reduce computing time, only voxels with density values above -600 Hounsfield Units (HU) are considered as potentially belonging to the liver. The partition $A = (A_1, \dots, A_p, \dots, A_{|P|})$ with $A_p \in \{\text{"obj"}, \text{"bkg"}\}$ can be used to represent the segmentation of P into object (“obj”) and background (“bkg”) voxels. Let N be the set of unordered neighboring pairs $\{p, q\}$ in set P according to the used neighborhood relation. In our case, a 6-neighborhood relation is used to save memory. The cost of a given graph cut segmentation A is defined as $E(A) = B(A) + \lambda R(A)$ where $R(A) = \sum_{p \in P} R_p(A_p)$ takes region properties into account and $B(A) = \sum_{p, q \in N} B_{p, q} \delta_{A_p \neq A_q}$, with $\delta_{A_p \neq A_q}$ equals 1 if $A_p \neq A_q$ and 0 if $A_p = A_q$, being boundary properties. The parameter λ with $\lambda \geq 0$ allows to tradeoff the influence of both cost terms. Using the s - t cut algorithm, a partition A can be found which globally minimizes $E(A)$.

Region term The region term $R(A)$ specifies the costs of assigning a voxel to a label based on its gray-value similarity to object and background regions. For this purpose, user defined seed regions are utilized. The region cost $R_p(\cdot)$ for a given voxel p is defined for labels “obj” and “bkg” as negative log-likelihoods $R_p(\text{"obj"}) = -\ln(\text{Pr}(I_p | \text{"obj"}))$ and $R_p(\text{"bkg"}) = -\ln(\text{Pr}(I_p | \text{"bkg"}))$ with $\text{Pr}(I_p | \text{"obj"}) = e^{-(I_p - m_{obj})^2 / (2\sigma_{obj}^2)}$ and $\text{Pr}(I_p | \text{"bkg"}) = 1 - \text{Pr}(I_p | \text{"obj"})$, respectively. From an object seed region placed inside the liver, the mean m_{obj} and standard deviation σ_{obj} are calculated. Clearly, in the above outlined approach, a simplification is made since liver gray-value appearance is usually not homogeneous. However, in combination with the other processing steps this simplification works quite well. Further, the specified object seeds are incorporated as hard constraints, and the boundary of the scene is used as background seeds.

Boundary term The basic idea is to utilize a surfaceness measure as boundary term which is calculated in four steps:

1. *Gradient tensor calculation:* First, to reduce the effect of unrelated structures on the gradient, the gray value range of the image is adapted:

$$\tilde{I}_f = \kappa \left(I_f \right) = \begin{cases} v_{low} & \text{if } I_f < t_{low} \\ v_{high} & \text{if } I_f > t_{high} \\ I_f & \text{otherwise} \end{cases} . \text{ Second, a gradient vector } \nabla f =$$

$(f_x, f_y, f_z)^T$ is calculated for each voxel f on the with κ gray-value transformed data volume \mathbf{V} by means of Gaussian derivatives with the kernel

$g_\sigma = 1/(2\pi\sigma^2)^{\frac{3}{2}} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}$ and standard deviation σ . The gradient tensor $\mathbf{S} = \nabla f \nabla f^T$ is calculated for each voxel after gray-value transformation.

2. *Spatial non-linear filtering*: To enhance weak edges and to reduce false responses, a spatial non-linear averaging of the gradient tensors is applied. The non-linear filter kernel consists of a Gaussian kernel which is modulated by the local gradient vector ∇f . Given a vector \mathbf{x} that points from the center of the kernel to any neighboring voxel, the weight for this voxel is calculated

$$\text{as: } h_{\sigma', \rho}(x, \nabla f) = \begin{cases} \frac{1}{N} e^{-\frac{r}{2\sigma'^2}} e^{-\frac{\tan(\phi)^2}{2\rho^2}} & \text{if } \phi \neq \frac{\pi}{2} \\ 0 & \text{if } \phi \neq \frac{\pi}{2} \text{ and } r = 0, \text{ with } r = \mathbf{x}^T \mathbf{x} \text{ and} \\ \frac{1}{N} & \text{otherwise} \end{cases}$$

$\phi = \frac{\pi}{2} - |\arccos(\nabla f^T \mathbf{x} / (|\nabla f| |\mathbf{x}|))|$. Parameter ρ determines the strength of orientedness, and σ' determines the strength of punishment depending on the distance. N is a normalization factor that makes the kernel integrate to unity. The resulting structure tensor is denoted as \mathbf{W} .

3. *Surfaceness measure calculation*: Let $\mathbf{e}_{1W(\mathbf{x})}$, $\mathbf{e}_{2W(\mathbf{x})}$, $\mathbf{e}_{3W(\mathbf{x})}$ be the eigenvectors and $\lambda_{1W(\mathbf{x})} \geq \lambda_{2W(\mathbf{x})} \geq \lambda_{3W(\mathbf{x})}$ the corresponding eigenvalues of $\mathbf{W}(\mathbf{x})$ at position \mathbf{x} . If \mathbf{x} is located on a plane-like structure, we can observe that $\lambda_1 \gg 0$, $\lambda_2 \approx 0$, and $\lambda_3 \approx 0$. Thus, we define the surfaceness measure as $t(\mathbf{W}(\mathbf{x})) = \sqrt{\lambda_{1W(\mathbf{x})} - \lambda_{2W(\mathbf{x})}}$ and the direction of the normal vector to the surface is given by $\mathbf{e}_{1W(\mathbf{x})}$.

4. *Boundary weight calculation*: In liver CT images, objects are often separated only by weak boundaries, with higher gray level gradients present in close proximity. To take these circumstances into account, we propose the following weighting function $\xi(t) =$

$$\begin{cases} c_1 & \text{if } t < t_1 \\ c_2 & \text{if } t > t_2 \\ (t - t_2) \frac{c_2 - c_1}{t_2 - t_1} + c_1 & \text{otherwise} \end{cases} \quad \text{which models an uncertainty zone between}$$

t_1 and t_2 (note: $t_1 < t_2$ and $c_1 > c_2$). Ideally, the graph cut segmentation should follow the ridges of the gradient magnitude. Therefore, we punish non-maximal responses in the gradient magnitude volume by adjusting the weighting function as follows: $\xi_{nonmax}(t) = \min\{\xi(t) + c_k, 1\}$, where c_k is a constant.

Summing up, the boundary cost term is determined by

$$B_{p,q} = \min\{\xi_{nonmax}(t(\mathbf{W}(\mathbf{x}_p))), \xi_{nonmax}(t(\mathbf{W}(\mathbf{x}_q)))\}$$

2.2 Chunk-based Segmentation Refinement

After initial segmentation, objects with a similar gray-value range in close proximity may be merged or tumors with different gray-value appearance might be missing. Therefore, a refinement may be needed in some cases. The first refinement stage

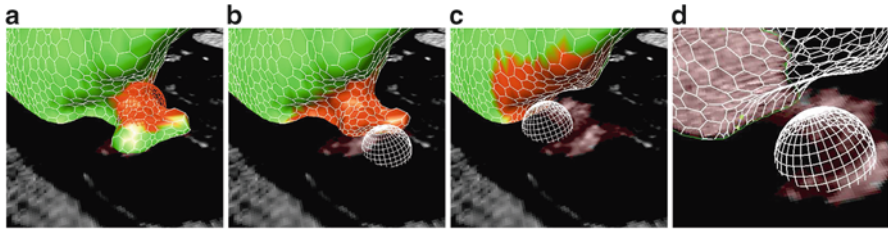


Fig. 1 Mesh-based refinement using a sphere deformation tool. In this case the segmentation error is a leak. (a) Marking the region containing the segmentation error. (b) Refinement using the sphere tool. (c) After some time using the sphere tool the error is fixed. (d) The corrected region in wire frame mode highlighting the mesh contour

is based on volume chunks, which subdivide the graph cut segmentation result (object) as well as the background into disjunct subregions; the segmentation can be represented by a set of chunks. The presented approach partitions the image based on constrictions in the initial segmentation and based on boundary information. It allows to fix larger errors (e.g. due to high contrast tumors) in a time efficient manner by altering the initial segmentation.

By thresholding $t(\mathbf{W})$, a binary boundary volume (threshold t_b) representing boundary/surfaces parts is generated and merged with the boundary from the graph cut segmentation by using a logical “or” operation. Then the distance transformation is calculated. Inverting this distance map results in an image that can be interpreted as a height map. To avoid over-segmentation, all small local minima resulting from quantization noise in the distance map are eliminated. After running a watershed segmentation, boundary voxels are merged with the neighboring chunks containing the most similar adjacent voxels. Since the method can handle gaps in the edge scene, the threshold t_b can be set very conservatively to suppress background noise. Refinement can be done very efficiently, since the user has to select/deselect predefined chunks, which does not require a detailed border delineation. This step requires adequate tools for interactive data inspection and selection methods. For this purpose, a hybrid user interface was developed, which is described in Sect. 2.4.

2.3 Simplex-Mesh-based Refinement

After the first refinement step, selected chunks are converted to a simplex mesh representation. Different tools allow then a deformation of the mesh representation. One example is shown in Fig. 1. More details regarding this mesh-based refinement step can be found in [4].

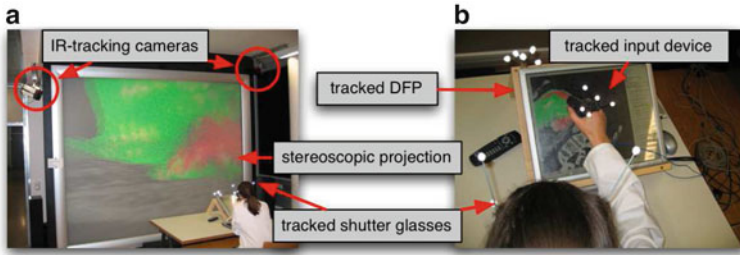


Fig. 2 Components of the hybrid desktop/virtual reality user interface

2.4 Hybrid Desktop/Virtual Reality User Interface

To facilitate segmentation refinement, a hybrid user interface consisting of a desktop part and a virtual reality (VR) part was developed as shown in Fig. 2 (see [3] for details). The VR system part provides stereoscopic visualization on a large screen projection wall, while the desktop part of the system uses a touch screen for monoscopic visualization. The tracked input device can be used for interaction in 3D, or it can be used as a stylus in combination with the touch screen.

This hybrid user interface allows to solve individual refinement tasks using the best suited interaction technique, either in 2D or 3D. Inspection of the segmentation, localization of errors, and correction of larger segmentation errors is facilitated best in 3D as the VR system allows interaction in 3D directly. Refinement of smaller segmentation errors, interaction with single image slices, or selecting the appropriate refinement tools is facilitated best in 2D.

3 Experimental Setup, Evaluation Results, and Performance Comparison

For evaluation of the segmentation approach, ten liver CT data sets with undisclosed manual reference segmentation were provided by the workshop organizers [7]. Segmentation results were sent to the organizers, which provided in return evaluation results (see [7] for details). For evaluation, the following parameters have been used: Gaussian derivative kernel: $\sigma = 3.0$; non-linear filtering: $\sigma' = 6.0$, $\rho = 0.4$; graph cut: $\lambda = 0.05$; weighting function: $t_1 = 2.0$, $t_2 = 10.0$, $c_1 = 1.0$, $c_2 = 0.001$, $c_k = 0.75$; Threshold for chunk generation: $t_b = 10.0$; gray-value transformation: $t_{low} = -50$, $v_{low} = -150$, $t_{high} = 200$, and $v_{high} = 60$. To simulate the clinical workflow, the initial seed regions were provided manually and the graph cut segmentation as well as the chunk generation were calculated automatically. Based on the initial segmentation, a medical expert was asked to perform: (a) chunk-based (CBR) and (b) mesh-based refinement (MBR). Intermediate results and task completion times

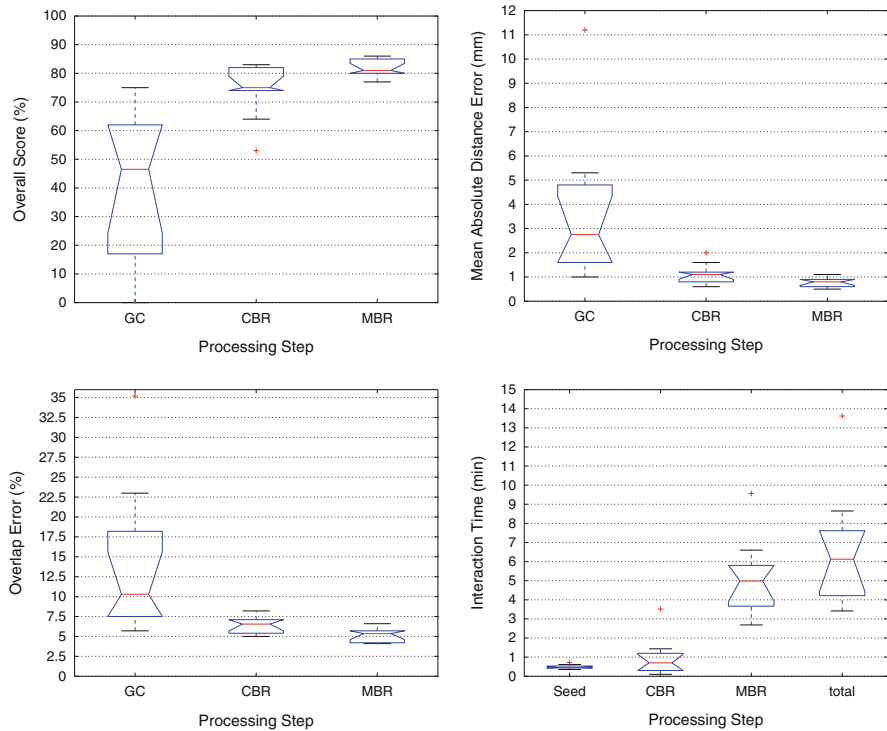


Fig. 3 Segmentation quality of the initial graph cut segmentation result, the CBR result, the MBR result and required user interaction time. See text for details

were recorded. Prior to evaluation, the expert was introduced to the system by an instructor.

Results of our method for each processing step are summarized in form of box-and-whisker plots in Fig. 3, and tables for each test case can be found in [1]. Figure 3(a) depicts the overall segmentation score derived from the volumetric overlap error, the relative absolute volume difference, the average symmetric surface distance, the RMS symmetric surface distance, and the maximum symmetric surface distance [7]. A higher overall score implies a better segmentation performance. The effectiveness of both refinement steps is clearly demonstrated. This is also reflected in the plots for the mean distance error (Fig. 3(b)) and the overlap error (Fig. 3(c)). Time required for initial seed placement, CBR, and MBR as well as the total interaction time is plotted in Fig. 3(d). Results on one CT scan are depicted in Fig. 4.

As reported in [7], the best overall average segmentation score for automated segmentation methods was 73, for semi-automated segmentation methods 75, and for interactive segmentation methods, excluding our method, 76. In comparison, our method reaches a overall mean segmentation score of 74 after the CBR, requiring

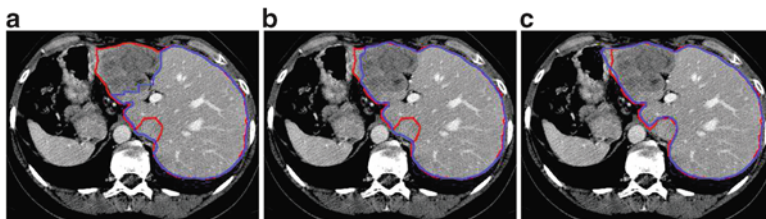


Fig. 4 Visual comparison of segmentation results for initial graph cut result (a), CBR (b), and MBR (c). The segmentation result is shown in blue and the manual reference in red. The improvement in segmentation quality with each refinement step can be seen clearly

less than one minute of interaction on average. The CBR result can be significantly improved by the MBR step, which leads to a mean overall score of 82, the best result of all the fifteen methods evaluated in [7]. For MBR, an additional 5.1 minutes were required on average. Both results for CBR and MBR are well within score values gained by comparing the manual reference to an additional independent human observer, which yield a score of 75 for the liver test cases.

4 Discussion and Future Work

Several additional experiments with different physicians have shown that the system can be used after a short learning phase (typically less than one hour), because of the intuitive 3D user interface. The proposed refinement method can easily be integrated into clinical work-flow. The CT volume together with the manual generated start region is sent by a radiology assistant to a computing node which performs the automated segmentation steps. As soon as a result is available, a radiologist is notified that data are ready for further processing. After inspection, possible refinement, and approval of correctness, the segmentation can be used for clinical investigations or planning of treatment. For our experiments, we have used a full-blown VR setup which is quite expensive. However, a fully functional scaled-down working setup can be built for a reasonable price, comparable to the costs of a radiological workstation.

The evaluation of our method on ten test CT data sets shows that a high segmentation quality (mean average distance of less than 1 mm) can be achieved by using this approach. In addition, the interaction time needed for refinement is quite low (approx. 6.5 minutes). Thus, the presented refinement concept is well suited for clinical application in the context of liver surgery planning. Future work will focus on making the MBR step faster by incorporating local data terms into the user-steered deformation process. In general, our approach is not limited to a specific organ or modality, and therefore, it is very promising for other medical segmentation applications.

Acknowledgments This work was supported in part by the Austrian Science Fund (FWF) under Grants P17066-N04 and Y193 and the Doctoral Program Confluence of Vision and Graphics W1209-N15

References

1. R. Beichel, C. Bauer, A. Bornik, E. Sorantin, and H. Bischof. Liver segmentation in CT data: a segmentation refinement approach. In T. Heimann, M. Styner, and B. van Ginneken, editors, *3D Segmentation in The Clinic: A Grand Challenge*, pages 235–245, 2007.
2. R. Beichel and et al. Shape- and appearance-based segmentation of volumetric medical images. In *Proc. of ICIP 2001*, volume 2, pages 589–592, 2001.
3. A. Bornik, R. Beichel, E. Kruijff, B. Reitinger, and D. Schmalstieg. A hybrid user interface for manipulation of volumetric medical data. In *Proceedings of IEEE Symposium on 3D User Interfaces 2006*, pages 29–36. IEEE Computer Society, Mar. 2006.
4. A. Bornik, R. Beichel, and D. Schmalstieg. Interactive editing of segmented volumetric datasets in a hybrid 2D/3D virtual environment. In *VRST '06: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 197–206, Nov. 2006.
5. Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. In *International Journal of Computer Vision (IJCV)*, 70(2):109–131, 2006.
6. L. Grady and G. Funka-Lea. An energy minimization approach to the data driven editing of presegmented images/volumes. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, volume 4191, pages 888–895. Springer, 2006.
7. T. Heimann, M. Styner, and B. van Ginneken. 3D segmentation in the clinic: A grand challenge. In *3D Segmentation in The Clinic: A Grand Challenge*.
8. T. Heimann, I. Wolf, and H.-P. Meinzer. Active shape models for a fully automated 3D segmentation of the liver - an evaluation on clinical data. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4191 of *Lecture Notes in Computer Science*, pages 41–48. Springer Berlin / Heidelberg, 2006.
9. M. Jackowski and A. Goshtasby. A computer-aided design system for revision of segmentation errors. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2, pages 717–724, Oct. 2005.
10. Y. Kang, K. Engelke, and W. A. Kalender. Interactive 3D editing tools for image segmentation. *Medical Image Analysis*, 8(1):35–46, 2004.
11. H. Lamecker and et al. Segmentation of the liver using a 3D statistical shape model. Technical report, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2004.
12. S. Pan and M. Dawant. Automatic 3D segmentation of the liver from abdominal CT images: A level-set approach. In M. Sonka and K. M. Hanson, editors, *Medical Imaging: Image Processing*, volume 4322 of *Proc. SPIE*, pages 128–138, 2001.
13. B. Reitinger, A. Bornik, R. Beichel, and D. Schmalstieg. Liver surgery planning using virtual reality. *IEEE Comput. Graph. Appl.*, 26(6):36–47, 2006.
14. B. Reitinger and et al. Tools for augmented reality-based liver resection planning. In R. L. Galloway, editor, *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, volume 5367, pages 88–99. SPIE, 2004.
15. A. Schenk, G. P. M. Prause, and H.-O. Peitgen. Efficient semiautomatic segmentation of 3D objects in medical images. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 186–195. Springer, 2000.

Part V
Emerging Modalities & Domains

Intra and inter subject analyses of brain functional Magnetic Resonance Images (fMRI)

J.B. Poline, P. Ciuciu, A. Roche, and B. Thirion

Abstract This chapter proposes a review of the most prominent issues in analysing brain functional Magnetic Resonance data. It introduces the domain for readers with no or little knowledge in the field. The introduction places the context and orients the reader in the many questions put to the data, and summarizes the currently most commonly applied approach. The second section deals with intra subject data analysis, emphasizing hemodynamic response estimation issues. The third section describes current approaches and advances in analysing group data in a standard coordinate system. The last section proposes new spatial models for group analyses. Overall, the chapter gives a brief overview of the field and details some specific advances that are important for application studies in cognitive neurosciences.

J.B. Poline (✉)

CEA, DSV, I2BM, Neurospin, Batiment 145, 91191 Gif-sur-Yvette, France

Henry H. Wheeler Jr. Brain Imaging Center, University of California at Berkeley, CA, US

e-mail: jean-baptiste.poline@cea.fr

P. Ciuciu • B. Thirion

CEA, DSV, I2BM, Neurospin, Batiment 145, 91191 Gif-sur-Yvette, France

Parietal project-team, INRIA, 1 rue Honore d'Estienne d'Orves, 91120 Palaiseau, France

e-mail: philippe.ciuciu@cea.fr; bertrand.thirion@inria.fr

A. Roche

Advanced Clinical Imaging Technology group, Siemens Healthcare/CHUV Dept

Radiology/EPFL LTS5 Lausanne, Switzerland

e-mail: alexis.roche@epfl.ch

1 Introduction and overview

The goal of this section is to present a synthetic view of the principles, goals and techniques of fMRI data analysis. It does not try to be exhaustive but proposes a specific view on this domain of research. It may therefore interest readers with some knowledge of the field as well as naïve readers looking for entry points. As the first functional Magnetic Resonance Images were acquired in the early nineties, the domain is still young and the current techniques are evolving quickly. Nevertheless, the research questions and directions described in this chapter are likely to still be of interest for some time even with the anticipated evolution of acquisition and processing techniques.

1.1 *Background*

1.1.1 Principle of an activation study

The principle of an activation study with fMRI is simple. It consists in placing a subject or a patient in the scanner and acquiring data that reflect the hemodynamic activity of the brain. While the acquisition occurs, subjects are submitted to various stimuli or asked to perform certain tasks. These will influence the regional neural activity of the brain which in turn will modulate locally the hemodynamic (blood flow, blood volume and oxygenation) of some brain regions. The data acquired is a series of slices, each of which can be recorded typically in less than 80ms, with an in plane resolution and thickness of the order of the mm (1-3 mm). Therefore, the whole brain volume can be covered in a few seconds with current scanners but equipment is improving fast and these numbers are getting down. Data acquisition is continuous for 5 to 20 min, and is called a “run”, an MR session usually consists of several runs for a subject, together with the acquisition of other kind of images reflecting the anatomy of the brain (e.g. T1 images with grey white matter contrast, diffusion images from which fibre bundles connecting brain regions can be reconstructed). The data acquired for one subject is of the order of the Giga byte, and an actual experiment usually involves 15 to 40 subjects (for instance using different groups of subjects such as control versus patients). During a run, an experimental paradigm designed to elicit a certain component of perception or cognition is proposed to the subject. It consists in several tasks or stimuli (the “conditions”) and fMRI data is commonly used to detect brain regions in which the hemodynamic response varies between conditions. Other data such as reaction time, eye movement, cardiac or breathing rhythms can be conjointly acquired. A comparison of 2 or more experimental conditions is called a contrast.

1.1.2 A short perspective of functional neuroimaging

As previously described, fMRI data are often used for detecting brain regions whose hemodynamic varies across experimental conditions. Applications in humans' studies can be roughly located on two axes: group versus single subject studies and normal versus patient studies. On the latter axis, the understanding of human brain functions is opposed to dysfunctions in psychiatric or neurological diseases. On the former axis, the specific information obtained from a particular subject is contrasted to the description of the information obtained at the population level. Using groups of normal subjects, cognitive neuroscientists that use fMRI to probe brain functions are seriously challenged by philosophers or psychologist such as J. Fodor [1] who claim that localizing brain regions that respond to certain stimuli does not help to understand how the brain functions. Fodor uses the mechanistic analogy of a car engine, and asks how the knowledge of the localisation of pieces such as the piston or carburettor helps to understand the function of these pieces in the engine. It does not, unless one is interested in mending some parts, as the neurosurgeon might be for pathologies involving brain surgery. While the argument is potent it does not account for the numerous occasions where the spatial organisation is a reflection of the functioning, such as the retinotopic organisation of the early visual cortex, and that brain region characterisation and localisation might be a necessary first step in the process of defining models of brain functioning.

1.1.3 What we need to know about the data

As described previously, the origin of the data is hemodynamic. Following an increase of synaptic and spiking activity, neurons require more energy conveyed through the blood in the form of oxygen and glucose [2]. Through a mechanism not fully known, the result is first an increase of oxygen extraction followed by an increase of blood flow that over compensates the loss of oxy-haemoglobin. The Blood Oxygen Level Dependant (BOLD) contrast can be measured with MRI because the oxy-haemoglobin is diamagnetic while the deoxy-haemoglobin is paramagnetic, so that the MR signal increases with the ratio oxy over deoxy-haemoglobin [3]. The timing of this effect is relatively slow and picks around 4-6 seconds after a short stimulation and returns to baseline in about 25s. Hemodynamic responses are further described and studied in Sect. 2. It is interesting to note that even if the absolute timing may be considered as poor, the differential timing between two conditions can be precise to the order of 30-50 ms. The BOLD contrast does not show many anatomical details, and is sensitive to artefacts due to magnetic field variation around air/tissue interfaces inducing loss of signal or volume deformation in these areas.

1.2 Data analysis of fMRI: a view of today research axes or how not to be lost in the literature

In this section, we propose a particular view of the current research in the analysis of fMRI data. While the domain is complex and rapidly growing, the classification that we describe suggests a certain view on the domain and should help the reader to orient himself or herself in the current techniques proposed.

1.2.1 Standard analysis: a short description. Massively univariate techniques

The most common approach that has dominated the past decade can be decomposed in the following steps. Its success is linked to freely available tools such as SPM (see www.fil.ucl.ac.uk/SPM) or FSL (fsl.fmrib.ox.ac.uk/fsl).

Step 1. Data pre-processing: temporal and spatial realignments. Subjects are never completely still in the scanner and movement needs to be corrected with a strong impact on the signal obtained. Movements correlated to the experimental paradigm are particularly difficult to correct if not impossible. Most current techniques assume a rigid body movement between two brain scans. Temporally, the slices of one brain volume are not acquired at the same time, and all voxel time series are usually interpolated to impose a unique timing for all voxels of one brain volume.

Step 2. (can also be done after step 3) If a group analysis is to be performed, the data of different subjects need to be placed in a standard coordinate system. While there are many different techniques to perform this, the most usual procedure is to first realign the functional volumes to the anatomical volume acquired in the same scanning session, and use this more detailed image to derive a deformation field that warps the subject brain anatomy to a standard template (generally the so called ICBM152 volume image which represents the average of 152 healthy T1 brain images by reducing it to 2 mm isotropic resolution). This template corresponds (but only approximately) to a brain neurosurgical atlas, the Talairach and Tournoux atlas [4]. Because of anatomical or functional variability across subjects, the registration is not perfect and regional activity from different subjects is not located exactly at the same location in the standard coordinate system. Gaussian filtering is therefore often applied to fMRI data to enhance the possible overlap of regional activity across subjects.

Step 3. Modelling the BOLD signal and constructing statistical maps. To a first approximation the BOLD hemodynamic response function (HRF) can be considered as a linear phenomenon with respect to stimulation. A linear model is constructed that includes all experimental factors which are believed to impact the BOLD signal. For instance, three experimental conditions will be modelled by three regressors. Each regressor is constructed as the convolution of a standard HRF with a time series

representing the occurrence and duration of the experimental conditions (stimuli or task). The model may then include other experimental factors of no interest (low frequencies confound, cardiac rate, subject movement, etc). Remaining noise is modelled as AR processes. The model parameters (magnitude of effects, noise variance) are then estimated, and statistical maps of t-test or F-test are constructed for each subject. The technique has the advantages of simplicity and flexibility, and is computationally very efficient, but it makes a series of strong assumptions that are not necessarily valid (linearity of the model, known and constant HRF, etc). In Sect 2, methods for estimating the HRF are reviewed (see for instance Ciuciu and coll. Section 2).

Step 4. Modelling at the group level. To infer results at the population level, hierarchical models are constructed, in which the estimated individual effects at the first level become the data of a simpler model at the second level. The second level models are usually much simpler, modelling for instance only the mean of a condition or contrast of conditions and assuming iid errors. So called mixed effect models can account for the variability of the estimation at the first level and estimation can be performed with expectation maximisation techniques. Section 3 describes this in more details.

Step 5. Thresholding. Once a statistical map is constructed at the individual or group level, it remains to be decided which brain regions are worth reporting, hence to derive a threshold defining brain regions with significant activity. The principle on which this threshold is chosen varies across studies. Three types of statistical threshold are used. Uncorrected thresholds are defined to control for the risk of false positive at the voxel level. If only one region or voxel is tested, this is a valid approach. Often however, a priori localisation is not known, and because a brain volume contains several tens of thousands of voxels, a number of regions are likely to be reported by chance with this kind of threshold. To prevent false positive results, a number of techniques have been derived for controlling the risk of error family wise (across voxels or regions). Worsley and colleagues in particular have based this thresholding on the notion of the expected number of above threshold regions approximated by the Euler characteristic of random fields. These approximations have been extended to a number of statistical fields (T, F, X², etc) and geometry (any volume, surface, 4D data, etc) [5]. These tests are computationally efficient but also rely on several assumptions, including a heavy smoothing of the volume to be thresholded which contradicts the MRI physicist efforts to improve image resolution. Other approaches (see Sect. 2 and the work of Roche and colleagues) have been developed using permutation tests which are shown to be exact and often more sensitive techniques. The third approach uses the false detection rate (FDR) threshold that controls for the false positive rate as a function of the number of reported voxels.

As usual, thresholding is a tradeoff between sensitivity and specificity, which may depend on the application purpose although this is rarely acknowledged in the literature. The defined regions are then reported using their local maxima in the standard space.

1.2.2 The problem of localising functional activity in the brain

Localising the functional activity, while crucial to the field, is a difficult issue that meets several challenges. At the single subject level, reporting the localisation appropriately relies on a good correspondence between the anatomical and functional images, as well as a good identification of the individual brain structures. The more difficult problem arises when reporting group results. Indeed, the algorithms that warp a subject anatomy to a template do not and cannot perform a perfect match. The information used for the warping is the main anatomical features (deep sulci, ventricles) but the variation of the anatomy between subjects is such that there is no obvious point to point correspondence between subjects. The identification of individual structures is not easy either (see [6]). The relation between sulco-gyral anatomy and functional activity is still to be further studied. Secondly, activity of different subjects may not be localised exactly at the same location within an anatomical structure. For instance, the Fusiform Face Area (responding more to face than to objects) may be localised more anteriorly in the fusiform gyrus in one subject compared to another. This prompts to other solutions than the standard stereotactic space to detect and report the localisation of functional activity (see Sect. 4). A number of laboratories now consider that the appropriate localisation technique is to define subject per subject functional regions using a first experiment, then study the activity of those well defined regions in a second step [7]. This also makes statistical inference trivial.

1.2.3 Detecting modules versus establishing functional connections

While the techniques described previously aim at localizing the activity in the brain, and therefore defining spatially functional modules, an increasingly large part of the literature is now devoted to establishing functional connections between brain regions. The original observation by Biswal et al. show that even during no motor activity (resting state or other such as visual stimulation), BOLD signal of a series of regions that respond to motor tasks are correlated. Since then, two main approaches are concurrently explored. The first one tries to extract networks of correlated activity (or sharing some information) and techniques such as principal component analysis, independent component analysis (ICA), probabilistic ICA, partial least square (PLS), various clustering techniques, self organizing maps, etc. To summarize, those methods aim at defining the various functional networks that underlie brain activity and their relation to external tasks or stimulations. The alternative approach consists in defining a specific network, a graphical model, choosing a priori the nodes and the structure of the graph, and to estimate the functional links given the experimental paradigm. This led to the development of structural equation models, dynamical causal models, etc. With the current approaches, graphical models are generally not able to identify network structures without strong a priori knowledge, while exploratory approaches often suffer from a lack of interpretability. Furthermore, the steady states used in many functional

connectivity studies are not always well-defined states, hence it would be of interest to extend the notion of functional connectivity to states that are controlled to a larger extent by the experimenter and that follow a predefined dynamic. Finally, signal similarity does not only come about by functional interaction, but can be influenced by confounding physiological effects of no interest like heart beat or artefacts like subject motion (Dodel et al. 2004).

1.3 Analysing fMRI data: where does the future lie?

In this last subsection, we review briefly three important axes that should get an increasing attention in the future.

- Adapting techniques to inter individual variability. The development of spatial models able to account for (limited) inter individual variability of the activity localisation is currently being established. This research relates to parcellation techniques, that define parcels with similar functional activity and close spatial activity, or hierarchical models of the spatial localisation in which the second level model parameter corresponds to the group location. Variability is also to be modelled for the magnitude of the activity once a location is defined, for which non parametric modelling and testing show promising results. The identification of individual structures such as sulci or fibre tracks will play an important role in this by providing better coordinate systems based on individual anatomical or functional landmarks.
- Decoding versus detecting? Recent works originating from Haxby et al. REF have reversed the usual data processing by considering how fMRI data can predict the experimental conditions. This was thought by some to be a step forward an actual decoding of the brain activity. Often, these works try to extract the regions or voxels that have the best predictive power. Methods to select those voxels or regions are still under development and the specificity of those voxels or regions with respect to the task or condition (their importance for the prediction) is still an open question. Those techniques are sometime called Multivariate Pattern Analysis (MPVA) and have hoped to shift the focus from where the processing occurs in the brain to how that processing represents information. These methods are likely to be much more sensitive compared to massively univariate approaches but should lose localisation information.
- Databasing and datamining. The need to store, organise and share the large amount of information that is acquired and processed in functional neuroimaging has been acknowledge early with various initiatives. We cite here a few significant attempt. Brainmap from P. Fox and coll. proposes to store bibliographic information, experimental descriptors and 3D Talairach coordinates. The fMRIDC database contains fMRI scanning data and summary images. The BIRN initiative has created a repository of anatomical and functional brain images accessible to a large network of hospitals. Clearly, the results obtained from a database of hundreds or thousands of brain scan can reach a level of sensitivity

that is not comparable to the results obtained with ten or twenty subjects. However, those are yet early attempts limited in their use and in their proposals that lack query and search systems as well as stable ontology of domains such as brain localisation and neuropsychology or cognitive neuroscience. Nevertheless, the need for sharing and exploiting large amounts of imaging, behavioural, physiological and genetic data is likely to put pressure on neuroimaging as the human genome project has a decade before.

2 Intra subject analysis and HRF estimation

Within-subject analysis in fMRI is usually addressed using a hypothesis-driven approach that actually postulates a model for the HRF and enable voxelwise inference in the General Linear Model (GLM) framework.

2.1 Standard GLM-based approach

In this formulation, the modelling of the BOLD response *i.e.* the definition of the *design* matrix \mathbf{X} is crucial. In its simplest form, this matrix relies on a spatially invariant temporal model of the BOLD signal across the brain meaning that the expected response to each stimulus is modelled by a single regressor. Assuming the neurovascular system as *linear* and *time-invariant* (LTI), this regressor is built up as the convolution of the stimulation signal \mathbf{x}^m associated to the m^{th} stimulus type with the canonical HRF \mathbf{h}_c , *i.e.* a composition of two gamma functions which reflects the BOLD signal best in the visual and motor cortices [8]. The GLM therefore reads:

$$[\mathbf{y}_1, \dots, \mathbf{y}_J] = \mathbf{X} [\mathbf{a}_1, \dots, \mathbf{a}_J] + [\mathbf{b}_1, \dots, \mathbf{b}_J] \quad (1)$$

where \mathbf{y}_j is the fMRI time series measured in voxel V_j at times $(t_n)_{n=1:N}$ and $\mathbf{a}_j \in \mathbb{R}^M$ defines the vector of BOLD effects in V_j for all stimulus type $m = 1 : M$. Noise \mathbf{b}_j is usually modelled as a first-order autoregressive (*i.e.* AR(1)) process in order to account for the spatially-varying temporal correlation of fMRI data [5]: $b_{j,t_n} = \rho_j b_{j,t_{n-1}} + \varepsilon_{j,t_n}$, $\forall j, t$, with $\varepsilon_j \sim \mathcal{N}(\mathbf{0}_N, \sigma_{\varepsilon_j}^2 \mathbf{I}_N)$, where $\mathbf{0}_N$ is a null vector of length N , and \mathbf{I}_N stands for the identity matrix of size N . Then, the estimated BOLD magnitudes $\hat{\mathbf{a}}_j$ in V_j are computed in the maximum likelihood sense by:

$$\hat{\mathbf{a}}_j = \arg \min_{\mathbf{a} \in \mathbb{R}^M} \|\mathbf{y}_j - \mathbf{X}\mathbf{a}_j\|_{\hat{\sigma}_{\varepsilon_j}^{-2} \hat{\mathbf{\Lambda}}_j}^2,$$

where $\hat{\sigma}_{\varepsilon_j}^{-2} \hat{\mathbf{\Lambda}}_j$ defines the inverse of the estimated autocorrelation matrix of \mathbf{b}_j ; see for instance [9] for details about the identification of the noise structure. Later, extensions that incorporate prior information on the BOLD effects $(\mathbf{a}_j)_{j=1:J}$ have been developed in the Bayesian framework [10]. In such cases, vectors $(\hat{\mathbf{a}}_j)_{j=1:J}$

are computed using more computationally demanding strategies [10]. However, all these GLM-based contributions consider a unique and global model of the HRF shape while intra-individual differences in its characteristics have been exhibited between cortical areas [11].

2.2 Flexible GLM models

Intra-individual differences in the characteristics of the HRF have been exhibited between cortical areas in [11, 12, 13, 14]. Although smaller than inter-individual fluctuations, this regional variability is large enough to be regarded with care. To account for these spatial fluctuations at the voxel level, one usually resorts to hemodynamic function basis. For instance, the canonical HRF \mathbf{h}_c can be supplemented with its first and second time derivatives ($[\mathbf{h}_c \mid \mathbf{h}'_c \mid \mathbf{h}''_c]$) to model eg. differences in time-to-peak. Although powerful and elegant, flexibility is achievable at the expense of fewer effective degrees of freedom and decreased sensitivity in any subsequent statistical test. Importantly, in a GLM involving several regressors per condition, the BOLD effect becomes multivariate (*i.e.* $\mathbf{a}_j^m \in \mathbb{R}^P$) and the Student-t statistic can no longer be used to infer on differences $\mathbf{a}_j^m - \mathbf{a}_j^n$ between the m^{th} and n^{th} stimulus types. Rather, an *unsigned* Fisher statistic has to be computed, making direct interpretation of activation maps more difficult. Indeed, the null hypothesis is actually rejected whenever any of the contrast components deviates from zero and not specifically when the difference of the response magnitudes is far from zero.

2.3 Beyond parametric modelling

The localisation of brain activation strongly depends on the modelling of the brain response and thus of its estimation. Of course, the converse also holds: HRF estimation is only relevant in voxels that elicit signal fluctuations correlated with the paradigm. Hence, detection and estimation are intrinsically linked to each other. The key point is therefore to tackle the two problems in a common setting, *i.e.* to set up a formulation in which *detection* and *estimation* enter naturally and simultaneously. This setting cannot be the classical hypothesis testing framework. Indeed, the sequential procedure which consists in first estimating the HRF on a given dataset and then building a specific GLM upon this estimate for detecting activations in the same dataset, entails statistical problems in terms of sensitivity and specificity: the control of the false positive rate actually becomes hazardous due to the use of an erroneous number of degrees of freedom. Instead, we explore a Bayesian approach that provides an appropriate framework, called the Joint Detection Estimation (JDE) framework in what follows, to address both detection and estimation issues in the same formalism.

Fig. 1 Sagittal views of a color-coded multi-subject parcellation. Left: Subject 1. Right: Subject 2



2.3.1 Regional non-parametric modelling of the BOLD signal

A spatially varying HRF model is necessary to keep a *single* regressor per condition, and thus enable direct statistical comparison ($\widehat{\beta}_j^m - \widehat{\beta}_j^n$). The JDE framework proposed in [15, 16] allows us to introduce such a spatially adaptive GLM in which a local estimation of \mathbf{h} is performed. To conduct the analysis efficiently, HRF estimation is carried out at a regional scale coarser than the voxel level. To properly define this scale, the functional brain mask is divided in K functionally homogeneous *parcels* using the parcellation technique proposed in [17]. This algorithm relies on the minimisation of a compound criterion reflecting both the spatial and functional structures and hence the topology of the dataset. The spatial similarity measure favours the closeness in the Talairach coordinates system. The functional part of this criterion is computed on parameters that characterise the functional properties of the voxels, for instance the fMRI time series themselves.

The number of parcels K is set by hand. The larger the number of parcels, the higher the degree of within-parcel homogeneity but potentially the lower the signal-to-noise ratio (SNR). To objectively choose an adequate number of parcels, Bayesian information criterion (BIC) and cross validation techniques have been used in [18] on an fMRI study of ten subjects. The authors have shown converging evidence for $K \approx 500$ for a whole brain analysis leading to typical parcel sizes around a few hundreds voxels. Importantly, since the parcellation is derived at the group level, there is a one-to-one correspondance of parcels across subjects, as shown in Fig. 1.

The parcel-based model of the BOLD signal introduced in [15, 16] is illustrated in Fig. 2. As shown, this model supposes that the HRF shape \mathbf{h} is constant within a parcel, while the magnitudes of activation a_j^m can vary in space and across stimulus types. Moreover, the model is said *non-parametric* since no model is assumed for the impulse response \mathbf{h} , which has therefore to be identified in each parcel. Let $\mathcal{P} = (V_j)_{j=1:J}$ be the current parcel. Then, the generative BOLD model reads:

$$\mathbf{y}_j = \sum_{m=1}^M a_j^m \mathbf{X}^m \mathbf{h} + \mathbf{P} \boldsymbol{\ell}_j + \mathbf{b}_j, \quad \forall j, V_j \in \mathcal{P}. \quad (2)$$

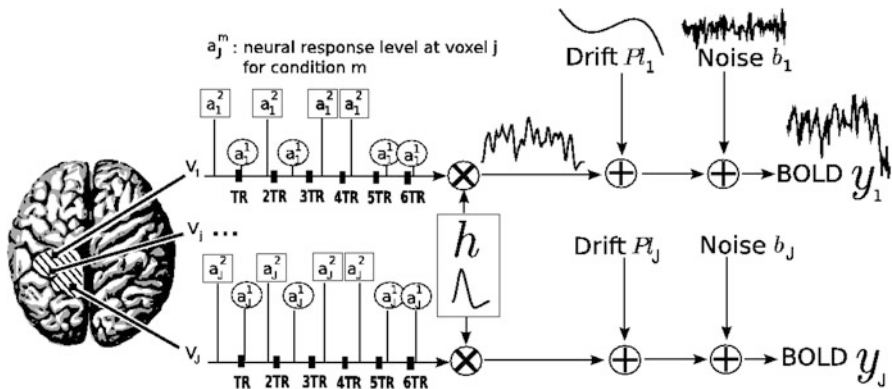


Fig. 2 Regional model of the BOLD signal in the JDE framework. The neural response levels a_j^m match with the BOLD effects a_j^m

\mathbf{X}^m denotes the $N \times (D + 1)$ binary matrix that codes the onsets of the m^{th} stimulus. Vector $\mathbf{h} \in \mathbb{R}^{D+1}$ represents the unknown HRF shape in \mathcal{P} . The term Pl_j models a low-frequency trend to account for physiological artifacts and noise $b_j \sim \mathcal{N}(\mathbf{0}_N, \sigma_{\epsilon_j}^2 \mathbf{\Lambda}_j^{-1})$ stands for the above mentioned AR(1) process.

2.3.2 Spatial mixture modelling and Bayesian inference

The HRF shape \mathbf{h} and the associated BOLD effects $(a_j)_{j=1:J}$ are jointly estimated in \mathcal{P} . Since no parametric model is considered for \mathbf{h} , a smoothness constraint on the second order derivative is introduced to regularise its estimation; see [15]. On the other hand, our approach also aims at detecting which voxels in \mathcal{P} elicit activations in response to stimulation. To this end, prior mixture models are introduced on $(a^m)_{m=1:M}$ to segregate activating voxels from the non-activating ones in a stimulus-specific manner (i.e. for each m). In [16], it has been shown that Spatial Mixture Models (SMMs) make it possible to recover clusters of activation instead of isolated spots and hence to account for spatial correlation in the activation detection process without smoothing the data. As our approach stands in the Bayesian framework, other priors are formulated upon every other sought object in model (2). The reader is referred to [15, 16] for their expressions. Finally, inference is based upon the full posterior distribution $p(\mathbf{h}, (a_j), (\ell_j), \Theta | \mathbb{Y})$, which is sampled using a Gibbs sampling scheme [16]. Posterior mean (PM) estimates are therefore computed from these samples according to: $\hat{x}^{\text{PM}} = \sum_{k=L_0}^{L_1} x^{(k)} / L, \quad \forall x \in \{\mathbf{h}, (a_j), \Theta\}$ where $L = L_1 - L_0 + 1$ and L_0 stands for the length of the burn-in period. Note that this estimation process has to be repeated over each parcel of each subject's brain. Since the fMRI data are considered spatially independent across parcels, parallel implementation makes the computation faster: whole brain analysis is achievable in about 60 mn for $N = 125$ and $K = 500$.

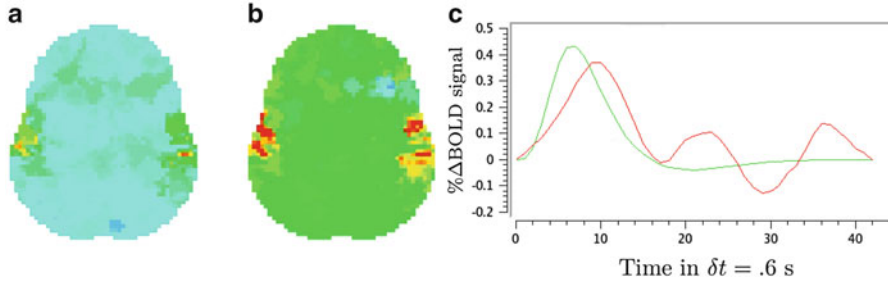


Fig. 3 BOLD effects estimates $(\hat{d}_j^{A.-V.})_j$ in a given subject for the A. – V. contrast. (a): SPM-based results obtained with the canonical HRF h_c . (b): JDE-based results considering model (2). c: Comparison of HRF shapes in the mostly activated parcel \mathcal{P} : h_c and $\hat{h}_{\mathcal{P}}$ appear in red and, respectively

2.4 Illustration on a real fMRI dataset

2.4.1 Data acquisition

Real fMRI data were recorded in fifteen volunteers during an experiment, which consisted of a single session of $N = 125$ scans lasting $TR = 2.4$ s each. The main goal of this experiment was to quickly map several brain functions such as motor, visual and auditory responses, as well as higher cognitive functions like computation. Here, we only focus on the auditory and visual experimental conditions and so on the *auditory-visual* contrast of interest (referenced as A. – V.).

2.4.2 Results

We compare the BOLD effect estimates for the two within-subject analyses under study. Fig. 3 clearly emphasizes for the A. – V. contrast that the JDE method achieves a better sensitivity (bilateral activations) in comparison with GLM-based inference when processing *unsmoothed* data. Indeed, the BOLD effects $\hat{d}_j^{A.-V.}$ have higher values in Fig. 3(b) and appear more enhanced. This is partly due to the modeling of spatial correlation using SMM in the JDE framework. As shown in Fig. 3(c)-[red line], notice that the HRF estimate $\hat{h}_{\mathcal{P}}$ computed in the mostly activating parcel deviates from the canonical shape depicted in Fig. 3(c)[green line].

3 Inference in group analyses: non-parametric and permutation tests

3.1 Classical parametric population-based inference

To clarify the context, assume that S subjects are selected randomly in a population of interest and submitted to the same fMRI experiment. As shown in previous sections, the two types of within-subject analyses produce, in one particular voxel V_j of the standardized space (usually, the MNI/Talairach space) and for each subject s , BOLD effect estimates $\hat{a}_{j,s}$. Comparison between experimental conditions is usually addressed through contrast definition. For mathematical convenience, we restrict ourselves to scalar contrasts. Hence we focus on signed differences $\hat{d}_{j,s}^{m-n} = \hat{\beta}_{j,s}^m - \hat{\beta}_{j,s}^n$ of the BOLD effect relative to the m^{th} and n^{th} stimulus types. For notational convenience, we will drop index j and the contrast under study $m - n$ in what follows.

While the estimated difference \hat{d}_s generally differs from the true but unobserved effect d_s , assume for now perfect within-subject estimation so that $\hat{d}_s = d_s$ for $s = 1 : S$. We thus are given a sample (d_1, \dots, d_S) drawn from an unknown probability density function (pdf) $f(d)$ that describes the distribution of the effects in the population. Here, we are concerned with inferences about a location parameter (mean, median, mode, ...). Assume for instance we wish to test the null hypothesis that the population mean is negative:

$$H_0 : \quad \mu_G = \int d f(d) dd \leq 0$$

where G stands for the group. To that end, we may use the classical one-sample t test. We start with computing the t statistic:

$$t = \frac{\hat{\mu}_G}{\hat{\sigma}_G / \sqrt{S}}, \text{ with } : \hat{\mu}_G = \frac{\sum_s d_s}{S}, \hat{\sigma}_G^2 = \frac{\sum_s (d_s - \hat{\mu}_G)^2}{S - 1} \tag{3}$$

Next, we reject H_0 , hence accept the alternative $H_1: \mu_G > 0$, if the probability under H_0 of attaining the observed t value is lower than a given false positive rate. Under the assumption that $f(d)$ is normal, this probability is well-known to be obtained from the Student distribution with $S - 1$ degrees of freedom. In this parametric context, the t statistic can be proved to be optimally sensitive (technically, in the sense of the uniformly most powerful unbiased test, see [19]).

3.2 *Non-Gaussian populations*

If normality is not tenable, however, the Student distribution is valid only in the limit of large samples, and may thus lead to inexact control over the false positive rate in small samples. This problem can be worked around using non-parametric calibration schemes such as sign permutations [19], which allow exact inferences under a milder assumption of symmetry regarding $f(d)$. Although we recommend permutation tests, they only provide an alternative strategy of thresholding a given statistic and, as such, address a *specificity* issue.

The fact that the sampling pdf $f(d)$ may not be normal also raises a *sensitivity* issue as the t statistic may no longer yield optimal power when normality does not hold. Without prior knowledge of the shape of $f(d)$, a reasonable default choice for the test statistic is one that maintains good detection performance over a wide range of pdfs. Such a statistic is robust, not quite in the classical sense of being resistant to outliers, but in the looser sense of being resistant to distributions that tend to produce outliers, such as heavy-tailed, or multimodal distributions. In the following, we use Wilcoxon's signed rank (WSR) statistic while other robust choices could be envisaged (Fisher's sign test, empirical likelihood ratio). As a matter of fact, such statistics have been used previously in fMRI group analyses [20], most often combined with permutation tests.

3.3 *Illustration on real fMRI dataset*

To enforce the coherence of our group level comparison with actual pipelines for fMRI data processing (SPM, FSL), the fMRI images that enter in model (1) were spatially filtered using isotropic Gaussian smoothing at 5 mm. In the JDE formalism, we still consider unsmoothed but normalized data to build the group parcellation as described in Fig. 4. Note that both approaches will be available in the next release of BrainVisa (<http://brainvisa.info>) in March, 2008.

Fig. 5 provides us with the WSR statistical maps, corrected for multiple comparisons in the permutation testing framework. The displayed slices matched with the place of most significant activations. Activation clusters appear larger in Fig. 5(b-d-f), *i.e.* using the GLM based approach, as a direct consequence of smoothing. The statistical map derived at the group level from the JDE analyses seems to have a lesser extent while being more significant at the cluster level than the GLM counterpart in the right hemisphere (left side). Moreover, the JDE formalism allows us to detect a gain in sensitivity since activations in Broca's area can be seen in the front of Fig. 5(a), right side. Table 1 confirms quantitatively these results and emphasizes that GLM-based inference systematically reports clusters of larger size (see col. 3). However, in terms of significance, the situation appears more contrasted since cluster level p -value is lower in the right hemisphere for

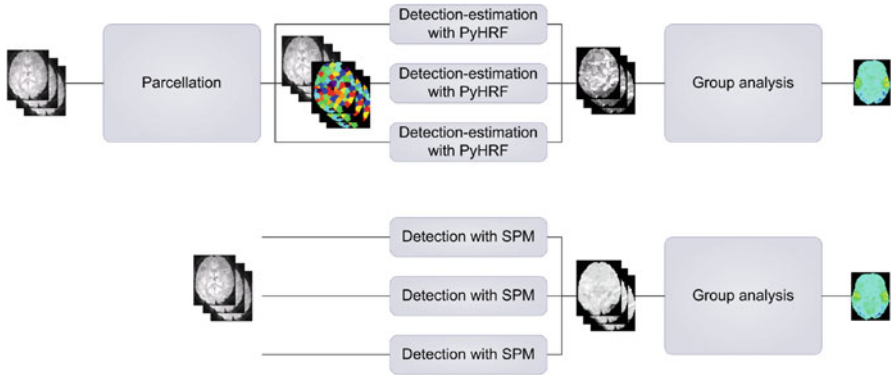


Fig. 4 Pipelines associated to the two fMRI group analyses

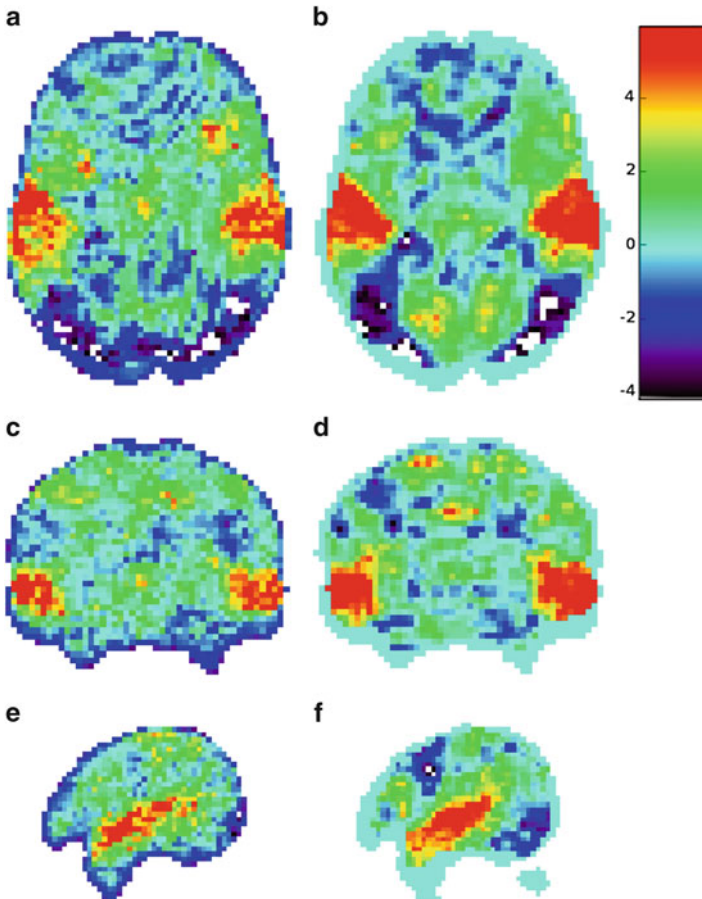
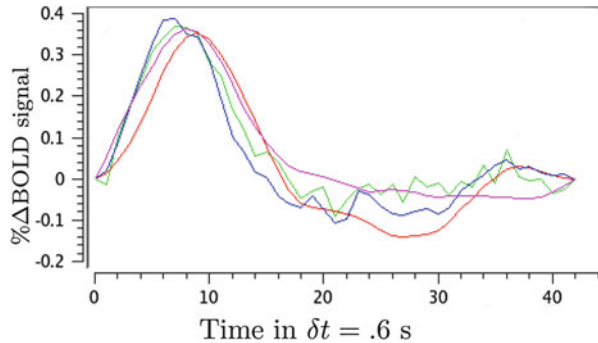


Fig. 5 RFX analysis maps based on the WSR statistics in the slice corresponding to the mostly activated cluster. **Radiological** convention: left is right. (a)-(c)-(e) and (b)-(d)-(f): results obtained using the JDE and SPM analyses at the subject level, respectively

Table 1 Suprathreshold clusters summary for the WSR statistic

	Cluster level	Cluster size (voxels)	Voxel level	Peak coords.		
	p_{corr}		p_{corr}	x	y	z
JDE	0.002	1151	1e-06	8	30	26
	0.003	876	0.0007	47	27	30
SPM	0.0022	1788	0.0001	5	29	28
	0.0028	1680	0.0001	45	27	27

Fig. 6 Subjects are color-coded: HRF estimates computed over the parcel associated to the voxel of maximal WSR value

JDE (top line in Table 1) in one cluster over two and thus provides most significant activation. This might be a consequence of the between-subject variability that we observed in the HRF estimate as reported in Fig. 6.

4 Spatial models for group analyses: accounting for spatial intersubject variability

In group analyses of fMRI data, the question is to decide which regions show a positive effect in average across subjects for a given functional contrast. This can be assessed in mass univariate framework through a normalization of the images to a common template, which in turns coregisters the data across subjects. However, as discussed in Sect. 4.1, this procedure aligns neither anatomical landmarks nor functional regions very accurately. Some solutions are thus proposed, which fall into two categories: *i*) a prior subdivision of the brain into putatively homogeneous regions (parcels) that may be better matched across subjects than voxels (Sect. 4.2) and *ii*) structural approaches that try to extract individual patterns and compare them at the group level (Sect. 4.3).

4.1 Taking into account the absence of a brain reference

Brain normalization consists in warping anatomical MR images in order to match a template, usually the MNI template [25]. Such warping procedures typically use an affine coregistration followed by a diffeomorphic, low-frequency spatial deformation model. More recent approaches use a segmentation of the grey matter to have a more accurate matching of the cortical surface [21]. Functional MR images, which are assumed to be coregistered with the individual anatomy, are then warped accordingly. Assuming a correct correspondence with the template, this allows a massively univariate analysis of group fMRI data using a general linear model (typically a mixed effects model).

However, this approach suffers from important shortcomings: *i*) From an anatomical point of view, sulco-gyral patterns, but also cyto-architectonic areas [22] are not -and probably cannot be- matched exactly through a diffeomorphism across individuals. A straightforward example is given by sulci that cannot be found in some subjects, or have unusual structure [23]. In other terms, there is yet no common template of the brain structures. *ii*) From an empirical point of view, it is commonly found that the variability that remains after normalization for some anatomical landmarks or functional areas is about 1cm. The commonly adopted solution to this problem is a blurring of the data (using 8-12 mm FWHM Gaussian kernel), which reduces the mismatch that remains after normalization. This solution is not fully satisfactory, because volume-based smoothing mixes the signals from heterogeneous anatomical compartments. More generally, it is doubtful that the resulting group maps yield an unbiased picture of the activation maps.

When considering specifically cortical structure, a partial solution consists in making the analysis on the cortical surface instead of the brain volume. Although the detailed correspondence of sulco-gyral patterns across subjects remains an issue, the constraints introduced by the surface-based representation benefit to the specificity of the group-level activation maps [22,24]. However, one-to-one mapping is still not possible.

A practical solution to this normalization issues has been to identify some regions based on a separate experiment (*functional localizer*) in each individual, and then to perform group inference at the region level. This approach is justified by the fact that functional information itself may be the most reliable landmark to discriminate and identify different brain regions, given the residual cyto-architectonic and sulco-gyral variability [25]. Moreover, this procedure also bypasses a traditional difficulty in fMRI-based inference, the multiple comparison problem, since only few statistical tests are performed -one for each region of interest (ROI). However, one might be concerned by the objectivity or reproducibility of these kind of procedure, as well as the bias induced by the restriction to these ROIs.

4.2 *Parcelling the brain*

A parcellation of the brain is a division of the brain into entities which are thought to correspond to well-defined anatomical or functional regions. In the context of group inference for neuroimaging, basing the analysis on parcels amounts to sacrificing spatial resolution to obtain a more reliable as well as interpretable matching of functional regions. Although atlas-based divisions are quite frequently used, it should be pointed out that these procedures do not adapt to the individual anatomy, and thus do not address the problem raised here.

Parcellations can be based on anatomical or functional features. Anatomical parcellations are usually based on sulco-gyral anatomy [26], and yield some segmentations of the main cortical lobes into a few tens of connected regions. Basal ganglia and the cerebellum are handled with specific (usually atlas-based) procedures. However, these procedures yield extremely coarse divisions of the brain, and thus cannot be used straightforwardly for brain mapping. Sulci-based parcellations can be performed at a much finer scale in individual datasets [27] but then the correspondence of the parcels between subject can be difficult to guarantee.

As we have noticed, functional information itself could ultimately be a very useful feature to segment the brain into small homogeneous regions. Brain parcellation can thus be driven by both functional and anatomical information. Such procedures are usually based on clustering algorithms that segment small clusters of voxels so that *i*) each parcel should be defined in each subject of the group under study, *ii*) the parcels should be functionally homogeneous and spatially connected within subjects *iii*) they should have similar functional properties and position across subjects. Ideally, parcellations could be performed on some functional localizer experiment, and then used to perform group inference in some experiment of interest. Alternatively, the same data can be used for parcellation and inference, but then the inference procedure should be based on a costly non-parametric test which involves the computation of parcellation and statistic [17].

Parcellations are an especially interesting tool if one is interested in describing the modular structure of the brain. This perspective raises important questions on how many brain modules could be delineated at the population level based on some functional information. There is clearly a compromise between the accuracy of the description, which would favor small parcels, and the inter-subject reproducibility of the structures, which is better assessed with coarse descriptions.

4.3 *Structural approaches*

The model underlying anatomo-functional parcellations is that the brain consists of elementary units or modules. In many cases, this point of view may be considered as limited. For instance, the activated regions for some tasks may be more general patterns than simple cortical patches, and it may be more useful to describe them in

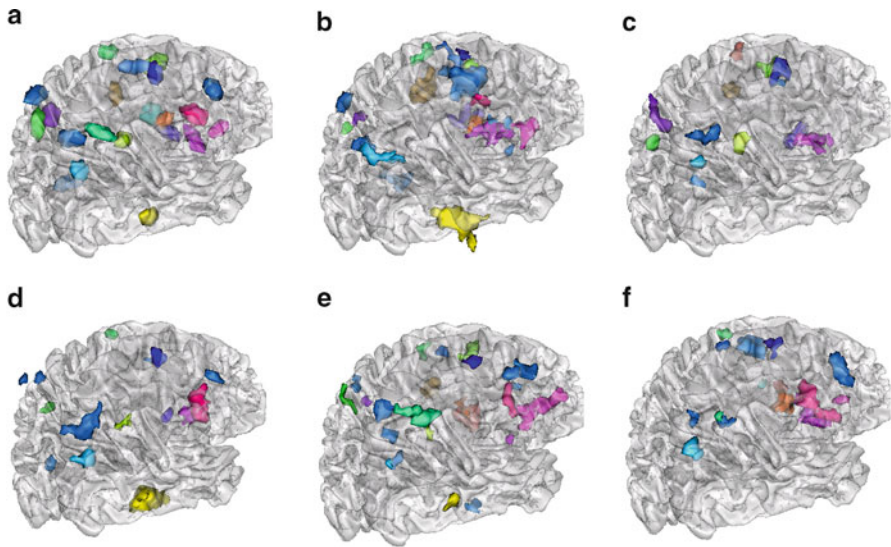


Fig. 7 structural analysis of the regions involved in a computation task: (a) Group template built from datasets from 10 subjects. (b-f) corresponding regions in 5 representative subjects. It is important to notice that there is no one-to-one correspondence, but a global pattern similarity between subjects. Moreover, the group template should be taken as an abstract model of the individual maps, while each individual data represent true activations; see [29] for more details

terms of blobs, peaks or gradients. Finding which patterns of that kind are present in a dataset and how frequent/typical they are in a population of subjects is what we call here a structural approach to understanding brain functional anatomy. By contrast with traditional approaches this kind of inference follows bottom-up strategy, where objects extracted individually are compared at a high-level of description.

Typically, structural features or patterns that are relevant for descriptions are local maxima of activity, regions segmented by watershed methods or blob models. The emphasis may be either on the structure of the active regions (bifurcations pattern), or merely the peaks (local maxima).

Whatever the kind of pattern extracted from the data, the most difficult questions are *i*) to decide how likely these patterns represent true activity or noise-related artifacts; *ii*) to infer a pattern common to the population of subjects. Several approaches have been discussed in the literature

- The description in terms of scale-space blobs embedded in a Markov-Random field was introduced in [28] in order to yield a kind of threshold-free analysis, where inter-subject reproducibility plays a key role for inference. A stepwise re-formulation of this approach was proposed in [29]. See Fig. 7.
- The idea that peaks could be a reliable feature to describe the information carried by an activity map and implicitly align the subjects yielded the concept of brain functional landmark [30].

The benefit of this kind of structural method is that regions extracted from individual datasets can then be compared across subjects or groups from the point of view of their position, size or shape, which is not possible in traditional voxel-, cluster- or even parcel-based inference framework.

Conclusion

One of the challenges of functional neuroimaging analysis methods and softwares will be to enable neuroscientists to capitalize on their past experiments to build finer or more complex models of information processing in the brain. Treating the brain as an homogeneous three-dimensional space may not be the well suited for that purpose. If the real matter is to understand the specialization of functional regions, and what characterizes them both on anatomical and functional aspects, it will be important not to rely solely on voxel-based descriptions, but also on region-based or structure-based approaches.

References

1. J. Fodor, "Let your brain alone.", *London review of books*, 1999.
2. R. M.E., M. AM, S. AZ, P. WJ, G. DA, and S. GL., "Blood flow and oxygen delivery to human brain during functional activity: theoretical modeling and experimental data", *Proc Natl Acad Sci*, vol. 98(12), pp. 6859–64, June 2001.
3. S. Ogawa, T. Lee, A. Kay, and D. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation", *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 9868–9872, 1990.
4. J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System : An Approach to Cerebral Imaging*, Thieme Medical Publishers, Inc., Georg Thieme Verlag, Stuttgart, New York, 1988.
5. K. Worsley, C. Liao, J. Aston, V. Petre, G. Duncan, F. Morales, and A. Evans, "A general statistical analysis for fMRI data", *Neuroimage*, vol. 15, pp. 1–15, Jan. 2002.
6. J.-F. Mangin, D. Rivière, O. Coulon, C. Poupon, A. Cachia, Y. Cointepas, J.-B. Poline, D. L. Bihan, J. Régis, and D. Papadopoulos-Orfanos, "Coordinate-based versus structural approaches to brain image analysis", *Artificial Intelligence in Medicine*, vol. 30, pp. 177–197, 2004.
7. G. Yovel and N. Kanwisher, "Face perception: domain specific, not process specific.", *Neuron*, vol. 44, pp. 747–8, Dec. 2004.
8. G. H. Glover, "Deconvolution of impulse response in event-related BOLD fMRI", *Neuroimage*, vol. 9, pp. 416–429, 1999.
9. W. D. Penny, S. Kiebel, and K. J. Friston, "Variational Bayesian inference for fMRI time series", *Neuroimage*, vol. 19, pp. 727–741, 2003.
10. M. Woolrich, M. Jenkinson, J. Brady, and S. Smith, "Fully Bayesian spatio-temporal modelling of fMRI data", *IEEE Trans. Med. Imag.*, vol. 23, pp. 213–231, Feb. 2004.
11. D. A. Handwerker, J. M. Ollinger, , and M. D'Esposito, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses", *Neuroimage*, vol. 21, pp. 1639–1651, 2004.
12. G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human BOLD hemodynamic responses", *Neuroimage*, vol. 7, pp. 574, 1998.

13. F. M. Miezin, L. Maccotta, J. M. Ollinger, S. E. Petersen, and R. L. Buckner, "Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing", *Neuroimage*, vol. 11, pp. 735–759, 2000.
14. J. Neumann and G. Lohmann, "Bayesian second-level analysis of functional magnetic resonance images", *Neuroimage*, vol. 20, pp. 1346–1355, 2003.
15. S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline, "Joint detection-estimation of brain activity in functional MRI: a multichannel deconvolution solution", *IEEE Trans. Signal Processing*, vol. 53, pp. 3488–3502, Sep. 2005.
16. T. Vincent, P. Ciuciu, and J. Idier, "Spatial mixture modelling for the joint detection-estimation of brain activity in fMRI", in *32th Proc. IEEE ICASSP*, Honolulu, Hawaii, Apr. 2007, vol. I, pp. 325–328.
17. B. Thirion, G. Flandin, P. Pinel, A. Roche, P. Ciuciu, and J.-B. Poline, "Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets", *Hum. Brain Mapp.*, vol. 27, pp. 678–693, Aug. 2006.
18. B. Thyreau, B. Thirion, G. Flandin, and J.-B. Poline, "Anatomo-functional description of the brain: a probabilistic approach", in *Proc. 31th Proc. IEEE ICASSP*, Toulouse, France, May 2006, vol. V, pp. 1109–1112.
19. P. Good, *Permutation, Parametric, and Bootstrap Tests of Hypotheses*, Springer, 3rd edition, 2005.
20. S. Mériaux, A. Roche, B. Thirion, and G. Dehaene-Lambertz, "Robust statistics for nonparametric group analysis in fMRI", in *Proc. 3th Proc. IEEE ISBI*, Arlington, VA, Apr. 2006, pp. 936–939.
21. J. Ashburner, "A fast diffeomorphic image registration algorithm.", *Neuroimage*, vol. 38, pp. 95–113, Oct 2007.
22. B. Fischl, N. Rajendran, E. Busa, J. Augustinack, O. Hinds, B. T. T. Yeo, H. Mohlberg, K. Amunts, and K. Zilles, "Cortical folding patterns and predicting cytoarchitecture.", *Cereb Cortex*, Dec 2007.
23. D. Rivière, J.-F. Mangin, D. Papadopoulos-Orfanos, J.-M. Martinez, V. Frouin, and J. Régis, "Automatic recognition of cortical sulci of the human brain using a congregation of neural networks", *Medical Image Analysis*, vol. 6, pp. 77–92, 2002.
24. B. Fischl, M. I. Sereno, R. B. Tootell, and A. M. Dale, "High-resolution intersubject averaging and a coordinate system for the cortical surface.", *Hum Brain Mapp*, vol. 8, pp. 272–284, 1999.
25. M. Brett, I. Johnsrude, and A. Owen, "The problem of functional localization in the human brain.", *Nature Reviews Neuroscience*, vol. 3, pp. 243–249, Mar. 2002.
26. R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, and R. J. Killiany, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest.", *Neuroimage*, vol. 31, pp. 968–980, July 2006.
27. G. Flandin, F. Kherif, X. Pennec, G. Malandain, N. Ayache, and J.-B. Poline, "Improved detection sensitivity of functional MRI data using a brain parcellation technique", in *Proc. 5th MICCAI*, Tokyo, Japan, Sep. 2002, LNCS 2488 (Part I), pp. 467–474, Springer Verlag.
28. O. Coulon, J.-F. Mangin, J.-B. Poline, M. Zilbovicius, D. Roumenov, Y. Samson, V. Frouin, and I. Bloch, "Structural group analysis of functional activation maps", *Neuroimage*, vol. 11, pp. 767–782, 2000.
29. B. Thirion, P. Pinel, A. Tucholka, A. Roche, P. Ciuciu, J.-F. Mangin, and J.-B. Poline, "Structural analysis of fMRI data revisited: Improving the sensitivity and reliability of fMRI group studies", *IEEE Trans. Med. Imag.*, vol. 26, pp. 1256–1269, Sep. 2007.
30. B. Thirion, P. Pinel, and J.-B. Poline, "Finding landmarks in the functional brain: detection and use for group characterization.", *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, vol. 8, pp. 476–483, 2005.

Diffusion Tensor Estimation, Regularization and Classification

R. Neji, N. Azzabou, G. Fleury, and N. Paragios

Abstract In this chapter, we explore diffusion tensor estimation, regularization and classification. To this end, we introduce a variational method for joint estimation and regularization of diffusion tensor fields from noisy raw data as well as a Support Vector Machine (SVM) based classification framework.

In order to simultaneously estimate and regularize diffusion tensor fields from noisy observations, we integrate the classic quadratic data fidelity term derived from the Stejskal-Tanner equation with a new smoothness term leading to a convex objective function. The regularization term is based on the assumption that the signal can be reconstructed using a weighted average of observations on a local neighborhood. The weights measure the similarity between tensors and are computed directly from the diffusion images. We preserve the positive semi-definiteness constraint using a projected gradient descent.

The classification framework we consider in this chapter allows linear as well as non linear separation of diffusion tensors using kernels defined on the space of symmetric positive definite matrices. The kernels are derived from their counterparts on the statistical manifold of multivariate Gaussian distributions with zero mean

R. Neji
Siemens Healthcare, UK
e-mail: radhouene.neji@gmail.com

N. Azzabou (✉)
Institute of Myology 47 Boulevard Hôpital, 75013 Paris, France
e-mail: n.azzabou@institut-myologie.org

G. Fleury
Ecole Centrale Pékin
No. 37 Xueyuan Street, Haidian District Beijing, 100191, P.R. China
e-mail: gilles.fleury@supelec.fr

N. Paragios
Center for Visual Computing, Department of Applied Mathematics, Ecole Centrale Paris, Paris, France
e-mail: nikos.paragios@ecp.fr

or from distance substitution in the Gaussian Radial Basis Function (RBF) kernel. Experimental results on diffusion tensor images of the human skeletal muscle (calf) show the potential of our algorithms both in denoising and SVM-driven Markov random field segmentation.

1 Introduction

Diffusion tensor imaging (DTI) is an emerging non-invasive modality allowing the quantitative investigation of water protons diffusion within biologic tissues. Such a modality offers measurements of the amount of diffusion of water molecules in several different directions. One then can infer the estimation of a tensor which is a 3×3 symmetric positive definite matrix representing the uncertainty on the position of water protons with a Gaussian model of displacement. It can be visualized as an ellipsoid: the axes correspond to the principal directions of diffusion (eigenvectors) and the radii correspond to the amount of diffusion along each principal direction (eigenvalues). Diffusion in the presence of organized anatomical structures can reveal informative properties inherent to the architecture of the imaged organs. More explicitly, diffusion is hindered along some directions and facilitated along others that often coincide with fiber trajectories. Therefore, DTI has been used mostly in brain studies and has become a tool to infer white matter connectivity [4]. It is also starting to be used in the study of the architecture of the muscles of the lower leg [10, 11].

We address two issues that arise in the processing of diffusion tensor images: joint estimation and denoising of tensors as well as tensor classification. In fact, the DTI experimental protocol yields noisy observations due to the diffusion-sensitizing magnetic gradient. Furthermore, the clinical protocols refer to relatively low magnet strength, or a rather low signal-to-noise ratio. Therefore, signal reconstruction is crucial to obtain an appropriate estimate of the tensor field and for subsequent use of this estimate in applications like fiber tractography. Besides, since some anatomical structures show distinct fiber orientations, classifying diffusion tensors in order to learn orientation and eigenvalue distributions from manually segmented examples is an appealing task that can be useful to guide segmentation algorithms.

The remainder of this chapter is structured as follows: in Sect. 2, we review the previous work that addressed diffusion tensor regularization and classification. In Sect. 3, we detail the proposed framework of joint estimation and regularization. The SVM classification of tensors along with three classes of kernels on the manifold of symmetric positive definite matrices are treated in Sect. 4, where we also introduce the SVM-driven Markov random fields for segmentation purposes. Sect. 5 is dedicated to experimental results both on synthetic and clinical data of the calf muscle. In Sect. 6, we discuss the perspectives of this work.

2 Previous work

Several methods have been proposed to address diffusion tensor regularization. In [7], a two-step regularization was proposed consisting of the restoration of the principal diffusion directions using a total variation-model followed by the smoothing of the eigenvalues using an anisotropic tensor-driven formulation. In [2], the maximization of a log-posterior probability based on the Rician noise model is considered to smooth directly the diffusion-weighted images. A Bayesian model based on a Gaussian Markov Random Field was used in [20] to smooth the diffusion tensors. In [6], the authors consider the tensors as lying on a Riemannian manifold and use the corresponding distance to derive a local weighted averaging for DTI denoising. Tensors are assumed to be positive-definite matrices which was taken into account in [8] where an anisotropic filtering of the L^2 norm of the gradient of the diffusion tensor was considered and their proposed PDE scheme constrains the estimation to lie on this space. Such a concept was further developed in [28] where a variational method was proposed that aimed to minimize the L^p norm of the spatial gradient of the diffusion tensor under a constraint involving the non-linear form of Stejskal-Tanner equation. A non linear diffusion scheme is described in [29] where smoothing is made direction-dependent using a diffusion matrix in the PDE system. More recently, in [9] a joint reconstruction and regularization was proposed in the context of an energy minimization in a Log-Euclidean framework. The existing variational methods focused disproportionately on enforcing the positive-definiteness constraint, with the regularization term usually chosen as a function of the norm of the gradient. The main limitation of most of the above-mentioned methods is the nature of the cost function (non-convex) that entails a preliminary initialization step, while little attention was paid to defining appropriate smoothness components that account for the expected nature of tensors.

As far as diffusion tensor classification is concerned, very few previous works tried to include a priori knowledge about the diffusion tensors or fiber tracts [19, 21], mainly because they focused on brain anatomy and connectivity which are less well-known than the fiber architecture of the human skeletal muscle for example, where each muscle group has specific pennation angles. Common machine learning techniques were therefore mainly applied to separate populations of healthy and ill patients on a voxel-wise basis rather than to learn distributions of diffusion tensors within a specific anatomical region in order to guide segmentation algorithms [15, 27]. Besides, the non linear techniques used in the literature relied mainly on kernels that are not specific to the space of symmetric positive definite matrices, mainly the well known Gaussian and polynomial kernels.

3 DTI Estimation and Regularization

Let us assume that n DTI acquisitions $(S_k)_{k=1 \dots n}$ with respect to different magnetic gradient directions $(\mathbf{g}_k)_{k=1 \dots n}$ are available. Ideally, the expected signal at a voxel \mathbf{x} for the direction k as explained in [24] should respect the following condition

$$S_k(\mathbf{x}) = S_0(\mathbf{x}) \exp(-b \mathbf{g}_k^t \mathbf{D}(\mathbf{x}) \mathbf{g}_k)$$

with the tensor \mathbf{D} being the unknown variable and b a value that depends on the acquisition settings. The estimation of the tensors in the volume domain Ω can be done through direct inference (6 acquisitions are at least available), which is equivalent to minimizing:

$$E_{data}(\mathbf{D}) = \int_{\Omega} \sum_{k=1}^n (\log(S_k(\mathbf{x})/S_0(\mathbf{x})) + b \mathbf{g}_k^t \mathbf{D}(\mathbf{x}) \mathbf{g}_k)^2 d\mathbf{x}$$

This energy is based on the linearized diffusion tensor model which is reasonable for moderate values of SNR [23]. Such a direct estimation is quite sensitive to noise, on the other hand, it refers to a convex term, which is rather convenient when seeking its lowest potential. The most common approach to account for noise is through the use of an additional regularization term which constrains the estimation of \mathbf{D} to be locally smooth. One of the most prominent uses of DTI is fiber extraction. Therefore it is natural to assume that locally these fibers do have similar orientations. In such a context, the tensor can be expressed as a linear combination of the tensors lying in its neighborhood since they are likely to represent the same population of fibers. Such a regularization constraint was introduced in the case of image restoration in [1]. This assumption still holds at the boundaries between different groups of fibers as long as the linear combination is thoroughly chosen to ensure that the contribution of tensors belonging to a different fiber population is negligible. It is also more accurate than the underlying assumption of total-variation based approaches where the tensor field is considered piecewise constant. This leads us to define the following regularization component:

$$E_{smooth}(\mathbf{D}) = \int_{\Omega} \left\| \mathbf{D}(\mathbf{x}) - \frac{1}{Z(\mathbf{x})} \int_{y \in \mathcal{N}_x} w(\mathbf{x}, \mathbf{y}) \mathbf{D}(\mathbf{y}) d\mathbf{y} \right\|_F^2 d\mathbf{x}$$

where $w(\mathbf{x}, \mathbf{y})$ reflects the similarity between tensors $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$, $\|A\|_F$ being the Frobenius norm $\|A\|_F = \sqrt{\text{tr}(A^t A)}$ and $Z(\mathbf{x})$ is a normalization factor, i.e $Z(\mathbf{x}) = \int_{y \in \mathcal{N}_x} w(\mathbf{x}, \mathbf{y}) d\mathbf{y}$. The most critical aspect of such an approximation model is the definition of weights, measuring the similarity between tensors within the local neighborhood. The use of Gaussian weights is a common weight's selection, i.e $\left[w(\mathbf{x}, \mathbf{y}) = e^{-\frac{d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y}))}{2\sigma^2}} \right]$, where $d(\cdot, \cdot)$ is a distance between tensors and σ a scale

factor. In the context of direct estimation and regularization it is more appropriate to define similarities directly on the observation space rather than the estimation space. Such a choice will lead to a tractable estimation, while preserving the convexity of the cost function. Our distance definition as well as our minimization step are based on the representation of symmetric positive semi-definite matrices S_+^3 as a convex closed cone in the Hilbert space of symmetric matrices S^3 , where the standard scalar product is defined by $\langle A, B \rangle_F = tr(A^T B)$ which induces the corresponding Frobenius norm.

3.1 Measuring Similarities from diffusion weighted images

We aim at simultaneously estimating and smoothing the tensor field, therefore the weights $w(\mathbf{x}, \mathbf{y})$ in E_{smooth} should be precalculated using the raw data. The most straightforward estimation of the distances can be done through the algebraic distance between the $\log(S_k/S_0)$ for two neighborhood voxels in any direction

$$d(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})) = \frac{1}{b} \sqrt{\sum_{k=1}^N \left(\log \left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})} \right) - \log \left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})} \right) \right)^2}$$

One can easily show that such an expression does not reflect similarity between tensors according to the norm $\| \cdot \|_F$. In fact, this leads to

$$d(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})) = \sqrt{\sum_{k=1}^N \langle \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \mathbf{G}_k \rangle_F^2}$$

where $\mathbf{G}_k = \mathbf{g}_k \mathbf{g}_k^t$ do not form necessarily an orthonormal basis. We use a Gram-Schmidt orthogonalization scheme to calculate an orthonormal basis $\tilde{\mathbf{G}}_k$ such that $\tilde{\mathbf{G}}_k = \sum_l \alpha_{kl} \mathbf{G}_l$ (each new vector of the new basis is a linear combination of the vectors of the initial basis). This procedure allows us to have an approximation of $\|\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})\|_F$ directly from the raw data S_k and S_0 as follows

$$\begin{aligned} \|\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})\|_F &= \sqrt{\sum_{k=1}^N \langle \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \tilde{\mathbf{G}}_k \rangle_F^2} \\ \frac{1}{b} &= \sqrt{\sum_{k=1}^N \left(\sum_l \alpha_{kl} \left(\log \left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})} \right) - \log \left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})} \right) \right) \right)^2} \end{aligned}$$

3.2 Semi-Definite Positive Gradient Descent

One now can seek the lowest potential of the cost function towards recovering the optimal solution on the tensor space. The present framework consists of a convex energy with a unique minimum which can be reached using a projected gradient descent on the space of semi-definitive positive matrices. The projection from S^3 onto S^3_+ denoted by $\Pi_{S^3_+}$ is well defined and has an explicit expression. Indeed, projecting M amounts to replacing the negative eigenvalues in its spectral decomposition by 0 [8, 12]. Note that we minimize over the set of semi-definite positive matrices because it is topologically closed, as opposed to the set of definite positive matrices. In the current setting, the problem is well posed and the projected gradient descent algorithm is convergent for a suitable choice of the time step dt . Using a weighting factor λ between the data attachment term and the regularization energy, the gradient descent can be expressed as the following equation

$$\begin{aligned} \mathbf{D}^{t+1}(\mathbf{x}) &= \Pi_{S^3_+} \left(\mathbf{D}^t(\mathbf{x}) - dt \frac{\partial E}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t) \right) \\ &= \Pi_{S^3_+} \left(\mathbf{D}^t(\mathbf{x}) - dt \lambda \frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t) - dt \frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t) \right) \end{aligned}$$

where

$$\begin{aligned} \frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2\mathbf{D}(\mathbf{x}) - 2 \int_{\mathbf{y} \in \mathcal{N}_x} \frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{x})} \mathbf{D}(\mathbf{y}) d\mathbf{y} \\ &\quad - 2 \int_{\mathbf{y} \in \mathcal{N}_x} \frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{x})} \left(\mathbf{D}(\mathbf{y}) - \int_{\mathbf{z} \in \mathcal{N}_y} \frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{y})} \mathbf{D}(\mathbf{z}) d\mathbf{z} \right) d\mathbf{y} \\ \frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2b \sum_{k=1}^N (\log(s_k(\mathbf{x})/s_0(\mathbf{x})) + b\mathbf{g}'_k D(\mathbf{x}) \mathbf{g}_k) \mathbf{G}_k \end{aligned}$$

Let us define the norm $\|\cdot\|_{TF}$ over the whole tensor field \mathbf{D} as $\|\mathbf{D}\|_{TF} = \int_{\Omega} \|\mathbf{D}(\mathbf{x})\|_F d\mathbf{x}$. Considering two tensor fields \mathbf{D}_1 and \mathbf{D}_2 , we show in the following that the gradient of our energy functional is L -Lipschitz. The constant L will allow us to choose automatically a time step that insures the convergence of the algorithm.

$$\begin{aligned} \left\| \frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_1) - \frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_2) \right\|_F &= 2b^2 \left\| \sum_{k=1}^N \langle \mathbf{G}_k, \mathbf{D}_1 - \mathbf{D}_2 \rangle_F \mathbf{G}_k \right\|_F \\ &\leq 2b^2 \sum_{k=1}^N \|\mathbf{G}_k\|_F^2 \|\mathbf{D}_1 - \mathbf{D}_2\|_F \end{aligned}$$

Therefore $\|\nabla E_{data}(\mathbf{D}_1) - \nabla E_{data}(\mathbf{D}_2)\|_{TF} \leq 2b^2 \sum_{k=1}^N \|\mathbf{G}_k\|_F^2 \|\mathbf{D}_1 - \mathbf{D}_2\|_{TF}$. Besides, we can easily show the following inequality

$$\| \nabla E_{smooth}(\mathbf{D}_1) - \nabla E_{smooth}(\mathbf{D}_2) \|_{TF} \leq 2(1 + 2|\mathcal{N}_x| + |\mathcal{N}_x|^2) \| \mathbf{D}_1 - \mathbf{D}_2 \|_{TF}$$

where $|\mathcal{N}_x|$ is the number of the considered neighbors. Thus the gradient of the objective function is L -Lipschitz with $L = 2b^2 \sum_{k=1}^N \| \mathbf{G}_k \|_F^2 + 2\lambda(|\mathcal{N}_x| + 1)^2$. Choosing $0 < dt < \frac{1}{b^2 \sum_{k=1}^N \| \mathbf{G}_k \|_F^2 + \lambda(|\mathcal{N}_x| + 1)^2}$ makes the projected gradient descent convergent [3].

We can give an interpretation of our regularization energy in terms of diffusion-weighted images smoothing. It can be easily verified that for each direction k

$$\int_{\Omega} \langle \mathbf{D}(\mathbf{x}) - \int_{y \in \mathcal{N}_x} \frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{x})} \mathbf{D}(\mathbf{y}) d\mathbf{y}, \mathbf{G}_k \rangle_F^2 d\mathbf{x} = \frac{1}{b^2} \int_{\Omega} \left[\log \left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})} \right) - \log \left(\prod_{y \in \mathcal{N}_x} \left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})} \right)^{\frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{x})}} \right) \right]^2 d\mathbf{x}$$

Using Cauchy-Schwartz inequality we obtain:

$$\frac{1}{b^2} \int_{\Omega} \left[\log \left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})} \right) - \log \left(\prod_{y \in \mathcal{N}_x} \left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})} \right)^{\frac{w(\mathbf{x}, \mathbf{y})}{Z(\mathbf{x})}} \right) \right]^2 d\mathbf{x} \leq E_{smooth} \| \mathbf{G}_k \|_F^2$$

We can see that minimizing E_{smooth} has a direct implication on the normalized diffusion weighted images $\frac{S_k}{S_0}$. Reconstructing the tensors using a linear combination of the tensors in its neighborhood leads to the reconstruction of the normalized signals using a weighted geometric mean of the neighboring signals where the weights are not calculated only with a single volume S_k but also with the volumes obtained from the other magnetic gradient directions.

4 SVM classification and kernels on tensors

4.1 Two-class Support Vector Machines

We briefly review the principles of two class SVMs [26]. Given N points x_i with known class information y_i (either $+1$ or -1), SVM training consists in finding the optimal separating hyperplane described by the equation $\mathbf{w}^t \mathbf{x} + b = 0$ with the maximum distance to the training examples. It amounts to solving a dual convex quadratic optimization problem and each data point \mathbf{x} is classified using the SVM output function $f(x) = (\sum_i^N \alpha_i y_i \mathbf{x} \mathbf{x}_i) + b$. The algorithm is extended to achieve non linear separation using a kernel function $K(\mathbf{x}, \mathbf{y})$ (symmetric, positive definite) that is used instead of the standard inner product.

4.2 Information diffusion kernel

In order to define a kernel on the set of definite positive matrices, we can propagate class and structure information using its geometry as a Riemannian manifold [22]. Intuitively, we can see the construction of this kernel as diffusing the labels of the training set to the whole set of definite positive matrices. Therefore, similarly to heat diffusion on a Euclidean space, where the solution is given by the convolution of the initial condition by a Gaussian kernel, heat diffusion on a Riemannian manifold is driven by a kernel function K_t and given by the following asymptotic series expansion [17]:

$$K_t(\mathbf{D}_1, \mathbf{D}_2) \propto \exp\left(-\frac{d^2(\mathbf{D}_1, \mathbf{D}_2)}{4t}\right) \sum_{n=0}^{\infty} a_n(\mathbf{D}_1, \mathbf{D}_2)t^n$$

where d corresponds to the geodesic distance induced by the Riemannian metric, a_n are the coefficients of the series expansion and t is the diffusion time, which is a parameter of the kernel. We use a first order approximation in t of the previous expression that yields

$$K_t(\mathbf{D}_1, \mathbf{D}_2) \propto \exp\left(-\frac{d^2(\mathbf{D}_1, \mathbf{D}_2)}{4t}\right)$$

In our case, d has an explicit expression given by $d(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{\sum_i (\log(\lambda_i))^2}$ where λ_i are the generalized eigenvalues of \mathbf{D}_1 and \mathbf{D}_2 [22].

4.3 Bregman divergence kernels

Instead of using the geodesic distance in the information diffusion kernel, one can instead use the Bregman divergence framework [25] to define a rich class of kernels parametrized by a convex scalar function $\phi : \mathcal{S}_+^3 \rightarrow \mathcal{R}$ that extend in a natural way the Euclidean distance and therefore the standard Gaussian radial basis function kernel. Knowing ϕ , one can define the corresponding Bregman divergence \mathcal{D}_ϕ between two matrices \mathbf{D}_1 and \mathbf{D}_2 as follows:

$$\mathcal{D}_\phi(\mathbf{D}_1, \mathbf{D}_2) = \phi(\mathbf{D}_1) - \phi(\mathbf{D}_2) - tr(\nabla\phi(\mathbf{D}_2)^t(\mathbf{D}_1 - \mathbf{D}_2))$$

The symmetrization of the divergence gives the following similarity measure $\widehat{\mathcal{D}}$:

$$\widehat{\mathcal{D}}_\phi(\mathbf{D}_1, \mathbf{D}_2) = tr((\nabla\phi(\mathbf{D}_1) - \nabla\phi(\mathbf{D}_2))^t(\mathbf{D}_1 - \mathbf{D}_2))$$

It is clear that choosing $\phi(\mathbf{D}) = \frac{1}{2}\|\mathbf{D}\|_F^2$ yields the standard Euclidean distance. Therefore, we extend the Gaussian radial basis function (RBF) kernel using the exponential embedding:

$$K(\mathbf{D}_1, \mathbf{D}_2) = \exp\left(-\gamma \widehat{\mathcal{D}}_\phi(\mathbf{D}_1, \mathbf{D}_2)\right)$$

Two interesting cases of ϕ are the Burg entropy [Eq. (1)] and the Von Neumann Entropy [Eq. (2)]:

$$\phi_1(\mathbf{D}) = -\log(\det(\mathbf{D})) \quad (1)$$

$$\phi_2(\mathbf{D}) = \text{tr}(\mathbf{D} \log(\mathbf{D}) - \mathbf{D}) \quad (2)$$

They induce the following kernels:

$$\begin{aligned} K_1(\mathbf{D}_1, \mathbf{D}_2) &= \exp(-\gamma \text{tr}((\mathbf{D}_1 - \mathbf{D}_2)(\mathbf{D}_1^{-1} - \mathbf{D}_2^{-1}))) \\ K_2(\mathbf{D}_1, \mathbf{D}_2) &= \exp(-\gamma \text{tr}((\mathbf{D}_1 - \mathbf{D}_2)(\log(\mathbf{D}_1) - \log(\mathbf{D}_2)))) \end{aligned}$$

These kernels provide global similarity measures that quantify the differences between tensors both in eigenvalues and eigenvectors. Note that the divergence that derives from Burg entropy can also be obtained from a Kullback-Leibler divergence between Gaussian distributions with zero mean [28].

4.4 Probability product kernels

The third kernel we will study is a probability product kernel [13] to derive a class of kernels over the set of positive definite matrices. We consider again the set of Gaussian probability distributions of zero mean, two elements p_1 and p_2 of this set with \mathbf{D}_1 and \mathbf{D}_2 their covariance matrices and the corresponding probability product kernel:

$$K(p_1, p_2) = \int p_1(\mathbf{x})^\rho p_2(\mathbf{x})^\rho d\mathbf{x} = \langle p_1^\rho, p_2^\rho \rangle_{L^2}$$

where ρ is a positive constant. Note that the special case $\rho = 1/2$ coincides with the well known Bhattacharyya kernel. Replacing the probabilities p_1 and p_2 by their expressions gives the following kernel:

$$K(\mathbf{D}_1, \mathbf{D}_2) = \det(\mathbf{D}_1)^{-\rho/2} \det(\mathbf{D}_2)^{-\rho/2} \det(\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-\rho}$$

This defines a Mercer kernel as the probability product kernel on Gaussian distributions is based on the scalar product in L^2 which is a Hilbert space. This is

an advantage over the two above-cited classes of kernels which are not necessarily positive definite. Note however that in practice, a thorough choice of the parameter γ will ensure the positive definiteness in a statistical sense, i.e. the property will hold with high probability [5].

4.5 SVM-driven Markov random fields

The goal behind the use of an MRF model is two-fold: we aim at including spatial information, i.e. tensors along the same fiber should belong to the same class and we also try to minimize the effect of noise during segmentation. Besides, the MRF framework allows to use all the scores given by the SVMs, instead of making the labeling decision by simply taking the maximum score. Therefore, we define the following energy E to minimize:

$$E = \sum_{i \in \Omega} u^s(l(i)) + \lambda \sum_{i \in \Omega, j \in \mathcal{N}(i)} u^p(l(i), l(j)) \quad (3)$$

where Ω is the image domain, $l(i)$ is the label of the voxel i , $\mathcal{N}(i)$ is the considered neighborhood, u^s is the potential given by the SVM scores and u^p is a pairwise potential that imposes spatial regularization. We choose $u^s(l(i)) = \exp(-\alpha f_l(\mathbf{D}(i)))$ which is a decreasing potential in the score given by a one-against-all SVM classifier f_l . The choice of the pairwise potential u^p is application-dependent and should include prior knowledge about the structure of the anatomical region that is to be segmented. For the case of the muscles of the calf, we know that the fibers have a privileged orientation since they follow the direction of the leg with a pennation angle, accordingly we propose two different costs for neighboring voxels: if the voxels i and j belong to the same axial slice, the pairwise potential u^p is set to

$$u^p(l(i), l(j)) = c(1 - \delta_{l(i), l(j)})$$

where c is a constant and δ is the Kronecker delta, otherwise we choose

$$u^p(l(i), l(j)) = \left(1 - \frac{1}{2} \frac{\mathbf{v}^t \mathbf{D}(i) \mathbf{v}}{\lambda_{\max}(i)} - \frac{1}{2} \frac{\mathbf{v}^t \mathbf{D}(j) \mathbf{v}}{\lambda_{\max}(j)}\right) (1 - \delta_{l(i), l(j)})$$

where $\mathbf{v} = \frac{i-j}{\|i-j\|}$ and $\lambda_{\max}(i)$ is the largest eigenvalue of $\mathbf{D}(i)$, $\lambda_{\max}(i) = \max_{\|z\|=1} \mathbf{z}^t \mathbf{D}(i) \mathbf{z}$. This potential is low for tensors belonging to the same fiber. To minimize the energy defined in [Eq. (3)], we use the optimization algorithm proposed in [16].

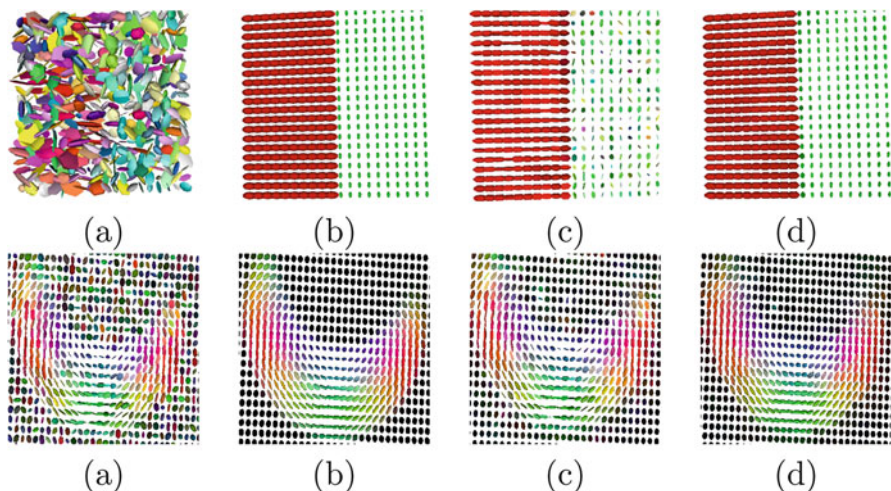


Fig. 1 Tensors on a volume slice: (a) Noisy tensors (b) Ground-truth (c) Result obtained with [8] (d) Result obtained with our method

5 Experiments and results

In order to validate the performance of the proposed algorithms, we (i) have generated artificial tensors volumes corrupted with synthetic noise and compared the obtained result with the output of the constrained anisotropic smoothing algorithm (ii) used manual segmentation on T1 muscle images and tried to improve the separability of classes in the DTI space after regularization (iii) compared the performance of kernels in the training and testing phases for two muscle groups (iv) run the SVM-driven MRF algorithm for the segmentation of the calf muscles in three distinct regions.

5.1 Denoising of Artificially Corrupted Tensors

Let us consider two volumes, one that consists of two classes with orthogonal axes on a $20 \times 20 \times 20$ lattice and a helix in which the internal voxels are anisotropic and the external ones are spheric [Fig. 1-b]. For the first volume, the tensor fields for each region are $\mathbf{D}_1 = 0.001 \times [1 \ 0.5 \ 0.5 \ 0 \ 0 \ 0]$ and $\mathbf{D}_2 = 0.001 \times [0.2 \ 0.4 \ 0.2 \ 0 \ 0 \ 0]$ where \mathbf{D} is presented in the form of $\mathbf{D} = [D_{xx} \ D_{yy} \ D_{zz} \ D_{xy} \ D_{xz} \ D_{yz}]$. We considered for both datasets a field strength $b = 700 \text{ s.mm}^{-2}$, a constant value for $S_0 = 60$ for all volume voxels and twelve different directions for diffusion gradient, which are used to generate the DTI corresponding to such tensor estimations.

The images were corrupted with a white zero-mean Gaussian noise forming a data set where ground-truth on the tensor are available. An estimation of the tensor

Table 1 Average Sum of Square Differences (SSD) $\times 10^4$. Comparisons between our method and the one in [8]

σ_n	Helix dataset			Homogeneous regions		
	0.5	1.2	3	1.5	4	9
Noisy Tensor	1.08	6.24	39.54	9.82	71.25	393.38
Method in [8]	0.33	1.60	10.57	3.32	22.47	120.70
Our Method	0.41	1.38	3.78	0.44	4.23	18.30

field relative to the noisy images provides the noisy tensors data. Then, to perform comparisons we considered the regularization algorithm on noisy tensors presented in [8]. The following parameters were used for our method: $\lambda = 50$, $\mathcal{N}_x = 3 \times 3 \times 3$, $dt = 10^{-7}$ with 50 iterations. To evaluate the performance of these methods, we considered the average sum of squared differences (SSD) between the regularized tensors and ground truth ones. In [Table 1], we can see that our estimation and regularization approach achieves better results and produces a tensor close to the ground truth. Our method outperforms the one of [8] when the level of noise is relatively important. In fact, it considers a more robust resemblance degree between voxels. Such a criterion insures a better selection of neighboring tensors involved in the estimation of a given tensor. On the other hand, the anisotropic diffusion based regularization relies on gradient information which is not robust in case of high noise. In order to assess qualitatively our algorithm, we reported in [Fig. 1] the resulting tensors using our regularization method and the constrained anisotropic one. We can observe that our method achieves a better direction preservation, even in the presence of a strong noise.

5.2 Denoising of Diffusion Tensors of the Calf Muscle

In order to perform validation using real data, the following experiment was considered. DTI acquisitions of human skeletal muscle (calf) using 12 directions were carried out on a 1.5 T MRI scanner with the following parameters : repetition time (TR) = 3600 ms, echo time(TE) = 70 ms, slice thickness = 7 mm and b value of 700 s.mm^{-2} . In order to improve the signal-to-noise ratio, the acquisition was repeated thirteen times (one can use the average of the measurements) while a high resolution T1-weighted volume was also obtained and manually segmented [Fig. 2]. The muscles that were considered in our study were the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

In order to proceed with an evaluation of the proposed method, the following scenario was considered: Using the manual segmentation, and the observed measurements of a given acquisition (12 directions), we have constructed seven linear classifiers (a multi-class linear SVM [14]) separating each class of muscle versus

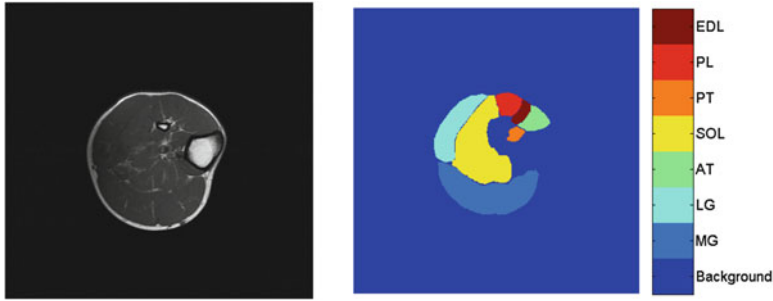


Fig. 2 A slice of the T1-weighted volume, different muscle groups segmented manually

Table 2 Correct classification rates for the different methods and for each muscle group. The first and third row show the average correct classification rates for the set of 13 volumes

	Overall	MG	LG	SOL
DE	78.1 %	86.16 %	51.1 %	84.43 %
ADE	84.46 %	90.47 %	65.72 %	88.43 %
DER	86.45 %	91.82 %	69.76 %	89.97 %

all others. Then, the success rate (percentage of voxels being attributed to the right class) from the classifier with respect to the ground truth was determined. We remark that linear separation is hardly achieved for PT, PL, EDL and AT while it yields quite satisfactory results for the MG, LG and to a lesser extent SOL which form the major part of the muscle. We have performed this test thirteen times for: (i) direct estimation (**DE**), (ii) direct estimation and regularization (**DER**), as well as using direct estimation of the average measurements of the thirteen acquisitions (**ADE**). One would expect that since muscles consist of myo-fibers of the same nature, the classification should be improved if the estimation of the tensors is properly done, i.e. with appropriate regularization. However, it is important to note that the aim of this paper is not automatic classification of voxels in different muscle regions using DTI (in such a case more advanced classifiers can be used).

In [Table 2], we present quantitative validation of the present framework for the linearly separable muscles. One can see that our method leads to an improvement in the correct classification rates with respect to a plain direct estimation. We also obtain better results when compared to the averaging + estimation method. We also show the result of our regularization on a slice of the volume in [Fig. 3].

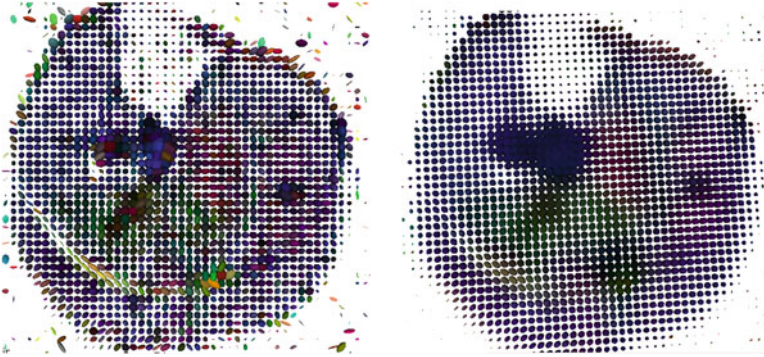


Fig. 3 Estimated tensors without regularization, tensors obtained with our method

5.3 *Kernels comparison*

In order to assess the behavior of the defined kernels, we consider the SOL and MG muscle groups. SVM classification was performed both in a linear and a non linear fashion using the above-defined kernels.

We motivate the use of kernel-based SVMs by focusing on the architecture of the soleus muscle. While the medial gastrocnemius is a unipennate muscle (the fibers have one line of action), the soleus is a bipennate one (the fibers have two lines of action) and exhibits a richer structure. As can be seen in [Fig. 4] where the principal directions of diffusion in the MG and SOL muscles are displayed as points on the unit sphere, it is more natural and mathematically more sound to trace the decision boundary while respecting the manifold structure, rather than using a hyperplane to separate the different classes or flattening the manifold using the Gaussian RBF kernel. We compared the behavior of the defined kernels in separating the MG from the SOL using a set of 9904 diffusion tensors with approximately the same number of tensors for each class (4976 tensors belonging to MG, 4928 tensors from the SOL muscle). We subdivided this set into a training set and a testing one to evaluate the generalization errors (50 % of the set for the training and 50 % for the testing). As shown in [Table 3], the kernels that are specific to the space of symmetric positive definite matrices perform better than the linear classification both in the training and testing phases. The best result was obtained for the information diffusion kernel with approximately 3 % of classification error. Note that the number of support vectors has not increased much with respect to a linear classification.

Fig. 4 Principal directions of diffusion of SOL and MG in red and blue respectively

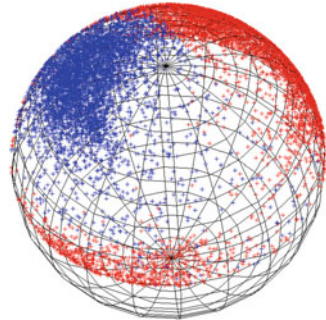


Table 3 Performance of the different kernels

Kernel	Training error	Support vectors	Testing error
Linear	5.27 %	741	5.81 %
Information diffusion	0.26 %	999	2.89 %
Probability product	3.77 %	547	4.94 %
Burg	3.23 %	482	3.94 %
V. Neumann	3.25 %	494	4.05 %

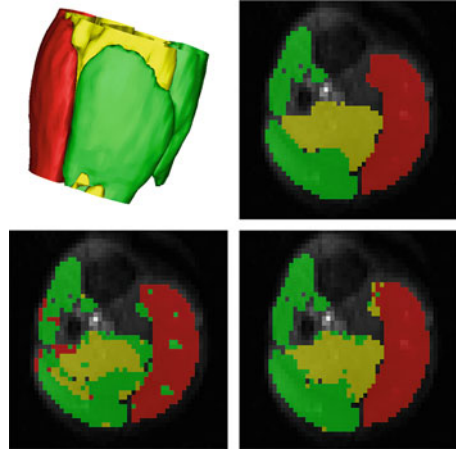
Table 4 Comparison of error rates with and without MRF regularization

	class 1	class 2	class 3
SVM	85.71 %	82.53 %	76.4 %
SVM+MRF	88.64 %	90.22 %	88.42 %

5.4 Segmentation of the calf muscle with SVM-driven MRFs

Two different diffusion datasets were considered to evaluate the segmentation algorithm, one for training and the other for testing. We used SVMs in a one against all fashion to learn the diffusions tensors of three major groups in the calf muscle: the soleus, the medial gastrocnemius and a third group consisting of the lateral gastrocnemius as well as the muscles of the anterior compartment of the calf. The learning was done on a manually segmented volume and we tested the performance of the SVM + MRF algorithm using the information diffusion kernel on another volume. As can be seen in [Table 4], the MRF regularization improves significantly the correct classification rates with respect to a labeling decision based on the maximum score, achieving a correct classification rate of approximately 90 % for each of the three classes. Qualitative results are provided in [Fig. 5]. We can see that while SVM classification can be misled by noise, the combination of SVMs and MRF provides a smoother result and decreases the number of misclassifications.

Fig. 5 Obtained 3D segmentation in three groups with some misclassifications visible, axial slice of the baseline image with overlaid segmentation for the ground truth, SVM classification and SVM + MRF algorithm respectively



6 Discussion

In this chapter a novel approach to direct estimation and regularization of diffusion tensor images was proposed along with an SVM framework for tensor classification. The main strength of the regularization algorithm is the smoothness term that assumes linear (convex) approximation of neighborhood tensors as well as the convex nature of the proposed cost function which can be easily optimized. Our method was compared and outperformed the anisotropic constrained regularization using generated data with known noise model, and improved human skeletal muscle segmentation/classification through DTI using real data.

A possible extension of the regularization framework is a better selection of the bandwidth σ which is a critical parameter of the process. Data-driven variable bandwidth models is a natural extension of the method. One would expect that the optimal bandwidth depends on the form of the observed anatomical structure which varies spatially. We can further extend this work by replacing the Frobenius norm in the energy functional by the Riemannian distance [6, 22] or the Log-Euclidean distance [9]. However this will be done at the expense of the convexity of the function and the computational time.

The SVM classification approach introduced in this paper allows linear and non-linear separation and learning of diffusion tensors. It can be incorporated in a Markov random field model for segmentation with prior knowledge. This framework can be generalized to classify fibers. In fact, a fiber can be considered as an ordered set of tensors and one can use the summation kernel or more elaborate kernels on structured data [18] to build kernels on entire fibers.

The use of DTI towards understanding the human skeletal muscle as well as providing means of diagnosis for muscular diseases is a more long-term objective of our research. The ability to understand the remodeling of myofibers due to muscular diseases using non-invasive means is a great perspective.

References

1. N. Azzabou, N. Paragios, F. Guichard, and F. Cao. Variable bandwidth image denoising using image-based noise models. In *CVPR*, 2007.
2. S. Basu, P. T. Fletcher, and R. T. Whitaker. Rician noise removal in diffusion tensor mri. In *MICCAI (1)*, pages 117–125, 2006.
3. D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1999.
4. D. L. Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.
5. S. Boughorbel, J.-P. Tarel, and F. Fleuret. Non-Mercer kernels for SVM object recognition. In *BMVC*, 2004.
6. C. A. Castano-Moraga, C. Lenglet, R. Deriche, and J. Ruiz-Alzola. A Riemannian approach to anisotropic filtering of tensor fields. *Signal Processing [Special Issue on Tensor Signal Processing]*, 87(2):263–276, 2007.
7. O. Coulon, D. C. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–67, March 2004.
8. R. Deriche, D. Tschumperle, C. Lenglet, and M. Rousson. Variational approaches to the estimation, regularization and segmentation of diffusion tensor images. In F. Paragios, Chen, editor, *Mathematical Models in Computer Vision: The Handbook*. Springer, 2005 edition, 2005.
9. P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, pages 786–789, Crystal Gateway Marriott, Arlington, Virginia, USA, Apr. 2006.
10. C. J. Galban, S. Maderwald, K. Uffmann, A. de Greiff, and M. E. Ladd. Diffusive sensitivity to muscle architecture: a magnetic resonance diffusion tensor imaging study of the human calf. *European Journal of Applied Physiology*, 93(3):253–262, Dec 2004.
11. C. J. Galban, S. Maderwald, K. Uffmann, and M. E. Ladd. A diffusion tensor imaging analysis of gender differences in water diffusivity within human skeletal muscle. *NMR in Biomedicine*, 2005.
12. J.-B. Hiriart-Urruty and C. Lemarechal. *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.
13. T. Jebara, R. Kondor, and A. Howard. Probability product kernels. *J. Mach. Learn. Res.*, 5:819–844, 2004.
14. T. Joachims. Making large-scale support vector machine learning practical. In A. S. B. Schölkopf, C. Burges, editor, *Advances in Kernel Methods: Support Vector Machines*. MIT Press, Cambridge, MA, 1998.
15. P. Khurd, R. Verma, and C. Davatzikos. Kernel-based manifold learning for statistical analysis of diffusion tensor images. In *IPMI*, pages 581–593, 2007.
16. N. Komodakis, G. Tziritas, and N. Paragios. Fast, approximately optimal solutions for single and dynamic MRFs. In *CVPR*, 2007.
17. J. Lafferty and G. Lebanon. Diffusion kernels on statistical manifolds. *J. Mach. Learn. Res.*, 6:129–163, 2005.
18. S. Lyu. Mercer kernels for object recognition with local features. In *CVPR*, 2005.
19. M. Maddah, W. E. L. Grimson, and S. K. Warfield. Statistical modeling and em clustering of white matter fiber tracts. In *ISBI*, pages 53–56, 2006.
20. M. Martin-Fernandez, C.-F. Westin, and C. Alberola-Lopez. 3D Bayesian regularization of diffusion tensor MRI using multivariate Gaussian Markov random fields. In *MICCAI (1)*, pages 351–359, 2004.
21. L. J. O’Donnell and C.-F. Westin. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Transactions on Medical Imaging*, 26(11):1562–1575, November 2007.

22. X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.
23. R. Salvador, A. Pea, D. K. Menon, T. A. Carpenter, J. D. Pickard, and E. T. Bullmore. Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping*, 24(2):144–155, 2005.
24. E. Stejskal and J. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288–292, 1965.
25. K. Tsuda, G. Ratsch, and M. Warmuth. Matrix exponentiated gradient updates for on-line learning and Bregman projection. *Journal of Machine Learning Research*, 6:995–1018, 06 2005.
26. V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
27. F. Vos, M. Caan, K. Vermeer, C. Majoie, G. den Heeten, and L. van Vliet. Linear and kernel fisher discriminant analysis for studying diffusion tensor images in schizophrenia. In *ISBI*, pages 764–767, 2007.
28. Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE Transactions on Medical Imaging*, 23(8):930–939, 2004.
29. J. Weickert, C. Feddern, M. Welk, B. Burgeth, and T. Brox. PDEs for tensor image processing. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 399–414. Springer, January 2006.

From Local Q-Ball Estimation to Fibre Crossing Tractography

M. Descoteaux and R. Deriche

Abstract Fibre crossing is an important problem for most existing diffusion tensor imaging (DTI) based tractography algorithms. To overcome limitations of DTI, high angular resolution diffusion imaging (HARDI) techniques such as q-ball imaging (QBI) have been introduced. The purpose of this chapter is to first give state of the art review of the existing local HARDI reconstruction techniques as well as the existing HARDI-based tractography algorithms. Then, we describe our analytical QBI solution to reconstruct the diffusion orientation distribution function (ODF) of water molecules and we propose a spherical deconvolution method to transform the diffusion ODF into a sharper fibre ODF. Finally, we propose a new deterministic and a new probabilistic algorithm based on this fibre ODF. We show that the diffusion ODF and fibre ODF can recover fibre crossing in simulated data, in a biological phantom and in real datasets. The fibre ODF improves angular resolution of QBI by more than 15° and greatly improves tractography results in regions of complex fibre crossing, fanning and branching.

1 Introduction & Problem Statement

At the current resolution of diffusion-weighted (DW) magnetic resonance imaging (MRI), research groups agree that there are between one and two thirds of imaging voxels in the human brain white matter that contain fibre crossing bundles [10].

M. Descoteaux (✉)

Department of computer Science, 2500 Blv. Université, Sherbrooke, Quebec J1K 2R1, Canada
e-mail: m.descoteaux@usherbrooke.ca

R. Deriche

Athena Project Team, INRIA Sophia Antipolis-Méditerranée, 2004 Route des Lucioles, BP 93
06902 Sophia Antipolis Cedex

We know that in these locations, the diffusion is non-Gaussian and the diffusion tensor (DT) [8] is limited due to its intrinsic Gaussian diffusion assumption. Hence, DT-based tractography algorithms can follow false tracts and produce unreliable tracking results. To overcome limitations of the DT, new high angular resolution diffusion imaging (HARDI) techniques [2, 65] have been proposed to estimate the diffusion orientation distribution function (ODF) [66] of water molecules or other high order spherical function estimate of the diffusion profile [4, 5, 20, 34, 37, 38, 40, 52, 58, 63, 64]. These HARDI techniques (see Fig. 1) were developed to deal with non-Gaussian diffusion process and to reconstruct spherical functions with maxima aligned with the underlying fibre populations.

In this chapter, we focus on deterministic and probabilistic tractography using state of the art reconstruction of the diffusion and fibre ODF from QBI. QBI is of interest because it is model-free and it can be computed analytically and robustly with low computational cost [23]. First, we review the existing HARDI reconstruction techniques and state of the art HARDI-based tractography algorithms to put our new methods into context. Then, we develop our analytical solution to reconstruct the diffusion ODF and the fibre ODF from QBI data. Finally, we describe a new deterministic tractography algorithm and a new probabilistic tractography algorithm able to recover complex fibre tracts with known crossing, fanning and branching configurations. Most current DTI based methods neglect these fibres, which might lead to wrong interpretations of the brain functions.

2 Prior Art

HARDI samples q-space along as many directions as possible in order to reconstruct estimates of the true diffusion probability density function (PDF) of water molecules. This true diffusion PDF is model-free and can recover the diffusion of water molecules in any underlying fibre population. HARDI depends on the number of measurements N and the gradient strength (b -value), which will directly affect acquisition time and signal to noise ratio in the signal. Typically, there are currently two strategies used in HARDI: 1) sampling of the whole q-space 3D Cartesian grid or 2) single shell spherical sampling.¹ In the first case, a large number of q-space points are taken over the discrete grid ($N > 200$) and the inverse Fourier transform of the measured DW signal is taken to obtain an estimate of the diffusion PDF. This is q-space imaging (QSI) and DSI [69]. The method requires very strong imaging gradients ($500 \leq b \leq 20000$ s/mm²) and a long time for acquisition (15-60 minutes) depending on the number of sampling directions. In the second case, a uniform sampling of a single sphere is done for a certain radius in q-space (given by the b -value). Typically, $60 \leq N \leq 200$, $b \geq 1000$ s/mm² is used and acquisition time is between 10 and 20 minutes.

¹There is very recent development in multiple-shell acquisition schemes (see [41]).

2.1 Review of HARDI Reconstruction Techniques

The goal of HARDI is to capture multiple fibre directions within the same imaging voxel. Some HARDI reconstruction techniques are model dependent, some model-free, some have linear solutions whereas others require non-linear optimization schemes. A schematic view of the major multiple fibre HARDI reconstruction algorithms is shown in Fig. 1. A good review of these methods up to 2005 can be found in [2]. We now summarize the major techniques.

It is a simple extension of the DTI model to assume that a mixture of Gaussians can describe the diffusion PDF. [65] proposed the initial solution and many other works [1, 13, 18, 48, 55, 65, 67] proposed variants of the multi-Gaussian with constraints such as forcing symmetry of eigenvalues, forcing certain magnitude and ratios of eigenvalues or imposing positive definiteness of the DT (see [2]). A similar approach to multi-Gaussian modeling is the *ball & stick* mixture model. It assumes that water molecules in an imaging voxel belong to one of two populations, a restricted population within or near fibre structures (*stick*), modeled with an anisotropic Gaussian distribution, and a free population that is not affected by fibre structure barriers (*ball*), modeling by an isotropic Gaussian distribution. The approach extends to a mixture of restricted compartments and is thus able to recover multiple fibre compartments [10, 33]. Another similar approach is the CHARMED technique [7]. This technique assumes a highly restricted compartment that is non-Gaussian and a hindered compartment that is approximately Gaussian. The approach can also be formulated as a mixture of restricted compartments and is thus able to recover multiple fibre compartments. The multi-Gaussian, ball & stick and CHARMED all suffer from the same shortcomings regarding model selection and numerical implementation. One must select the number of compartments a priori, one must use non-linear optimization to solve for the parameters and the methods are sensitive to noise and to the number of measurements.

Spherical deconvolution (SD) methods generalize the mixture modeling methods of the previous section by assuming a distribution of fibre orientations to overcome the limitation of the number of compartment selection n . The original SD method [64] was improved in [4, 20, 37, 40, 57] using non-linear optimization techniques that better deal with the SD instabilities, noise and negative diffusion appearing in the deconvolution process. A recent review SD methods can be found in [37]. In [38], the case of multiple fibre bundles is handled in a similar way to the SD methods. The novelty is that each fibre bundle is represented by a Wishart distribution. This leads to a reformulation of DTI in the presence of a single orientation but is also able to account for multiple fibre crossings. In contrast, in [49], the diffusion ODF is modeled with a mixture of von Mises-Fisher distributions, which allows for the definition of a closed-form Riemannian distance between diffusion ODFs.

Another model-independent method reconstructs the radially persistent angular structure (PAS) [2, 34] of the diffusion PDF. The reconstruction forces probabilities to be non-zero only on a spherical shell of a certain radius. The PAS reconstruction

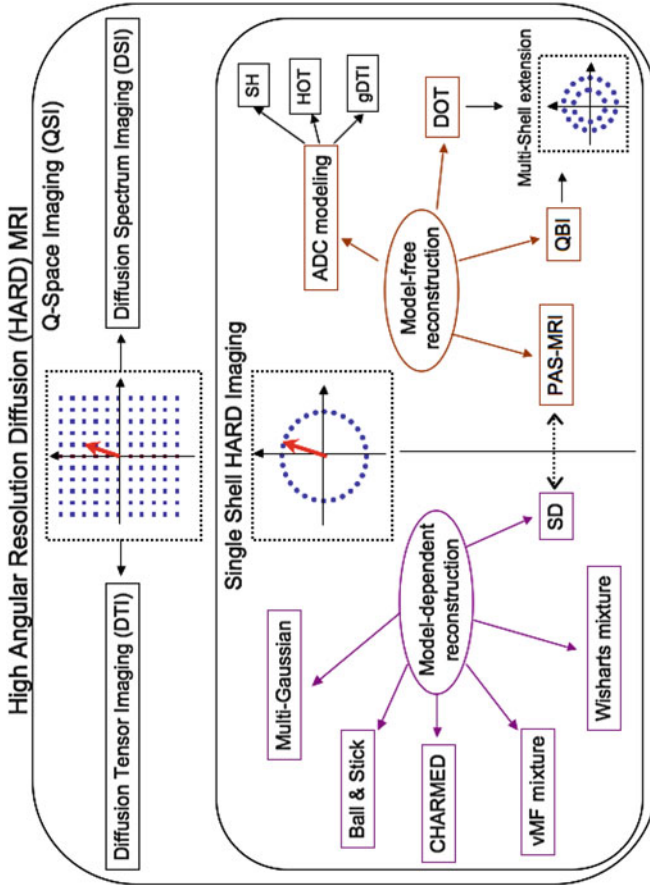


Fig. 1 Sketch of HARDI reconstruction techniques

is non-linear and computationally very heavy but recent efforts [61] have been done to propose a linearized solution to PAS-MRI based on the fact that it is a special case of SD methods (relation indicated by an arrow between SD and PAS-MRI in Fig. 1).

Finally, the diffusion orientation transform (DOT) proposed in [52] is yet another model-independent reconstruction algorithm. The DOT is a function that maps the apparent diffusion coefficient (ADC) profile to the diffusion PDF. Using similar techniques, [51] fit high-order tensors (HOT) to the HARDI measurements to model the ADC. ADC modeling is not discussed in this chapter because it is not an appropriate function for fibre tractography (see [22]) but it can also be modeled with spherical harmonics (SH) [3, 26] or generalized DTI (gDTI) [47]. Finally, similar to QBI, the DOT has a possible multiple shell HARDI extension with the bi and tri exponential fit [52].

2.2 Review of HARDI-Based Tractography Algorithms

In tractography, two families of algorithms exist: deterministic and probabilistic algorithms. Research groups have recently started to generalize both deterministic and probabilistic DT-based tracking algorithms to use HARDI reconstruction methods mentioned in the previous section. Some of these methods use the principal direction extracted from the diffusion ODF computed from DSI [30, 65], from a multi-tensor/Gaussian local model [12, 28, 43, 53], or from a q-ball diffusion ODF [14, 17]. In this chapter, we propose another extension to streamline tractography based on the multiple maxima information of the fibre ODF.

Deterministic tractography algorithms inherit the classical limitations of deterministic algorithms such as choice of initialization [39], sensitivity in estimated principal direction and lack of straightforward way to compute statistics on tracts and lack of connectivity information between regions of the brain [65]. To overcome limitations of deterministic tractography, DT-based probabilistic [11, 27, 42, 44, 53] and geodesic [35, 45] algorithms have been used. Probabilistic algorithms are computationally more expensive than deterministic algorithms but can better deal with partial volume averaging effects and noise uncertainty in underlying fibre direction and output a connectivity index measuring how probable two voxels are connected to one another.

HARDI-based probabilistic tractography have recently been published in the literature [10, 16, 31, 36, 40, 54, 56, 59, 62] to generalize several existing DT-based methods. First, in [40] parametric SD is used and in [10] a mixture of Gaussian model is used to extend the probabilistic Bayesian DT-based tracking [11]. Related to these techniques, [36] uses a Bayesian framework to do global tractography instead of tracking through local orientations. In [56], Monte Carlo particles move inside the continuous field of q-ball diffusion ODF and are subject to a trajectory regularization scheme. In [31, 54], an extension to their DT-based approach [53] is also proposed using a Monte Carlo estimation of the white matter geometry and recently, a Bingham distribution is used to model the peak anisotropy in the fibre distributions [62]. Finally, in [16], large number of M-FACT QBI streamlines are reconstructed and all pathways are reversed-traced from their end points to generate of map of connection probability. In this chapter, our new probabilistic algorithm is based the fibre ODF using a Monte Carlo random walk algorithm.

3 Analytical Solution to Q-Ball Imaging

QBI [66] is a model-independent method to estimate the diffusion ODF. This diffusion ODF contains the full angular information of the diffusion PDF and is defined as

$$\Psi(\theta, \phi) = \int_0^\infty P(\alpha \mathbf{r}) d\alpha, \quad (1)$$

where (θ, ϕ) obey physics convention ($\theta \in [0, \pi], \phi \in [0, 2\pi]$). [66] showed that it was possible to reconstruct a smoothed version of the diffusion ODF directly from single shell HARDI acquisition with the Funk-Radon transform (FRT). Intuitively, the FRT value at a given spherical point is the great circle integral of the signal on the sphere defined by the plane through the origin perpendicular to the point of evaluation. The original QBI has a numerical solution [66] and more recent methods [5, 23, 32] have introduced an analytical spherical harmonic reconstruction solution that is faster and more robust. To develop the analytical solution to QBI, we first need to estimate the HARDI signal with spherical harmonics (SH) and then, to solve the FRT analytically with SH.

Letting Y_ℓ^m denote the SH of order ℓ and degree m ($m = -\ell, \dots, \ell$), we define a modified SH basis that is real and symmetric. For even order ℓ , we define a single index j in terms of ℓ and m such that $j(\ell, m) = (\ell^2 + \ell + 2)/2 + m$. The modified basis is given by

$$Y_j = \begin{cases} \sqrt{2} \operatorname{Re}(Y_\ell^{|m|}), & \text{if } m < 0 \\ Y_\ell^m, & \text{if } m = 0 \\ \sqrt{2} \operatorname{Im}(Y_\ell^m), & \text{if } m > 0. \end{cases} \quad (2)$$

where $\operatorname{Re}(Y_\ell^m)$ and $\operatorname{Im}(Y_\ell^m)$ represent the real and imaginary parts of Y_ℓ^m respectively. The basis is designed to be symmetric, real and orthonormal. It is then possible to obtain an analytical diffusion ODF estimate, Ψ , with

$$\Psi(\theta, \phi) = \sum_{j=1}^L \underbrace{2\pi P_{\ell(j)}(0)c_j}_{c'_j} Y_j(\theta, \phi), \quad (3)$$

where $L = (\ell + 1)(\ell + 2)/2$ is the number of elements in the spherical harmonic basis, c_j are the SH coefficients describing the input HARDI signal, $P_{\ell(j)}$ is the Legendre polynomial of order $\ell(j)^2$ and c'_j the coefficients describing the ODF Ψ . Here, we estimate the c_j coefficients with the solution presented in [23] with a Laplace-Beltrami regularization of the SH coefficients c_j to obtain a more robust ODF estimation. The detailed implementation of the Laplace-Beltrami regularization and the comparison with the state of the art methods [5, 32] are presented in [22, 23].

² $\ell(j)$ is the order associated with the j^{th} element of the SH basis, i.e. for $j = 1, 2, 3, 4, 5, 6, 7, \dots$ $\ell(j) = 0, 2, 2, 2, 2, 2, 4, \dots$

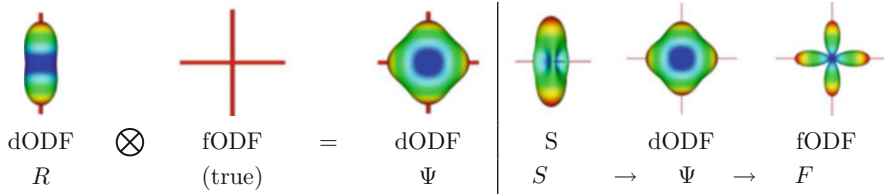


Fig. 2 Left: The convolution between the diffusion ODF kernel, R , and true fibre ODF produces a smooth diffusion ODF estimate, Ψ . Right: The Funk-Radon transform of the HARDI signal, S , produces a smooth diffusion ODF, Ψ , which is transformed into a sharper fibre ODF estimate, F , by the deconvolution

3.1 Fibre ODF reconstruction

The relation between the measured diffusion ODF and the underlying fibre distribution, the fibre ODF, is still an important open question in the field [56, 65]. The diffusion ODF is a blurred version of the “true” fibre ODF. Because of this blurring effect, the extracted maxima of the diffusion ODF are often used for fibre tractography. An alternative is to use spherical deconvolution methods that provide an estimate of the fibre ODF [5, 20, 25, 37, 40, 58, 63, 64]. These techniques have better angular resolution than QBI and produce sharper fibre ODF profiles than the q-ball diffusion ODF. Smaller fibre compartments with smaller volume fractions are visible with fibre ODF and not with the diffusion ODF. SD and fibre ODF estimation are currently subject to active research. Here, we use a simple linear transformation of our analytical QBI solution. A schematic view of our spherical deconvolution method is shown in Fig. 2.

The fibre ODF is reconstructed in three steps. 1) The regularized diffusion ODF coefficients c'_j are reconstructed using Eq. (3) of the last section, $c'_j = 2\pi P_{\ell(j)}(0)c_j/S_0$, where S_0 is the unweighted $b = 0$ diffusion image. 2) The single fibre diffusion ODF, R , used as deconvolution kernel is estimated from the real data. As in [5, 64], we assume an axially symmetric diffusion tensor model with eigenvalues (e_2, e_2, e_1) and $e_1 \gg e_2$ for the underlying single fibre diffusion model. The values of e_1 and e_2 are estimated from 300 voxels with highest FA value in our real dataset, as these voxels can each be assumed to contain a single fibre population. The single fibre diffusion ODF kernel has an analytical expression [25] and is given by

$$R(t) = \frac{(1 - \alpha t^2)^{-1/2}}{8\pi b \sqrt{e_1 e_2}}, \tag{4}$$

where $\alpha = (1 - e_2/e_1)$, b is the b -value of the real dataset and $t \in [-1, 1]$ is the variable that represents the dot product between the direction of the fibre and the point of evaluation (θ, ϕ) on the sphere.

3) The SH coefficient of the fiber ODF, f_j , are then obtained by a simple linear transformation,

$$f_j = c'_j / r_{\ell(j)}, \quad \text{with} \quad r_{\ell(j)} = 2\pi \int_{-1}^1 R(t) P_{\ell(j)}(t) dt, \quad (5)$$

which can be solved analytically by taking the power expansion of $P_{\ell(j)}(t)$ and integrating $r_{\ell(j)}$ term by term. As for the analytical diffusion ODF solution, the spherical deconvolution is obtained with the Funk-Hecke theorem [23]. Therefore, the fibre ODF in terms of the HARDI signal is

$$f_j = \frac{8\pi b \sqrt{e_1 e_2} P_{\ell(j)}(0)}{S_0 A_{\ell}(\alpha)} c_j, \quad (6)$$

where $A_{\ell}(\alpha) = \int_{-1}^1 (1 - \alpha t^2)^{-1/2} P_{\ell}(t) dt$. The final fibre ODF is reconstructed for any (θ, ϕ) and point p as $F(\theta, \phi)_p = \sum_{j=1}^R f_j Y_j(\theta, \phi)$. In [25], the fibre ODF is shown to be a valid choice of fibre ODF that is in close agreement with the classical SD method [64].

4 Q-Ball Tractography

4.1 Deterministic Tractography

We extend the classical streamline techniques [9, 19, 50] based on diffusion tensor principal direction to take into account multiple fibre ODF maxima at each step. We denote $p(s)$ as the curve parameterized by its arc-length. This curve can be computed as a 3D path adapting its tangent orientation locally according to vector field \mathbf{v} . Hence, for a given starting point p_0 , we solve $p(t) = p_0 + \int_0^t \mathbf{v}(p(s)) ds$. The integration is typically performed numerically with Euler or Runge-Kutta schemes of order 2 or 4. In the Euler case, we have the discrete evolution equation

$$p_{n+1} = p_n + \mathbf{v}(p_n) \Delta s, \quad (7)$$

where Δs is a small enough step size to obtain subvoxel precision.

For our deterministic algorithm [21, 25], we use a threshold on the GFA ([66]) or FA (typically, $FA \geq 0.1$) to prevent tracks to leak outside white matter. We set curving angle threshold $t_{\theta} = 75^\circ$ and $\Delta s = 0.1$ and we use Euler integration and classical trilinear interpolation to obtain diffusion ODF, fibre ODF and DT at subvoxel precision. The fibre ODF is reconstructed with order $\ell = 6$, regularization parameter $\lambda = 0.006$ and eigenvalues e_1, e_2 are estimated from our real dataset to be 13.9 and $3.55 \times 10^{-4} \text{ mm}^2 / \text{s}$. For the rest of the chapter, DT-STR refer to the streamline (STR) tracking using the DT principal eigenvector, dODF-STR and

fODF-STR refer to the STR tracking using a single diffusion ODF and fibre ODF maxima that is the closest to the incoming tangent direction of the curve and SPLIT-STR refers to STR tracking using the fibre ODF maxima with splitting if there are multiple maxima [21, 25].

4.2 Probabilistic Tractography

We propose an extension of the random walk method proposed in [42] to use the distribution profile of the fibre ODF. We start a large number of particles from the same seed point, let the particles move randomly according to our local fibre ODF estimate, F , and count the number of times a voxel is reached by the path of a particle. This yields higher transitional probabilities along the main fibre directions. The random walk is stopped when the particle leaves the white matter mask.

For each elementary transition of the particle, the probability for a movement from the seed point x to the target point y in direction \mathbf{u}_{xy} is computed as the product of the local fibre ODFs in direction \mathbf{u}_{xy} , i.e.

$$P(x \rightarrow y) = F(\mathbf{u}_{xy})_x \cdot F(\mathbf{u}_{xy})_y \quad (8)$$

where $P(x \rightarrow y)$ is the probability for a transition from point x to point y , $F(\mathbf{u}_{xy})_x$ is the fibre ODF at point x in direction xy (by symmetry, direction xy and yx are the same).

The transition directions in the local model are limited to 120 discrete directions corresponding to the angular sampling resolution of the acquired brain data and the step size of the particle step was fixed to 0.5 times the voxel size. We used trilinear interpolation of the fibre ODF for the subvoxel position and we use a white matter mask computed from a minimum FA value of 0.1 and a maximum ADC value of 0.0015. A total of 100000 particles were tested for each seed voxel. The connectivity of any given voxel with the seed voxel is estimated by the number of particles that reach the respective voxel, called a *tractogram*.

4.3 Data Acquisition

Our synthetic HARDI data is generated with the multi-tensor model [3, 22, 23, 66], which can control the separation angle, anisotropy, volume fraction of each fibre compartment as well as the signal to noise ratio (SNR) and number of gradient directions N . Then, we use a biological phantom dataset obtained from a 1.5 T scanner with 90 gradient directions and a $b = 3000$ s/mm² [15]. We also use a human brain dataset obtained on a 3 T scanner, which has 1.7 mm³ cubic grid and contains 116, 93 x 93 slices with 60 gradient directions and a $b = 1000$ s/mm² [6].

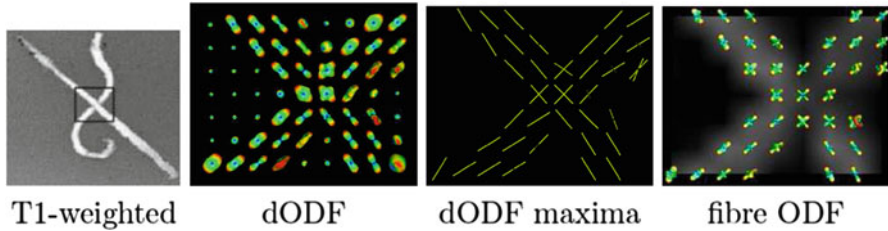


Fig. 3 Diffusion ODF (dODF) and fibre ODF recover multiple fibre crossing in the rat biological phantom [15]

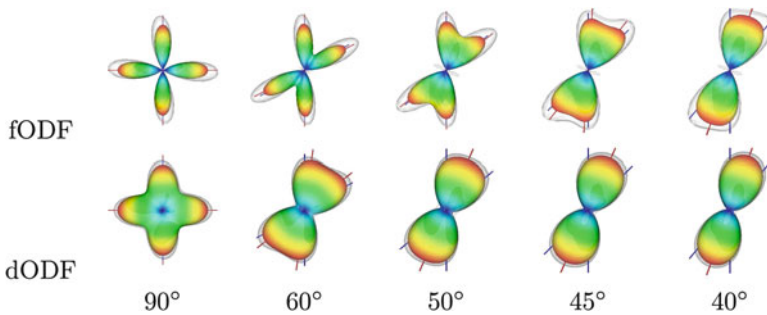


Fig. 4 The fibre ODF (fODF) improves angular resolution of the diffusion ODF (dODF) by more than 15° . The signal is generated with fibres equal volume fraction and $FA = 0.7$, $N = 60$ data points, $b = 3000$ s/mm² and SNR 30. The opaque surface is the mean fibre ODF over 100 noise trials, whereas the transparent surface corresponds to the mean $+ 2$ standard deviations. Blue and red lines correspond to ground truth fibre directions and detected maxima respectively

5 Evaluation With the State of the Art

5.1 Analytical QBI Results

The analytical QBI reconstruction has several advantages over the classical numerical QBI reconstruction [66]. Overall, the analytical QBI reconstruction of the diffusion ODF has the following four major advantages. (1) It is up to 15 times faster than the numerical QBI implementation. (2) It is more robust to noise than the numerical QBI solution. (3) It allows for more precise diffusion ODF reconstruction for lower number of gradient directions N in the acquisition. (4) Most of the information is contained in harmonic orders of order 6 and less. Higher order harmonics contain small perturbations due to noise. To illustrate some of these properties, Figs. 3 and 5 show that diffusion ODFs can recover multiple fibre crossings in real HARDI datasets. In Fig. 3, the diffusion ODFs have multiple peaks that agree with the known underlying fibre populations. For extensive details and discussion, we refer the reader to [23] and [32].

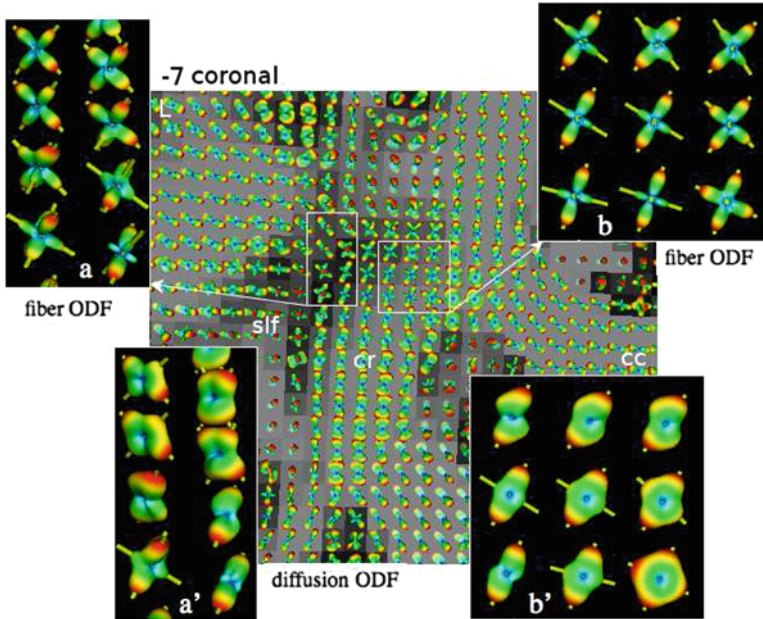


Fig. 5 The fibre ODF improves fibre detection of QBI. There are more crossings detected using the fibre ODF (a,b) than diffusion ODF (a',b')

5.2 Fibre ODF Deconvolution Results

In [25], we show that the fibre ODF obtained from QBI is a valid choice which agrees with the classical SD method [64]. Overall, the fibre ODF has a striking angular resolution gain over the q-ball diffusion ODF of more than 15° . This is seen in Fig. 4, where the fibre ODF is able to better discriminate the two fibre compartments at a separation angle of 45° whereas the diffusion ODF is limited to the separation angle of 60° . In general, as also seen in Figs. 3 and 5, fibre ODF can recover fibre crossings more easily while noise effect is kept under control.

Figure 5 shows the multi-directional information coming from the diffusion ODF and the fibre ODF on a region of interest in a coronal slice (Talairach -4) of the human brain dataset. In this ROI, the corpus callosum (CC) forms the roof of the lateral ventricles and fans out in a massive collateral radiation, the corticospinal tract (CST) lies lateral to the ventricle and is directed vertically and the SLF crosses the base of the precentral gyrus in anterior-posterior direction. The lateral projections of the CC cross the CST and the SLF. Fibres of the SLF partly intersect with the fibres of the CST and the CC.

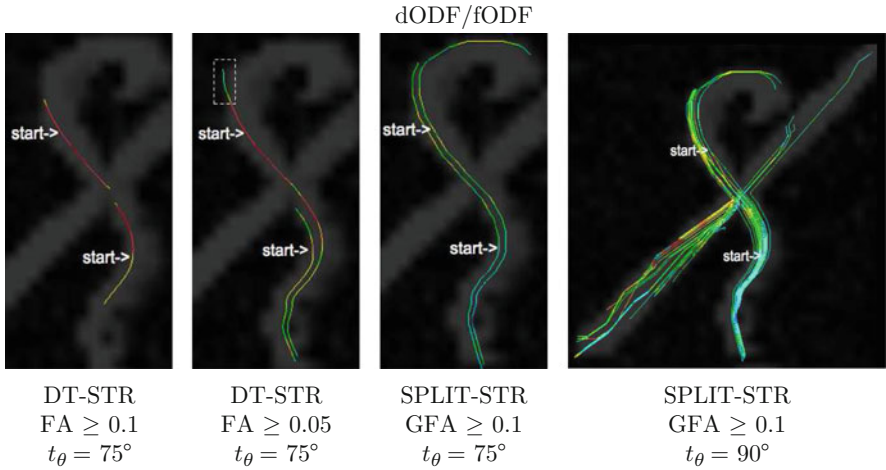


Fig. 6 Deterministic tracking on the biological phantom

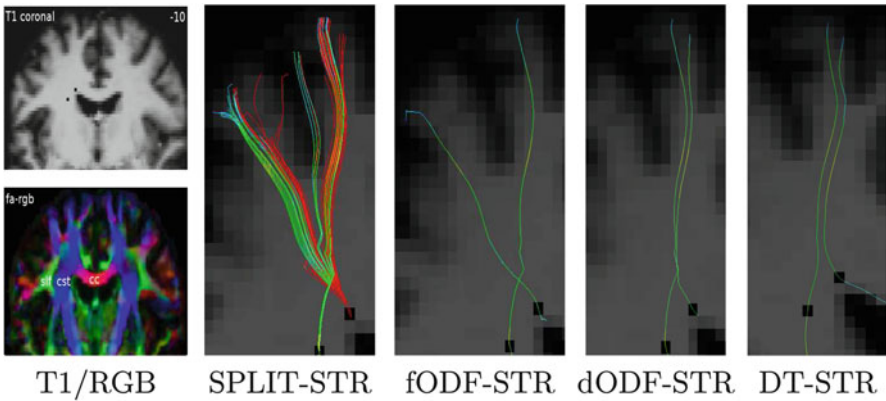


Fig. 7 SPLIT-STR recovers known fanning/crossing configurations to the two motor gyri from both seed points

5.3 Tractography Results

Figure 6 shows deterministic tracking on the biological phantom. DT-STR cannot track through the crossing area with a standard FA threshold of 0.1 whereas q-ball tracking can. It can go through the crossing if the threshold is lowered to 0.05 but then, the tracking steps out of the fibre bundle at the top. Deterministic tracking from the diffusion ODF and the fibre ODF produce the same qualitative result, even if allowed to split. Not only can q-ball tracking go through the crossing area, it can also recover part of the curving section of the spinal cord at the top. Finally, if there is no curvature constraint, i.e. $t_\theta = 90^\circ$, SPLIT-STR recovers parts of both spinal cord bundles.

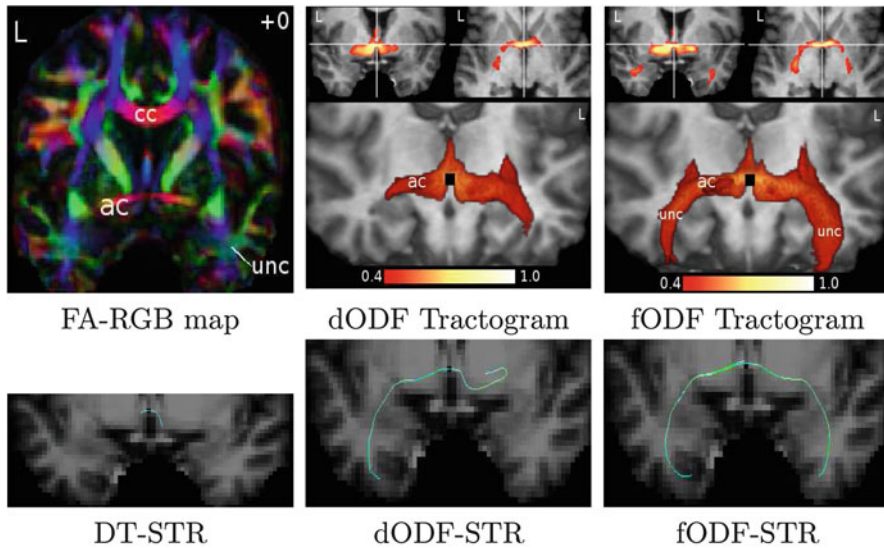


Fig. 8 Deterministic and probabilistic tracking of the anterior commissure (AC) fibres

Figure 7 shows a fanning/branching fibre configuration in the same ROI as seen in Fig. 5. One set of tracts (red fibres) are started from a voxel in the CC and another set of tracts are started from a voxel in the CST (green/yellow fibres). As expected, SPLIT-STR recovers the branching configuration of both fibre tracts and recovers fibres projecting in motor areas in two gyri. On the other hand, fODF-STR is able to step through the crossings whereas dODF-STR and DT-STR are limited.

Figure 8 shows the reconstruction of the fibres passing through the anterior commissural (AC) fibres. A seed voxel was placed in the mid-sagittal cross section of the AC. The tracking results of the AC fibres shows the advantages of the fODF-PROBA tracking over dODF-PROBA and DT-STR tracking. dODF-PROBA and DT-STR are blocked close to the seed point by low FA areas. Particles of dODF-PROBA cannot propagate to the temporal poles because the paths are diffusive and leak outside the anterior commissural bundle, which is only a few voxel wide around the seed point. As a consequence, dODF-PROBA mostly recovers shorter parts of the fibre bundles. However, with a multiple seeding approach (shown in [25]), DT-STR and dODF-STR are able to recover both paths to the temporal poles. In contrary, deterministic fODF-STR and SPLIT-STR tracking can reconstruct the fibres connecting the temporal pole via the AC from a single seed point.

Figure 9 shows the reconstruction of the commissural fibres connecting the contralateral inferior and middle frontal gyrus. A seed voxel was defined in the mid-sagittal section of the rostral body of the CC (Talairach 0, 18, 18). DT-STR and fODF-STR tracking can only find the commissural fibres connecting the medial parts of the frontal lobe. Fanning of the fibre bundle to the inferior and middle frontal gyrus is found with the SPLIT-STR method on the left hemisphere and

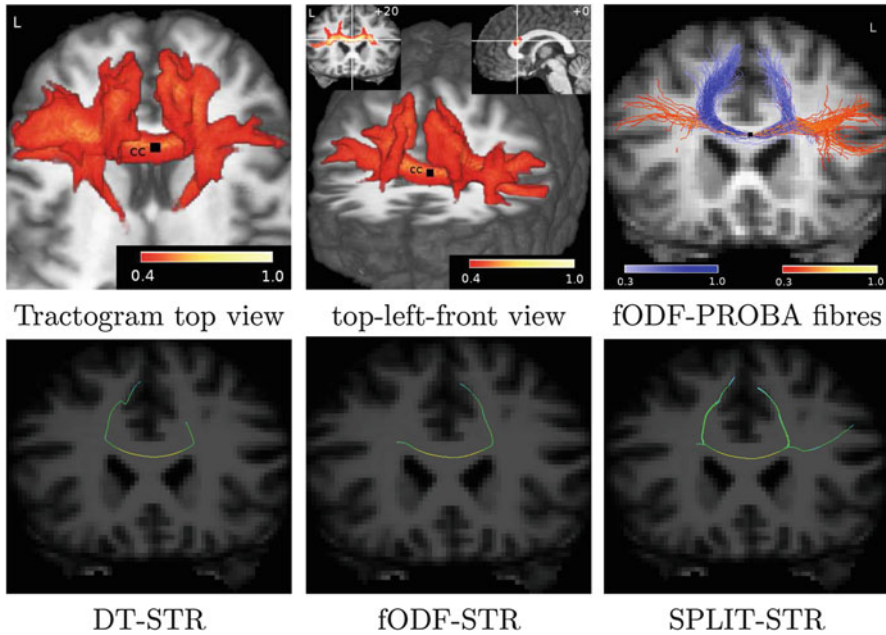


Fig. 9 Deterministic and probabilistic tracking of the projections of the corpus callosum

to a lower extent on the right. The tractogram computed with the fODF-PROBA method reveals a strong interhemispheric connection of the lateral parts of the frontal lobe. Additional fibres are found branching to the anterior thalamic radiation. The tractogram shows asymmetry with stronger connections to the left inferior and middle frontal gyrus than to the homologue area. We also show a selection of the probabilistic fibres colored differently depending on their end point projections to the lateral or medial areas. From the deterministic methods only SPLIT-STR can reconstruct this complex structure.

6 Discussion and Future Work

We have proposed an integral concept for tractography of crossing and splitting fibre bundles in the brain based on HARDI and QBI data. The fibre ODF shows great potential for fibre tractography. The better angular resolution of the fibre ODF allows the tracking to follow multiple maxima and recover fanning and branching structures.

Although more sensitive to initialization, streamline tracking is able to recover similar bundles as the probabilistic method in most cases. SPLIT-STR is thus an efficient and easy way to obtain a good idea of fibre tracts starting from only a few seeds. The underlying assumption of SPLIT-STR is that all multiple peaked fibre

ODF have an underlying fanning/branching structure, which makes it reasonable to follow all available maxima at each step. This is the reason for using a curvature threshold of the tracts of 75° instead of 90° . This threshold avoids following tracts through “pure” crossing configurations, where we know that we are then stepping into other fibre bundles. This is seen in the biological phantom tracking in Fig. 6. When the curvature threshold is set to 90° , both fibre cords are recovered even if the tracking is initialized in only one of the two single fibres. This raises questions for multi-directional deterministic tractography: Should the tracking algorithm split as much as possible to recover as much fibre structure as possible before clustering and post-processing the tracts to separate them into bundles? Or, should the tracking have a built-in scheme to differentiate the different sub-voxel crossing possibilities and decide whether or not a tract should be split in the tracking process? For instance, split in the case of a fanning/branching bundle but not split in the case of a crossing fibre because it is assumed that it steps into a different fibre bundle. More investigation and better characterization of crossing, kissing, fanning and branching fibre configurations remains to be done in the human brain [54, 56, 59]. Information about the local geometry, curvature and torsion of tracts can help for this problem [60]. In [59], preliminary results are obtained discriminating and labeling crossing and fanning sub-voxel fibre configurations.

To deal with the uncertainty in the fibre ODF maxima, a probabilistic approach is more robust and gives more complete results. In probabilistic tracking, tractograms produced by the diffusion ODF are diffusive, stop prematurely and leak into unwanted bundles. In our method, we use a fibre ODF computed from QBI and this function is sampled directly to account for the fact that the fibres in a bundle are not all strictly parallel. Our algorithm follows all possible fibre directions from this fibre ODF. However, it is still unknown how to disentangle the uncertainty and the actual spreading of fibre orientations. We are therefore conservative in our approach and track all possible directions that are given by our data and the assumed model assumption. Other methods use calibration [53, 62], statistical techniques like Bayesian modeling [27], Markov Chain Monte Carlo [11, 40], bootstrap [31] to infer a peak uncertainty of the fibre distributions. It is now important to compare these different HARDI-based probabilistic approaches and see how the different integration of the local reconstruction information impacts the resulting tractograms.

Although, the standard q-ball diffusion ODF reconstruction is smooth, it can be useful as such for other applications than fibre tractography. For instance, in segmentation of fibre bundles, we find that the statistics across diffusion ODF are more stable than with fibre ODFs [24]. Moreover, the diffusion ODF can also be used reliably for clustering. In [68], we show that fibre architecture can be grouped and clustered in different fibre bundles and fibre crossing area. It is now important to understand what information can be extracted precisely by HARDI/QBI/DTI clustering, HARDI/QBI/DTI segmentation and HARDI/QBI/DTI tractography and how they can improve one another. What is in common and complementary about these techniques? To answer this question and compare the algorithms, better tools for validation will be crucial. Ex-vivo and biological phantoms such as [15, 46, 56]

are useful but their configurations are not complex enough. We believe that further development of realistic and complex phantoms will greatly help the validation problem.

References

1. A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda, and D. L. Parker. Analysis of partial volume effects in diffusion-tensor mri. *Magnetic Resonance in Medicine*, 45(4):770–780, 2001.
2. D. Alexander. *An Introduction to Diffusion MRI: the Diffusion Tensor and Beyond*. Springer, 2006.
3. D. Alexander, G. Barker, and S. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.
4. D. C. Alexander. Maximum entropy spherical deconvolution for diffusion mri. In *Image Processing in Medical Imaging*, pages 76–87, 2005.
5. A. Anderson. Measurements of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54:1194–1206, 2005.
6. A. Anwender, M. Tittgemeyer, D. Y. von Cramon, A. D. Friederici, and T. R. Knosche. Connectivity-based parcellation of broca’s area. *Cerebral Cortex*, 17(4):816–825, 2007.
7. Y. Assaf and P. Basser. Composite hindered and restricted model of diffusion (charmed) mr imaging of the human brain. *NeuroImage*, 27(1):48–58, 2005.
8. P. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267, 1994.
9. P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.
10. T. E. J. Behrens, H. Johansen-Berg, S. Jbabdi, M. F. S. Rushworth, and M. W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations. what can we gain? *NeuroImage*, 34(1):144–155, 2007.
11. T. E. J. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnetic Resonance in Medicine*, 50:1077–1088, 2003.
12. Ø. Bergmann, G. Kindlmann, S. Peled, and C.-F. Westin. Two-tensor fiber tractography. In *4th International Symposium on Biomedical Imaging*, pages 796–799, Arlington, Virginia, USA, 2007.
13. R. Blyth, P. Cook, and D. Alexander. Tractography with multiple fibre directions. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 240, Toronto, Canada, 2003. International Society for Magnetic Resonance in Medicine.
14. J. S. W. Campbell, P. Savadjiev, K. Siddiqi, and B. G. Pike. Validation and regularization in diffusion mri tractography. In *Third IEEE International Symposium on Biomedical Imaging (ISBI): from Nano to Macro*, pages 351–354, Arlington, Virginia, USA, 2006.
15. J. S. W. Campbell, K. Siddiqi, V. V. Rymar, A. Sadikot, and G. B. Pike. Flow-based fiber tracking with diffusion tensor q-ball data: Validation and comparison to principal diffusion direction techniques. *NeuroImage*, 27(4):725–736, Oct. 2005.
16. Y.-P. Chao, C.-Y. Yang, K.-H. Cho, C.-H. Yeh, K.-H. Chou, J.-H. Chen, and C.-P. Lin. Probabilistic anatomical connection derived from qbi with mfact approach. In *International Conference on Functional Biomedical Imaging*, Hangzhou, China, October 2007.
17. Y.-P. Chao, C.-H. Yeh, K.-H. Cho, J.-H. Chen, and C.-P. Lin. Multiple streamline tractography approach with high angular resolution diffusion imaging data. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 1550, Berlin, Germany, June 2007.

18. Y. Chen, W. Guo, Q. Zeng, G. He, B. Vemuri, and Y. Liu. Recovery of intra-voxel structure from hard dwi. In *ISBI*, pages 1028–1031. IEEE, 2004.
19. T. Conturo, N. Lori, T. Cull, E. Akbudak, A. Snyder, J. Shimony, R. McKinstry, H. Burton, and M. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences*, 96:10422–10427, Aug. 1999.
20. F. Dell’Acqua, G. Rizzo, P. Scifo, R. Clarke, G. Scotti, and F. Fazio. A model-based deconvolution approach to solve fiber crossing in diffusion-weighted mr imaging. *Transactions in Biomedical Engineering*, 54(3):462–472, 2007.
21. R. Deriche and M. Descoteaux. Splitting tracking through crossing fibers: Multidirectional q-ball tracking. In *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI’07)*, pages 756–759, Arlington, Virginia, USA, April 2007.
22. M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine*, 56:395–410, 2006.
23. M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.
24. M. Descoteaux and R. Deriche. Segmentation of q-ball images using statistical surface evolution. In Springer, editor, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume LNCS 4792, pages 769–776, Brisbane, Australia, 2007.
25. M. Descoteaux, R. Deriche, and A. Anwander. Deterministic and probabilistic q-ball tractography: from diffusion to sharp fiber distributions. Technical Report 6273, INRIA Sophia Antipolis, July 2007.
26. L. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47(6):1083–1099, 2002.
27. O. Friman, G. Farneback, and C.-F. Westin. A bayesian approach for stochastic white matter tractography. *IEEE Transactions in Medical Imaging*, 25(8), 2006.
28. W. Guo, Q. Zeng, Y. Chen, and Y. Liu. Using multiple tensor deflection to reconstruct white matter fiber traces with branching. In *Third IEEE International Symposium on Biomedical Imaging: from Nano to Macro*, pages 69–72, Arlington, Virginia, USA, Apr. 2006.
29. P. Hagmann, L. Jonasson, P. Maeder, J.-P. Thiran, V. J. Wedeen, and R. Meuli. Understanding diffusion mr imaging techniques: From scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *RadioGraphics*, 26:S205–S223, 2006.
30. P. Hagmann, T. G. Reese, W.-Y. I. Tseng, R. Meuli, J.-P. Thiran, and V. J. Wedeen. Diffusion spectrum imaging tractography in complex cerebral white matter: an investigation of the centrum semiovale. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 623. International Society for Magnetic Resonance in Medicine, 2004.
31. H. A. Haroon and G. J. Parker. Using the wild bootstrap to quantify uncertainty in fibre orientations from q-ball analysis. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 903, Berlin, Germany, 19–25th May 2007.
32. C. Hess, P. Mukherjee, E. Han, D. Xu, and D. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine*, 56:104–117, 2006.
33. T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientation in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54:1480–1489, 2005.
34. K. M. Jansons and D. C. Alexander. Persistent angular structure: new insights fom diffusion magnetic resonance imaging data. *Inverse Problems*, 19:1031–1046, 2003.
35. S. Jbabdi, P. Bellec, R. Toro, J. Daunizeau, M. Pelegriani-Issac, and H. Benali. Accurate anisotropic fast marching for diffusion-based geodesic tractography. *International Journal of Biomedical Imaging*, in press, 2007.
36. S. Jbabdi, M. Woolrich, J. Andersson, and T. Behrens. A bayesian framework for global tractography. *NeuroImage*, 37:116–129, 2007.
37. B. Jian and B. C. Vemuri. A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted mri. *IEEE Transactions on Medical Imaging*, 26(11):1464–1471, 2007.

38. B. Jian, B. C. Vemuri, E. Ozarslan, P. R. Carney, and T. H. Mareci. A novel tensor distribution model for the diffusion-weighted mr signal. *NeuroImage*, 37:164–176, 2007.
39. D. K. Jones and C. Pierpaoli. Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach. *Magnetic Resonance in Medicine*, 53:1143–1149, 2005.
40. E. Kaden, T. R. Knosche, and A. Anwender. Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion mr imaging. *NeuroImage*, 37:474–488, 2007.
41. M. H. Khachaturian, J. J. Wisco, and D. S. Tuch. Boosting the sampling efficiency of q-ball imaging using multiple wavevector fusion. *Magnetic Resonance in Medicine*, 57:289–296, 2007.
42. M. Koch, D. Norris, and M. Hund-Georgiadis. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage*, 16:241–250, 2002.
43. B. W. Kreher, J. F. Schneider, J. Mader, E. Martin, H. J., and K. Il'yasov. Multitensor approach for analysis and tracking of complex fiber configurations. *Magnetic Resonance in Medicine*, 54:1216–1225, 2005.
44. M. Lazar and A. L. Alexander. Bootstrap white matter tractography (boot-tract). *NeuroImage*, 24:524–532, 2005.
45. C. Lenglet. *Geometric and Variational Methods for Diffusion Tensor MRI Processing*. PhD thesis, Universite de Nice-Sophia Antipolis, 2006.
46. C. Lin, V. Wedeen, J. Chen, C. Yao, and W. I. Tseng. Validation of diffusion spectrum magnetic resonance imaging with manganese-enhanced rat optic tracts and ex vivo phantoms. *NeuroImage*, 19:482–495, 2003.
47. C. Liu, R. Bammer, B. Acar, and M. E. Moseley. Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51:924–937, 2004.
48. S. E. Maier, S. Vajapeyam, H. Mamata, C.-F. Westin, F. A. Jolesz, and R. V. Mulkern. Biexponential diffusion tensor analysis of human brain diffusion data. *Magnetic Resonance in Medicine*, 51:321–330, 2004.
49. T. McGraw, B. Vemuri, B. Yezierski, and T. Mareci. Von mises-fisher mixture model of the diffusion odf. In *3rd IEEE International Symposium on Biomedical Imaging (ISBI): Macro to Nano*, 2006.
50. S. Mori and P. C. M. van Zijl. Fiber tracking: principles and strategies - a technical review. *NMR in Biomedicine*, 15:468–480, 2002.
51. E. Ozarslan and T. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution imaging. *Magnetic Resonance in Medicine*, 50:955–965, 2003.
52. E. Ozarslan, T. Shepherd, B. Vemuri, S. Blackband, and T. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (dot). *NeuroImage*, 31(3):1086–1103, 2006.
53. G. J. M. Parker and D. C. Alexander. Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *IPMI*, pages 684–695, 2003.
54. G. J. M. Parker and D. C. Alexander. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transactions of the Royal Society, Series B*, 360:893–902, 2005.
55. S. Peleda, O. Friman, F. Jolesz, and C.-F. Westin. Geometrically constrained two-tensor model for crossing tracts in dwi. *Magnetic Resonance Imaging*, 24:1263–1270, 2006.
56. M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Riviere, A. Constantinesco, D. L. Bihan, and J.-F. Mangin. Fiber tracking in q-ball fields using regularized particle trajectories. In *Information Processing in Medical Imaging*, pages 52–63, 2005.
57. A. Ramirez-Manzanares, M. Rivera, B. Vemuri, P. Carney, and T. Mareci. Diffusion basis functions decomposition for estimating white matter intra-voxel fiber geometry. *IEEE Transactions on Medical Imaging*, page in press, 2007.
58. K. E. Sakaie and M. J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted mri. *NeuroImage*, 34:169–176, 2007.

59. P. Savadjiev, J. S. W. Campbell, M. Descoteaux, R. Deriche, G. B. Pike, and K. Siddiqi. Disambiguation of complex subvoxel fibre configurations in high angular resolution fibre tractography. In *Joint Annual Meeting ISMRM-ESMRMB*, page 1477, Berlin, Germany, 19-25th May 2007. International Society of Magnetic Resonance in Medicine.
60. P. Savadjiev, J. S. W. Campbell, B. G. Pike, and K. Siddiqi. 3d curve inference for diffusion mri regularization and fibre tractography. *Medical Image Analysis*, 10:799–813, 2006.
61. K. K. Seunarine and D. C. Alexander. Linear persistent angular structure mri and non-linear spherical deconvolution for diffusion mri. In *International Society for Magnetic Resonance in Medicine*, page 2726, 2006.
62. K. K. Seunarine, P. A. Cook, M. G. Hall, K. V. Embleton, G. J. M. Parker, and D. C. Alexander. Exploiting peak anisotropy for tracking through complex structures. In *Mathematical Methods in Biomedical Image Analysis (MMBIA 2007)*, 2007.
63. J.-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, 2007.
64. J.-D. Tournier, F. Calamante, D. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, 23:1176–1185, 2004.
65. D. Tuch. *Diffusion MRI of Complex Tissue Structure*. PhD thesis, Harvard University and Massachusetts Institute of Technology, 2002.
66. D. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, 2004.
67. D. Tuch, T. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582, 2002.
68. D. Wassermann, M. Descoteaux, and R. Deriche. Recovering cerebral white matter structures with spectral clustering of diffusion mri data. Technical Report 6351, INRIA Sophia Antipolis-Méditerranée, November 2007.
69. V. Wedeen, T. Reese, D. Tuch, M. Wiegel, J.-G. Dou, R. Weiskoff, and D. Chessler. Mapping fiber orientation spectra in cerebral white matter with fourier-transform diffusion mri. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 82. International Society for Magnetic Resonance in Medicine, 2000.

Segmentation of Clustered Cells in Microscopy Images by Geometric PDEs and Level Sets

A. Kuijper, B. Heise, Y. Zhou, L. He, H. Wolinski, and S. Kohlwein

Abstract With the huge amount of cell images produced in bio-imaging, automatic methods for segmentation are needed in order to evaluate the content of the images with respect to types of cells and their sizes. Traditional PDE-based methods using level-sets can perform automatic segmentation, but do not perform well on images with clustered cells containing sub-structures. We present two modifications for popular methods and show the improved results.

1 Introduction

Automatic cell segmentation and cell scan analysis belong to the challenging tasks in image processing. Due to different types of microscopy, e.g. fluorescence, transmission or phase contrast microscopy, no general solution is applicable. Whereas

A. Kuijper (✉)

Fraunhofer IGD, Institute for Computer Graphics Research, Department of Computer Science, TU Darmstadt, Darmstadt, Germany
e-mail: arjan.kuijper@igd.fraunhofer.de

B. Heise

Department of Knowledge-Based Mathematical Systems,
Johannes Kepler University, Linz, Austria

Y. Zhou

Department of Virtual Design, Siemens AG, Munich, Germany

L. He

Luminescent Technologies Inc., Palo Alto, USA

H. Wolinski • S. Kohlwein

SFB Biomembrane Research Center, Institute of Molecular Biosciences,
Department Biochemistry, University of Graz, Graz, Austria

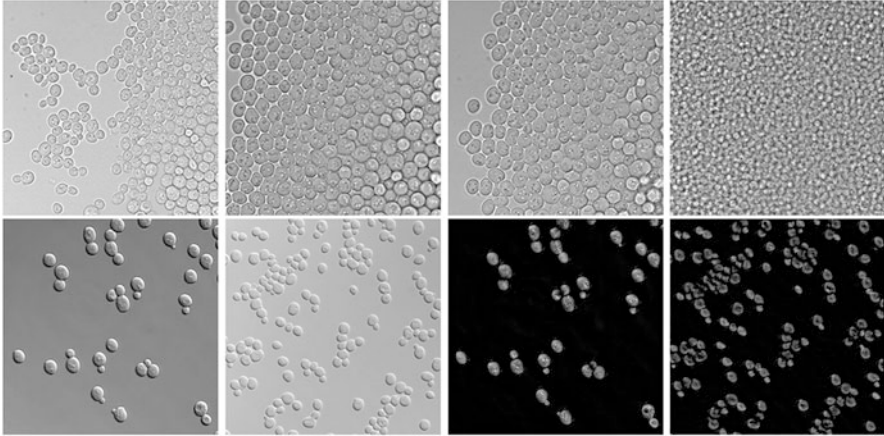


Fig. 1 Top row: Yeast cells in different concentrations recorded by transmission microscopy. Bottom row: Yeast cells in different concentrations recorded by DIC microscopy and the reconstructed OPL map. Note the wide variations in image quality

for fluorescence microscopy some image processing software exists, transmission and phase contrast microscopy are still slightly neglected in automatic bio-image analysis. In recent years, however, biologists renewed their interest in the latter and used microscopical setups avoiding additional staining to reduce stress reactions in cells.

The tasks of cell population analysis can be described by counting cells, determining the distribution of cell areas and distinguishing between different states as e.g. mother and daughter cells. Especially yeast cells act as a well-established model for several microbiological investigations in the field of proteomics. Often automatic segmentation and analysis of cell scans are connected with a previous screening of hundreds of cell scans automatically performed by a robot [24]. Due to this automated high-throughput imaging we have to cope in reality with several additional difficulties as various cell concentrations, defocused scans or overlapping cells. In this chapter we describe automatic cell segmentation methods for transmission microscopy (top row of Fig. 1) and Differential Interference Contrast (DIC) microscopy (bottom row left of Fig. 1) as two examples of light microscopy without any fluorescence staining.

In DIC microscopy the phase gradient is measured. For image processing this scan type is difficult to analyse directly. By a deconvolution method adapted to the DIC microscope [5] we can transform the DIC image into an intermediate image (bottom row right of Fig. 1) with an appearance similar to a fluorescence image. The deconvolution replaces the measured phase gradient by an optical path length (OPL) map. We will further use the OPL image for segmentation.

1.1 PDE methods

The task of automatic segmentation is to find coherent regions that relate to objects the observer is interested in. Typically, these regions are bounded and the boundary of objects are often found as the location where the intensity changes significantly. However, methods that use such edge information only, perform unsatisfactory as locally this edge information may be missing. In order to find edges at these locations, iterative schemes can be used that are based on Partial Differential Equations (PDEs). Roughly speaking, these PDEs are designed using driving forces that evolve curves towards edges. Usually, the PDE contains an edge detector which depends on the gradient of the image. Alternatively, a homogeneous area detector can be used which favours areas without much intensity changes. To minimise the influence of noise and missing information, constraints that penalise (for instance) the length of edges are added.

Examples of such approaches relate for instance to the active contour model [6], and modifications of it. The initial curve is to be placed around the object to be detected. The curve moves in the normal dimension to itself and ideally stops at the boundary of the object. The main drawbacks of the original snakes are their sensitivity to initial conditions and the difficulties associated with topological transformations.

Other approaches arise from the Mumford-Shah model [11]. It minimises a proposed energy of an image. This energy depends on piecewise smooth approximations for the given image. Although the model has its mathematical peculiarities, it formed the basis of many segmentation algorithms. For an overview of the ample literature on geometric pde models in image analysis, one can consult the collections of papers in e.g. [9, 13, 15].

1.2 Level set methods

A drawback of many PDEs approaches is the computation time. Often, the discretisation scheme requires small time steps in order to maintain a stable evolution. Secondly, PDE methods require an initial guess, which may heavily influence the final result. Thirdly, as one doesn't know the number of objects at forehand, the number of available edge contours has to be flexible.

A framework that overcomes these problems to a large extent use level sets, initiated by Osher and Sethian [14, 18]. It exploits mathematical properties of equations of the form

$$\phi_t + F \cdot |\nabla\phi| = 0, \tag{1}$$

where F is a function that is user-defined. It can, for example, contain edge information. This model is applied in a wide variety of applications, see e.g. [9, 12, 13, 19].

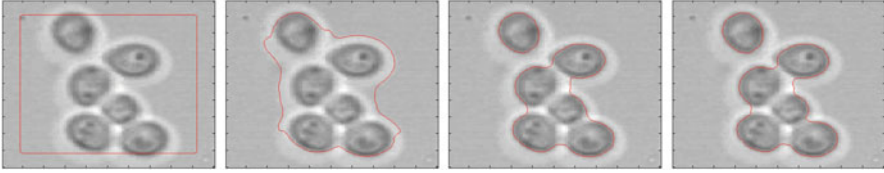


Fig. 2 Level Set implementations may fail. From left to right: The initial contour; results after 100, 1000, and 2000 iterations

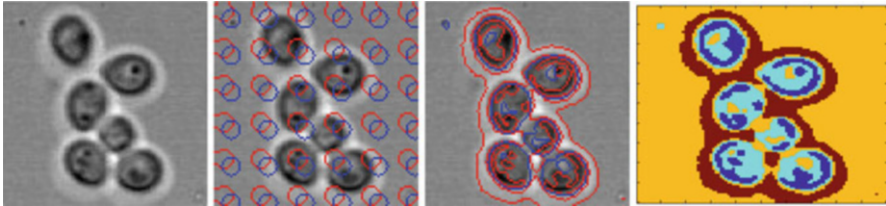


Fig. 3 Segmentation result of the Multiphase Chan-Vese model. From left to right: original image; initial condition; final contour after 2000 iterations, $\mu = 0.0015 \cdot 255^2$; final phase

In image segmentation, one can think of ϕ being (a level line of) the image that evolves as a function of its gradient, or, more explicitly: as the evolution of edges towards their mathematically optimal position. The level set framework allows for a flexible setup, as it is able to split and merge curves during the process. To some extent, this also overcomes the initialisation problem. Finally, fast methods are available for implementation [8, 10].

1.3 Cell segmentation

In cell segmentation, impressive results using this methodology can be reached [3, 20, 21] when the cells are not too clustered and when they do not contain too much sub-structure. However, when they do, even these methods fail. An example is given in Fig. 2, where the implementation by Li et al. is used [8]. The optimal solution finds a set of clustered cells, since in this case the length of the edge curve is smallest. Relaxing the edge-length constraint results in segmenting dark sub-structures. An example of this can be seen in Fig. 3.

In the next two sections we present extensions of two general methodologies that use PDEs in a level-set framework. Firstly, we use a pre-processing step in Sect. 2 to get useful results for the Chan-Vese model [27]. Secondly, in Sect. 3 we combine a voting method for the cell centre with geometric PDEs [7]. Full details on the methodology and comparison with related methods can be found in [28].

2 The Chan-Vese Level Set Method

Chan and Vese [1] proposed a model for image segmentation based on the Mumford-Shah functional and level sets. This model does not explicitly depend on the gradient of the image. It therefore performs well in images with ambiguous boundaries. The model is defined as the following minimisation problem [1, 2]: Given the original image u_0 in Ω and two unknown constants c_1 and c_2 in the inside and outside region, respectively, which are separated by the curve $\delta(\phi)|\nabla\phi|$, minimise the following energy with respect to c_1, c_2 and ϕ :

$$F(c_1, c_2, \phi) = \int_{\Omega} \lambda_1(u_0 - c_1)^2 H(\phi) + \lambda_2(u_0 - c_2)^2 (1 - H(\phi)) + \mu \delta(\phi) |\nabla\phi| d\Omega, \quad (2)$$

where H denotes the Heaviside function and δ the Dirac measure. The first and second term penalise the L_2 distance in the inside and outside regions, and the third term measures the length of the curve.

2.1 The multiphase level set method

The traditional Chan-Vese model can segment the image into two regions - background and foreground. In order to segment images into more regions, a Multiphase Level Set Method has been developed [4, 23]. It is motivated by the Four Colour Theorem which states that only four colours are enough to dye all the regions in a partition. Therefore, only two level set functions will suffice to represent any partition. The multiphase level set method is the direct extension of Eq. (2) and is defined as follows:

$$F(c, \phi) = \int_{\Omega} (u_0 - c_{11})^2 H(\phi_1) H(\phi_2) d\Omega + \int_{\Omega} (u_0 - c_{10})^2 H(\phi_1) (1 - H(\phi_2)) d\Omega + \int_{\Omega} (u_0 - c_{01})^2 (1 - H(\phi_1)) H(\phi_2) d\Omega + \int_{\Omega} (u_0 - c_{00})^2 (1 - H(\phi_1)) (1 - H(\phi_2)) d\Omega + \mu \int_{\Omega} |\nabla H(\phi_1)| + \mu \int_{\Omega} |\nabla H(\phi_2)|, \quad (3)$$

where the first four integrals compute the L_2 distance of a constant c_{ij} to u_0 in each of the four difference regions ($c_{11} \in \{(x, y) : \phi_1 > 0, \phi_2 > 0\}$, etc.). The Euler-Lagrange equations obtained by minimising Eq. (3) with respect to c and ϕ are:

$$\frac{\partial \phi_1}{\partial t} = \delta_{\epsilon}(\phi_1) \left\{ \mu \cdot \operatorname{div} \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) - [(u_0 - c_{11})^2 - (u_0 - c_{01})^2] H(\phi_2) + [(u_0 - c_{10})^2 - (u_0 - c_{00})^2] (1 - H(\phi_2)) \right\}, \quad (4)$$

$$\begin{aligned} \frac{\partial \phi_2}{\partial t} = & \delta_\epsilon(\phi_2) \left\{ \mu \cdot \operatorname{div} \left(\frac{\nabla \phi_2}{|\nabla \phi_2|} \right) \right. \\ & - [((u_0 - c_{11})^2 - (u_0 - c_{10})^2) H(\phi_1) \\ & \left. + ((u_0 - c_{01})^2 - (u_0 - c_{00})^2) (1 - H(\phi_1)) \right\}. \end{aligned} \quad (5)$$

Furthermore, $c_{ij} = |u_0|$ in the regions where they are defined.

The only varying parameter in the model is the coefficient of the length term μ . This term affects the length of the final contour. The smaller μ is, the longer the final contour is. So if we want to detect as many objects as possible (for example, segment all cells), the μ should be small. If we are only interested in large objects and do not want to detect small objects (for example, the sub-structure inside the cell, cf. Fig. 2), μ should be large. However, these two requirements contradict each other as we want both.

Figure 3 shows the result applying the Multiphase Chan-Vese model with a small value of μ . It manages to find the individual cells, but only at the cost of sub-segmenting the cells. The segmentation criterion for Chan-Vese based models is the average grey level of an object, cf. Eq. (3). For piecewise constant objects, the performance of this model is therefore excellent. However, real cells always contain sub-structures inside them, which will cause a false segmentation. Furthermore, in some cell images, the grey level difference between the cells and background is not distinct. Therefore, a pre-processing method is needed to solve these problems.

2.2 Pre-processing

In order to remove the substructures, one can try to detect these small parts and remove them [25]. This is not only computationally intensive, but also assumes that the locations of the cells (including boundaries!) are more or less known. We therefore proceed with a PDE-based pre-processing method that smoothes regions with small variability and maintains edges. The most popular among such approaches is the one by Perona and Malik [17]:

$$I_t = \nabla \cdot (c(|\nabla I|) \nabla I). \quad (6)$$

Here $c(|\nabla I|)$ is a decreasing function of the gradient. In this work, we choose the rather standard function $c(\xi) = (1 + \frac{\xi^2}{k^2})^{-1}$.

In Fig. 4 we show the result of the Perona - Malik pre-processing, followed by the multiphase Chan-Vese model. Indeed, the sub-structures are removed and the cells and background are properly segmented.

The Perona - Malik pre-processing method performs well for images with sufficiently contrast between cells and background, cf. Figs. 3 and 4. In most images of our data base, cf. Fig. 1, this is the case. However, in some cases the final result is not satisfactory due to the merging of cells. This happens when the cells in the

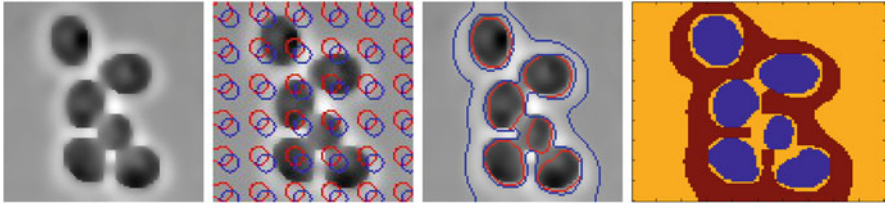


Fig. 4 Segmentation result after pre-processing. From left to right: Result of Perona-Malik processing ($K=10$, 30 iterations) of Fig. 3a; initial contours; final contour after 2000 iterations, $\mu = 0.0015 \cdot 255^2$; final phase

image are too close to each other or even overlap each other. We therefore consider a different approach in the next section which enforces the non-merging of cells.

3 Seed Detection and Geometric PDEs

The first step for this segmentation method is the detection of the centres of the cells. In the second step these locations are used as seed points for a PDE-based region growing method that stops at the cell boundaries.

3.1 Iterative voting

The shape of a cell has some characteristic properties, like symmetry, continuity, and closure. Therefore, we apply an iterative method using oriented kernels to detect the candidate position for the centre of the cell [3, 16, 26]. The basic idea of this algorithm is a voting method. It defines a series of cone-shaped kernels that vote iteratively along radial directions the likeness of the point as being the middle of a rather homogeneous structure. Applying this kernel along the gradient direction, at each iteration and each grid location, the orientation of the kernel is updated. The shape of the kernel is also refined and focused as the iterative process continues. Finally, the point of interest is selected by certain threshold.

Numerical results

The algorithm is tested on a data base with different types of images, cf. Fig. 1. We present the results of four typical images, showing different types of cell clustering for which results can be verified by human observers. The complete data base also contains images with completely overlapping and out-of-focus cells. In these cases, human observers cannot distinguish individual cells. In Fig. 5, the first column shows the original images. The second row shows the last image of the iterative

voting method. These images are thresholded in order to get the binary images in the last column, yielding the location of the seeds, located approximately in the middle of each cell, as expected.

3.2 Geometric PDE evolution using level sets

In the next step, the detected centre is used to define the initial condition for detecting the boundary. A sequence of three level set functions are set up for each cell. See [9, 20, 28] for more details.

Phase 1: Initial expansion

Motivated by the traditional level set model, the following level set equation is defined [20]:

$$\phi_t + g \cdot (1 - \epsilon\kappa) \cdot |\nabla\phi| - \beta\nabla g \cdot \nabla\phi = 0, \quad (7)$$

where $g = e^{-\alpha|\nabla(G*I_0(x))|}$, $\alpha > 0$, $G*I_0(x)$ is the convolution of the original image $I_0(x)$ with a Gaussian function G . The function $g \cdot (1 - \epsilon\kappa)$ contains an inflationary term (+1), which determines the direction of evolution to be outward. The curvature term ($\epsilon\kappa$) regularises the surface by accelerating the movement of those parts of surface behind the average of the front and slowing down the advanced parts of flow. The parameter ϵ determines the strength of regularisation, if ϵ is large, it will smooth out front irregularities. For extreme cases, the final front will become a circle. If ϵ is small, the front will maintain sharp corners. In practise, an intermediate value of ϵ is chosen that allows the front to have concavities (concavities are possible for the cell border), while small gaps and noise are smoothed. The effect of g is to speed up the flow in those areas where the image gradient is low and slowing down where the image gradient is high. Because of this term, the front slows down almost to stop when it reaches the internal cell boundary. The parameter α determines the sensitivity of the flow to the gradient. The extra term $\beta\nabla g \cdot \nabla\phi$ is a parabolic term that enhances the edge effect.

In this step, there is another restriction condition for the equation: *The growing front may not invade (and merge with) other cells' region when seed grows.* By using this level set equation, the internal boundary of the cells is detected.

Phase 2: Free expansion

The initial expansion level set function usually causes underestimation of cell area due to the thick boundary. So a second and third step are added to compensate the result. The second step is the free expansion, in which the front is allowed to expand freely and the speed of evolution doesn't rely on the gradient of original image. The level set equation is simply defined as below:

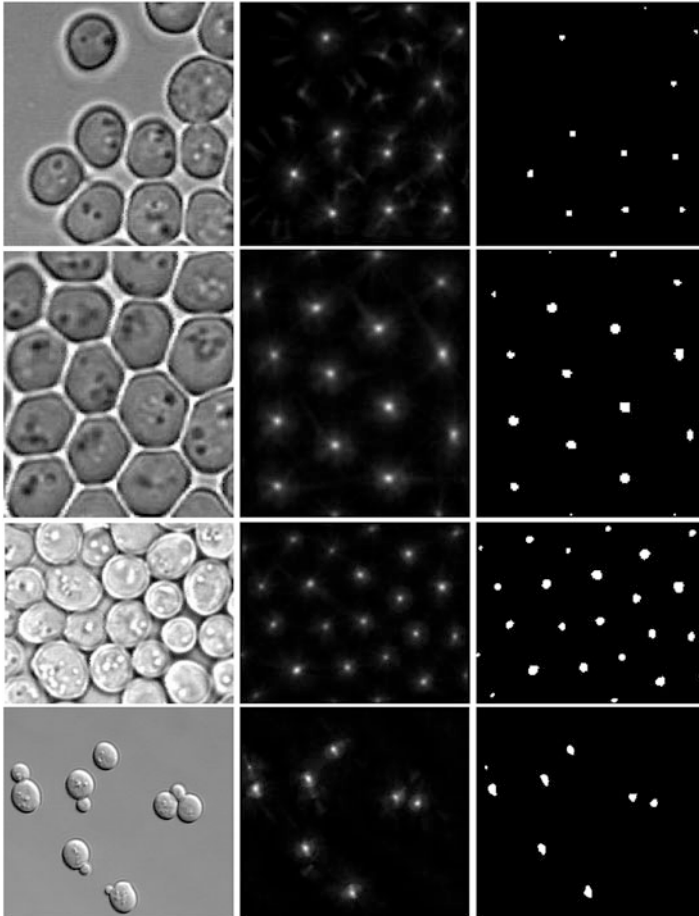


Fig. 5 Seed Detection Result for several types of images. For visualisation purposes we selected a smaller part of the original images. From left to right: Initial image; final voting result; resulting seed point

$$\phi_t + |\nabla\phi| = 0. \quad (8)$$

Similar as the first step, the growing fronts may not penetrate each other when the front expands. This expansion only needs a small number of steps to ensure that all the fronts move beyond the external boundary of cells. The number of iteration depends on the thickness of the cell boundary.

Phase 3: Surface wrapping

After the free expansion step, the fronts are located outside the cells' boundary. The last step is to move the front inwards to get the exact location of the external cell

boundary using

$$\phi_t + g \cdot (-1 - \epsilon\kappa)|\nabla\phi| = 0. \quad (9)$$

This is similar to initial expansion except for the shrinking term (-1) , which determines the direction of evolution to be inward.

Reinitialisation

For the three level set phases described above, reinitialisation is necessary. The purpose of reinitialisation is to keep the evolving level set function close to a signed distance function during the evolution. It is a numerical remedy for maintaining stable curve evolution [22]. The reinitialisation step is to solve the following evolution equation:

$$\begin{cases} \psi_\tau &= \text{sign}(\phi(t))(1 - |\nabla\psi|), \\ \psi(0, \cdot) &= \phi(t, \cdot). \end{cases} \quad (10)$$

Here, $\phi(t, \cdot)$ is the solution ϕ at time t . This equation is solved by an iterative method. Five iterations are sufficient. The result ψ will be the new ϕ used in the evolution.

Numerical results

Figure 6 shows the results on the images used in Fig. 5. The columns show the initial condition for level set function, the results after the three level set phases, and the final segmentation. Note that in the DIC image (the last row), the mother cells are segmented successfully.

4 Summary and Conclusions

We discussed two PDE-based methods to segment cells with level sets: the Chan-Vese model and geometric PDEs. In case of clustered cells with sub-structures, these models perform unsatisfactory. We presented possible pre-processing methods, viz. Perona-Malik smoothing and voting, which provide a modified image and starting conditions, respectively.

The methods are tested on different types of cell images, with both regular and irregular patterns, and overlapping cells. The modified Chan-Vese model is fast and performs well in most cases. The performance of the geometric PDEs combined with voting performs better, but requires more computation time, see [28] for full details. As long as the cells are distinguishable by the human eye, the method can segment individual cells properly. Both methods require parameter settings that need

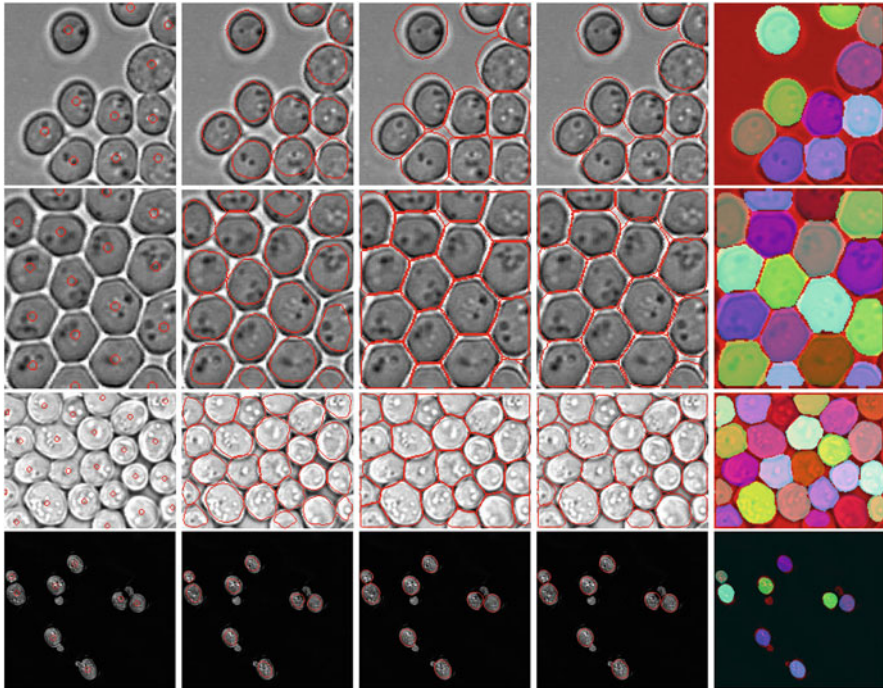


Fig. 6 Results for the seeded images of Fig. 5. From left to right: Initial Contour; Result after Initial Expansion; Result after Free Expansion; Result after Surface Wrapping; Final Result

to be determined once, as long as the intrinsic properties of the cell images are stable, i.e. when the images are not too much out of focus. In the latter case, even human observers are unable to segment the cells.

Acknowledgements The work was partially supported by the mYeasty pilot-project by the Austrian GEN_AU research program (www.gen-au.at). It was carried out when A. Kuijper, Y. Zhou, and L. He were with the Johann Radon Institute for Computational and Applied Mathematics (RICAM), Linz, Austria.

References

1. T. Chan and L. Vese. Active contours without edges. *IEEE Trans. on Image Processing*, 10:266–277, 2001.
2. T. Chan and L. Vese. Active contour and segmentation models using geometric pde's for medical imaging. In R. Malladi, editor, *Geometric Methods in Bio-Medical Image Processing*, chapter 4, pages 63–76. Springer, 2002.
3. H. Chang, Q. Yang, and B. Parvin. Segmentation of heterogeneous blob objects through voting and level set formulation. *Pattern Recognition Letters*, 28(13):1781–1787, 2007.

4. L. He and S. Osher. Solving the chan-vese model by a multiphase level set method algorithm based on the topological derivative. In *1st International Conference on Scale Space and Variational Methods in Computer Vision*, pages 777–788, 2007.
5. B. Heise and B. Armingier. Some aspects about quantitative reconstruction for differential interference contrast (dic) microscopy. In *PAMM 7(1): (Special Issue: Sixth International Congress on Industrial Applied Mathematics (ICIAM07) and GAMM Annual Meeting, Zürich 2007)*, pages 2150031–2150032, 2007.
6. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. of Comp. Vision*, 1:321–331, 1988.
7. A. Kuijper, Y. Zhou, and B. Heise. Clustered cell segmentation - based on iterative voting and the level set method. In *3rd International Conference on Computer Vision Theory and Applications (VISAPP, Funchal, Portugal, 22 - 25 January 2008)*, pages 307–314, 2008.
8. C. Li, C. Xu, C. Gui, and M. Fox. Level set evolution without re-initialization: A new variational formulation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 430–436, 2005.
9. R. Malladi. *Geometric Methods in Bio-Medical Image Processing*. Springer, 2002.
10. R. Malladi and J. A. Sethian. Fast methods for shape extraction in medical and biomedical imaging. In R. Malladi, editor, *Geometric Methods in Bio-Medical Image Processing*, chapter 1, pages 1–18. Springer, 2002.
11. D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math*, 42:577–685, 1989.
12. S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, New York, 2003.
13. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer, 2003.
14. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
15. N. Paragios, Y. Chen, and O. Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer, 2006.
16. B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. Barcellos-Hoff. Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE Trans Image Process.*, 16(3):615–623, 2007.
17. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *PAMI*, 12(7):629–639, 1990.
18. J. Sethian. Curvature and the evolution of fronts. *Comm. In Math. Phys.*, 101:487–499, 1985.
19. J. Sethian. *Level set methods and fast marching methods: Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge University Press, Cambridge, UK, 1999.
20. C. Solorzano, R. Malladi, S. Lelievre, and S. Lockett. Segmentation of nuclei and cells using membrane related protein markers. *Journal of Microscopy*, 201:404–415, 2001.
21. C. Solorzano, R. Malladi, and S. Lockett. A geometric model for image analysis in cytology. In R. Malladi, editor, *Geometric Methods in Bio-Medical Image Processing*, chapter 2, pages 19–42. Springer, 2002.
22. M. Sussman and E. Fatemi. An efficient, interface preserving level set redistancing algorithms and its application to interfacial incompressible fluid flow. *SIAM J.Sci. Comp.*, 20:1165–1191, 1999.
23. L. Vese and T. Chan. A multiphase level set framework for image segmentation using the Mumford and Shan Model. *Int. J. of Comp. Vision*, 50(3):271–293, 2002.
24. H. Wolinski and S. Kohlwein. Microscopic analysis of lipid droplet metabolism and dynamics in yeast. In *Membrane Trafficking*, volume 457 of *Methods in Molecular Biology*, chapter 11, pages 151–163. Springer, 2008.
25. Q. Yang and B. Parvin. Harmonic cut and regularized centroid transform for localization of subcellular structures. *IEEE Transactions on Biomedical Engineering*, 50(4):469–475, April 2003.

26. Q. Yang, B. Parvin, and M. Barcellos-Hoff. Localization of saliency through iterative voting. In *ICPR (1)*, pages 63–66, 2004.
27. Y. Zhou, A. Kuijper, and L. He. Multiphase level set method and its application in cell segmentation. In *5th International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2008, Innsbruck, Austria, February 13 - 15, 2008)*, pages 134–139, 2008.
28. Y. Zhou, A. Kuijper, B. Heise, and L. He. Cell segmentation using the level set method. Technical Report 2007-17, RICAM, 2007. <http://www.ricam.oeaw.ac.at/publications/reports/07/rep07-17.pdf>.

Atlas-based whole-body registration in mice

M. Baiker, J. Dijkstra, J. Milles, C.W.G.M. Löwik, and B.P.F. Lelieveldt

Abstract In this chapter, we present a fully automated approach for whole-body segmentation of mice in CT data, based on articulated skeleton registration. The method uses an anatomical animal atlas where position and degrees of freedom for each joint have been specified. Based on the registration result of the skeleton, a set of corresponding landmarks on bone and joint locations is used to derive further correspondences on surface representations of the lung and the skin. While atlas-based registration is applied to the former, a local geodesic shape context is employed for the latter. Subsequently, major organs are mapped from the atlas to the subject domain using Thin-Plate-Spline approximation, constrained by the landmarks on the skeleton, the lung and the skin. Accuracy and precision of the skeleton registration as well as organ approximation results in a follow-up study are demonstrated.

1 Introduction

In recent years, two widely applied imaging modalities in clinical practice, Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), have been adapted for small animal applications. Due to their non-invasive nature and the imaging field of view, the new modalities (μ CT and μ MRI) can be used for monitoring dynamic processes under realistic conditions *in-vivo* and in the entire animal. This adds a new dimension to animal experiments, since it enables studying

M. Baiker • J. Dijkstra • J. Milles • C.W.G.M. Löwik • B.P.F. Lelieveldt (✉)
Division of Image Processing, Department of Radiology, Leiden University Medical Center,
Albinusdreef 2, Leiden 2333ZA, Netherlands
e-mail: b.lelieveldt@lumc.nl

the effect of e.g. genetic manipulations or drug administration within the same subject, at subsequent points in time. Therefore, the traditional cross-sectional studies using different animals can be extended to follow-up studies.

2 Problem Statement

Efficient comparison, compilation or assessment of processes in entire animals requires capturing and matching the body as a whole. While a large amount of methods aim at rigid or non-rigid registration of objects with equal or similar intrinsic structural properties like single bones or organs [12,20], only little attention has been paid so far to the matching of a system that may consist of many different structures with significantly different properties. An animal body is an example of such a system because it contains rigid structures like bones next to elastic (non-rigid) structures like internal organs. In addition, it contains many articulated parts. As a result, the shape and the posture can vary significantly among animals and among acquisition time points.

The strategy to capture an animal body as a whole is dependent on the amount of Degrees of Freedom (DoFs) to be accounted for by the kind of study (intra-subject, cross-sectional or longitudinal). Also the way the data is acquired determines how prominent the expected articulations are (e.g. usage of an animal holder). Besides that, the imaging modality is an important factor because certain anatomical structures may not show sufficient contrast for registration. Examples are the lack of bone information in MRI or the lack of soft tissue contrast in CT data. Dependent on these issues, a method requires including a certain amount of *a-priori* information. This can e.g. be realized by adding knowledge about structural properties of single elements of the system, about kinematics of elements relative to each other or about position and spatial extent of anatomical objects.

2.1 *Prior art*

In the literature, there are two approaches for the problem of matching anatomical structures with heterogeneous and potentially articulated parts:

1. Data-driven registration of the entire body or body parts, based on the data directly or on extracted features like points or surfaces.
2. Registration based on an underlying model of the relation between subparts of the body (articulated registration), applied to body parts. Again, the registration can be based on the data or extracted features like points or surfaces.

An example of a data-driven approach is presented in Chaudhari et al. [6]. The authors perform a surface-based registration between an entire mouse body and the Digimouse atlas (Dogdas et al. [8]). They segment the animal interior by surface-constrained warping of the atlas volume to the subject using harmonic maps.

The method does not take anatomical heterogeneity into account. Li et al. [11] present a whole-body intra-modality approach (CT) for mice that besides the skin uses the skeleton for registration. They put additional constraints on a point matching framework to account for rigidity of bones. Kovacevic et al. [10] demonstrate whole-body registration in mice based on a basic model, the “part-of” concept. This is a hierarchical intra-modality approach that first separates the main organ compound and refines that division as the registration progresses, down to single bones and organs. While the method integrates inter-structure relationships inside the body, these are only exploited for initializing the registration of low-hierarchical elements by the result from high-hierarchical elements.

Martin-Fernandez et al. [13] make use of an anatomically realistic articulation model to register 2D hand radiographs. The bones are thereby represented by a wire-frame where individual ‘rods’ are registered imposing kinematic constraints. Papademetris et al. [14] register the legs of a mouse by modeling the joints. After registration of the leg bones, they pay special attention on the propagation of the deformations to soft tissue parts by focusing on the folding problem at interfaces of articulated parts. Du Bois d’Aische et al. [9] register a human head, based on a model of the jugular part of the spine. Articulated cervices are registered to the target image and the deformation is propagated to the rest of the head using a linear elastic model.

In summary, some available methods can be used either for whole-body applications, as long as differences in posture and shape are small, or for registration of subparts of a body. However, most methods need a significant amount of user interaction e.g. to define joint locations or manually segment bones prior to registration.

3 Methodology

In this section, a method is proposed to segment an entire mouse body from Micro-CT data using an anatomical mouse atlas that contains the skeleton and major organs (Segars et al. [16]). To this end, the skeleton is registered as a first step, because it 1) forms the rigid, articulated frame of a body, 2) is the main determinant of whole-body posture and 3) can be robustly and automatically extracted from the data. Combining the “part-of” concept with a hierarchical anatomical model and articulated registration enables fully automated registration of the atlas to the skeleton of a given mouse.

Due to the lack of soft-tissue contrast in Micro-CT data, intensity-based organ registration is not possible. However, the mouse atlas contains all major organs which can be mapped from the atlas domain to the subject domain using Thin-Plate-Spline (TPS) interpolation [5]. While the necessary set of corresponding anatomical landmarks is mainly determined by the result of the skeleton registration, more lateral landmarks are needed in the ventral part of the animal abdomen to sufficiently constrain the TPS mapping. Besides bone, also the lung and the skin show sufficient contrast for robust segmentation from CT data. Therefore, corresponding lung

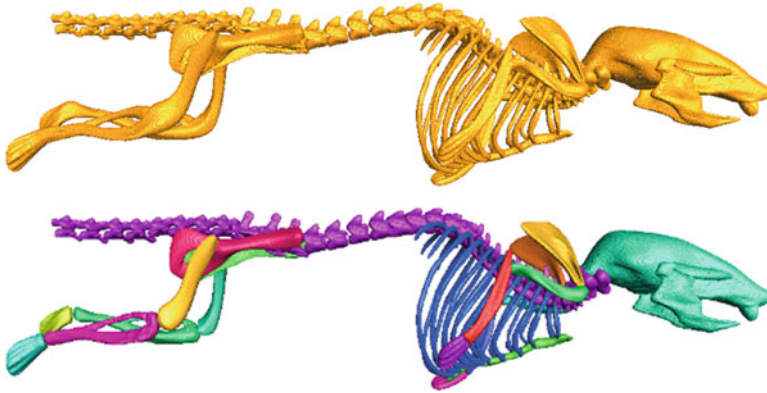
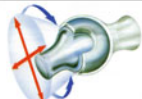



Fig. 1 The mouse skeleton as included in the atlas (top) and after segmentation of single bones and adding joints (bottom)(adapted from [2], ©2007 IEEE)

Table 1 Joint types of the atlas skeleton and the DoFs for the registration of the distal articulated bone (Pictograms from [7])

Joint types	Modeled joint	DoFs of the articulated bone
 Ball joint	Shoulder Wrist Hip Ankle	0 translations 3 rotations 3 scalings
 Hinge joint	Elbow Knee	0 translations 1 rotation 3 scalings

landmarks can be determined using atlas-based registration. Moreover, a sparse set of skin correspondences can be derived since at many locations in the animal, skin is very close to the skeleton. This serves to determine a denser set of skin correspondences. Parts of this work have been published in [1, 2]

3.1 Modeling articulation kinematics

The used atlas skeleton (Segars et al. [16]) does not distinguish between individual bones. Therefore, these were manually segmented (Fig. 1). Second, the position and the DoFs were specified for each joint.

Three types of joints have to be modeled: ball joints, hinge joints and the shoulder complex (both shoulders combined). Table 1 shows the DoFs for the ball and hinge

joints. In addition to these anatomically relevant DoFs, three translations for both joint types and two rotations for the hinge joints are allowed to a small extent, to be able to compensate for potential errors that have been made during previous registration of another bone that is rigidly connected. Due to the large number of DoFs in the shoulder, an additional motion constraint has to be introduced for the shoulder by allowing only a coupled, symmetric displacement of both front upper limbs, with a varying distance between the shoulders and a rotation towards and away from each other. Subsequently, the left and the right front upper limb can be decoupled.

3.2 Hierarchical anatomical model of the skeleton

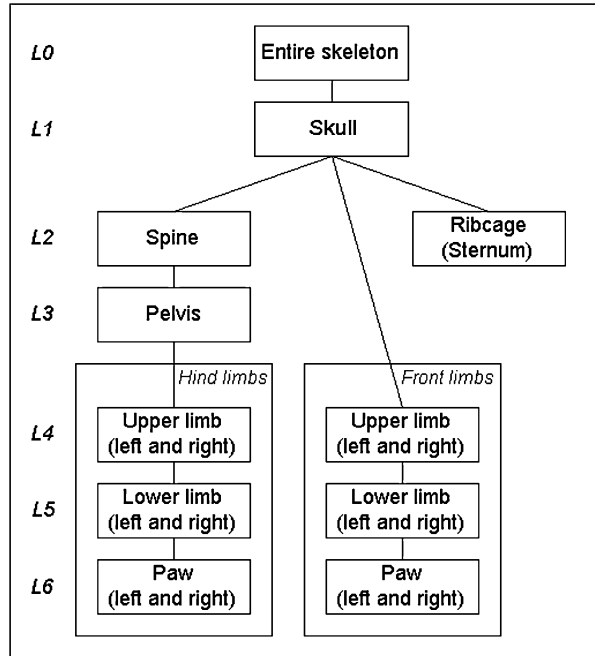
The hierarchical anatomical tree used for this work is shown in Fig. 2. The strategy to incorporate the entire bone structure is to first align the atlas and the subject skeleton coarsely and then to apply an articulated registration scheme traversing the hierarchical tree. In this way, lower tree levels are initialized and constrained by higher level transformations. The highest hierarchical level is the entire mouse skeleton itself ($L0$). The skull is placed on the next lower level ($L1$), because its registration result initializes the matching of all other parts. The rest of the skeleton is divided into three subparts, consisting either of single bones or bone compounds ($L2 - L6$).

These are the rear part consisting of spine, pelvis, upper and lower hind limbs and the paws, the front part consisting of upper and lower front limbs and the paws and the ribcage, represented by the sternum. The relations between all elements in the three subparts are fully determined by rigid connections (joints). Including the shoulder blade into the tree or making further distinctions such as refining the paws is not relevant for the goal of capturing the animal posture and therefore can be omitted. Assuming that the spine and the sternum sufficiently constrain the shape of the ribcage, the ribs can be left out as well.

3.3 Resolving global degrees of freedom

For initialization of the articulated registration, the mouse atlas needs to be coarsely aligned to the skeleton segmented from CT i.e. global DoFs have to be removed. Taking the nodes of a skeleton surface representation as a 3D point set enables to resolve several global DoFs by aligning the principal axes. Subsequent extraction and analysis of a 3D curve that represents the skeleton allows to resolve all other possible global DoFs (refer to Baiker et al. [2] for details).

Fig. 2 Hierarchical anatomical tree for the skeleton. The connections depict relations between single bones or bone compounds such that a part on a lower level is initialized by the registration result on a high level (figure adapted from [2], ©2007 IEEE)



3.4 3D point matching algorithm for bones and the lung

Two problems arise in matching two point sets: correspondence and transformation. A method that solves for both simultaneously is the Iterative Closest Point (ICP) algorithm (Besl et al. [4]). While originally been developed for incorporating rigid transformations only, pilot experiments have shown that non-isotropic scaling can be integrated too as long as it is moderate. In this way, it is possible to account for inter-subject variability. The articulated registration of the skeleton is performed by stepwise traversing the hierarchical anatomical tree (Fig. 2) in a top-down manner. If the correspondence, i.e. the Euclidean distance between the point sets representing the atlas bone and the target bone surfaces respectively does not improve any more within an iteration step, the final transformation function is used to initialize the registration of a bone at a lower hierarchical level. Depending on the joint type, the DoFs for this node level are kinematically constrained. Traversing the tree, the overall correspondence improves gradually. The lung is registered in the same way as the bones including non-isotropic scaling, to account for shape variations due to breathing.

There is one exception to the registration scheme. Due to its high flexibility, the spine is not determined by registration but by binning the bone point set along its longitudinal axis and applying three dimensional region growing using the head-spine connection as seed point after the head is registered. The number of bone voxels per bin enables to determine the spine-pelvis connection and therefore to initialize the pelvis registration.

3.5 Determination of corresponding landmarks on the skin

Establishing correspondence between two point sets representing skin surfaces in the most general case is very difficult. This is because depending on how the animal is placed during acquisition, the shape can be almost rotation-symmetric to the longitudinal axis of the body, symmetric to the sagittal plane or even almost symmetric to the transverse plane. Furthermore, the shape of an animal can differ significantly e.g. if two mice are positioned in prone and supine position respectively.

The first issue can be resolved directly, if a registered skeleton is at hand. Then it is possible to determine a sparse set of corresponding landmarks on the surface of the animal skin by calculating the nodes with the smallest Euclidean distance from a set of bone landmarks. Having defined a sparse landmark set allows removing ambiguity during matching. The second issue can be resolved by taking surface information into account and relying on a local shape context to identify corresponding nodes. The idea of matching shapes based on global shape contexts has been introduced by Belongie et al. [3]. However to be able to take local shape deformation into account, the representation has to be bending invariant. Using geodesic instead of Euclidean distances in a local context allows rendering the representation rotation, translation and bending invariant. Scaling invariance can be achieved as well, by normalization of the geodesic distances.

Let $P = \{p_1 \dots p_n\}$ and $Q = \{q_1 \dots q_m\}$ be the nodes of two surfaces to be matched and h_i and h_j be the histograms of geodesic distances from nodes $p_i \in P$ and $q_j \in Q$ to other nodes of the surface. The method to find dense correspondence on the skin is as follows:

1. Initialize a list with corresponding skin nodes that are known (landmarks derived from skeleton registration)

Repeat steps 2-7 for all elements in the correspondence list:

2. Select all nodes on P and Q in the vicinity (i.e. between a maximum and minimum geodesic distance and with a minimum distance to already known correspondences of the next element on the list)
3. Calculate a local shape context based on geodesic distances towards K known corresponding nodes for the selected nodes in P and Q

Repeat steps 4-7 until no selected nodes are left:

4. Calculate the cost for matching two nodes p_i and q_j by calculating the distance of their histograms:

$$C_{ij} = \sum_{k=1}^K |h_i(k) - h_j(k)|$$

5. Find the best match as $\min(C_{ij})$

6. Add the found match to the list with correspondences
7. Remove the selected nodes whose geodesic distance to the newly found correspondence is too small

Due to the discretization of the surfaces, detected correspondences are generally not exact. To avoid the accumulation of small localization errors while progressing over the surface, determination of corresponding nodes should start at a coarse scale (in terms of inter-node distance) and, dependent on the amount of detail to be captured, continue at a smaller scale e.g. as proposed in Wang et al. [19]).

3.6 *Atlas organ interpolation*

Based on landmarks on bone, lung and skin, organs can be warped from the atlas domain to the subject domain. In its original form i.e. if used as an interpolant, the TPS does force landmarks in the source domain to fit landmarks in the target domain exactly. However in general small spatial errors may occur and this can cause local distortions of the mapping. A remedy is to allow small landmark localization errors and relax the constraint of interpolation towards approximation (thin-plate smoothing spline [18]).

4 Implementation details/Evaluation

For evaluation, 26 data volumes of 22 animals in prone and supine position and with arbitrary limb position were acquired with a Skyscan 1178 Micro-CT scanner (Kontich, Belgium). In a follow-up study, one mouse was scanned five times within five weeks.

The data was subsampled and smoothed, yielding a voxel size of $320 \times 320 \times 320 \mu\text{m}^3$. Subsequently, the skeleton was segmented through isodata thresholding (Ridler et al. [15]). The skin and the lung were extracted using object-background thresholding and 3D region growing respectively, with a seed point relative to selected points on the spine and sternum. Triangular meshes of the skeleton, the lung and the skin were determined from the volume labels, smoothed and subsampled. The atlas elements were represented as triangular meshes as well.

The registration of the bones and the lung was done in two iterations, using ICP together with Levenberg-Marquardt minimization to optimize correspondence with respect to the parameters (DoFs). The first iteration was used for coarse rigid alignment allowing 6 DoFs. The second iteration incorporated scaling as well (9 DoFs). The minimization scheme was terminated if the difference between subsequently estimated parameters was below a certain threshold. This was 0.01 degrees for the rotation, $3.2 \mu\text{m}$ for the translation and 0.001 for the scaling parameters. For determining correspondences on the skin, a triangulated surface

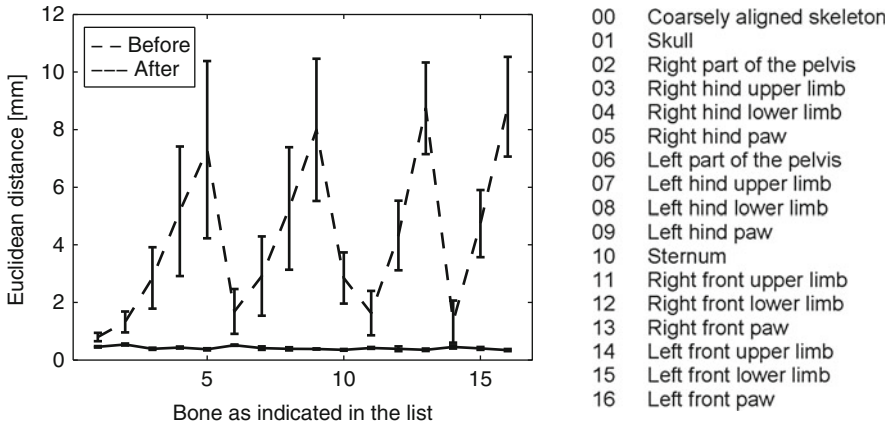


Fig. 3 Mean error of specific bones for 26 datasets ([2] ©2007 IEEE)

representation and a sparse set of 32 landmarks, derived from the joints, the spine and the skull were used. Geodesic distances were determined using the Fast Marching Algorithm [17] and the error criterion was based on the eight closest landmarks (i.e. $K=8$). The initial set was replenished by ≈ 120 landmarks from the skin all over the torso and 30 landmarks on the lung, yielding a total set of ≈ 182 corresponding nodes to constrain the TPS approximation.

During registration, the error decreased from an average of 2.93 ± 0.63 mm to 0.58 ± 0.04 mm for the skeleton and from an average of 1.76 ± 0.49 mm to 0.42 ± 0.068 mm for the lung, including all 26 cases. A detailed overview of the registration error for specific bones is given in Fig. 3. The mean Euclidean distance between atlas and subject skin decreased from 1.73 ± 0.4 mm to 0.34 ± 0.036 mm. Two examples of the skeleton registration and subsequent organ approximation are shown in Fig. 4. Results of the follow-up study are given in Fig. 5 and Table 2.

5 Discussion/Future Work

In this chapter, we presented an atlas-based method to segment bone and organs from CT data of mice that are placed arbitrarily during acquisition. The approach is based on registration of body parts that allow robust intensity-based segmentation from the data and approximation of other body parts that do not show sufficient contrast. The chosen ICP algorithm together with Levenberg-Marquardt minimization has shown to lead to a registration result of the skeleton and the lung with an average node distance within two times the resolution of the data, for all datasets. Due to the non-rigid shape difference between the atlas bones and the subject bones as well as between the atlas lung and the subject lung, the proposed method has limited accuracy. This could be faced by applying a subsequent non-

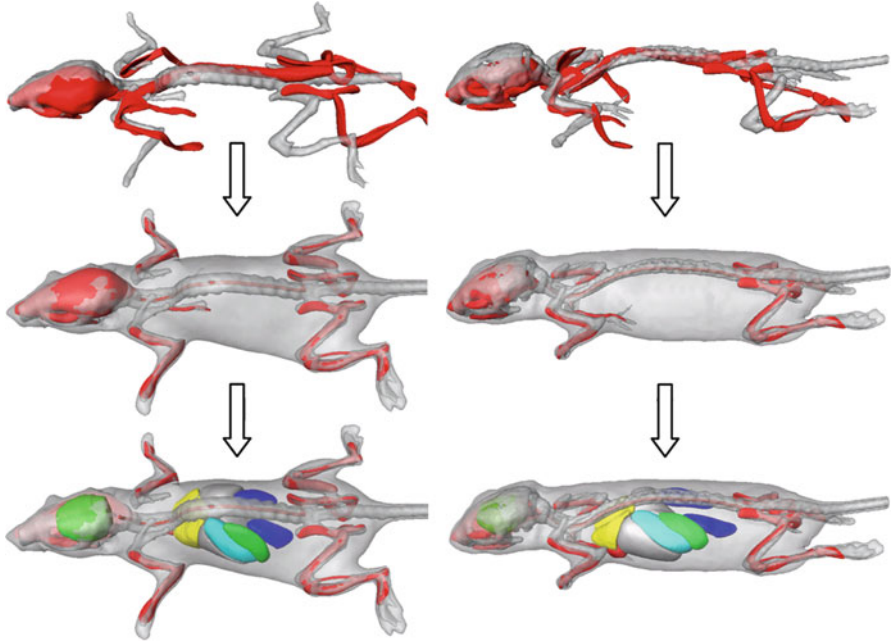


Fig. 4 Registration results between the atlas (red) and two different subjects (gray) after coarsely aligning the skeleton (top), after the articulated registration (middle) and after organ approximation (bottom) (adapted from [2], ©2007 IEEE)

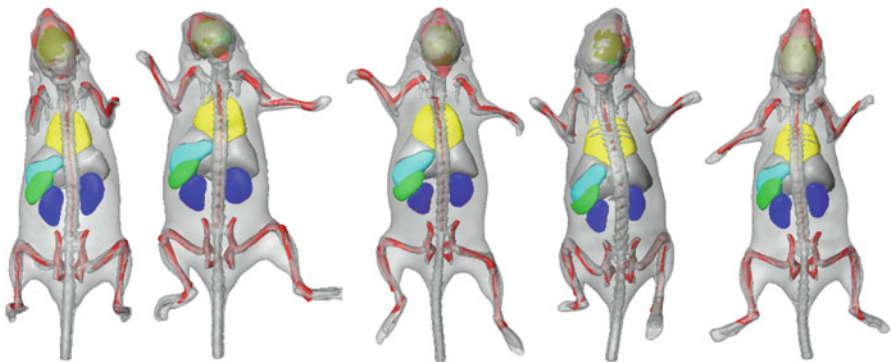


Fig. 5 Skeleton registration and organ approximation for a follow-up study over five weeks (t1-t5) with the subject in supine (t1-t3) and prone (t4-t5) position respectively

rigid registration to individual bones of interest or by repeating the registration at a higher data resolution, using the low-resolution result as initialization. Furthermore we show that starting from a sparse set of initial landmarks allows to robustly and non-ambiguously derive dense correspondence on the skin. This provides sufficient constraints for a TPS based approximation of the skin and major organs, leading to a

Table 2 Organ volumes (in mm³) of the original atlas and after mapping for a follow-up study of a mouse acquired at 5 subsequent time points (t1-t5)

	Brain	Heart	Lungs	Liver	Kidneys
Atlas	415,05	200,93	330,67	1779,87	257,24
Subject t1	449,09	196,25	380,40	1624,68	246,99
Subject t2	363,50	206,72	392,08	1902,58	267,40
Subject t3	389,99	201,12	363,25	1797,59	262,07
Subject t4	384,69	239,10	426,37	1822,90	194,83
Subject t5	448,88	242,03	425,04	1970,06	245,92
Mean (t1-t5)	407,23	217,04	397,43	1823,56	243,44
Std (t1-t5)	39,39	21,81	27,78	130,32	28,73
%	9,67	10,05	6,99	7,15	11,80

mean skin distance between the atlas and a subject within the data resolution range. The results of the organ approximation for the follow-up study (Fig. 5 and Table 2) reveal, that the method allows consistent localization and shape approximation of the brain, the heart, the lungs, the liver and the kidneys with a low variability in organ volume estimation (standard deviation <12%). The stomach and the spleen are shown for referencing, but no volume data is given due to the very large environmentally dependent variability of shape, location and volume. The same holds for the small and large intestine and the bladder, which have therefore not been included.

In conclusion, the presented method is applicable for referencing of internal processes in molecular imaging research or whole-body segmentation (e.g. to provide a heterogeneous tissue model for bioluminescence tomography). Furthermore, the approximation could serve to initialize a subsequent highly accurate registration of specific bones or organs, as long as the image data shows sufficient contrast. For CT data this might be realized using a contrast agent. In return, the approximation could be improved using organ registration results. We expect that this method generalizes well towards other rodents, provided that an anatomical atlas is available.

Acknowledgements The authors gratefully acknowledge Dr Paul Segars for providing the mouse atlas, Ivo Que for generating the CT datasets and Elke van de Castele from Skyscan for providing the Micro-CT scanner used in this research.

References

1. M. Baiker, J. Dijkstra, I. Que, C. W. G. M. Löwik, J. H. C. Reiber, and B. P. F. Lelieveldt. Organ approximation in micro-ct data with low soft tissue contrast using an articulated whole-body atlas. In *Proc. IEEE Intl. Symp. on Biomedical Imaging*, pages 1267–1270, 2008.
2. M. Baiker, J. Milles, A. M. Vossepoel, I. Que, E. L. Kaijzel, C. W. G. M. Löwik, J. H. C. Reiber, J. Dijkstra, and B. P. F. Lelieveldt. Fully automated whole-body registration in mice using an articulated skeleton atlas. In *Proc. IEEE Intl. Symp. on Biomedical Imaging*, pages 728–731, 2007.

3. S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.
4. P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
5. F. L. Bookstein. Principal warps - Thin-Plate Splines and the decomposition of deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
6. A. J. Chaudhari, A. A. Joshi, F. Darvas, and R. M. Leahy. A method for atlas-based volumetric registration with surface constraints for optical bioluminescence tomography in small animal imaging. In *Proc. SPIE Medical Imaging*, volume 6510, pages 651024–1–651024–10, 2007.
7. G. Cheers. *Anatomica*. Weltbild Verlag, 2004. ISBN-10: 3828920683.
8. B. Dogdas, D. Stout, A. F. Chatzioannou, and R. M. Leahy. Digimouse: a 3D whole body mouse atlas from CT and cryosection data. *Physics in Medicine and Biology*, 52(3):577–587, 2007.
9. A. du Bois d’Aische, M. De Craene, B. Macq, and S. K. Warfield. An articulated registration method. In *Proc. IEEE Intl. Conf. on Image Processing*, volume 1, pages 21–24, 2005.
10. N. Kovacevic, G. Hamarneh, and M. Henkelman. Anatomically guided registration of whole body mouse MR images. In *Proc. MICCAI*, pages 870–877, 2003.
11. X. Li, T. E. Yankeelov, T. E. Peterson, J. C. Gore, and B. M. Dawant. Constrained non-rigid registration for whole body image registration: method and validation. In *Proc. SPIE Medical Imaging*, volume 6512, pages 651202–1–651202–8, 2007.
12. J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
13. M. A. Martin-Fernandez, E. Munyoz-Moreno, M. Martin-Fernandez, and C. Alberola-Lopez. Articulated registration: elastic registration based on a wire model. In *Proc. SPIE Medical Imaging*, volume 5747, pages 182–191, 2005.
14. X. Papademetris, D. P. Dione, L. W. Dobrucki, L. Staib, and S. A. J. Articulated rigid registration for serial lower-limb mouse imaging. In *Proc. MICCAI*, pages 919–926, 2005.
15. T. W. Ridler and S. Calvard. Picture thresholding using an iterative selection method. *IEEE Trans. on Systems, Man and Cybernetics*, 8(8):630–632, 1978.
16. W. P. Segars, B. M. W. Tsui, E. C. Frey, G. A. Johnson, and S. S. Berr. Development of a 4D digital mouse phantom for molecular imaging research. *Molecular Imaging and Biology*, 6(3):149–159, 2004.
17. J. Sethian. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision and materials science*. Cambridge University Press, 1999. ISBN-13: 978-0521645577.
18. G. Wahba. *Spline models for observational data*. SIAM, Philadelphia, 1990. ISBN-13: 978-0898712445.
19. Y. M. Wang, B. S. Peterson, and L. H. Staib. 3D brain surface matching based on geodesics and local geometry. *Computer Vision and Image Understanding*, 89(2-3):252–271, 2003.
20. B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

Potential carotid atherosclerosis biomarkers based on ultrasound image analysis

S. Golemati, J. Stoitsis, and K.S. Nikita

Abstract It has been shown that computerized analysis of ultrasound images of the carotid artery may provide quantitative disease biomarkers and can potentially serve as a "second opinion" in the diagnosis of carotid atherosclerosis. Extending the findings of previous work on the subject, a set of methodologies are presented in this chapter, suitable for application on two-dimensional B-mode ultrasound images. More specifically, a Hough-Transform-based technique for automatic segmentation of the arterial wall allows the estimation of the intima-media thickness and the arterial distension waveform, two widely used determinants of arterial disease. Texture features extracted from Fourier-, wavelet-, and Gabor-filter-based methods can characterize symptomatic and asymptomatic atheromatous plaque. Finally, a methodology based on least-squares optical flow is proposed for the analysis and quantification of motion of the arterial wall.

The suggested methodologies allow the extraction of useful biomarkers for the study of (a) the physiology of the arterial wall and (b) the mechanisms of carotid atherosclerosis.

1 Introduction

The carotid arteries are responsible for supplying blood to the brain. Each common carotid artery divides into an external and an internal branch at the carotid bifurcation. The external carotids supply blood to the neck, pharynx, larynx, lower jaw and face, whereas the internal carotids enter the skull delivering blood to the brain. The presence of an atheromatous lesion, or plaque, in the carotid arteries, also

S. Golemati (✉)
School of Medicine, University of Athens, Greece

J. Stoitsis • K. Nikita
Electrical and Computer Engineering, Technical University of Athens, Greece

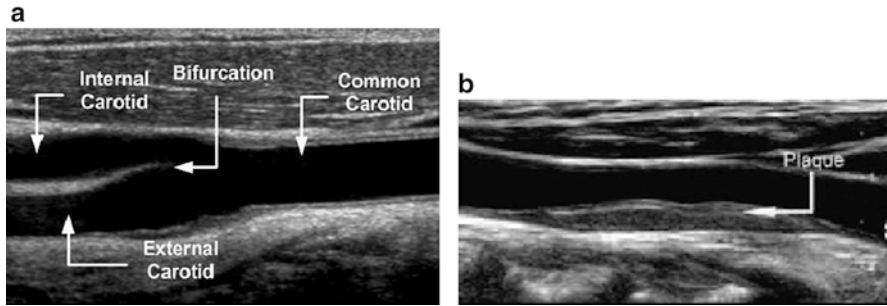


Fig. 1 Examples of B-mode ultrasound images of (a) a healthy (non-atherosclerotic) and (b) a diseased (atherosclerotic) carotid artery

known as carotid atherosclerosis, may disturb the normal circulatory supply to the brain. Carotid atherosclerosis may produce total occlusion of a specific arterial site or cause a thromboembolic event. In advanced stages of the disease, cerebrovascular symptoms, such as transient ischaemic attack, amaurosis fugax (temporal blinding) or stroke, may occur.

Ultrasound imaging of the carotid artery is the most widely used modality in the diagnosis of carotid atherosclerosis due to its noninvasiveness, non-ionizing nature and low cost. In particular, B (Brightness)-mode imaging, i.e. the reproduction of the amplitude of the reflected waves by their brightness, is commonly used to assess arterial wall morphology. B-mode images exhibit a granular appearance, called speckle pattern, which is caused by the constructive and destructive interference of the wavelets scattered by the tissue components. In B-mode ultrasound, blood reflects very little and the vessel lumen appears as a hypoechoic band. Figure 1 shows examples of B-mode ultrasound images of (a) a healthy (non-atherosclerotic) and (b) a diseased (atherosclerotic) carotid artery.

2 Quantitative assessment of carotid atherosclerosis

Currently, severity of carotid atherosclerosis and selection of patients to be considered for endarterectomy, i.e. surgical removal of plaque, are based (a) on the degree of stenosis caused by the plaque, in asymptomatic subjects, and (b) on both the degree of stenosis and previous occurrence of clinical symptoms, in symptomatic subjects. However, there is evidence that atheromatous plaques with relatively low stenosis degree may produce symptoms and that highly stenotic atherosclerotic plaques can remain asymptomatic. Because not all carotid plaques are necessarily harmful and because carotid endarterectomy carries a considerable risk for the patient, the crucial task of optimized selection of patients for operation may be greatly facilitated by the use of novel biomarkers. B-mode ultrasonic images, in combination with appropriate image processing methods for segmentation, texture

and motion analysis, may be used to extract useful diagnostic indices of the geometry, echogenicity and strain, respectively, of the carotid artery wall.

2.1 Prior Art

Previous work on computerized analysis of ultrasound images of the carotid arteries includes automatic segmentation of the arterial lumen, plaque texture analysis and tissue motion analysis. The use of deformable models [12], including snakes [3], allows automatic identification of the random-shaped carotid artery wall from static ultrasound images. In addition to this, the Hough Transform (HT) has been used to segment the arterial wall from sequences of images [15, 17]. In this case, the arterial distension waveform can be estimated facilitating the study of the dynamic arterial geometry. Plaque echogenicity, estimated from B-mode ultrasound images using texture analysis techniques, may be used to characterize atheromatous plaque and differentiate between symptomatic and asymptomatic cases. Plaque echogenicity has been analyzed using a number of statistical, model-based and Fourier-based methods [4, 16]. Motion of the arterial wall and atheromatous plaque has recently gained attention as a determinant of carotid atherosclerosis. It has been shown that plaque strain, expressed as relative motion between different parts of the plaque, may be related to plaque instability, i.e. to the risk for cerebrovascular complications, such as stroke [13]. Temporal sequences of ultrasound images can be used to estimate movement of the carotid artery wall by tracking the speckle patterns generated by the tissue [5, 7, 13].

The purpose of this chapter is to suggest a set of methodologies which extend the findings of previous work on computerized analysis of ultrasound images of the carotid artery in an attempt to identify quantitative indices of diagnostic value. Such indices may be useful for early and valid diagnosis of atherosclerosis as well as for the study of the physiology of the normal and diseased arterial wall.

3 Computerized analysis of ultrasound images of carotid artery

The methodologies described in this chapter are suitable for two-dimensional (2D) B-mode ultrasound imaging. This is the most widely used modality for the assessment of the carotid artery. The methodologies aim at (a) automatic segmentation of the arterial wall from longitudinal and transverse sections, (b) texture analysis of atheromatous plaque and (c) analysis and quantification of motion of the arterial wall. It is recommended that the methodologies be applied to normalized ultrasound images, according to widely accepted specifications [6], to minimize variability introduced by different equipment, operators and gain settings and facilitate imaged

tissue comparability. The techniques are designed to be applied to sequences of images, thus allowing the extraction of quantitative information at different phases of the cardiac cycle, e.g. systole or diastole. Obviously, it is possible to apply them to individual (static) images, if this is required by the clinical application.

4 Early disease biomarkers

Early disease stages may be assessed by interrogating the arterial wall on which focal lesions (plaques) may not yet have become obvious. The suggested HT technique allows the automatic extraction of straight lines and circles to approximate the wall-lumen boundary in longitudinal and transverse sections, respectively. HT can be used to detect parametric curves of the form $v(c, p_i) = 0$ in digital images, where c is the vector of coordinates, p the vector of parameters and $i = 1..n$, the number of parameters required to define the curve. HT transforms the image to an n -dimensional parametric space, called the accumulator array. Operating on a binary image of edge pixels, all possible curves $v(c, p_i) = 0$ through a pixel with vector coordinates c are transformed to a combination of parameters p_i , which then increment the corresponding cell of the accumulator array. The main steps of the methodology, which are described in detail in [8], include:

Reduction of image area This may be achieved by automatically isolating a rectangular area containing the vessel lumen. To this end, four points may be defined to delimit the area to be investigated. This is an important step because it minimizes the possibility to detect unwanted structures, which may be present biasing the representation of the arterial lumen and, thus, reduces the computational cost and the time required to perform the segmentation task.

Image pre-processing This step includes removal of high frequency noise using a symmetric Gaussian lowpass filter and morphological closing to suppress small ‘channels’ and ‘openings’ of the image.

Edge detection The image is first transformed into binary through the application of a global threshold and then a Sobel gradient operator may be applied.

Hough Transform Longitudinal sections are searched for lines defined as $z = x \cos \theta + y \sin \theta$, where z is the distance from the left upper corner of the image and θ is the angle with the x-axis. Transverse sections are searched for circles defined as $(x - a)^2 + (y - b)^2 = r^2$, where (a, b) are the coordinates of the center and r is the radius of the circle.

Selection of dominant lines/circle Two lines in longitudinal sections and one circle in transverse sections with the maximal values in the corresponding accumulator arrays are eventually selected, representing the boundaries of the wall-lumen interface.

Figure 2 shows examples of the application of the HT technique in longitudinal and transverse sections of ultrasound images of the carotid artery. Arterial diameters can be calculated through the application of the previously described methodology,

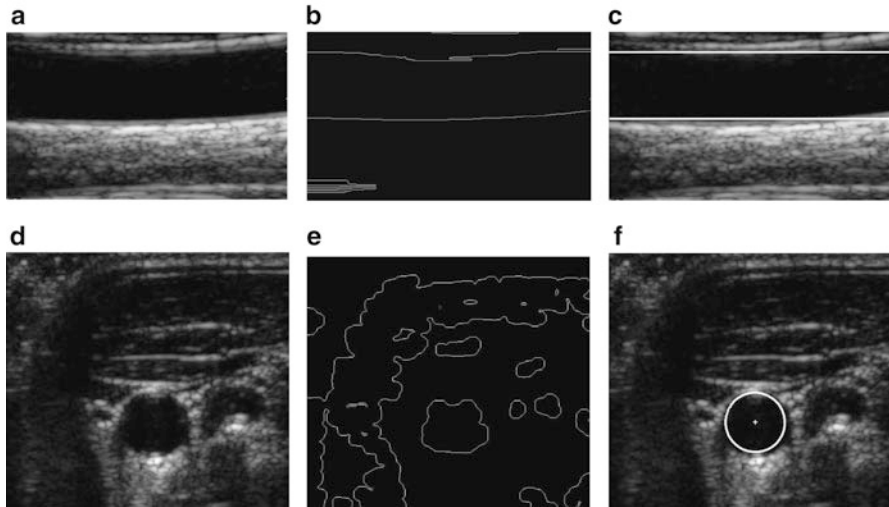


Fig. 2 Examples of the application of the HT technique in longitudinal and transverse sections of ultrasound images of the carotid artery. (a), (d) original images, (b), (e) images after morphological closing, thresholding and edge detection. (c), (f) HT technique result shown on original image

namely from the distance of the two lines in longitudinal sections and the circle diameter in transverse sections. The arterial distension waveform can then easily be estimated by recording the diameter values in all images of the sequences. Furthermore, in longitudinal sections, the HT methodology can be applied to the far wall alone, to extract two dominant lines corresponding to the boundaries of the wall, from which the intima-media thickness can be evaluated.

A methodology combining HT and active contours is also suggested, in an attempt to achieve a more accurate approximation of the arterial wall geometry in transverse sections. Departure from the strict geometrical shape indicated by HT is more evident in these sections. The methodology is based on the generation of a gradient vector flow field [18], an approach attempting to overcome conventional active contours constraints. The main steps include:

HT technique Application of the HT methodology described above allows the estimation of a circle, which is subsequently used for initializing the active contour.

Image pre-processing This step includes a number of tasks (calculation of gradient field, thresholding, morphological closing and smoothing, and gradient operator application) to estimate the image edge map.

Calculation of gradient vector flow field

Contour estimation Deformation of initial curve (circle) based on gradient vector flow field.

Figure 3 shows an example of the application of the combined HT-active-contours methodology in a transverse section of the carotid artery. As we can see, the methodology results in a random shaped boundary which follows more closely the

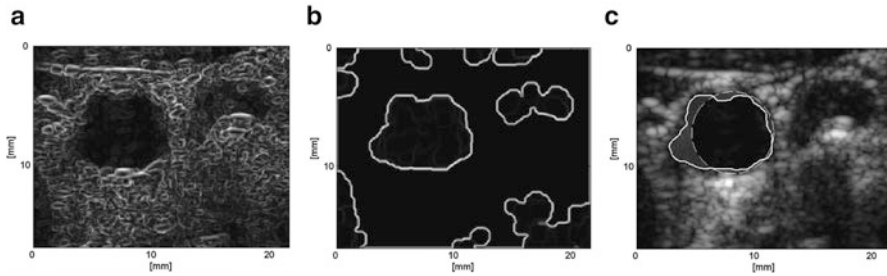


Fig. 3 Examples of the application of the combined HT-active-contours technique in a transverse section of the carotid artery. (a) illustration of gradient field of original image. (b) image after application of gradient operator on gradient field . (c) combined HT-active-contours methodology result shown on original image (solid line); dashed line indicates result (circle) of HT methodology

actual wall-lumen interface than the circle. However, widely used physiological indices, such as the arterial distension waveform, may be more easily estimated from the latter.

5 Assessment of advanced stages of disease

The severity of the carotid atherosclerotic plaque, an advanced disease stage, can be assessed through its echogenicity estimated by a number of texture analysis techniques. The use of transform-based texture analysis is suggested here, which has not been previously applied in ultrasound images of the carotid artery. The Fourier Transform (FT), the Wavelet Transform (WT) and Gabor filters allow the estimation of texture features capable of characterizing symptomatic and asymptomatic plaques.

The discrete 2D FT [9] can be used to quantify image texture in the frequency domain. The radial distribution of values in Fourier Power Spectrum (FPS) is sensitive to texture coarseness in an image, whereas their angular distribution is sensitive to the directionality of the texture. Power concentration in low spatial frequencies indicates coarse texture, while power concentration in high frequencies indicates fine texture. Texture with strong directional characteristics produces a power spectrum concentrated along lines perpendicular to the texture direction. A total of nine texture features can be extracted from the FPS, five corresponding to the radial and four to the angular distribution of the FPS.

The 2D WT can be used to analyze the frequency content of an image within different scales [1] and, thus, to extract information about the low and high frequencies of an image at different resolutions. The resulting wavelet coefficients are called the sub-images at different resolutions and consist of an approximation image and three detail images, namely the horizontal, vertical and diagonal detail images. Quantitative texture measures can be extracted from the wavelet coefficients. Each

plaque image can be decomposed up to five scales using an orthogonal, near symmetric and compactly supported mother wavelet, the symlet20 [2]. The choice of the mother wavelet is critical and, because the interrogated images exhibit rapid intensity fluctuations, a tight wavelet should be preferred. We suggest avoiding the use of the approximation sub-images for texture analysis because they represent a rough estimate of the original image and capture variations induced by lighting and illumination. A total of one hundred features can be extracted for each plaque image based on this methodology.

Gabor filters are a group of wavelets. They can be obtained by the dilation and rotation of a Gaussian function modulated by a complex sinusoid. A set of filtered images can be obtained by convolving Gabor filters with the original plaque image. Each of these filtered images represents texture information of the image at a specific scale and orientation. From each filtered image two features, namely the mean and standard deviation of the magnitude of the transformed coefficients, can be extracted [10]. These represent the energy content at different scales and orientations of the image. We suggest using five scales and four orientations resulting in a total of twenty Gabor-based texture features.

6 Motion analysis of normal and diseased arterial wall

To address the problem of arterial wall motion estimation, a methodology based on weighted least-squares optical flow (WLSOF) [11] is suggested. Compared to conventional optical flow, which was previously used in similar applications [13, 14], WLSOF allows smoothing out of large velocity differences between adjacent sites. WLSOF is used to recover velocity from the following equation, known as the gradient constraint equation relating velocity to the space and time derivatives at any one point of the image:

$$\mathbf{u} \cdot \nabla I(\mathbf{x}, t) + I_t(\mathbf{x}, t) = 0$$

where $\mathbf{u} = \frac{\Delta u}{\Delta t} = (u_x, u_y)$, and $\nabla I(\mathbf{x}, t)$ and I_t denote spatial and temporal partial derivatives, respectively, of the image I .

A common way to constrain velocity is to use gradient constraints from neighboring pixels, assuming that they share the same 2D velocity. In reality, there may be no single velocity value that simultaneously satisfies all pixels of the region, so the velocity that minimizes the constraint errors is found instead. The least-squares estimator that minimizes the squared errors is:

$$E(\mathbf{u}) = \sum_{\mathbf{x}} g(\mathbf{x}) \cdot [\mathbf{u} \cdot \nabla I(\mathbf{x}, t) + I_t(\mathbf{x}, t)]^2$$

where $g(\mathbf{x})$ is a Gaussian weighted function. It is used to enhance constraints in the center of the neighborhood, thus increasing their influence. The minimum of $E(\mathbf{u})$

can be found from the critical points, where derivatives with respect to \mathbf{u} are equal to zero:

$$\frac{\partial E(\mathbf{u})}{\partial u_x} = \sum_{\mathbf{x}} g(\mathbf{x}) \cdot [u_x \cdot I_x^2 + u_y \cdot I_x \cdot I_y + I_x \cdot I_t]$$

$$\frac{\partial E(\mathbf{u})}{\partial u_y} = \sum_{\mathbf{x}} g(\mathbf{x}) \cdot [u_y \cdot I_y^2 + u_x \cdot I_x \cdot I_y + I_y \cdot I_t]$$

The above equations can be written in matrix form and the resulting linear system can be solved using the Gaussian elimination method.

To compute spatial (I_x, I_y) and temporal (I_t) gradients, the images can first be smoothed using a Gaussian 7×7 kernel with standard deviation 0.8. Use of the Gaussian lowpass filter also allows removal of high frequency noise inherent in ultrasound images.

To estimate arterial wall motion, the previous method can be applied to appropriately selected image areas. These include pixels of the normal and/or diseased wall and exclude pixels of the lumen and the surrounding tissue. Specifically, pixels at a distance of 1.5 mm along the interface, i.e. in the longitudinal direction, and at a distance of 0.5 mm through the tissue, i.e. in the radial direction, can be selected. Pixel density is lower in the longitudinal direction because less relative motion is expected compared to the radial direction. Reduction of the image area to a set of individual pixels for further investigation reduces significantly the computational cost without compromising the related physiological information.

Figure 4 shows examples of velocity fields of the far arterial wall of a normal artery, as well as for a symptomatic and an asymptomatic case. In the same figure, examples of longitudinal and radial displacement waveforms for two points on the wall are also shown. The points distance is approximately 12.5 mm; in the case of the diseased (atherosclerotic) wall one point is on the plaque and the other on the adjacent normal part of the wall. As we can see, axial displacement, i.e. displacement along the arterial wall, exhibits a periodic pattern of frequency equal to almost double the frequency of the radial displacement. This finding agrees with recently reported results [5].

7 Experimental results and discussion

The methodologies presented here are useful for the quantitative assessment of carotid atherosclerosis. In combination with the experience of a specialized physician, they may improve the diagnostic power of ultrasound imaging. Their integration into clinical practice depends not only on their performance but also on how well the physician performs a task when the computer output is used as an aid.

More specifically, the suggested HT technique provides a simple, fast and accurate way to identify the arterial wall in longitudinal and transverse sections of the carotid artery and can be used in clinical practice to estimate indices of

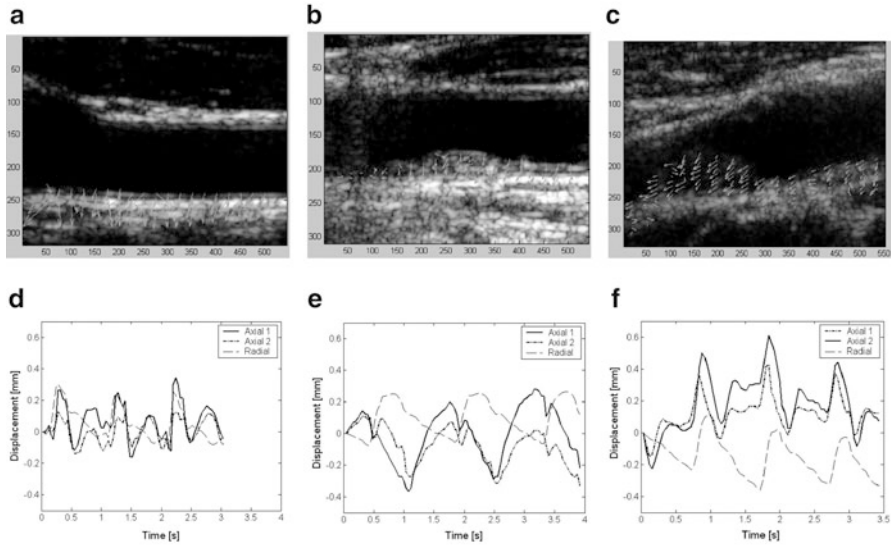


Fig. 4 Examples of velocity fields and displacement waveforms obtained by the application of the WLSOF methodology in a healthy (non-atherosclerotic) arterial wall (a, d), an asymptomatic plaque (b, e) and a symptomatic plaque (c, f). Illustrated vectors were enhanced by a factor of 10. Velocities correspond to beginning of systole

arterial wall physiology, such as the IMT and the ADW. In ten normal subjects, the specificity and accuracy of HT-based segmentation were on average higher than 0.96 for both sections, whereas the sensitivity was higher than 0.96 in longitudinal and higher than 0.82 in transverse sections. The corresponding validation parameters for IMT estimation were generally higher than 0.90. The HT technique was also applied to 4 subjects with atherosclerosis, in which sensitivity, specificity and accuracy were comparable to those of normal subjects; the low values of sensitivity in transverse sections may reflect departure from the circular model due to the presence of plaque. For these cases, the combined HT-active-contours technique was found to increase the sensitivity values.

Texture features using the three transform-based methods described previously were extracted from a limited number of symptomatic (ten plaques) and asymptomatic (nine plaques) cases. Differences between the two case types were estimated using bootstrapping. Both the WT- and the Gabor-filter-based methodologies resulted in significantly different features, which characterized texture at low resolutions and in the horizontal direction. Features at low resolutions are indicative of fine texture; finer texture was found in symptomatic compared to asymptomatic plaques. Horizontal texture patterns are an interesting finding, especially when one combines this information with arterial wall biomechanics. Mechanical stresses due to blood blow may be responsible for such texture patterns; compared to blood pressure, the other main cause of stress on the arterial wall, the effect of blood flow is more pronounced around a plaque. The discriminative ability of the transform-based

texture features was found superior to that of the gray-scale median, a widely used texture parameter of carotid plaque, for the small population that was interrogated, emphasizing the need for using advanced techniques to efficiently characterize atheromatous plaque.

Reliable estimation of arterial wall motion is a challenging task and is believed to provide a powerful tool in the study of the physiology and biomechanics of atheromatous plaque. The strain experienced by the arterial wall is a crucial biomarker of carotid atherosclerosis and can be assessed through motion analysis. In combination with information of the exerted stresses, it can prove useful for the study of the mechanical behavior of cardiovascular tissue.

8 Conclusion

The methodologies presented in this chapter are expected to provide powerful tools in the diagnosis of carotid atherosclerosis because they can assist interpretation of ultrasound images. Individual techniques facilitate the diagnostic tasks of vessel wall identification, plaque characterization and strain estimation of normal and diseased arterial wall.

References

1. S. Arivazhagan and L. Ganesan. Texture classification using wavelet transform. *Pattern Recogn Lett*, 24:1513–1521, 2003.
2. M. H. Bharati, J. J. Liu, and M. J. F. Image texture analysis: methods and comparisons. *Chemometr Intell Lab*, 72:57–71, 2004.
3. D. C. Cheng, A. Schmidt-Trucksäss, K. S. Cheng, and H. Burkhardt. Using snakes to detect the intimal and adventitial layers of the common carotid artery wall in sonographic images. *Comput Meth Prog Bio*, 67:27–37, 2002.
4. C. I. Christodoulou, C. S. Pattichis, M. Pantziaris, and A. Nicolaides. Texture-based classification of atherosclerotic carotid plaques. *IEEE Trans Med Imag*, 22:902–912, 2003.
5. M. Cinthio, A. R. Ahlgren, J. Bergkvist, T. Jansson, H. W. Persson, and K. Lindström. Longitudinal movements and resulting shear strain of the arterial wall. *Am J Physiology - Heart and Circulatory Physiology*, 291:H394–H402, 2006.
6. A. Elatrozy, T. and Nicolaides, T. Tegos, A. Zarka, M. Griffin, and M. Sabetai. The effect of b-mode ultrasonic image standardization of the echodensity of symptomatic and asymptomatic carotid bifurcation plaque. *Int Angiol*, 7:179–186, 1998.
7. S. Golemati, A. Sassano, M. J. Lever, A. A. Bharath, S. Dhanjil, and A. N. Nicolaides. Motion analysis of carotid atherosclerotic plaque from b-mode ultrasound. *Ultrasound Med Biol*, 29:387–399, 2003.

8. S. Golemati, J. Stoitsis, E. G. Sifakis, T. Balkizas, and K. S. Nikita. Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery. *Ultrasound Med Biol*, 33:1918–1932, 2007.
9. D. C. He and L. Wang. Texture features based on texture spectrum. *Pattern Recogn*, 24:391–399, 1991.
10. A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recogn*, 24:1167–1186, 1991.
11. B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereoscopic vision. In *Int Conf Artificial Intelligence*, pages 674–679, 1981.
12. F. Mao, J. Gill, D. Downey, and A. Fenster. Segmentation of carotid artery in ultrasound images: Method development and evaluation technique. *Med Phys*, 27:1–10, 2000.
13. S. Meairs and M. Hennerici. Four-dimensional ultrasonographic characterization of plaque surface motion in patients with symptomatic and asymptomatic carotid artery stenosis. *Stroke*, 30:1807–1813, 1999.
14. M. Mokhtari-Dizajl, M. Montazeri and H. Saberi. Differentiation of mild and severe stenosis with motion estimation in ultrasound images. *Ultrasound Med Biol*, 10:1493–1498, 2006.
15. J. M. Nash, J. N. Carter, and M. S. Nixon. Dynamic feature extraction via the velocity hough transform. *Pattern Recogn Lett*, 18:1035–1047, 1997.
16. J. E. Wilhjelm, M. L. M. Grønholdt, B. Wiebe, S. K. Jespersen, L. K. Hansen, and H. Sillesen. Quantitative analysis of ultrasound b-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination. *IEEE Trans Med Imag*, 17:910–922, 1998.
17. S. M. Wu, Y. W. Shau, F. C. Chong, and F. J. Hsieh. Non-invasive assessment of arterial distension waveforms using gradient-based hough transform and power doppler ultrasound imaging. *Med Biol Eng Comput*, 39:627–632, 2001.
18. C. Xu and J. L. Prince. Snakes, shapes and gradient vector flow. *IEEE Trans Imag Process*, 7:359–369, 1998.