Novel Methods for Multidimensional Image Segmentation

A Thesis
Presented to
The Academic Faculty

by

Eric Pichon

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

School of Electrical and Computer Engineering
Georgia Institute of Technology
December, 2005

Novel Methods for Multidimensional Image Segmentation

Approved by:

Dr. Allen Tannenbaum, Advisor
School of Electrical and Computer Engineering
Georgia Instituite of Technology

Dr. Magnus Egerstedt
School of Electrical and Computer Engineering
Georgia Instituite of Technology

Dr. Anthony Yezzi
School of Electrical and Computer Engineering
Georgia Instituite of Technology

Dr. Chin-Hui Lee
School of Electrical and Computer Engineering
Georgia Instituite of Technology

Dr. Laurence Jacobs
School of Civil and Environmental Engineering
Georgia Instituite of Technology

Date Approved: October 5th, 2005

ACKNOWLEDGEMENTS

This thesis is the visible, final result of almost four years of work at the Georgia Insitute of Technology. During that time I was fortunate to meet and exchange with many extraordinary people. The document you are about to read would not have been possible without them. In particular, I am forever indebted to:

- Allen Tannenbaum, my Jedi-master. You taught me much more than computer vision...
- Marc Niethammer (your help was very precious, your friendship even more), Patricio Vela (you kept my body fit and my heart joyfull), Eli Herskovitz (your humorous and poetic view of things enriched my days me from the Appalachian forests to the jungles of Guatemala), Delphine Nain (I benefited much from your energy and our endless discussions) and all the other members of our group.
- my Atlanta friends, Bob McElroy, Philippe Karam, Carole Montarou, Giovanni Iachello and all the others. You counted a lot for me during these years and you will certainly count in the years to come.
- Céline Cravatte
- Désirée Johnson
- my grand-parents, my parents, my brothers

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Artificial vision is the problem of creating systems capable of extracting information from digital images in a way that would mimic or even surpass the human eye-brain system. A fundamental sub-problem of artificial vision is *image segmentation*, the problem of detecting a structure from a digital image. Examples of segmentation problems include the detection of a road from an aerial photograph or the determination of the boundaries of the brain's ventricles from medical imagery. The extraction of structures allows for subsequent higher-level cognitive tasks. One of them is *shape comparison*. For example, if the brain ventricles of a patient are segmented, can their shapes be used for diagnosis? That is to say, do the shapes of the extracted ventricles resemble more those of healthy patients or those of patients suffering from schizophrenia?

Because the human eye-brain system is still not very well understood, it cannot be directly simulated. Therefore, computer programs (and the underpinning mathematical theories) are proposed that are useful for analyzing visual information even though they might bear no direct relation to biological vision.

This thesis deals with the problem of image segmentation and shape comparison in the mathematical framework of *partial differential equations*. The contribution of this thesis is threefold:

1. A technique for the segmentation of regions is proposed. A cost functional is defined for regions based on a non-parametric functional of the distribution of image intensities inside the region. This cost is constructed to favor regions that are homogeneous. Regions that are optimal with respect to that cost can be determined with limited user interaction.

2. The use of direction information is introduced for the segmentation of open curves and closed surfaces. A cost functional is defined for structures (curves or surfaces)

by integrating a local, direction-dependent pattern detector along the structure. Optimal structures, corresponding to the best match with the pattern detector, can be determined using efficient algorithms. The problem of learning the pattern detector from some expert's segmentation (and thus capturing external knowledge) is also considered.

3. A technique for shape comparison based on the Laplace equation is proposed. Given two surfaces, one-to-one correspondences are determined that allow for the characterization of local and global similarity measures. The local differences among shapes (resulting for example from a segmentation step) can be visualized for qualitative evaluation by a human expert. It can also be used for classifying shapes into, for example, normal and pathological classes. Finally, another important application is the quantification of the difference between the output of two different segmentation algorithms or one algorithm and a human expert.

CHAPTER 1

INTRODUCTION

1.1 Artificial Vision

Artificial intelligence (AI) was initiated as a field in the 1950's with the ambitious (and so-far unrealized) goal of creating systems with human-like "intelligence" (the exact definition of intelligence is still very problematic). Whereas classical AI had been mostly concerned with symbolic representation and reasoning, new subfields were created as researchers embraced the complexity of the goal and realized the importance of perception and lower-level cognitive tasks that deal with sub-symbolic information.

In particular, *artificial vision* [1, 2, 3, 4] emerged in the 1970's with the more limited goal to mimic human vision with man-made systems (in practice, computers).

Vision is such a basic aspect of human cognition that it may superficially appear somewhat trivial, but, after decades of research, the scientific understanding of biological vision remains extremely fragmentary. Even though the general problem of artificial vision is far from being solved, many techniques have been devised and successfully solve specific problems in fields such as Earth observation, industrial automation, video surveillance, robotics and different military applications [4].

Medical imaging, which emerged with the discovery of X-rays in the late 19th century offers many examples of the use of artificial vision [5]. Medical imaging needs highly trained technicians and clinicians to determine the details of image acquisition (e.g. choice of modality, of patient position, of an optional contrast agent, etc.), as well as to analyze the results. The dramatic increase in availability, diversity, and resolution of medical imaging devices over the last 50 years threatens to overwhelm these human experts. Computer vision techniques have therefore become indispensable. Artificial systems must be designed to analyze medical datasets either in a partially or even a fully automatic manner.

The human eye-brain system evolved over tens of millions of years and at this point no artificial system is as versatile and powerful for everyday tasks. In the same way that a chess-playing program is not directly modeled after a human player, many mathematical techniques are employed in artificial vision that do not pretend to simulate biological vision. Artificial vision systems will not be set within the natural limits of human perception. For example, human vision is inherently two dimensional (stereoscopic vision does not allow us to see *inside* objects. It is sometimes described as "2.1 *dimensional perception.*"). To accommodate this limitation, radiologists must resort to visualizing only two-dimensional planar slices of three-dimensional medical images. An artificial system is free of that limitation and can "see" the image in its entirety. Other advantages are that artificial systems can work on very large image datasets, are fast, do not suffer from fatigue, and produce repeatable results. Because artificial vision system designers have so far been unsuccessful in incorporating high level understanding of real-life applications, artificial systems typically complement rather than replace human experts.

1.2 Image Segmentation

When looking at an image, a human observer cannot help seeing structures which can often be identified with objects. Digital images, however, as the raw retinal input of local intensities, are the sum of local color information. Information on global structures is not directly present and has to be extracted from local information. Segmentation is the process of creating a structured visual representation from an unstructured one. Visual information is thus dramatically condensed, which makes further processing and higher level cognitive tasks tractable.

The problem was first studied in the 1920s by psychologists of the Gestalt[1] school [6] and later by psychophysicists [7, 8].

Segmentation has been defined as the partitioning of an image into homogeneous regions that are semantically meaningful, i.e., that correspond to objects we could name. A typical

[1]Segmentation was then called "perceptual grouping."

example is partitioning some medical imagery of the brain into the white and gray matters and tumor.

The intrusion of linguistics, through the adverb "semantically", into an otherwise sub-symbolic problem is rather problematic. In this work we will define segmentation as *the determination of structures from a digital image*. We will consider that such structures can be regions ; the proposed definition therefore encompasses the previous one while removing the coupling with linguistics. Another important class of structures is open space curves. Some physical objects and their projections on an image are more naturally described by curves than by regions. Rivers, roads, blood vessels or neural tracts are all examples of structures that can appear in different types of imagery and be represented by curves.

Segmentation can be seen as an extreme form of smoothing or information reduction. Segmentation is also related to registration in the sense that if an atlas[2] can be perfectly registered to a dataset at hand then the registered atlas labels are the segmentation (see Section 2.1 for a definition of these terms). Segmentation is useful for visualization,[3] allows for quantitative shape analysis and provides indispensable anatomical context for virtually any subsequent automatic analysis.

It is important to understand that segmentation is not concerned with actually determining what the structures are. In that sense it is a lower level problem than object recognition.

Finally, a very important and often overlooked aspect of segmentation is *validation*. How much does the output of a given segmentation process differ from reality? Given two different segmentation processes, which one is more robust or performant? Very often the ground truth, i.e., the real boundaries of the object to be detected, is not readily accessible. That is usually the reason why imaging is being employed instead of direct measurement. In that case, the designer of a segmentation algorithm can only compare his system against human experts. If an artificial system produces exactly the same result as a human expert,

[2] An atlas is an image from a typical patient that has been manually segmented by some human expert.

[3] As discussed previously, human vision is inherently two-dimensional. In order to "see" an organ we therefore have to resort to visualizing its outside boundary. Determining this boundary is equivalent to segmenting the organ.

certainly it will be deemed satisfactory. This is along the lines of the famous Turing test [9] of artificial intelligence. The problem, however, is that human experts typically do not agree exactly on the same segmentation. In fact, given the same image, the same expert may not produce exactly the same segmentation. These shortcomings of human segmentation are one of the reasons why artificial vision is being pursued. At the same time, they make its validation difficult. Segmentation, being the mimicking of human perception, is highly subjective. Finally another difficulty is how to quantify the discrepancy between two segmentation results. The problem can be defined as the comparison of two similar shapes. Different metrics are possible (volume overlap, maximum minimum distance, etc.) and the choice will affect the results of the validation process.

1.3 Organization of this Thesis

This thesis is divided into four main chapters that are meant to be relatively self-contained. If part of the work has been made publicly available, references to the publications are given in the introduction of the corresponding chapter.

In Chapter 2, we review the application of partial differential equations (PDEs) to computer vision and image processing. As explained previously, while the ultimate goal of computer vision is mimicking and eventually surpassing the human visual cognitive abilities, computer vision systems are typically not built to simulate exactly each and every aspect of the human eye-brain system. This is due to the fact that the human eye-brain system is not understood very well. Moreover, to the extent that it is understood, its information processing architecture does not lead itself very easily to simulation on a computer. Therefore, many mathematical approaches have been investigated for applications in artificial vision that do not pretend to simulate accurately the inner functioning of the human eye-brain system but that reproduce some of its processing at a much coarser level.[4] Methods based on PDEs have proved extremely powerful and will constitute the mathematical basis of this work.

[4]To give another example, consider planes and birds. While they both fly, the design of airplanes is rather remotely inspired by birds.

In Chapter 3, we present a new algorithm for the segmentation of regions. The algorithm is versatile, fast, relatively simple to implement, and semi-automatic. It is based on maximizing a global energy defined from a learned non-parametric estimation of the statistics of the region to be segmented. Implementation details are discussed and source code is freely available as part of the 3D Slicer project. In addition, a new unified set of validation metrics is proposed. Results on artificial and real magnetic resonance imagery datasets show that the algorithm performs well on large brain structures both in terms of accuracy and robustness to noise.

In Chapter 4, a technique is proposed for the segmentation of open curves and closed hypersurfaces. Structures are obtained as the minimizers of some functional that involves a local cost dependent on structure direction We show that these local costs can be based on a model of a class of visual neurons. A general connection to pattern detection is also proposed. The framework is also applied to the detection of neural fibers in diffusion magnetic resonance imagery. Minimization techniques are proposed along with implementation details and results on synthetic as well as real images.

In Chapter 5, a principled approach for shape (hypersurface) comparison is proposed. Given two surfaces, one-to-one correspondences are determined based on the Laplace equation. The distance between corresponding points is then used to define both global and local dissimilarity statistics between the surfaces. This technique provides a powerful method to compare shapes both locally and globally. It can also be used to validate quantitatively the result of a region segmentation algorithm by comparing it to some known (or expert-defined) ground truth or for more general shape analysis tasks.

CHAPTER 2

PARTIAL DIFFERENTIAL EQUATIONS IN ARTIFICIAL VISION

In this chapter, we review the use of partial differential equations (PDEs) for computer vision. This chapter is based on [10].

Many mathematical approaches have been investigated for applications in artificial vision (e.g., fractals and self-similarity, wavelets, pattern theory, stochastic point process, random graph theory; see [11]), In Section 2.1 we outline how computer vision problems can be set in a PDE framework.

We will then sequentially consider the following sub-problems:

- *Smoothing*, the problem of simplifying the image while retaining important information (Section 2.2).
- *Registration*, the problem of fusing images of the same region acquired from different modalities or putting in correspondence images of one patient at different times or of different patients (Section 2.3).
- *Segmentation*, the problem of identifying structures such as regions or curves from an image (Section 2.4).

The processing and analysis of medical images offers scores of interesting and challenging examples for all these problems and will be used to illustrate this chapter. The techniques, however, are quite general and could be used for many other applications.

2.1 General Mathematical Framework

Here, we briefly outline the major concepts involved in using PDEs for image processing and computer vision. More detailed reviews include [12] and [13].

As explained in detail in [14], one can think of an image as a map $I : \mathfrak{D} \to \mathfrak{C}$, i.e., to any point $\mathbf{p}$ in the domain $\mathfrak{D}$, I associates a "color" $I(\mathbf{p})$ in a color space $\mathfrak{C}$. For ease of presentation, we will mainly restrict ourselves to the case of a two-dimensional grayscale image, which we can think of as a function from a domain $\mathfrak{D} = [0,1] \times [0,1] \subset \mathbb{R}^2$ to the unit interval $\mathfrak{C} = [0,1]$.

The algorithms typically involve solving the initial value problem for some PDE for a given amount of time. The solution to this PDE can be either the image itself at different stages of modification, or some other object (such as a closed curve delineating object boundaries), whose evolution is driven by the image.

For example, introducing an artificial time t, the image can be deformed according to

$$\frac{\partial I}{\partial t} = \mathcal{F}[I], \tag{1}$$

where $I(\mathbf{p}, t) : \mathfrak{D} \times [0, T) \to \mathfrak{C}$ is the evolving image, $\mathcal{F}$ is an operator that characterizes the given algorithm, and the initial condition is the input image I_0. The processed image is the solution $I(\mathbf{p}, t)$ of the differential equation at time t. The operator $\mathcal{F}$ usually is a differential operator, although its dependence on I may also be nonlocal.

Similarly, one can evolve a closed curve $\Gamma \subset \mathfrak{D}$ representing the boundaries of some planar shape (Γ need not be connected and could have several components). In this case, the operator $\mathcal{F}$ specifies the normal velocity of the curve that it deforms. In many cases this normal velocity is a function of the curvature κ of Γ, and of the image I evaluated on Γ. A flow of the form

$$\frac{\partial \Gamma}{\partial t} = \mathcal{F}(I, \kappa)\mathbf{N} \tag{2}$$

is obtained, where $\mathbf{N}$ is the unit normal to the curve Γ.

Very often, the deformation is obtained as the steepest descent for some functional. In what follows, we will use the terms "functional", "energy" or "cost" interchangeably.

For example, the energy

$$\mathcal{E}(I) = \frac{1}{2} \int \|\nabla I\|^2 \ \mathrm{d}x \ \mathrm{d}y \tag{3}$$

and its associated steepest descent, the heat equation,

$$\frac{\partial I}{\partial t} = \Delta I \tag{4}$$

correspond to the classical Gaussian smoothing (see Section 2.2.1).

The use of PDEs allows for the modeling of the crucial but poorly understood interactions between top-down and bottom-up vision.[1] In a variational framework, for example, an energy $\mathcal{E}$ is defined globally, while the corresponding operator $\mathcal{F}$ will influence the image locally. Algorithms defined in terms of PDEs treat images as continuous rather than discrete objects. This simplifies the formalism, which becomes grid independent. On the other hand, models based on nonlinear PDEs may be be much harder to analyze and implement rigorously.

Digital images typically suffer from one or more of the following imperfections:

- low resolution (in the spatial and spectral domains);
- high level of noise;
- low contrast;
- geometric deformations;
- presence of imaging artifacts.

These imperfections can be inherent to the imaging modality or the result of a deliberate trade-off during acquisition. In medical imaging, for example, X-rays offer low contrast for soft tissues, ultrasound produces very noisy images, and metallic implants cause imaging artifacts in MRI. Finer spatial sampling may be obtained through a longer acquisition time. However, that would also increase the probability of patient movement and therefore blurring. In this work, we are only interested in the processing and analysis of images and we are not be concerned with the challenging problem of designing optimal procedures for

[1] "Top-down" and "bottom-up" are loosely defined terms from computer science, computation and neuroscience. The bottom-up approach can be characterized as searching for a general solution to a specific problem (e.g. obstacle avoidance), without using any specific assumptions. The top-down approach refers to trying to find a specific solution to a general problem, such as structure from motion, using specific assumptions (e.g., rigidity or smoothness).

their acquisition. The ideal computer vision application would be fast, robust with regard to image imperfections, simple to use, and as automatic as possible.

Note that for ease of presentation, the techniques we present below are applied mostly to two-dimensional grayscale images. The majority of them, however, can be extended to higher dimensions, i.e., vector-valued volumetric images $I : \mathbb{R}^n \to \mathbb{R}^p$.

2.2 Image Smoothing

Smoothing is the action of simplifying an image while preserving important information. The goal is to reduce noise or useless details without introducing too much distortion so as to simplify subsequent analysis.

It was realized that the process of smoothing is closely related to that of pyramiding[2] which led to the notion of *scale space*. This was introduced by Witkin [15] and formalized in such works as [16, 17]. Basically, a *scale space* is a family of images $\{S^t \mid t \geq 0\}$ in which $S^0 = I$ is the original image and $S^t,\ t > 0$ represent the different levels of smoothing for some parameter t. Larger values of t correspond to more smoothing.

In [18], Alvarez et al. propose an axiomatic description of scale space. These axioms, which describe many of the methods in use, require that the process T^t, which computes the image $S^t = T^t[I]$ from I, have the following properties:

Causality / Semi-group $T^0[I] \equiv I$ and for all $t, s \geq 0$, $T^t\left[T^s[I]\right] = T^{t+s}[I]$. (In particular, if the image S^t has been computed, all further smoothed images $\{S^s \mid s \geq t\}$ can be computed from S^t, and the original image is no longer needed.)

Generator The family $S^t = T^t[I]$ is the solution of an initial value problem $\partial_t S^t = A[S^t]$, in which A is a nonlinear elliptic differential operator of second order.

Comparison Principle If $S_1^0(\mathbf{p}) \leq S_2^0(\mathbf{p})$ for all $\mathbf{p} \in \mathfrak{D}$, then $T^t[S_1^0] \leq T^t[S_2^0]$ pointwise on $\mathfrak{D}$. This property is closely related to the Maximum Principle for parabolic PDEs.

[2] An initial image is recursively reduced in size by a constant factor. If all images are stacked on top of each other a "pyramid" is obtained. Images near the top contain very few details.

Euclidean invariance The generator A and the maps T^t commute with Euclidean transformations[3] (i.e., compositions of translations and rotations) acting on the image S^0.

The requirement that the generator A of the semi-group be an elliptic differential operator may seem strong and even arbitrary at first, but it is argued in [18] that the semi-group property, the Comparison Principle, and the requirement that A act locally make this axiom quite natural. One should note that already in [16], it is shown that in the linear case a scale space must be defined by the linear heat equation. (See our discussion below.)

2.2.1 Naive, Linear Smoothing

If a given image $I : \mathfrak{D} \to \mathfrak{C}$ contains a certain amount of noise, then the most straightforward way of removing this noise is to approximate I by a mollified function S, i.e. one replaces the image function I by the convolution $S^\sigma = G_\sigma * I$, where

$$G_\sigma(\mathbf{p}) = \left(2\pi\sigma^2\right)^{-n/2} e^{-\|\mathbf{p}\|^2/2\sigma^2} \tag{5}$$

is a Gaussian kernel of covariance the diagonal matrix $\sigma^2\mathrm{Id}$. This mollification will smear out fluctuations in the image on scales of order σ and smaller. This technique had been in use for quite a while before it was realized[4] by Koenderink [16] that the function $S^{2\sigma} = G_{2\sigma} * I$ satisfies the linear diffusion equation

$$\frac{\partial S^t}{\partial t} = \Delta S^t, \quad S^0 = I. \tag{6}$$

Thus, to smooth the image I, the diffusion equation (6) is solved with initial data $S^0 = I$. The solution S^t at time t is then the smoothed image.

The method of smoothing images by solving the heat equation has the advantage of simplicity. There are several effective ways to compute the solution S^t from a given initial image $S^0 = I$, e.g., using the fast Fourier transform. Linear Gaussian smoothing is Euclidean invariant and satisfies the Comparison Principle. However, in practice one finds that

[3]Because an image is contained in a finite region $\mathfrak{D}$, the boundary conditions which must be imposed to make the initial value problem $\partial_t S^t = A[S^t]$ well-posed will generally keep the T^t from obeying Euclidean invariance even if the generator A does so.

[4]This was of course common knowledge among mathematicians and physicists for at least a century. The fact that this was not immediately noticed shows how disjoint the imaging/engineering and mathematics communities were.

Gaussian smoothing blurs edges. For example, if the initial image $S^0 = I$ is the characteristic function of some smoothly bounded set $\Omega \subset \mathfrak{D}$, so that it represents a black and white image with no gray regions, then for all but very small $t > 0$ the image S^t will resemble the original image in which the sharp boundary $\partial\Omega$ has been replaced with a fuzzy region of varying shades of gray. (See Section 2.4.1 for a discussion on edges in computer vision.)

Figure 1(a) is a typical MRI brain image. Specular noise is usually present in such images, so edge-preserving noise removal is essential. The result of Gaussian smoothing implemented via the linear heat equation is shown on Figure 1(b). The edges are visibly smeared. Note that even though two-dimensional slices of the three-dimensional image are shown to accommodate human perception, the processing was actually performed in three dimensions, and not independently on each two-dimensional slice.

We now discuss several methods that have been proposed to avoid this edge blurring effect while smoothing images.

2.2.2 Anisotropic Smoothing

Perona and Malik [19] replaced the linear heat equation with the nonlinear diffusion equation

$$\frac{\partial S}{\partial t} = \operatorname{div}\{g(|\nabla S|)\nabla S\} = \sum_{i,j} a_{ij}(\nabla S)\nabla^2_{ij} S \tag{7}$$

with

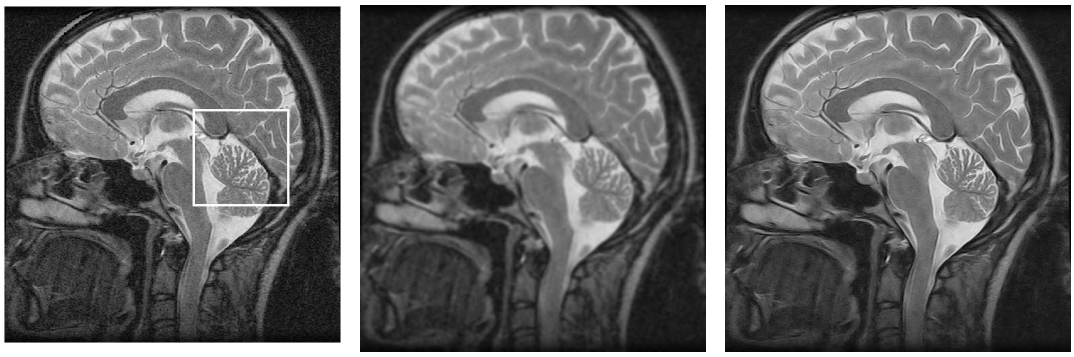
$$a_{ij}(\nabla S) = g(|\nabla S|)\delta_{ij} + \frac{g'(|\nabla S|)}{|\nabla S|}\nabla_i S \nabla_j S.$$

Here g is a nonnegative function for which $\lim_{p\to\infty} g(p) = 0$. The idea is to slow down the diffusion near edges, where the gradient $|\nabla S|$ is large. (See Sections 2.4.1 and 2.4.2 for a description of edge detection techniques.)

The matrix a_{ij} of diffusion coefficients has two eigenvalues, one $\lambda^{\|}$ for the eigenvector ∇S, and one $\lambda^{\perp}$ for all directions perpendicular to ∇S. They are

$$\lambda^{\|} = g(|\nabla S|) + g'(|\nabla S|)|\nabla S|, \text{ and } \lambda^{\perp} = g(|\nabla S|).$$

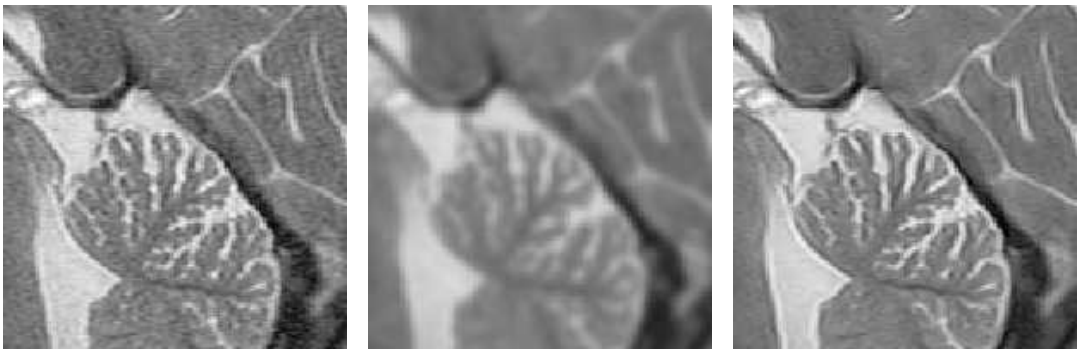
While $\lambda^{\perp}$ is always nonnegative, $\lambda^{\|}$ can change sign. Thus the initial value problem is ill-posed if $sg'(s) + g(s) < 0$, i.e., if $sg(s)$ is decreasing, and one actually wants $g(s)$ to vanish



(a) Original brain MRI image

(b) Linear (Gaussian) smoothing

(c) Affine smoothing



(d) Original (zoom)

(e) Linear (Gaussian) smoothing (zoom)

(f) Affine smoothing (zoom)

Figure 1: Linear smoothing smears the edges.

very quickly as $s \to \infty$ (e.g. $g(s) = e^{-s}$). Even if solutions S^t could be constructed in the ill-posed regime, they would vary strongly and unpredictably under tiny perturbations in the initial image $S^0 = I$, while it is not clear if the Comparison Principle would be satisfied.

Despite these objections, numerical experiments [19] have indicated that this method actually does remove a significant amount of noise before edges are smeared out very much.

2.2.3 Regularized Anisotropic Smoothing

Alvarez, Lions and Morel [20] proposed a class of modifications of the Perona-Malik scheme that results in well-posed initial value problems. They replaced (7) with

$$\frac{\partial S}{\partial t} = h(|G_\sigma * \nabla S|)\ |\nabla S|\ \mathrm{div}\ \frac{\nabla S}{|\nabla S|}, \tag{8}$$

which can be written as

$$\frac{\partial S}{\partial t} = h(|G_\sigma * \nabla S|)\ |\nabla S| \sum_{i,j} b_{ij}(\nabla S)\nabla^2_{ij} S, \tag{9}$$

with

$$b_{ij}(\nabla S) = \frac{|\nabla S|^2 \delta_{ij} - \nabla_i S\ \nabla_j S}{|\nabla S|^3}.$$

Thus, the stopping function g in (7) has been set equal to $g(s) = 1/s$, and a new stopping function h is introduced. In addition, a smoothing kernel G_σ, which averages ∇S in a region of order σ, is introduced. One could let G_σ be the standard Gaussian (5), but other choices are possible. In the limiting case $\sigma \searrow 0$, in which $G_\sigma * \nabla S$ simply becomes ∇S, a PDE is obtained.

2.2.4 Level Set Flows

Level set methods for the implementation of curvature driven flows were introduced by Osher and Sethian [21] and have proved to be a powerful numerical technique with numerous applications; see [22, 23] and the references therein.

Equation (8) can be rewritten in terms of the level sets of the image S. If S is smooth, and if c is a regular value of $S^t : \mathfrak{D} \to \mathfrak{C}$ (i.e., $\forall \mathbf{p} \in \mathfrak{D}, S^t(\mathbf{p}) = c \Rightarrow\ \mathrm{d}S^t(\mathbf{p}) \neq 0$), then $\Gamma^t(c) = \{\mathbf{p} \in \mathfrak{D} \mid S^t(\mathbf{p}) = c\}$ is a smooth curve in $\mathfrak{D}$. It is the boundary of the region with gray level c or less. As time goes by, the curve $\Gamma^t(c)$ will move about, and as long as it is

a smooth curve one can define its normal velocity V by choosing any local parametrization $\Gamma : [0,1] \times (t_0, t_1) \to \mathfrak{D}$ and declaring V to be the normal component of $\partial_t \Gamma$.

If the normal is chosen to be in the direction of ∇S (rather than $-\nabla S$), then

$$V = -\frac{\partial_t S}{|\nabla S|}.$$

The curvature of the curve $\Gamma^t(c)$ (also in the direction of ∇S) is

$$\kappa = -\operatorname{div} \frac{\nabla S}{|\nabla S|} = -\frac{S_y^2\, S_{xx} - 2 S_x S_y\, S_{xy} + S_x^2\, S_{yy}}{\left(S_x^2 + S_y^2\right)^{3/2}} \tag{10}$$

Thus, (8) can be rewritten as

$$V = h(|G_\sigma * \nabla S|)\kappa,$$

which, in the special case $h \equiv 1$, reduces to the curve shortening equation[5]

$$V = \kappa. \tag{11}$$

So, if $h \equiv 1$ and if $S : \mathfrak{D} \times [0, T) \to \mathfrak{C}$ is a family of images that evolves by (8), then the level sets $\Gamma^t(c)$ evolve independently of each other.

This leads to the following simple recipe for noise removal: given an initial image $S^0 = I$, let it evolve so that its level curves $(S^t)^{-1}(c)$ satisfy the curve shortening equation (11). For this to occur, the function S should satisfy the Alvarez-Lions-Morel equation (8) with $h \equiv 1$, i.e.,

$$\frac{\partial S}{\partial t} = |\nabla S| \operatorname{div} \frac{\nabla S}{|\nabla S|} = \frac{S_y^2\, S_{xx} - 2 S_x S_y\, S_{xy} + S_x^2\, S_{yy}}{S_x^2 + S_y^2} \tag{12}$$

It was shown by Evans and Spruck [30] and Chen, Giga and Goto [31] that, even though this is a highly degenerate parabolic equation, a theory of *viscosity solutions* can still be constructed.

The fact that level sets of a solution to (12) evolve independently of each other turns out to be desirable in noise reduction since it eliminates the edge blurring effect of the linear smoothing method. For example, if I is a characteristic function, then S^t will also be a characteristic function for all $t > 0$.

[5]There is a substantial literature on the evolution of immersed plane curves, or immersed curves on surfaces by curve shortening and variants thereof. We refer to [24, 25, 26, 27, 28, 29], knowing that this is a far from exhaustive list of references.

The independent evolution of level sets also implies that besides obeying the axioms of Alvarez, Lions and Morel [18] mentioned above, this method also satisfies the additional axiom

Gray scale invariance For any initial image $S^0 = I$ and any monotone function $\phi : \mathfrak{C} \to \mathfrak{C}$, one has $T^t[\phi \circ I] = \phi \circ T^t[I]$.

One can verify easily that (12) formally satisfies this axiom, and it can in fact be rigorously proven to be true in the context of viscosity solutions. See [31, 30].

2.2.5 Affine Invariant Smoothing

There are several variations on curve shortening that will yield comparable results. Given an initial image $I : \mathfrak{D} \to \mathfrak{C}$, one can smooth it by letting its level sets move with normal velocity given by some function of their curvature:

$$V = \Phi(\kappa) \tag{13}$$

instead of directly setting $V = \kappa$ as in curve shortening. Using (10), one can turn the equation $V = \Phi(\kappa)$ into a PDE for S. If $\Phi : \mathfrak{C} \to \mathfrak{C}$ is monotone, then the resulting PDE for S will be degenerate parabolic, and the existence and uniqueness of viscosity solutions have been shown [31, 18].

A particularly interesting choice of Φ leads to *affine curve shortening* [18, 32, 33, 34, 35]

$$V = (\kappa)^{1/3}, \tag{14}$$

(where we agree to define $(-\kappa)^{1/3} \triangleq -(\kappa)^{1/3}$).

While the evolution equation (13) is invariant under Euclidean transformations, the affine curve shortening equation (14) has the additional property that it is invariant under the action of the Special Affine group $\mathrm{SL}(2, \mathbb{R})$, i.e., compositions of rotations, translations, dilations, and shears. If Γ^t is a family of curves evolving by (14) and $A \in \mathrm{SL}(2, \mathbb{R})$, $b \in \mathbb{R}^2$, then $\tilde{\Gamma}^t = A \cdot \Gamma^t + b$ also evolves by (14).

Affine smoothing respects edges much better than Gaussian smoothing; see Figure 1(c). The affine smoothing filter was implemented based on a level set formulation using the numerics suggested in [36].

2.2.6 Total Variation

Rudin et al. presented an algorithm for noise removal based on the minimization of the total first variation of S, given by

$$\int_{\mathfrak{D}} |\nabla S| \, \mathrm{d}\mathbf{p}. \tag{15}$$

(See [37] for the details and the references therein for related work on the total variation method.) The minimization is performed under certain constraints and boundary conditions (zero flow on the boundary). The constraints they employed are zero mean value and given variance σ^2 of the noise, but other constraints can be considered as well. More precisely, if the noise is additive, the constraints are given by

$$\int_{\mathfrak{D}} S \, \mathrm{d}\mathbf{p} = \int_{\mathfrak{D}} I \, \mathrm{d}\mathbf{p}, \quad \int_{\mathfrak{D}} (S - I)^2 \, \mathrm{d}\mathbf{p} = 2\sigma^2. \tag{16}$$

Noise removal according to this method proceeds by first choosing a value for the parameter σ, and then minimizing $\int |\nabla S|$ subject to the constraints (16). For each $\sigma > 0$ there exists a unique minimizer $S^\sigma \in \mathrm{BV}(\mathfrak{D})$ satisfying the constraints.[6] The family of images $\{S^\sigma \mid \sigma > 0\}$ does not form a scale space and does not satisfy the axioms set forth by Alvarez et al. [18]. Furthermore, the smoothing parameter σ does not measure the "length scale of smoothing" in the way the parameter t in scale spaces does. Instead, the assumptions underlying this method of smoothing are more statistical. One assumes that the given image I is actually an ideal image to which a "noise term" $N(\mathbf{p})$ has been added. The values $N(\mathbf{p})$ at each $\mathbf{p} \in \mathfrak{D}$ are assumed to be independently distributed random variables with average variance σ^2.

The Euler-Lagrange equation for this problem is

$$\mathrm{div}\left(\frac{\nabla S}{|\nabla S|}\right) = \lambda(S - I), \tag{17}$$

[6] $\mathrm{BV}(\mathfrak{D})$ is the set of functions of bounded variation on $\mathfrak{D}$. See [38].

where λ is a Lagrange multiplier. In view of (10) we can write this as $\kappa = -\lambda(S - I)$, i.e., we can interpret (17) as a prescribed curvature problem for the level sets of S. To find the minimizer of (15) with the constraints given by (16), start with the initial image $S^0 = I$ and modify it according to the L^2 steepest descent flow for (15) with the constraint (16), namely,

$$\frac{\partial S}{\partial t} = \text{div}\left(\frac{\nabla S}{|\nabla S|}\right) - \lambda(t)(S - I), \tag{18}$$

where $\lambda(t)$ is chosen so as to preserve the second constraint in (16). The solution to the variational problem is given when S achieves steady state. This computation must be repeated ab initio for each new value of σ.

2.3 Image Registration

Image registration is the process of bringing two or more images into spatial correspondence (aligning them). In the context of medical imaging, image registration allows for the concurrent use of images taken with different modalities (e.g., MRI and CT), at different times or with different patient positions. In surgery, for example, images are acquired before (pre-operative) as well as during (intra-operative) surgery. Because of time constraints, the real-time intra-operative images have a lower resolution than the pre-operative images obtained before surgery. Moreover, deformations, which occur naturally during surgery, make it difficult to relate the high-resolution pre-operative image to the lower-resolution intra-operative anatomy of the patient. Image registration attempts to help the surgeon relate the two sets of images.

Registration has an extensive literature. Numerous approaches have been explored ranging from statistics to computational fluid dynamics and various types of warping methodologies. See [39, 40] for a detailed description of the field as well as an extensive set of references, and [41, 42] for recent papers on the subject.

Registration typically proceeds in several steps. First, one decides how to measure similarity between images. One may include the similarity among pixel intensity values, as well as the proximity of predefined image features such as implanted fiducials or anatomical

landmarks.[7] Next, one looks for a transformation that maximizes similarity when applied to one of the images. Often this transformation is given as the solution of an optimization problem where the transformations to be considered are constrained to be of a predetermined class $\mathcal{C}$. In the case of *rigid registration* (Section 2.3.1), $\mathcal{C}$ is the set of Euclidean transformations. Soft tissues in the human body typically do not deform rigidly. For example, physical deformation of the brain occurs during neurosurgery as a result of swelling, cerebrospinal fluid loss, hemorrhage and the intervention itself. Therefore, a more realistic and more challenging problem is *elastic registration* (Section 2.3.2), where $\mathcal{C}$ is the set of smooth diffeomorphisms. Because of anatomical variability, non rigid deformation maps are also useful when comparing images from different patients.

2.3.1 Rigid Registration

Given some similarity measure S on images and two images I and J, the problem of rigid registration is to find a Euclidean transformation $T^*\mathbf{p} = R\mathbf{p} + a$ (with $R \in \mathrm{SO}(3, \mathbb{R})$ and $a \in \mathbb{R}^3$) that maximizes the similarity between I and $T^*(J)$, i.e.,

$$T^* = \text{maximizer of } S(I, T(J)) \text{ over all Euclidean transformations } T. \tag{19}$$

A number of standard multidimensional optimization techniques are available to solve (19).

Many similarity measures have been investigated [43], e.g., the *normalized cross correlation*:

$$S(I, J) = \frac{(I, J)_{L^2}}{\|I\|_{L^2}\|J\|_{L^2}}. \tag{20}$$

Information-theoretic similarity measures are also popular. By selecting a pixel $\mathbf{p}$ at random with uniform probability from the domain $\mathfrak{D}$ and computing the corresponding gray value $I(\mathbf{p})$, one can turn any image I into a random variable. If we write p_I for the probability density function of the random variable I and p_{IJ} for the joint probability density of I and J, then one can define the *entropy of the difference* and the *mutual information* of

[7] Registration or alignment is most commonly achieved by marking identifiable points on the patient with a tracked pointer and in the images. These points, often called *fiducials,* may be anatomical or artificially attached landmarks, i.e. "implanted fiducials".

two images I and J:

$$S(I,J) = \inf\left\{\sum_c p_K(c)\log p_K(c) \mid K = I - sJ, s \in \mathbb{R}\right\} \tag{21}$$

$$S(I,J) = \sum_{a,b} p_{IJ}(a,b)\log\frac{p_{IJ}(a,b)}{p_I(a)p_J(b)}. \tag{22}$$

The normalized cross correlation (20) and the entropy of the difference (21) are maximized when the intensities of the two images are linearly related. In contrast, the mutual information measure (22) only assumes that the co-occurrence of the most probable values in the two images is maximized at registration. This relaxed assumption explains the success of mutual information in registration [44, 45].

2.3.2 Elastic Registration

In this section we describe a method of Haker et al. [46] for elastic registration in which the similarity between two images is measured by their L^2 Kantorovich–Wasserstein distance. Finding "the best map" from one image to another then leads to an optimal transport problem. Optimal transport methods have appeared in computer vision [47] as well as in econometrics, fluid dynamics, automatic control, transportation, statistical physics, shape optimization, expert systems, and meteorology [48].

In the approach of [46] one thinks of a gray scale image as a Borel measure[8] μ on some open domain $\mathfrak{D} \subset \mathbb{R}^d$, where, for any $E \subset \mathfrak{D}$, the "amount of white" in the image contained in E is given by $\mu(E)$. Given two images, $(\mathfrak{D}_0, \mu_0)$ and $(\mathfrak{D}_1, \mu_1)$, one considers all maps $u : \mathfrak{D}_0 \to \mathfrak{D}_1$ that preserve the measures, i.e., that satisfy[9]

$$u_\#(\mu_0) = \mu_1, \tag{23}$$

and one is required to find the map (if it exists) that minimizes the Monge-Kantorovich transport functional (see the exact definition below). This optimal transport problem was introduced by Gaspard Monge in 1781 when he posed the question of moving a pile of soil

[8] This can be physically motivated. For example, in some forms of MRI the image intensity is the actual proton density.

[9] If X and Y are sets with σ-algebras $\mathcal{M}$ and $\mathcal{N}$, and if $f : X \to Y$ is a measurable map, then we write $f_\#\mu$ for the pushforward of any measure μ on $(X, \mathcal{M})$, i.e., for any measurable $E \subset Y$ we define $f_\#\mu(E) = \mu\big(f^{-1}(E)\big)$.

from one site to another in a manner that minimizes transportation cost. The problem was given a modern formulation by Kantorovich [49], and so now is known as the *Monge-Kantorovich problem*.

More precisely, assuming that the cost of moving a mass m from $\mathbf{p} \in \mathbb{R}^d$ to $\mathbf{q} \in \mathbb{R}^d$ is $m \cdot \Phi(\mathbf{p}, \mathbf{q})$, where $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ is called the *cost function*, Kantorovich defined the cost of moving the measure μ_0 to the measure μ_1 by the map $u : \mathfrak{D}_0 \to \mathfrak{D}_1$ to be

$$M(u) = \int_{\mathfrak{D}_0} \Phi(\mathbf{p}, u(\mathbf{p}))\ \mathrm{d}\mu_0(\mathbf{p}). \tag{24}$$

The infimum of $M(u)$ taken over all measure preserving $u : (\mathfrak{D}_0, \mu_0) \to (\mathfrak{D}_1, \mu_1)$ is called the *Kantorovich–Wasserstein distance* between the measures μ_0 and μ_1. The minimization of $M(u)$ constitutes the mathematical formulation of the Monge-Kantorovich optimal transport problem.

If the measures μ_i and Lebesgue measure $\mathrm{d}\mathcal{L}$ are absolutely continuous with respect to each other, so that we can write $\mathrm{d}\mu_i = m_i(\mathbf{p})\ \mathrm{d}\mathcal{L}$ for certain strictly positive densities $m_i \in L^1(\mathfrak{D}_i,\ \mathrm{d}\mathcal{L})$, and if the map u is a diffeomorphism from $\mathfrak{D}_0$ to $\mathfrak{D}_1$, then the mass preservation property (23) is equivalent to

$$m_0(\mathbf{p}) = \det\big(Du(\mathbf{p})\big) \cdot m_1(u(\mathbf{p})), \qquad \text{(for almost all } \mathbf{p} \in \mathfrak{D}_0\text{)}. \tag{25}$$

The Jacobian equation (25) implies that if a small region in $\mathfrak{D}_0$ is mapped to a larger region in $\mathfrak{D}_1$, there must be a corresponding decrease in density in order to comply with mass preservation. In other words, expanding an image darkens it.

The L^2 Monge–Kantorovich problem corresponds to the cost function $\Phi(\mathbf{p}, \mathbf{q}) = \frac{1}{2}|\mathbf{p} - \mathbf{q}|^2$. A fundamental theoretical result for the L^2 case [50, 51, 52] is that there is a unique optimal mass preserving u, and that this u is characterized as the gradient of a convex function p, i.e., $u = \nabla p$. General results about existence and uniqueness may be found in [53] and the references therein.

To reduce the Monge–Kantorovich cost $M(u)$ of a map $u^0 : \mathfrak{D}_0 \to \mathfrak{D}_1$, in [46] the authors consider a rearrangement of the points in the domain of the map in the following sense: the initial map u^0 is replaced by a family of maps u^t for which one has $u^t \circ s^t = u^0$ for some

family of diffeomorphisms $s^t : \mathfrak{D}_0 \to \mathfrak{D}_0$. If the initial map u^0 sends the measure μ_0 to μ_1, and if the diffeomorphisms s^t preserve the measure μ_0, then the maps $u^t = u^0 \circ (s^t)^{-1}$ will also send μ_0 to μ_1.

Any sufficiently smooth family of diffeomorphisms $s^t : \mathfrak{D}_0 \to \mathfrak{D}_0$ is determined by its velocity field, defined by $\partial_t s^t = v^t \circ s^t$.

If u^t is any family of maps, then, assuming $u^t_{\#}\mu_0 = \mu_1$ for all t, one has

$$\frac{\mathrm{d}}{\mathrm{d}t} M(u^t) = \int_{\mathfrak{D}_0} \left\langle \Phi_{\mathbf{p}}(\mathbf{p}, u^t(\mathbf{p})), v^t(\mathbf{p}) \right\rangle \ \ \mathrm{d}\mu_0(\mathbf{p}). \tag{26}$$

The steepest descent is achieved by a family $s^t \in \mathrm{Diff}^1_{\mu_0}(\mathfrak{D}_0)$, whose velocity is given by

$$v^t = -\frac{1}{m_0(\mathbf{p})} \mathcal{P}\left(\Phi_{\mathbf{p}}(\mathbf{p}, u^t(\mathbf{p}))\right). \tag{27}$$

Here, $\mathcal{P}$ is the Helmholtz projection, which extracts the divergence-free part of vector fields on $\mathfrak{D}_0$, i.e., for any vector field w on $\mathfrak{D}$ one has $w = \mathcal{P}[w] + \nabla p$.

From $u^0 = u^t \circ s^t$ one gets the transport equation

$$\frac{\partial u^t}{\partial t} + v^t \cdot \nabla u^t = 0 \tag{28}$$

where the velocity field is given by (27). In [54] it is shown that the initial value problem (27), (28) has short time existence for $C^{1,\alpha}$ initial data u^0, and a theory of global weak solutions in the style of Kantorovich is developed.

For image registration, it is natural to take $\Phi(\mathbf{p}, \mathbf{q}) = \frac{1}{2}|\mathbf{p} - \mathbf{q}|^2$ and $\mathfrak{D}_0 = \mathfrak{D}_1$ to be a rectangle. Extensive numerical computations show that the solution to the unregularized flow converges to a limiting map for a large choice of measures and initial maps. Indeed, in this case, we can write the minimizing flow in the following "nonlocal" form:

$$\frac{\partial u^t}{\partial t} = -\frac{1}{m_0} \left\{u^t + \nabla(-\Delta)^{-1} \operatorname{div}(u^t)\right\} \cdot \nabla u^t. \tag{29}$$

The technique has been mathematically justified in [54] to which we refer the reader for all of the relevant details.

2.3.3 Optimal Warping

Typically in elastic registration, one wants to see an explicit warping that smoothly deforms one image into the other. This can easily be done using the solution of the Monge–Kantorovich problem. Thus, we assume now that the gradient descent process has been applied as described above and that it has converged to the Monge–Kantorovich mapping $\tilde{u}_{MK}$.

Following the work of Benamou and Brenier [55], (see also [51]), we consider the following related problem:

$$\inf \int \int_0^1 m(t,x)\|v(t,x)\|^2 \, dt \, dx$$

over all time-varying densities m and velocity fields v satisfying

$$\frac{\partial m}{\partial t} + \operatorname{div}(mv) = 0, \tag{30}$$

$$m(0,\cdot) = m_0, \quad m(1,\cdot) = m_1. \tag{31}$$

It is shown in [55] that this infimum is attained for some $m_{\min}$ and $v_{\min}$, and that it is equal to the L^2 Kantorovich–Wasserstein distance between $\mu_0 = m_0 \, \mathrm{d}\mathcal{L}$ and $\mu_1 = m_1 \, \mathrm{d}\mathcal{L}$. Further, the flow $X = X(x,t)$ corresponding to the minimizing velocity field $v_{\min}$ via

$$X(x,0) = x, \quad X_t = v_{\min} \circ X$$

is given simply as

$$X(x,t) = x + t \, (\tilde{u}_{MK}(x) - x). \tag{32}$$

Note that when $t = 0$, X is the identity map and when $t = 1$, it is the solution $\tilde{u}_{MK}$ to the Monge–Kantorovich problem. This analysis provides appropriate justification for using (32) to *define* our continuous warping map X between the densities m_0 and m_1.

This warping is illustrated on MR heart images (Figure 2). Since the images have some black areas where the intensity is zero, and since the intensities define the densities in the Monge-Kantorovich registration procedure, we typically replace m_0 by some small perturbation $m_0 + \epsilon$ for $0 < \epsilon \ll 1$ to ensure that densities are strictly positive. Full details may be found in [56].

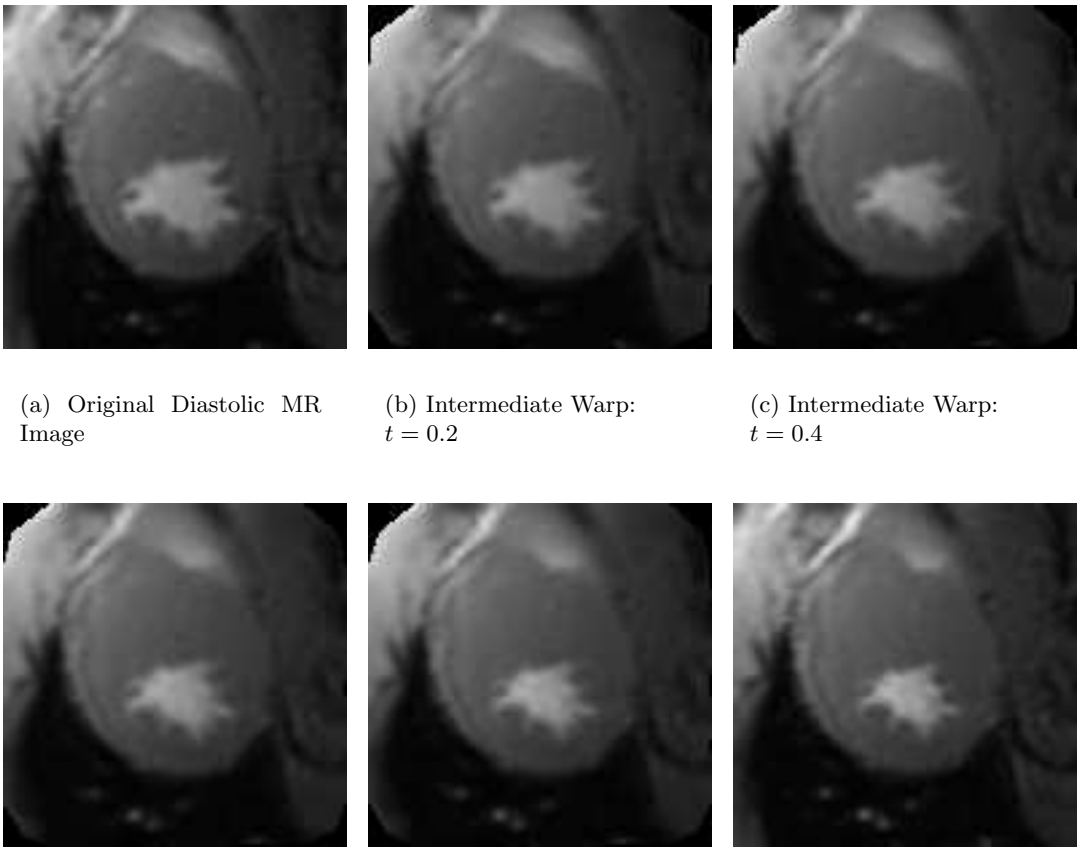
Specifically, two MR images of a heart are given at different phases of the cardiac cycle. In each image the white part is the blood pool of the left ventricle. The circular gray part is the myocardium. Since the deformation of this muscular structure is of great interest, the blood pool is masked and we only consider the optimal warp of the myocardium in the sense described above. Figure 2(a) is a diastolic image and Figure 2(f) is a systolic image.[10] *These are the only data given.* Figures 2(b) to Figure 2(e) show the warping using the Monge-Kantorovich technique between these two images. These images offer a very plausible deformation of the heart as it undergoes its diastole/systole cycle. This is remarkable considering that they were all *artificially created* by the Monge-Kantorovich mappings. The plausibility of the deformation demonstrates the quality of the resulting warping. Images become much more vivid when viewed as a short animation available at `http://www.bme.gatech.edu/groups/minerva/publications/papers/medicalBAMS2005.html`).

2.4 Image Segmentation

When looking at an image, a human observer cannot help seeing structures that often may be identified with objects. However, digital images (as well as the raw retinal input of local intensities) are not structured. Segmentation is the process of creating a structured visual representation from an unstructured one. The problem was first studied in the 1920's by psychologists of the Gestalt school (see Kohler [6] and the references therein) and later by psychophysicists[11] [7, 8]. In its modern formulation, image segmentation is the problem of partitioning an image into homogeneous regions that are semantically meaningful, i.e., that correspond to objects that can be identified by a human expert. Segmentation is not concerned with actually determining what the partitions are. In that sense, it is a lower level problem than object recognition.

In the context of medical imaging, these regions have to be anatomically meaningful. A typical example is partitioning an MRI image of the brain into the white and gray matter. Since it replaces continuous intensities with discrete labels, segmentation can be seen as an

[10] *Diastolic* refers to the relaxation of the heart muscle while *systolic* refers the contraction of the muscle.
[11] Segmentation was called "perceptual grouping" by the Gestaltists. The idea was to study the preferences of human beings for the grouping of sets of shapes arranged in the visual field.



(a) Original Diastolic MR Image

(b) Intermediate Warp: $t = 0.2$

(c) Intermediate Warp: $t = 0.4$

(d) Intermediate Warp: $t = 0.6$

(e) Intermediate Warp: $t = 0.8$

(f) Original Systolic MR Image

Figure 2: Optimal warping of myocardium from diastolic to systolic in cardiac cycle (see text).

extreme form of smoothing/information reduction. Segmentation is also related to registration in the sense that if an atlas[12] can be perfectly registered to a dataset at hand, then the registered atlas labels are the segmentation. Segmentation is useful for visualization,[13] it allows for quantitative shape analysis and provides an indispensable anatomical framework for virtually any subsequent automatic analysis.

Indeed, segmentation is perhaps the central problem of artificial vision, and accordingly many approaches have been proposed (for a nice survey of modern segmentation methods, see the monograph [57]). There are basically two dual approaches. In the first, one can start by considering the whole image to be the object of interest and then refine this initial guess. These "split and merge" techniques can be thought of as somewhat analogous to the top-down processes of human vision. In the other approach, one starts from one point assumed to be inside the object and adds other points until the region encompasses the object. Those are the "region growing" techniques and bear some resemblance to the bottom-up processes of biological vision.

The dual problem to segmentation is that of determining the boundaries of the segmented homogeneous regions. This approach has been popular for some time since it allows one to build upon the well-investigated problem of *edge detection* (Section 2.4.1). Difficulties arise with this approach because noise can be responsible for spurious edges. Another major difficulty is that local edges need to be connected into topologically correct region boundaries. To address these issues, it was proposed to set the topology of the boundary to that of a sphere and then deform the geometry in a variational framework to match the edges. In two dimensions, the boundary is a closed curve and this approach was named *snakes* (Section 2.4.2). Improvements of the technique include *geometric active contours* (Section 2.4.3) and *conformal active contours* (Section 2.4.4). All these techniques are generically referred to as *active contours*. Finally, as described in [57], most segmentation methods can be set in the elegant mathematical framework proposed by Mumford and Shah [58] (Section 2.4.6).

[12] An image from a typical patient that has been manually segmented by some human expert.

[13] As discussed previously, human vision is inherently two-dimensional. In order to "see" an organ radiologists therefore have to resort to visualizing its outside boundary. Determining this boundary is equivalent to segmenting the organ.

2.4.1 Edge Detectors

Consider the ideal case of a bright object $\mathcal{O}$ on a dark background. The physical object is represented by its projections on the image I. The characteristic function $1_{\mathcal{O}}$ of the object is the ideal segmentation, and since the object is contrasted on the background, the variations of the intensity I are large on the boundary $\partial\mathcal{O}$. It is therefore natural to characterize the boundary $\partial\mathcal{O}$ as the locus of points where the norm of the gradient $|\nabla I|$ is large. In fact, if $\partial\mathcal{O}$ is piecewise smooth, then $|\nabla I|$ is a singular measure whose support is exactly $\partial\mathcal{O}$.

This is the approach taken in the 60's and 70's by Roberts [59] and Sobel [60], who proposed slightly different discrete convolution masks to approximate the gradient of digital images. Disadvantages with these approaches are that edges are not precisely localized and may be corrupted by noise. Figure 3(b) is the result of a Sobel edge detector on a medical image. Note the thickness of the boundary of the heart ventricle as well as the presence of "spurious edges" resulting from noise. Canny [61] proposed adding a smoothing pre-processing step (to reduce the influence of the noise) as well as a thinning post-processing phase (to ensure that the edges are uniquely localized). See [62] for a survey and evaluation of edge detectors using gradient techniques.

A slightly different approach, initially motivated by psychophysics, was proposed by Marr and Hildreth [63, 64] where edges are defined as the zeros of $\Delta G_\sigma * I$, the Laplacian of a smooth version of the image. One can give a heuristic justification by assuming that the edges are smooth curves; more precisely, assume that near an edge the image is of the form

$$I(\mathbf{p}) = \varphi\left(\frac{S(\mathbf{p})}{\varepsilon}\right), \tag{33}$$

where S is a smooth function with $|\nabla S| = O(1)$ which vanishes on the edge, ε is a small parameter proportional to the width of the edge, and $\varphi : \mathbb{R} \to [0, 1]$ is a smooth increasing function with limits $\varphi_\pm = \lim_{t\to\pm\infty} \varphi(t)$.

The function φ describes the profile of I transverse to its level sets. We will assume that the graph of φ has exactly one inflection point.

The assumption (33) implies $\nabla I = \varphi'(S/\varepsilon)\nabla S$, while the second derivative $\nabla^2 I$ is given by

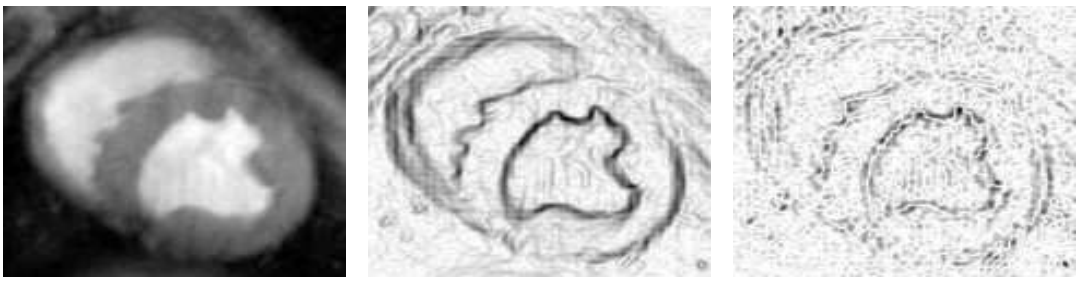
$$\nabla^2 I = \frac{\varphi''(S/\varepsilon)}{\varepsilon^2}\nabla S \otimes \nabla S + \frac{\varphi'(S/\varepsilon)}{\varepsilon}\nabla^2 S. \tag{34}$$

Thus, $\nabla^2 I$ will have eigenvalues of moderate size ($O(\varepsilon^{-1})$) in the direction perpendicular to ∇I, while the second derivative in the direction of the gradient will change from a large $O(\varepsilon^{-2})$ positive value to a large negative value as one crosses from small I to large I values.

From this discussion of $\nabla^2 I$, it seems reasonable to define the edges to be the locus of points where $|\nabla I|$ is large and where either $(\nabla I)^T \cdot \nabla^2 I \cdot \nabla I$, or ΔI, or $\det \nabla^2 I$ changes sign. The quantity $(\nabla I)^T \cdot \nabla^2 I \cdot \nabla I$ vanishes at $\mathbf{p} \in \mathfrak{D}$ if the graph of the function I restricted to the normal line to the level set of I passing through $\mathbf{p}$ has an inflection point at $\mathbf{p}$ (assuming $\nabla I(\mathbf{p}) \neq \mathbf{0}$). In general, zeros of ΔI and $\det \nabla^2 I$ do not have a similar description, but since the first term in (34) will dominate when ε is small, both $(\nabla I)^T \cdot \nabla^2 I \cdot \nabla I$ and $\det \nabla^2 I$ will tend to vanish at almost the same places.

The computation of second derivatives is numerically awkward, so one prefers to work with a smoothed out version of the image, such as $G_\sigma * I$ instead of I itself.

See Figure 3(c) for the result of the Marr edge detector. In these images the edge is located precisely at the zeroes of the Laplacian (a thin white curve).



(a) Original image. (b) Sobel edge detector. (c) Marr edge detector.

Figure 3: Edge detectors on a MRI image of a heart

While it is now understood that these local edge detectors cannot be used directly for the detection of object boundaries (local edges would need to be connected in a topologically

correct boundary), these techniques proved instrumental in designing the active contour models described below and are still being investigated [65, 66].

2.4.2 Snakes

A first step toward automated boundary detection was taken by Witkin, Kass and Terzopoulos [67], and later by a number of researchers (see [68] and the references therein). Given an image $I : \mathfrak{D} \subset \mathbb{R}^2 \to \mathfrak{C}$, they subject an initial parametrized curve $\Gamma : [0, 1] \to \mathfrak{D}$ to a steepest descent flow for an energy functional of the form[14]

$$\mathcal{E}(\Gamma) = \int_0^1 \left\{ \frac{1}{2} w_1(p) |\Gamma_{pp}|^2 + \frac{1}{2} w_2(p) |\Gamma_p|^2 + W(\Gamma(p)) \right\} \, \mathrm{d}p. \tag{35}$$

The first two terms control the smoothness of the active contour Γ. The contour interacts with the image through the potential function $W : \mathfrak{D} \to \mathbb{R}$, which is chosen to be small near the edges and large everywhere else (it is a decreasing function of some edge detector). For example, one could take

$$W(\mathbf{p}) = \frac{1}{1 + \|\nabla G_\sigma * I(\mathbf{p})\|^2}. \tag{36}$$

Minimizing $\mathcal{E}$ will therefore attract Γ toward the edges. The gradient flow is the fourth-order nonlinear parabolic equation

$$\frac{\partial \Gamma}{\partial t} = -\left(w_2(p)\Gamma_{pp}\right)_{pp} + \left(w_1(p)\Gamma_p\right)_p + \nabla W(\Gamma(p,t)). \tag{37}$$

This approach gives reasonable results. See [69] for a survey of snakes in medical image analysis. One limitation, however, is that the active contour or snake cannot change topology. It starts out being a topological circle and will always remain a topological circle and will not be able to break up into two or more pieces, even if the image would contain two unconnected objects and this would give a more natural description of the edges. Special ad hoc procedures have been developed to handle splitting and merging [70].

2.4.3 Geometric Active Contours

Another disadvantage of the snake method is that it explicitly involves the parametrization of the active contour Γ, while there is no obvious relation between the parametrization of

[14] We present Cohen's [68] version of the energy.

the contour and the geometry of the objects to be captured. Geometric models have been developed by [71, 72] to address this issue.

As in the snake framework, one deforms the active contour Γ by a velocity that is essentially defined by a curvature term, and a constant inflationary term weighted by a stopping function W. By formulating everything in terms of quantities that are invariant under reparametrization (such as the curvature and normal velocity of Γ), one obtains an algorithm that does not depend on the parametrization of the contour. In particular, it can be implemented using level sets.

More specifically, the model of [71, 72] is given by

$$V = W(\mathbf{p})(\kappa + c), \tag{38}$$

where both the velocity V and the curvature κ are measured using the normal $\mathbf{N}$ for Γ. Here, as previously, W is small at edges and large everywhere else, and c is a constant, called the *inflationary parameter*. When c is positive, it helps push the contour through concavities and can speed up the segmentation process. When it is negative, it allows expanding "bubbles," i.e., contours that expand rather than contract to the desired boundaries. We should note that there is no canonical choice for the constant c, which has to be determined experimentally.

In practice Γ is deformed using the Osher-Sethian level set method described in Section 2.2.4. One represents the curve Γ^t as the zero level set of a function $\Phi : \mathfrak{D} \times \mathbb{R}^+ \to \mathbb{R}$,

$$\Gamma^t = \{\mathbf{p} \in \mathfrak{D} : \Phi(\mathbf{p}, t) = 0\}. \tag{39}$$

For a given normal velocity field, the defining function Φ is then the solution of the Hamilton-Jacobi equation

$$\frac{\partial \Phi}{\partial t} + V\big|\nabla \Phi\big| = 0,$$

which can be analyzed using viscosity theory [73].

Geometric active contours have the advantage that they allow for topological changes (splitting and merging) of the active contour Γ. The main problem with this model is that the desired edges are not steady states for the flow (38). The effect of the factor $W(\mathbf{p})$ is

merely to slow the evolution of Γ^t down as it approaches an edge, but it is not the case that the Γ^t will eventually converge to anything like the sought-for edge as $t \to \infty$. Some kind of artificial intervention is required to stop the evolution when Γ^t is close to an edge.

2.4.4 Conformal Active Contours

In [74, 75], the authors propose a novel technique that is both geometric and variational. In their approach, one defines a Riemannian metric g^W on $\mathfrak{D}$ from a given image $I : \mathfrak{D} \to \mathbb{R}$, by conformally changing the standard Euclidean metric to

$$g^W = W(\mathbf{p})^2 \| \, \mathrm{d}\mathbf{p} \, \|^2. \tag{40}$$

Here, W is defined as above in (36). The length of a curve in this metric is

$$\mathcal{L}^W(\Gamma) = \int_\Gamma W(\Gamma(s)) \;\; \mathrm{d}s. \tag{41}$$

where s is Euclidean arc-length, i.e., the parametrization such that $\|\frac{\partial curve}{\partial s}(s)\| = 1, \forall s$. Curves that minimize this length will prefer to be in regions where W is small, which is exactly where one would expect to find the edges. So, to find edges, one should minimize the W-weighted length of a closed curve Γ, rather than some "energy" of Γ (which depends on a parametrization of the curve).

To minimize $\mathcal{L}^W(\Gamma)$, one computes a gradient flow in the L^2 sense. Since the first variation of this length functional is given by

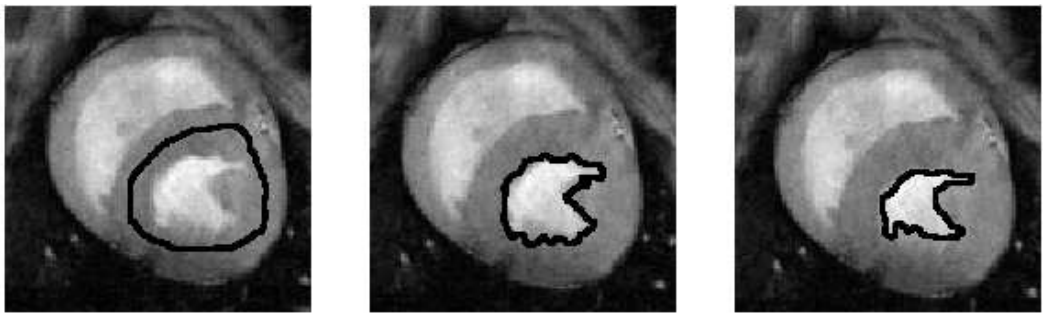
$$\frac{\mathrm{d}\mathcal{L}^W(\Gamma)}{\mathrm{d}t} = -\int_\Gamma V\big\{W\kappa - \mathbf{N}\cdot\nabla W\big\} \;\; \mathrm{d}s,$$

where V is the normal velocity measured in the Euclidean metric, and $\mathbf{N}$ is the Euclidean unit normal, the corresponding L^2 gradient flow is

$$V_{\mathrm{conf}} = W\kappa - \mathbf{N}\cdot\nabla W. \tag{42}$$

Note that this is not quite the curve shortening flow in the sense of [26, 76] on $\mathbb{R}^2$ given the Riemannian manifold structure defined by the conformally Euclidean metric g^W. Indeed, a simple computation shows that in that case one would have

$$V = W^{-2}\big(W\kappa - \mathbf{N}\cdot\nabla W\big). \tag{43}$$



(a) Initial active contour. (b) Evolving active contour. (c) Steady state.

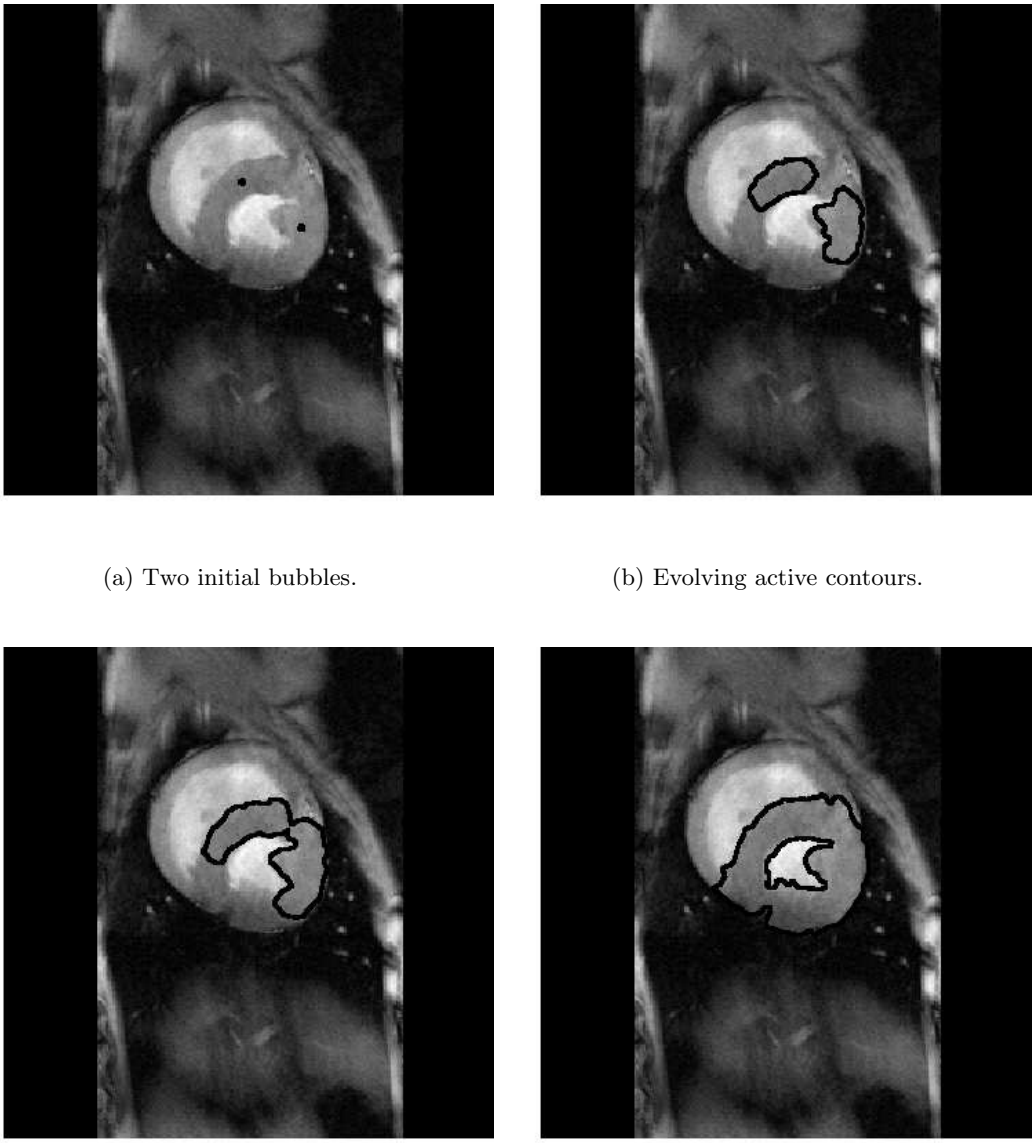
Figure 4: Ventricle segmentation in MRI heart image via shrinking conformal active contour.

Thus, the term "geodesic active contour" used in [75] is a bit of a misnomer, and the term "conformal active contour" as in [74] is prefered.

This technique can also be implemented in the level set framework to allow for the splitting and merging of evolving curves. Figure 4 shows the ventricle of an MR image of the heart being captured by a conformal active contour V_{conf}. Topological changes are illustrated in Figure 5. Two bubbles, which are evolved by the full flow V_{act} with $c < 0$, merge to capture the myocardium of the same image. Conformal active contours can be employed for the segmentation of images from many modalities. In Figure 6, the contour of a cyst was successfully captured in an ultrasound image. This again used the full flow V_{act} with $c < 0$. Because that modality is particularly noisy, the image had been pre-smoothed using the affine curve shortening nonlinear filter (see Section 2.2.5). In Figure 7, the contracting active contour splits to capture two disconnected osseous regions on a CT image.

2.4.5 Conformal Area Minimizing Flows

Typically, to get expanding bubbles, an inflationary term is added in the model (42) as in (38). Many times segmentations are more easily performed by seeding the image with bubbles rather than contracting snakes. The conformal active contours will not allow this



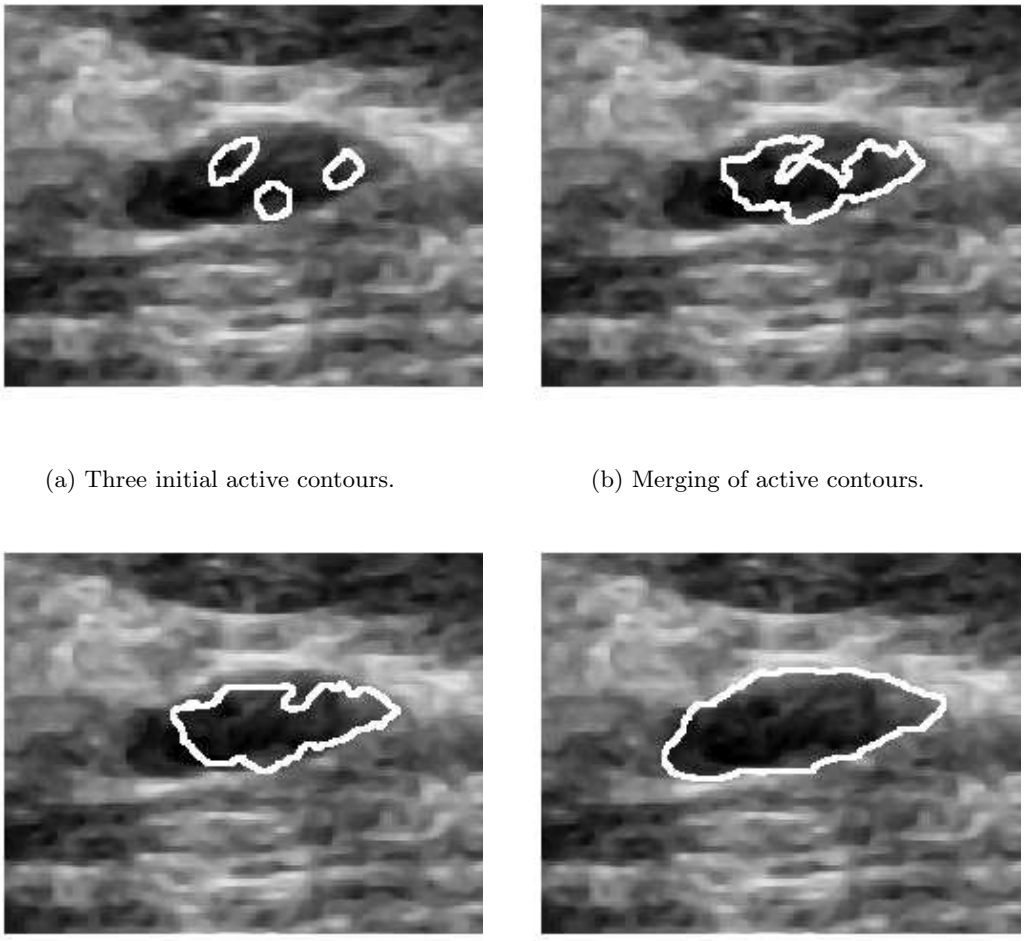
(a) Two initial bubbles.

(b) Evolving active contours.

(c) Merging of active contours.

(d) Steady state.

Figure 5: Myocardium segmentation in MRI heart image with two merging expanding conformal active contours.



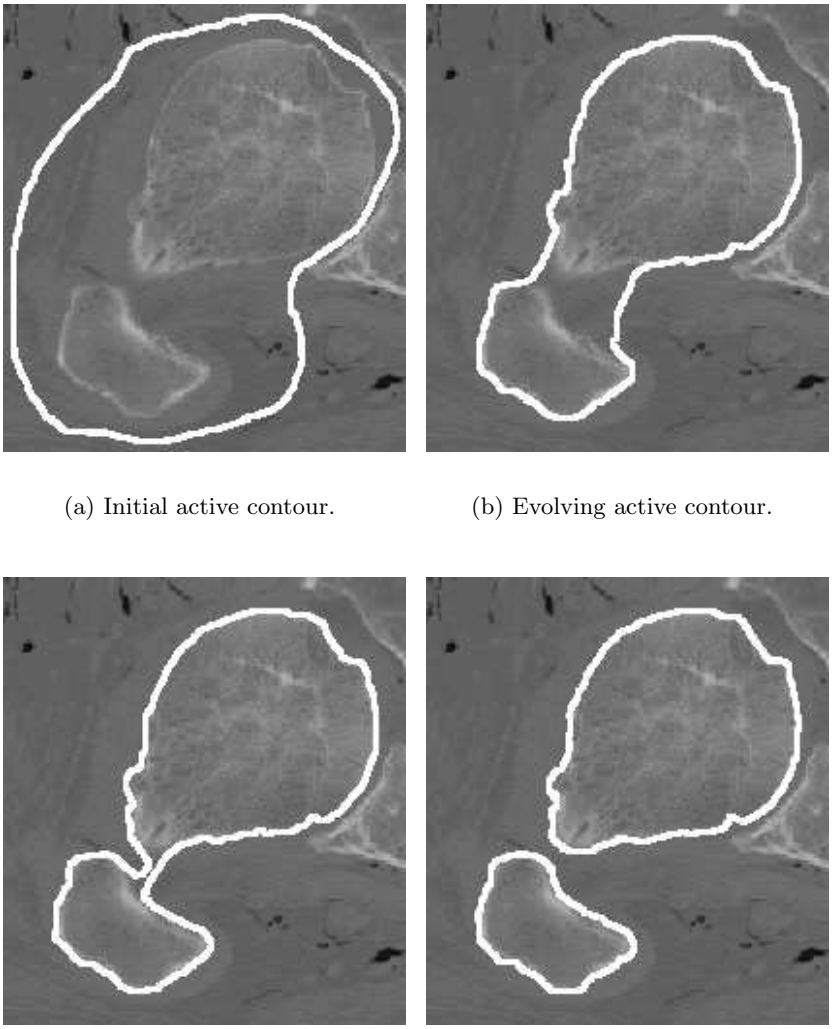
(a) Three initial active contours.

(b) Merging of active contours.

(c) Evolving active contour.

(d) Steady state.

Figure 6: Cyst segmentation in ultrasound breast image with three merging expanding conformal active contours.



(a) Initial active contour.

(b) Evolving active contour.

(c) Evolving active contour.

(d) Splitting of active contour and steady state.

Figure 7: Bone segmentation in CT image with splitting shrinking conformal active contour.

since very small curves will simply shrink to points under the flow (42). To get a curve evolution that will force small bubbles to expand and converge toward the edges, it is convenient to subtract a weighted area term from the length functional $\mathcal{L}^W$, namely,

$$\mathcal{A}^W(\Gamma) = \int_{R_\Gamma} W(\mathbf{p})\ \mathrm{d}\mathbf{p},$$

where $\mathrm{d}\mathbf{p}$ is the two-dimensional Lebesgue measure, and R_Γ is the region enclosed by the contour Γ.

The first variation of this weighted area is [77, 78, 79]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^W(\Gamma^t) = -\int_{\Gamma^t} W(\Gamma(s))V\ \mathrm{d}s \tag{44}$$

where, as before, V is the normal velocity of Γ^t.

The functional that one now tries to minimize is

$$\mathcal{E}^W(\Gamma) = \mathcal{L}^W(\Gamma) + c\mathcal{A}^W(\Gamma), \tag{45}$$

where $c \in \mathbb{R}$ is a constant called the *inflationary parameter*.

The steepest descent for $\mathcal{E}^W$ is obtained by setting

$$V_{\mathrm{act}} = V_{\mathrm{conf}} + cW = \big(\kappa + c\big)W(\mathbf{p}) - \mathbf{N}\cdot\nabla W. \tag{46}$$

For $c = 1$ this is a conformal length/area minimizing flow (see [78]). As in the model of [71, 72] the inflationary parameter c may be chosen as positive (snake or inward moving flow) or negative (bubble or outward moving flow).

In practice, for expanding flows (negative c, weighted area maximizing flow), one expands the bubble using only the inflationary part:

$$V = cW$$

until the active contour is sufficiently large, and then "turns on" the conformal part V_{conf}, which brings the contour to its final position. Again, as in [71, 72], the curvature part of V_{act} also acts to regularize the flow. Finally, [74] contains a detailed mathematical analysis of (46) as well as extensions to the 3-dimensional space (in which case the curvature κ is replaced by the mean curvature H).

2.4.6 Mumford-Shah Framework

Mumford and Shah [80] define segmentation as a joint smoothing and edge detection problem. In their framework, given an image $I(\mathbf{p}) : \mathfrak{D} \subset \mathbb{R}^2 \rightarrow \mathfrak{C}$, the goal is to find a set of discontinuities $K \subset \mathfrak{D}$, and a piecewise smooth image $S(\mathbf{p})$ on $\mathfrak{D} \setminus K$ that minimize

$$\mathcal{E}_I(S, K) = \int_{\mathfrak{D} \setminus K} \left\{ \|\nabla S\|^2 + (I - S)^2 \right\} \, \mathrm{d}\mathbf{p} + \mathcal{L}(K), \tag{47}$$

where $\mathcal{L}(K)$ is the Euclidean length, or more generally, the 1-dimensional Hausdorff measure of K.

The first term ensures that any minimizer S is smooth (except across edges), while the second term ensures that minimizers approximate I. The last term will cause the set K to be regular. It is interesting to note as argued in [57] that many segmentation algorithms (including some of the most common) can be formulated in the Mumford-Shah framework. Further, the Mumford-Shah functional can be given a natural Bayesian interpretation [81].

The functional itself is very difficult to analyze mathematically even though there have been some interesting results. The book [57] gives a nice survey on the mathematical results concerning the Mumford-Shah functional. For example, Ambrosio [82] has found a weak solution to the problem in the class of Special Bounded Variation functions. The functional itself has influenced several different segmentation techniques some connected with active contours including [83, 84].

CHAPTER 3

REGION-BASED IMAGE SEGMENTATION

In this chapter, we consider the problem of finding the boundaries of only one region with limited user interaction. Interactivity is desirable since the user is given the opportunity to make use of important implicit external knowledge to guide the algorithm toward a result that makes sense for her task. The segmentation process can be repeated in order to identify as many regions as necessary. This chapter is based on [85, 79].

After briefly reviewing techniques previously proposed for image segmentation (Section 3.1), we present a unified framework for a large class of region-based image segmentation techniques (Section 3.2). We then propose a new variational method for region based segmentation based on non-parametric statistics (Section 3.3), provide full mathematical details on the fundamental flow and the non-parametric estimation of image statistics (Sections 3.4 and 3.5) and discuss how this algorithm has been implemented in the open-source software 3D Slicer (Section 3.6). We propose a novel validation framework (Section 3.7) based on the Euclidean distance and use it to analyze the performance of the proposed algorithm both on simulated and manually segmented images (Section 3.8).

The problem of validating the result of a segmentation algorithm is analogous to the problem of comparing the output of that algorithm with some optimal answer called *ground truth*. This can be either known by construction if an artificial phantom is imaged in the sole purpose of validating the segmentation procedure or determined by some authoritative source, typically some human expert. Quantifying the differences between surfaces will be treated in more detail and sophistication in Chapter 5.

3.1 Previous Work

Many different approaches have been proposed to address the problem of region segmentation which can be dually considered as finding regions or finding boundaries (see [86] and

references therein). Focusing only on the boundaries is computationally less complex but also less robust since information inside the region is discarded. This is the approach of the snakes and active contours variational methods [87, 69, 88]. An approach that can be seen as intermediate (i.e., in between region-based and surface-based techniques will be proposed in Chapter 4).

While the original region-growing algorithm [1] formalism is somewhat crude, interesting extensions have been proposed by Adams [89] where some statistical information is derived from the region as it expands. These techniques have been applied to medical image analysis [90, 91]. The relation between region-growing and active contours has been studied by Zhu [92] and more recently active contours have been extended by Chan [84] and Paragios [93] to an elegant active regions formalism where regions boundaries are deformed according to an evolution equation derived to minimize an energy based on statistics on the regions.

We refer the interested reader to Section 2.4 for a more detailed description of variational segmentation algorithms.

3.2 Image Segmentation Techniques based on Volume, Mean and Variance

3.2.1 General Framework

Given a region $R \subset \mathfrak{D}$ and a scalar image $I : \mathfrak{D} \to \mathbb{R}$ defined on a domain $\mathfrak{D}$, we can define the volume (area for 2-dimensional domains, volume for 3-dimensional domains etc.) $v(R)$ of the region R as well as the volume $\bar{v}(R)$ of the complement $\bar{R} = \mathfrak{D} \setminus R$.

In a similar manner, the mean intensity $\mu(I, R)$ of the image *inside* the region can be defined as well as the mean intensity $\bar{\mu}(I, R)$ of the image *outside* the region.

Any higher order moment can also be defined inside and outside such as the variances $\sigma^2(I, R)$ and $\bar{\sigma^2}(I, R)$.

By combining these scalar quantities through a given function $\mathcal{F} : \mathbb{R}^6 \to \mathbb{R}$ we obtain a very general class of region energies of the form

$$\mathcal{E}(I, R) = \mathcal{F}(v, \mu, \sigma^2, \bar{v}, \bar{\mu}, \bar{\sigma^2}). \tag{48}$$

This kind of functionals have been prviously proposed by Chan et al. [84], Yezzi et al. [94] among others.

3.2.2 Minimizing Gradient Flow

The minimizing gradient flow for such an energy is readily obtained by composition.

$$\frac{\partial \Sigma}{\partial t} = -\nabla \mathcal{E} = -\{ \frac{\partial \mathcal{F}}{\partial v}\nabla v + \frac{\partial \mathcal{F}}{\partial \bar{v}}\nabla \bar{v} + \frac{\partial \mathcal{F}}{\partial \mu}\nabla \mu + \frac{\partial \mathcal{F}}{\partial \bar{\mu}}\nabla \bar{\mu} + \frac{\partial \mathcal{F}}{\partial \sigma^2}\nabla \sigma^2 + \frac{\partial \mathcal{F}}{\partial \bar{\sigma^2}}\nabla \bar{\sigma^2} \} \tag{49}$$

In order to compute the gradient flow of the volumes and moments, we introduce the notation $S(\alpha, R) = \int_R \alpha \, \mathrm{d}R$ where α is a scalar field defined on the same domain $\mathfrak{D}$ as the image.

As explained in Section 3.4, the *minimizing* gradient flow for $S(\alpha, R)$ is

$$\frac{\partial \Sigma}{\partial t} = -\nabla S(\alpha, R) = -\alpha(\Sigma)\mathbf{N}, \tag{50}$$

where the surface $\Sigma = \partial R$ is the boundary of the region R and $\mathbf{N}$ is its outward normal.

The minimizing gradient flows of the volumes and moments are then readily computed.

3.2.2.1 Volume v

The case of the volume corresponds to $\alpha = 1$, i.e., $v(R) = S(1, R)$. The gradient flow ∇v is therefore simply the evolution in the normal direction $\Sigma_t = -\nabla v = -\mathbf{N}$.

Symmetrically, the complement $\bar{R} = \mathfrak{D} \setminus R$ has the same boundary Σ but the normal is flipped. Using the normal $\mathbf{N}$ of R, we obtain $\Sigma_t = -\nabla \bar{v} = \mathbf{N}$.

3.2.2.2 Mean μ

The mean can be expressed by $\mu(I,R) = S(I,R)/S(1,R)$. Once again the gardient flow can be obtained by composition of known results.

$$\begin{aligned}
\frac{\partial \Sigma}{\partial t} &= -\nabla \mu(I,R) \\
&= -\nabla(\frac{S(I,R)}{S(1,R)}) \\
&= -\frac{S(1,R)\nabla S(I,R) - S(I,R)\nabla S(1,R)}{S(1,R)^2} \\
&= -\frac{vI(\Sigma)\mathbf{N} - S(I,R)\mathbf{N}}{v^2} \\
&= -\frac{I(\Sigma)-\mu}{v}\mathbf{N}
\end{aligned}$$

Symmetrically, $-\nabla\bar{\mu} = +(I(\Sigma)-\bar{\mu})/\bar{v}\mathbf{N}$.

3.2.2.3 Variance σ^2

The variance can be expressed by:

$$\begin{aligned}
\sigma^2(I,R) =& S((I-\mu)^2,R)/S(1,R) \\
=& S(I^2 - 2\mu I + \mu^2, R)/S(1,R)
\end{aligned}$$

using the linearity of the integration,

$$\begin{aligned}
=& [\ S(I^2,R) - 2\mu S(I,R) + \mu^2 S(1,R)\ ]/S(1,R) \\
=& S(I^2,R)/S(1,R) - 2\mu^2 + \mu^2 \\
=& S(I^2,R)/S(1,R) - \mu^2 \\
=& S(I^2,R)/S(1,R) - [\ S(I,R)/S(1,R)\ ]^2
\end{aligned}$$

Once again the gradient flow can be obtained by composition of known results.

$$\begin{aligned}\frac{\partial \Sigma}{\partial t} &= -\nabla\sigma^2(I,R) \\ &= -\nabla\frac{S(I^2,R)}{S(1,R)} + \nabla(\frac{S(I,R)}{S(1,R)})^2 \\ &= -\frac{S(1,R)\nabla S(I^2,R) - S(I^2,R)\nabla S(1,R)}{S(1,R)^2} + 2\frac{S(I,R)}{S(1,R)}\nabla\frac{S(I,R)}{S(1,R)} \\ &= -\frac{vI^2\mathbf{N} - S(I^2,R)\mathbf{N}}{v^2} + 2\mu\nabla\mu \\ &= -\frac{vI^2\mathbf{N} - S(I^2,R)\mathbf{N}}{v^2} + 2\mu\frac{(I(\Sigma)-\mu)}{v}\end{aligned}$$

using again $S(I^2,R)/v = \sigma^2 + \mu^2$,

$$\begin{aligned}&= -\frac{I^2 - (\sigma^2+\mu^2) - 2\mu I + 2\mu^2}{v}\mathbf{N} \\ &= -\frac{(I-\mu)^2 - \sigma^2}{v}\mathbf{N}\end{aligned}$$

Symmetrically, $-\nabla\bar{\sigma^2} = +[\ (I-\bar{\mu})^2 - \bar{\sigma^2}\ ]/\bar{v}\mathbf{N}$.

3.2.3 Example

As an illustration, consider the energy

$$\mathcal{E}(I,R) = f(v)\sigma^2 \tag{51}$$

corresponding to the function $\mathcal{F}(v,\mu,\sigma^2,\bar{v},\bar{\mu},\bar{\sigma^2}) = f(v)\sigma^2$.

The resulting minimizing gradient flow is

$$\frac{\partial \Sigma}{\partial t} = -(f'(v)\sigma^2 + f(v)\frac{(I-\mu)^2 - \sigma^2}{v})\mathbf{N} \tag{52}$$

Since the artificial time parameter is not meaningful (we are only interested in the steady state), this gradient can be multiplied by any scalar function $g(t)$ (which would correspond to a change of the parameter t). Therefore the minimizing surface evolution can be written:

$$\frac{\partial \Sigma}{\partial t} = -\frac{vf'(v)}{f(v)}\mathbf{N} - (\frac{(I-\mu)^2}{\sigma^2} - 1)\mathbf{N} \tag{53}$$

The first term will tend to expand the region if $f'(v) < 0$, i.e., if increasing the volume v would reduce $f(v)$. The second term will tend to expand the region to pixels $\mathbf{p}$ for which

$(I(\mathbf{p}) - \mu)^2/\sigma^2 - 1 < 0$, i.e., pixels whose intensity are less than one standard deviation away from the mean of the region.

3.2.4 Implementation

If the surface evolution is implemented in a levelset framework, it is possible to test if a pixel $\mathbf{p}$ is inside or outside of the region R using the sign of the levelset function $\phi(\mathbf{p})$ at this point. For example, here we assume that $\phi \leq 0$ in R.

Given an initial region, the inside and outside volumes and statistics can be easily computed. If narrow-band implementation of the levelset framework is used then not all pixels are visited at each iteration and the naive approach of recomputing the moments at each iteration would prove very expensive.

Instead, the moments can be updated at each iteration using only information from pixels that happen to change sign in the narrow band. In order to do that, the moments and volume have to be computed globally once and four accumulators $s_1, s_2, \bar{s}_1$ and $\bar{s}_2$ are initialized to 0. Then, as the levelset values are being recomputed for each pixel $\mathbf{p}$ of the narrow band, those accumulators are updated using Algorithm 1. Before starting a new iteration through the narrow band, the moments are updated and the accumulators reset to 0 using Algorithm 2.

While this technique is simple to implement and very robust with respect to image noise, two problems are that:

- it is based on parametric statistics (the mean and variance)
- it is based on the idea that "inside" and "outside" are binary characteristics

The use of parametric statistcs (such as the variance inside a region) can be justified by the idea that, in an ideal image, intensities inside the region R^* to be segmented would be uniform. If that image is corrupted by Gaussian white noise, then intensities inside the region will therefore be Gaussian distributed. The variance inside some region R will therefore be minimal when $R = R^*$ and the variance would then be the variance of the noise. This can be problematic if the region to be detected is not uniform (for example

Algorithm 1: Updating accumulators

Require: image I, current and previous levelset functions ϕ^{n+1} and ϕ^n at pixel $\mathbf{p}$

if $(\phi^n(\mathbf{p}) > 0)$ **and** $(\phi^{n+1}(\mathbf{p}) \leq 0)$ **then**

$t = I(\mathbf{p}) - \mu$

$s_1 \leftarrow s_1 + t$

$s_2 \leftarrow s_2 + (t^2 - \sigma^2)$

$v \leftarrow v + 1$

$\bar{t} = I(\mathbf{p}) - \bar{\mu}$

$\bar{s}_1 \leftarrow \bar{s}_1 - \bar{t}$

$\bar{s}_2 \leftarrow \bar{s}_2 - (\bar{t}^2 - \bar{\sigma^2})$

$\bar{v} \leftarrow \bar{v} - 1$

else if $(\phi^n(\mathbf{p}) \leq 0)$ **and** $(\phi^{n+1}(\mathbf{p}) > 0)$ **then**

$t = I(\mathbf{p}) - \mu$

$s_1 \leftarrow s_1 - t$

$s_2 \leftarrow s_2 - (t^2 - \sigma^2)$

$v \leftarrow v - 1$

$\bar{t} = I(\mathbf{p}) - \bar{\mu}$

$\bar{s}_1 \leftarrow \bar{s}_1 + \bar{t}$

$\bar{s}_2 \leftarrow \bar{s}_2 + (\bar{t}^2 - \bar{\sigma^2})$

$\bar{v} \leftarrow \bar{v} + 1$

end if

Algorithm 2: Updating moments

Require: accumulators $s_1, s_2, \bar{s}_1$ and $\bar{s}_2$

$\mu \leftarrow \mu + s_1/v$

$\sigma^2 \leftarrow \sigma^2 + (s_2/v - [s_1/v]^2)$

$s_1 \leftarrow 0$

$s_2 \leftarrow 0$

$\bar{\mu} \leftarrow \bar{\mu} + \bar{s}_1/\bar{v}$

$\bar{\sigma^2} \leftarrow \bar{\sigma^2} + (\bar{s}_2/\bar{v} - [\bar{s}_1/\bar{v}]^2)$

$\bar{s}_1 \leftarrow 0$

$\bar{s}_2 \leftarrow 0$

when trying to detect the countour of a person with a shirt and pants of different colors) or if the additive noise is not Gaussian. To aleviate this problem another technique is proposed in Section 3.3 that is based on the full distribution of intensities inside (and possibly also outside) the region and not only a few scalar parameters such as the mean and variance.

Another weakness of the technique is that it does not differentiate between image information that is, for example, "slightly outside" the region R and image information that is "completely outside" the region R. Inside and outside are just binary (mutually exclusive) properties. A different (surface-based) framework is proposed in Chapter 4 that handles that difficulty (see in particular the discussion in Section 4.4).

3.3 Non-parametric, Region-based Functional

In this section we briefly present a very general flow for image segmentation. This technique is based purely on the statistics of the image.

In particular it does does not necessitate external information (such as an atlas) nor makes extra anatomy or modality-based assumptions. This method is therefore extremely versatile and we will show in Section 3.8 that it can compete with more specialized approaches.

Given an image I and a region R, using Bayes' rule,

$$\begin{aligned} & P(\mathbf{x} \in R \mid I(\mathbf{x}), \|\nabla I(\mathbf{x})\|) \\ =& P(I(\mathbf{x}), \|\nabla I(\mathbf{x})\| \mid \mathbf{x} \in R) \\ & \times \frac{P(\mathbf{x} \in R)}{P(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)}. \end{aligned}$$

Assuming uniform priors $P(\mathbf{x} \in R)$ and $P(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)$ the likelihood $P(I(\mathbf{x}), \|\nabla I(\mathbf{x})\| \mid \mathbf{x} \in R)$ and the posterior distributions $P(\mathbf{x} \in R \mid I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)$ are proportional. We use the notation $P_R(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)$ for either the likelihood or the posterior. This can be justified by the facts that:

- the quantity $P(\mathbf{x} \in R)$ is not known.

- using $P(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)$ would introduce undesired global information. The existence in the dataset of another unrelated region with statistics similar to region R should not have any influence[1].

We accordingly define the energy functional

$$E(I, R) := \int_R P_R(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|) \, \mathrm{d}\mathbf{x}. \tag{54}$$

Here, $E(I, R)$ is the volume of the region R where each voxel is weighted by the probability $P_R(I(\mathbf{x}), \|\nabla I(\mathbf{x})\|)$ of the intensity and the norm of the gradient of I at this voxel. Likely voxels therefore contribute more to $E(I, R)$ than unlikely voxels and the energy of a region R will be high if and only if its voxels have consistent values in terms of intensity and norm of the gradient.

An initial region R_0 at $t = 0$ can be deformed into a region $R(t)$ to maximize $E(I, R)$. We show in Section 3.4 that the gradient ascent flow is

$$\frac{\partial S}{\partial t} = P_R(I, \|\nabla I\|)\mathbf{N}, \tag{55}$$

where $S(t) = \partial R(t)$ the boundary of R at time t and $\mathbf{N}$ is the unit outward normal.

As the region is deformed, P_R is estimated in a non-parametric fashion as detailed in Section 3.5.

3.4 Fundamental Flow

In what follows we only consider the three-dimensional case. The region R is an open connected bounded subset of $\mathbb{R}^3$ with smooth boundary $S = \partial R$ and $\mathbf{N}$ denotes the corresponding outward unit normal vector to S.

Given an image I, a non-negative weighting function $w(\cdot, \cdot)$ and a region R we define the energy

$$E(I, w, R) := \int_R w(\, I(\mathbf{x}),\ \|\nabla I(\mathbf{x})\|\ ) \ \mathrm{d}\mathbf{x}. \tag{56}$$

[1]That would be acceptable in the ideal case of a bimodal image. Most medical datasets are composed of more than two classes.

E is the weighted volume of the region R. The weight of a voxel $\mathbf{x}$ is determined by the function $w(\cdot,\cdot)$ of the local properties $I(\mathbf{x})$ and $\|\nabla I(\mathbf{x})\|$ of the image. Ideally, w should reflect the local properties of the region to be segmented. As this is not known a priori, w it is heuristically estimated as R is evolved in order to maximize E.

Proposition 3.1. *Notation as above. Then for a given weighting function w, the evolution in which the energy $E(I, w, R)$ is increasing as fast as possible (using only local information) is* $\dfrac{\partial S}{\partial t} = w\mathbf{N}$.

Proof. Let $\psi^t : R \to \mathbb{R}^n$ be a family of embeddings, such that ψ^0 is the identity. Let $w : \mathbb{R}^n \to \mathbb{R}$ be a positive C^1 function. We set $R(t) := \psi^t(R)$ and $S(t) := \psi^t(\partial R)$. We consider the family of w-weighted volumes

$$\begin{aligned} H(t) &:= \int_R w(\psi^t(\mathbf{x}))\ \mathrm{d}\psi^t(\mathbf{x}) \\ &= \int_{R(t)} w(\mathbf{y})\ \mathrm{d}\mathbf{y}. \end{aligned}$$

Set $X = \frac{\partial \psi^t}{\partial t}|_{t=0}$ then using the area formula [77] and then by the divergence theorem, the first variation is

$$\begin{aligned} \frac{dH}{dt}\,\big|_{\,t=0} &= \int_R \mathrm{d}iv(wX)\ \mathrm{d}\mathbf{x} \\ &= -\int_{\partial R} (wX)\cdot\mathbf{N}\ \mathrm{d}\mathbf{y}, \end{aligned}$$

where $\mathbf{N}$ is the inward unit normal to ∂R. Consequently the corresponding w-weighted volume maximizing flow is

$$\frac{\partial S}{\partial t} = w\mathbf{N}.$$

A different derivation of the same result has previously been proposed by Siddiqi [78]. □

Since w is a non-negative function, the flow is reversible. In particular, the flow in the reverse direction,

$$\frac{\partial S}{\partial t} = -w\mathbf{N}, \tag{57}$$

gives the direction in which the energy is increasing as fast as possible (using local information). In the context of segmentation, one may think of (57) as a *bubble* and of the original flow as a *snake*.

Given an approximation R_0 of the region to be segmented, a maximum likelihood-like approach can be used to determine the weighting function w_0, which would a posteriori justify the segmentation of R_0.

Proposition 3.2. *For a given fixed region R_0, the energy $E(I, w, R_0)$ is maximized by setting w to p_{R_0} the conditional probability on that region*

$$\begin{aligned} p_{R_0} &= \arg \max_{w, \|w\|_1 = 1} E(I, w, R_0) \\ &= P(\ I(\mathbf{x}),\ \|\nabla I(\mathbf{x})\| \mid \mathbf{x} \in R_0\ ). \end{aligned} \tag{58}$$

Proof. We can rewrite the energy as

$$E(I, w, R_0) = \int_I \int_{\|\nabla I\|} N_{R_0}(u, v).w(u, v)\ \ \mathrm{d}u\ \mathrm{d}v,$$

where $N_{R_0}(u, v)$ is the volume of the set of points $\mathbf{x} \in R_0$ such that $I(\mathbf{x}) = u$ and $\|\nabla I(\mathbf{x})\| = v$. But this is just a constant multiple of $p_{R_0} = P(\ I(\mathbf{x}),\ \|\nabla I(\mathbf{x})\| \mid \mathbf{x} \in R_0\ )$ which is therefore by the Schwartz inequality is the maximizer of E. □

As the region evolves, w is periodically updated according to (58). This changes the definition of the energy (56) and therefore (57) can only be considered a gradient flow for every time interval when w is fixed.

3.5 Non-parametric Estimation of Image Statistics

Instead of using the distribution $p_{R_0} = P(\ I(\mathbf{x}),\ \|\nabla I(\mathbf{x})\| \mid \mathbf{x} \in R_0\ )$ as described in (58) we use $p = p_M \cdot p_H$ where M and H are the median and interquartile range (the difference between the first and last quartile) operators on a $3 \times 3 \times 3$ neighborhood. M and H convey about the same information as I (gray level) and $\|\nabla I\|$ (local homogeneity). For example if $\|\nabla I\|$ is large then values in a $3 \times 3 \times 3$ neighborhood are very dispersed and therefore the interquartile range is large. These measures were chosen primarily because they are robust to noise[2] and they respect edges[3] of the image better than their linear counterparts.

[2] the interquartile range and the median are completely insensitive to isolated outliers

[3] see [1] for details on median filtering and its advantages

Parzen windows (see for example [95]) are used to estimate the probability density functions. This is a non-parametric technique and therefore no assumption is required on the shape of the distributions. Given a window function ϕ and N samples $m_1, \ldots, m_N$ and $h_1, \ldots, h_N$ the densities are estimated by:

$$\hat{p}_M(m) = \frac{1}{N} \sum_{i=1}^{N} \phi(m - m_i)$$
$$\hat{p}_H(h) = \frac{1}{N} \sum_{i=1}^{N} \phi(h - h_i).$$

This corresponds to convolving the samples histogram with ϕ. It can be shown that the estimates $\hat{p}_M$ and $\hat{p}_H$ converge toward the true estimates p_M and p_H with $n \to \infty$ and $\phi \to \delta$. A centered Gaussian kernel $\phi = g_\sigma$ of standard deviation $\sigma = \hat{\sigma}_H/10$ was used to estimate p_H and $\sigma = \hat{\sigma}_M/10$ to estimate p_M.

3.6 Implementation

We implemented our method as a module of the open-source software 3D Slicer. It is freely available at `http://www.slicer.org`. Thanks to the properties of our flow, we were able to use the very efficient Fast Marching method for evolving the surface. Segmenting a large structure typically[4] takes less than one minute.

The flow (55) is unidirectional (the surface can only expand since $p_R \geq 0$) any voxel $\mathbf{x}$ is eventually reached at a time $T(\mathbf{x})$. Knowing T is equivalent to knowing R or S since by definition

$$\begin{cases} R(t) = \{\ \mathbf{x},\ T(\mathbf{x}) \leq t\ \} \\ S(t) = \partial R(t). \end{cases} \tag{59}$$

Solving the flow (55) for $S(t)$ is equivalent to solving for $T(\mathbf{x})$ the Eikonal equation

$$\|\nabla T(\mathbf{x})\| \cdot\ p_R(I(\mathbf{x})) = 1. \tag{60}$$

This can be done very efficiently using the Fast Marching method [96, 23]. Starting from known seed points which define the initial surface, the algorithm marches outward by

[4] 3 GHz processor, 1 GB memory

considering neighboring voxels and iteratively computing arrival times T in increasing order. The seed points are set by the user inside the structure to be segmented. By construction, when computing $T(\mathbf{x})$, the surface contains the voxel $\mathbf{x}$ as well as all voxels for which T has already been computed. The algorithm terminates when T is known for all points. Then using (59), $S(t)$ can be determined for any t and let the user determine what time t_0 of the evolution corresponds best to the region she wants.

Note that our method is, in its implementation, reminiscent of region growing. The min-heap data structure which makes Fast Marching efficient is the direct equivalent of the sequentially sorted list in the seeded region growing algorithm [89]. In fact our algorithm could be made a direct non-parametric extension of seeded region growing simply by artificially forcing arrival times to zero for all points inside the surface S. Relations between region growing and variational schemes have been previously exposed by Zhu [92].

3.7 A Euclidean Distance-based Validation Framework

Objective and quantitative analysis of performance is absolutely crucial (but often overlooked) when proposing a segmentation algorithm. Since designing a segmentation method is challenging (lack of unifying formalism, high diversity in the applications, subjectivity, implicitness, etc.) it does not come as a surprise that the validation of such an algorithm is also challenging. Different methods have been studied (see [97] and references therein). We propose a unifying framework for discrepancy measures based on the number and the position of mis-segmented voxels and show how it relates to classical measures. We then apply it to the validation of segmentation of realistic synthetic images (for which the "ground truth", i.e. perfect segmentation is known) at different levels of noise for accuracy and robustness assessment as well as to manual expert segmentation of real datasets.

Quantifying the differences between surfaces will be treated in more detail and sophistication in Chapter 5.

3.7.1 Classical Discrepancy Measures

Different measures have been proposed to assess the resemblance between a proposed segmentation S and the corresponding ground truth G. The Dice Similarity Coefficient has

been widely used and it can be derived as an approximation of the kappa statistic (a chance-corrected measure of agreement, see [98]). It is defined as

$$DSC(S,G) := \frac{V_{S\cap G}}{\frac{1}{2}(V_S + V_G)},$$

where V_X is the volume (number of voxels) of set X.

One disadvantage of this coefficient is that it only takes into account the number of mis-segmented voxels and disregards their position and therefore the severities of errors. This was corrected in Yasnoff's discrepancy measure DM [99] and the Factor of Merit FOM [100]:

$$\mathrm{DM} := \frac{1}{N}\sum_{i=1}^{N} d(i)^2$$

$$\mathrm{FOM}_e := \frac{1}{N}\sum_{i=1}^{N} \frac{1}{1+d(i)^2},$$

where N is the number of mis-segmented voxels and $d(i)$ is the Euclidean distance from the i^{th} voxel to the ground truth. Another popular measure is the Hausdorff distance

$$H(S,G) := \max\{\ \max_{s\in S}\min_{g\in G}\|s-g\|,$$

$$\max_{g\in G}\min_{s\in S}\|s-g\|\ \}.$$

$H(S,G)$ is the maximum distance one would have to move the boundaries of one set so that it would encompass completely the other set. As this is extremely sensitive to extreme errors, the partial Hausdorff distance [101] $H_f(S,G)$ can be introduced as the maximum distance one would have to move the boundaries of one set so that it would cover $f\%$ of the other set.

3.7.2 Proposed Framework

Consider now the error-distance:

$$d(\mathbf{x}) := \begin{cases} 0 & \mathbf{x}\in S\cap G \\ \min_{s\in S}\|\mathbf{x}-s\| & \mathbf{x}\in G\backslash S \\ \min_{g\in G}\|\mathbf{x}-g\| & \mathbf{x}\in S\backslash G \end{cases}$$

Assuming that all points $\mathbf{x} \in S \cup G$ are equally likely d can be seen as a random variable D which describes completely the discrepancy between the segmentation S and the ground-truth G. Using basic statistical tools we can define the probability of error (PE), mean of errors[5] ($\mu_{D>0}$), standard deviation of errors ($\sigma_{D>0}$) and partial distance-error (D_f) by:

$$\begin{aligned} \text{PE} &:= Pr(D > 0) \\ \mu_{D>0} &:= \text{mean}(D \mid D > 0) \\ \sigma_{D>0} &:= \text{stdev}(D \mid D > 0) \\ D_f &:= f\text{-quantile}(D). \end{aligned}$$

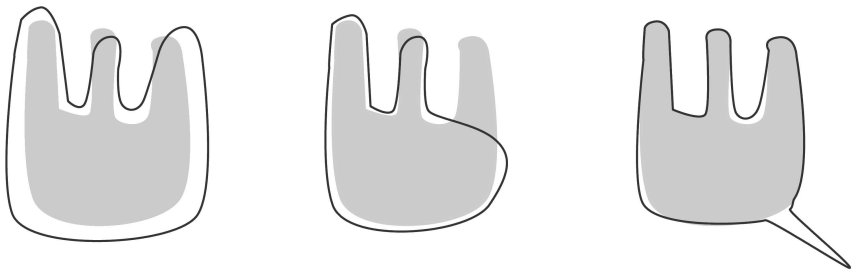
These measures have a natural intuitive interpretation:

- PE is the probability for a voxel $\mathbf{x} \in S \cup G$ to be misclassified (either over- or under-segmented, i.e. $(\mathbf{x} \in S \cup G) \backslash ((S \cap G))$).
- An erroneous voxel is on average $\mu_{D>0}$ pixels off. This value is very typical if the standard deviation $\sigma_{D>0}$ is small.
- D_{1-f} is the error distance of the worst $f\%$ voxels. For example $D_{0.5}$ is the median of errors. Equivalently the maximum distance one would need to move erroneous voxels for the error to be improved to PE $= f$.

As an example, PE $= 10\%$, $\mu_{D>0} = 3.1$, $\sigma_{D>0} = 0.3$ and $D_{0.99} = 14$ would mean that the overlap between the ground truth and the proposed segmentation is 90%. The 10% remaining pixels are either under-segmented or over-segmented pixels ("false positive" i.e. pixels that are in S and not in G). On average these pixels are 3.1 pixels off. This value is very typical since the standard variation is low (0.3). However there is no reason for the error to be Gaussian and, here, the tail probability is not negligible since the worst 1% pixels are at least 14 pixels off. This could be due to a thin, long finger of mis-segmented pixels.

[5]Note: this is the mean of errors and not the mean error. Valid points are not taken into account at all.

Figure 8 illustrates 3 different cases of mis-segmentation. Figures 8(a) and 8(b) have approximately the same probability of error PE (and therefore the same DSC (see Equation 61) but 8(a) has a lower $\mu_{D>0}$ and partial distance error $D_{0.95}$. This is due to the fact that even though 8(a) and 8(b) have roughly the same number of mis-segmented pixels, the errors tend to be more severe in 8(b). Figure 8(c) illustrate the case of a low probability of error PE and a high $\mu_{D>0}$. This might seem counter-intuitive. $\mu_{D>0}$ is the mean of mis-classified points. Here most points are correctly classified and the few points that are not are rather far off which explains a high $\mu_{D>0}$. Moreover the standard deviation of errors $\sigma_{D>0}$ is lower than in 8(b) since there are less small errors and therefore $\mu_{D>0}$ is typical. Depending on the end-task 8(a) might be a better segmentation than 8(b) or not and any of the above mentioned measures might be the most important metric.



(a) (b) (c)

Figure 8: These synthetic examples illustrate different kinds of mis-segmentation. (ground truth G in gray and segmentation S in black)

These measures are related to the measures presented in Section 3.7.1 according to:

$$1 - \mathrm{DSC} \leq \mathrm{PE} = \frac{1 - \mathrm{DSC}}{1 - \frac{\mathrm{DSC}}{2}} \tag{61}$$

$$\frac{1}{\mathrm{FOM}_e} - 1 \leq \mu_{D>0}^2 + \sigma_{D>0}^2 = \mathrm{DM} \tag{62}$$

$$H_{1-\frac{f}{1-\mathrm{PE}}} \leq D_{1-f} \leq H_{1-\frac{f}{2}} \tag{63}$$

$$\text{(in particular, } D_1 = H\text{)}$$

Proof.

- (61): let a, b, c, d be the true positive, true negative, false negative and false positive probabilities respectively then

$$\text{PE} = (c + d)/(a + c + d)$$

$$\text{DSC} = 2a/[(a + c) + (a + d)].$$

Simple algebra yields the equality. The inequality follows immediately from the fact that $0 < 1 - \text{DSC}/2 \leq 1$.

- (62): By definition and implying condition $D > 0$ for all expectancies

$$\begin{aligned}\text{DM} &= E\{D^2\} \\ &= E\{[D - E\{D\}]^2\} + E\{D\}^2 \\ &= \sigma^2_{D>0} + \mu^2_{D>0}\end{aligned}$$

which is the equality.

For the inequality consider $Y = D^2$ and $f(Y) = 1/(1 + Y)$. $f'' \geq 0$ and therefore f is convex (on $\mathbb{R}^+$). But FOM is $E\{\ f(Y)\ \}$. By convexity $E\{\ f(Y)\ \} \geq f(E\{\ Y\ \})$. Since $f^{-1}(x) = 1/x - 1$ is decreasing $f^{-1}(\text{FOM}) \leq E\{\ Y\ \} = (\mu^2_{D>0} + \sigma^2_{D>0})$.

- (63): let $D_S = \{d(x, G), x \in S\}$ and $D_G = \{d(x, S), x \in S\}$ be the distance of all points of one set to the other set. Consider that D_S and D_G are ordered such that value at rank 0 is the minimum. Then the partial Haussdorff distance H_{1-f} is the max of the values at index (fV_S) and (fV_G) in D_S and D_G. Consider $D_{S\cup G}$ to be the values of D_S and D_G where the points corresponding to $S \cap G$ (those values are all 0 by construction) are counted only once. We also consider that $D_{S\cup G}$ is sorted. Then the proposed metric D_{1-f} is the element at index $(fV_{S\cup G})$. We know that

$$fV_{S\cup G} \geq \frac{f}{2}(V_S + V_G).$$

This means that the value at index $(fV_{S\cup G})$ of $D_{S\cup G}$ has to be smaller or equal to the largest of the value at index $(\frac{f}{2}V_S)$ of D_S and the value at index $(\frac{f}{2}V_G)$ of D_G.

The equality occurs only when these two values are equal. This proves the right side of (63). For the left side, use

$$\begin{aligned} fV_{S\cup G} &= \frac{f}{1-\mathrm{PE}} V_{S\cap G} \\ &\leq \frac{f}{1-\mathrm{PE}} \min\{V_S, V_G\}. \end{aligned}$$

□

3.8 Validation Results

We evaluated the validation algorithm proposed in Section 3.3 using the validation framework proposed in Section 3.7 on 2 simulated and 10 real MRI brain datasets. It is fundamental to understand that the proposed algorithm is very general. In particular it is not designed or tuned for any particular structure. Other approaches have been proposed that necessitate (and take advantage of) prior information. For example in [102] a model of brain structures (i.e., an atlas) is deformed to match the dataset; in [69, 103], the white and gray matter of the brain are segmented using special geometric constraints based on the neuroanatomical knowledge that the thickness of the cortical mantle is nearly constant; in [104] the datasets are pre-processed to compensate for bias fields in the MR images and non-brain tissue is removed. Our framework is more general and does not require external information or make extra assumptions on the anatomical region to be segmented or the imaging modality. Because implementations of previously proposed segmentation techniques are typically not publicly available and these algorithms were typically not validated on publicly available datasets it is difficult to quantitatively compare performances. In contrast an implementation of our technique is freely available (as part of the open-source software 3D Slicer). The technique has been validated on publicly available simulated and real datasets.

The work of Shattuck [104] is a notable exception since it has been validated on publicly available images. We will show in Section 3.8.1 that even though it is very general and assumption free, the performance of the proposed technique is comparable to this more specialized approach.

3.8.1 Simulated Datasets (N=2)

The *Brain Web* datasets have been generated from a known ground truth using a physical modeling of the MRI process [105]. The performance of our method can be assessed in a perfectly objective way by comparing the result of our segmentation with the underlying ground truth. Note that even though these datasets are computer-generated they are very realistic (see figure 9(b)) Another interesting aspect of this project is that from the same ground truth, datasets with different levels of noise can be simulated which allows us to study the robustness of our method with respect to noise. Using the proposed framework, the authors segmented the lateral ventricle, white matter (WM) and white matter and gray matter (WM+GM) on 2 datasets:

- Normal brain, T1, $1 \times 1 \times 1$ mm ($181 \times 181 \times 217$ voxels), 3% noise, 20% intensity non-uniformity ("RF") (standard parameters of the *Brain Web* model).
- Normal brain, T1, $1 \times 1 \times 1$ mm ($181 \times 181 \times 217$ voxels), 9%, 40% (highest levels of noise available).

The results (Table 1) show that the proposed algorithm gives very good results on these structures (according to [98] $DSC > 0.7$ is regarded as good agreement in the literature). The complex structure of the white matter makes its segmentation more challenging and explains the somewhat mediocre performance (in the case of the maximum noise dataset, the cerebellum was not perfectly segmented).

In the highest level of noise, connectivity between the lateral and the third ventricles was lost (the intraventricular foramen of Monro disappeared in the noise). This increased the strength of the ventricle edges in the noisy dataset and, paradoxically, simplified the segmentation. Overall the algorithm appears extremely robust to noise.

On the same datasets, Shattuck [104] reports[6] $DSC = 93\%$ (standard noise) and $DSC = 81\%$ (maximum noise) for the white matter. These scores are slightly better than our own results ($DSC = 91.9\%$ and $DSC = 80.3\%$ respectively). However it is very important to keep in mind that there is a trade-off between performance on a specific problem and

[6] scores are the better of ML and MAP approaches, with no bias fields compensation

versatility. Our technique was not created specifically for white matter extraction and, unlike more specialized techniques, it can be used for a very wide variety of structures and modalities. In summary, the wide applicability of the proposed technique comes at a minor performance cost vis-à-vis the work of Shattuck [104].

Table 1: Performance measure of proposed technique on artificial dataset. Left bold, with standard noise, right, with maximum noise

	DSC	PE	$\mu_{D>0}$	$\sigma_{D>0}$	$D_{0.95}$	$D_{0.99}$
Ventricle	**92.0%** 95.1%	**14.9%** 9.4%	**1.07** 1.13	**0.48** 0.61	**1.00** 1.00	**1.00** 1.41
WM	**91.9%** 80.3%	**15.0%** 32.0%	**1.59** 2.03	**1.58** 1.94	**1.00** 2.83	**3.61** 8.25
WM+GM	**96.2%** 95.2%	**7.4%** 9.2%	**1.42** 1.40	**1.25** 1.15	**1.00** 1.00	**1.41** 2.00

3.8.2 Real Datasets (N=10)

In this section we use, as the ground truth, the expert manual segmentations of 10 full brains and brain tumors from the Brain Tumor Database [106]. The semi-automatic segmentation was performed by a student with no special medical training and no inside knowledge of the proposed algorithm.

The 10 patients' heads were imaged in the sagittal and axial plane with a 1.5 T MRI system[7] with a postcontrast three-dimensional sagittal spoiled gradient recalled (SPGR) acquisition with contiguous slices. The resolution is $0.975 \times 0.975 \times 1.5$ mm ($256 \times 256 \times 124$ voxels). Datasets where manually segmented into 3 regions:

- tumor
- white and gray matter
- other.

Because of inter- and intra-expert variability these results should be expected *not* to be as good as in the synthetic case. It should also be noted that the arbitrary conventions of the manual segmentations are responsible for a lot of the observed error since for example the ventricle was labeled as gray matter, the medulla oblongata and the spinal cord have been

[7]Signa, GE Medical Systems, Milwaukee, WI.

left out etc. (compare Figures 9(a) and 9(c)). Overall, nonetheless, results are consistent with the artificial case (Tables 2 and 3).

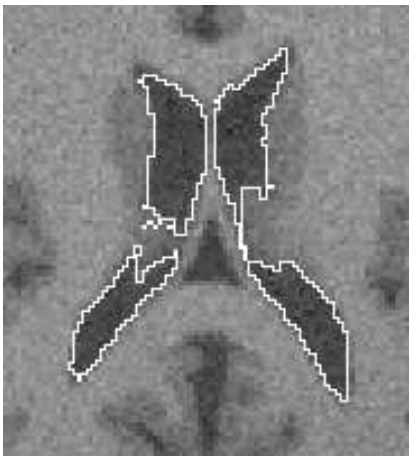
Table 2: Performance measure of proposed technique for white and gray matter segmentation on real datasets

	DSC	**PE**	$\boldsymbol{\mu}_{D>0}$	$\boldsymbol{\sigma}_{D>0}$	$\boldsymbol{D}_{0.95}$	$\boldsymbol{D}_{0.99}$
Mean	**88.9%**	**19.8%**	**1.79**	**1.47**	**2.03**	**4.91**
Std. Dev.	**4.0%**	**6.3%**	**0.58**	**0.79**	**1.52**	**2.82**
Case 1	89.5%	19.0%	1.45	0.87	1.41	3.00
Case 2	90.6%	17.1%	2.29	2.44	2.00	8.25
Case 3	90.9%	16.6%	1.39	0.80	1.41	2.83
Case 4	84.7%	26.5%	1.28	0.79	1.41	3.00
Case 5	93.3%	12.5%	1.58	1.48	1.00	3.00
Case 6	87.0%	23.0%	1.87	1.66	2.24	5.66
Case 7	86.5%	23.8%	1.89	1.68	2.24	6.08
Case 8	81.3%	31.6%	3.15	3.00	6.16	10.67
Case 9	91.4%	15.9%	1.83	0.48	1.00	1.73
Case 10	93.8%	11.8%	1.19	0.48	1.00	1.73

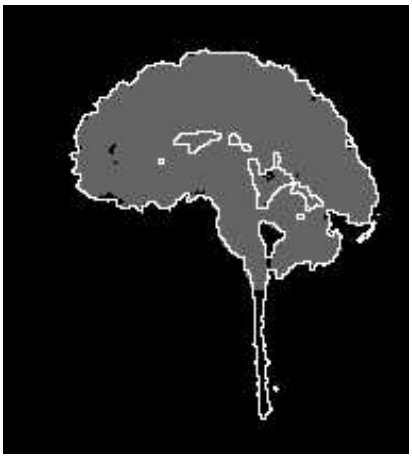
Table 3: Performance measure of proposed technique for tumor segmentation on real datasets

	Tumor Type	**DSC**	**PE**	$\boldsymbol{\mu}_{D>0}$	$\boldsymbol{\sigma}_{D>0}$	$\boldsymbol{D}_{0.95}$	$\boldsymbol{D}_{0.99}$
Mean	–	**83.1%**	**27.6%**	**1.54**	**0.88**	**2.35**	**3.65**
Std. Dev.	–	**10.9%**	**15.8%**	**0.53**	**0.72**	**1.73**	**2.70**
Case 1	meningioma	94.6%	10.2%	1.07	0.27	1.00	1.41
Case 2	meningioma	87.2%	22.8%	1.43	0.79	1.73	3.16
Case 3	meningioma	97.5%	4.9%	1.03	0.15	0.00	1.00
Case 4	low grade glioma	84.0%	27.6%	1.51	0.90	2.24	2.24
Case 5	astrocytoma	65.7%	51.1%	1.36	0.56	2.24	3.16
Case 6	low grade glioma	92.1%	14.7%	1.07	0.24	1.00	1.41
Case 7	astrocytoma	88.9%	20.0%	1.16	0.35	1.41	2.00
Case 8	astrocytoma	70.6%	45.4%	2.03	1.49	4.12	6.40
Case 9	astrocytoma	72.7%	42.8%	2.09	1.80	4.36	7.48
Case 10	low grade glioma	77.7%	36.4%	2.61	2.20	5.39	8.25

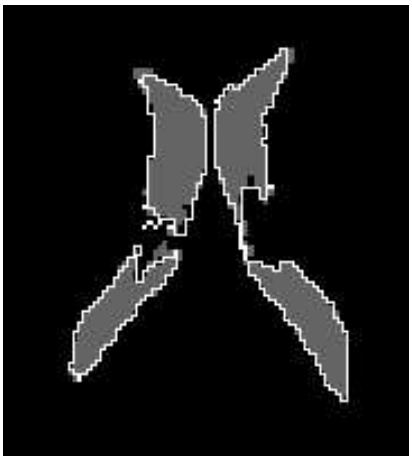
(a) Sagittal slice of real dataset and proposed segmentation (WM+GM)



(b) Axial slice of artificial dataset and proposed segmentation (ventricle)

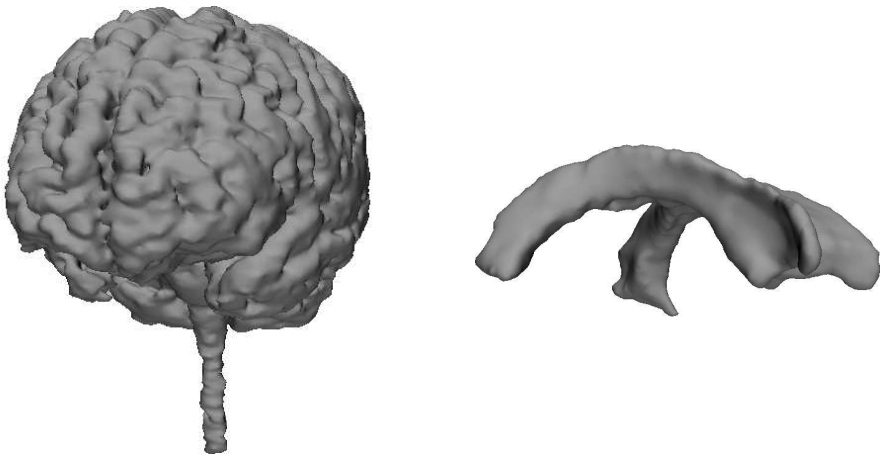


(c) Expert segmentation (gray) and proposed segmentation (white)



(d) Underlying ground truth (gray) and proposed segmentation (white)

Figure 9: Results of proposed technique on real and noisy simulated datasets (left and right respectively). Two-dimensional slices are shown, however the processing is performed on the three-dimensional images. See Figure 10 for corresponding surface renderings.



(a) Rendered surface of proposed segmentation (WM+GM)

(b) Rendered surface of proposed segmentation (ventricle)

Figure 10: Surface renderings of segmented structures. These results correspond to the slices shown on Figure 9.

CHAPTER 4

DIRECTION-BASED IMAGE SEGMENTATION

In this chapter we propose a direction based framework for the segmentation of open space curves as well as closed hypersurfaces. See also [107] and [108].

Image segmentation is set in an optimization framework. This had previously been proposed in the active contour and LiveWire techniques in which a local cost is defined based on image information and curves are sought that minimize a total cost obtained by integrating local costs (Section 4.1, see also Section 2.4 for a more general review). In these techniques, the cost depends only on position and is typically a decreasing function of some edge detector. We propose to augment this model with direction information. The global cost of a structure (a curve or surface) is then computed by integrating the local cost along the structure using direction information (i.e., the tangent of the curve or the normal of the surface) (Section 4.2). Optimal structures can be determined using continuous minimization techniques and efficient algorithms (Section 4.3).

It is known that some neurons in the mammalian (and in particular human) visual cortex function as detectors of oriented edges. This has been studied extensively by psychophysicists. We show in Section 4.4 that this knowledge of the biological eye-brain system can be used to define the local cost that will be integrated along the curve (or surface). This is illustrated on synthetic as well as real images.

Another way to define these local costs is to determine, for every position and every direction, how much the image locally looks like a known pattern. Structures are then determined that locally follow this pattern in an optimal way. Conversely, a pattern detector can be determined that would minimize the global cost of an expert-given structure on a sample image. After this learning phase, the pattern can be used to determine optimal structures on previously unseen images. This is detailed and illustrated on examples in Section 4.5.

Finally, another application for structure segmentation using direction-dependent local costs are images that inherently comprise direction information. Diffusion magnetic resonance imagery is one such example and we show in Section 4.6 how the proposed framework can be used to determine the location of plausible neural tracts from such datasets.

4.1 Related Previous Work

4.1.1 Conformal Active Contour

Active contours have been very popular for image segmentation. In particular, in the conformal (or geodesic) active contour model (see [74, 75] as well as Section 2.4.4) a local cost, $\psi : \mathbb{R}^n \to \mathbb{R}^{+*}$ is defined based on image information. For a given curve $\mathbf{\Gamma}$ the total cost $\mathcal{E}(\mathbf{\Gamma})$ is defined as the integration of local costs along the curve:

$$\mathcal{E}(\mathbf{\Gamma}) \triangleq \int_{\mathbf{\Gamma}} \psi(\mathbf{\Gamma}) \, \mathrm{d}s. \tag{64}$$

This energy can be interpreted as the ψ-weighted length of the curve. Minimal curves will therefore tend to go through regions where ψ is small while at the same time keeping their total length as small as possible (in particular since the local costs are strictly positive everywhere the curve cannot develop some sort of fractal behavior since that would result in a higher weighted-length and hence higher cost). It is important to note that s is the arc-length parameterization and this energy is purely geometric (see Section 2.4.4).

If the curve (or surface) is closed[1], a partial differential equation is obtained by calculus of variations that continuously deforms an initial curve $\mathbf{\Gamma}(t = 0)$ (or surface $\mathbf{\Sigma}(t = 0)$) in a way that optimally minimizes its total cost $\mathcal{E}$. This can be interpreted as a gradient descent on the infinite dimensional space of curves (or of surfaces).

In the case of the functional (64), the PDE that deforms a given curve in order to minimizes the energy as fast as possible in a L_2 sense is

$$\frac{\partial \mathbf{\Gamma}}{\partial t} = -\nabla\psi + \psi \mathbf{\Gamma}_{ss}. \tag{65}$$

[1] or, in the case of the curve, if the end points are fixed

The local cost is typically some decreasing function of some edge detector. A classical choice is

$$\psi(\mathbf{p}) = \frac{1}{1+|\nabla I|}. \tag{66}$$

4.1.2 LiveWire

Mortensen et al. [109] had previously proposed the Live-Wire segmentation technique that also determines optimal curves for the same kind of functional. Their framework is based on Dynamic Programming and is applicable to open curves with one end fixed on a given seed region S.

The underlying principle of Dynamic Programming is the *principle of optimality* verified by minimum-cost problems such as (64). The principle is that any sub-path p of an optimal path P is itself optimal (otherwise the P could be improved by following another sub-path p' instead of p). This leads to the definition of the *value function* $\mathcal{E}^*$ which is the minimal cost to reach the seed region S from any point $\mathbf{p}$ of the domain.

$$\mathcal{E}^*(\mathbf{p}) \triangleq \min\{\ \mathcal{E}(\mathbf{\Gamma}), \mathbf{\Gamma}(0) = \mathbf{p}, \mathbf{\Gamma}(1) \in S\ \}$$

In problems such as (64), the value functions satisfies the Eikonal equation $|\nabla \mathcal{E}^*(\mathbf{p})| = \psi(\mathbf{p})$ with boundary condition $\mathcal{E}^* = 0$ on S. This equation can be solved numerically using the Fast Marching algorithm [96, 23]. or can be discretely approximated using Dijkstra's algorithm.

From any point in the domain an optimal curve in the sense of (64) can then be determined by gradient descent on the scalar field $\mathcal{E}^*$.

4.2 Direction-based Image Segmentation Framework

In both techniques presented in Section 4.1, segmentation is set in an optimization framework and the cost is not directly defined using the image but rather using the positive scalar field ψ. The global cost of a curve is defined by the ψ-weighted length of curves (or ψ-weighted area of hypersurfaces in higher dimensions). If this conformal factor ψ is an edge detector then locally optimal curves will lay on top of the edges of the image. In fact,

ψ could be defined by any general pattern detector. A strong constraint however is that the previously described framework does not incorporate direction information and the pattern detector must therefore be independent from rotations of the image. This is for example the case in (66) where the norm of the gradient is used to determine the strength of the edge but the direction of the gradient (and therefore that of the edge) is ignored.

We propose to extend the conformal active contour and LiveWire technique to directional data by considering direction-dependent local costs of the form[2] $\psi : \mathbb{R}^n \times \mathbb{S}^{n-1} \to \mathbb{R}^+$.

As previously, the total cost of a curve can be determined by weighting its length with the local cost ψ, which will be a function not only of the position of the curve $\mathbf{\Gamma}$ but also of the direction of its unit tangent $\hat{\mathbf{T}}$:

$$\mathcal{E}(\mathbf{\Gamma}) \triangleq \int_{\mathbf{\Gamma}} \psi(\mathbf{\Gamma}, \hat{\mathbf{T}}) \, \mathrm{d}\mathbf{\Gamma}. \tag{67}$$

Similarly, the total cost of a closed hypersurface can be determined by weighting its surface-area with the local cost ψ as a function of the position of the surface $\mathbf{\Sigma}$ and the direction of its unit normal $\hat{\mathbf{N}}$:

$$\mathcal{E}(\mathbf{\Sigma}) \triangleq \int_{\mathbf{\Sigma}} \psi(\mathbf{\Sigma}, \hat{\mathbf{N}}) \, \mathrm{d}\mathbf{\Sigma}. \tag{68}$$

In the conformal active contour framework, the cost of a given curve (or surface) is defined to be the length of the curve (the area of the surface) locally weighted by an image-dependent scalar field. This scalar field is defined to be small at the edges of the image and therefore the total cost of a curve is small if and only if it lies as much as possible on the edges of the image. Starting from some user-determined curve, an evolution is obtained (by gradient descent, using calculus of variations) that deforms the curve to minimize the total cost. It is important to understand that using this technique, only a local minimum of the cost functional is obtained. This is actually desirable since the global minima of the cost functional are trivial (they are the degenerate curves of zero length, whose weighted length will also be zero). In order to guarantee the cost functional

[2] A position $\mathbf{p}$ in the n-dimensional space $\mathbb{R}^n$ is written in bold, and a hat denotes a direction $\hat{\boldsymbol{d}}$ of the unit sphere $\mathbb{S}^{n-1}$.

will not have too many local minima, the local edge-weighting is spatially smoothed. In this formulation orientation is not taken into account and therefore smoothing has to be performed isotropically.[3] This means that, when considering how to deform a curve based on local information, no difference can be made between the processing of image information that is a little bit further away along the curve and the processing of image information that is a little bit further away normal to the curve. In the case of a surface, image information inside and outside the surface (and in fact also along the surface) have to be treated in exactly the same manner. In contrast, in the proposed framework, the length of the curve (or area of the surface) is weighted based on the location and the direction (of the unit tangent in the case of the curve, of the unit normal in the case of the surface).

4.3 Minimization

In this section we present two different techniques for determining optimal structures. In the first technique (Section 4.3.1), the first variation of the energy with respect to the structure is determined and subsequently used to determine an optimal minimizing deformation. This technique can be applied to curves and surfaces. It necessitates some initial guess and can lead to a local minimum. This technique constitutes an extension of the one described in Section 4.1.1.

The second technique (Section 4.3.2), is based on the ideas presented in Section 4.1.2. The principle of optimality is used to define an auxiliary *value function* which can be solved for numerically and from which optimal curves can be determined. This techniques only applies to open space curves and will always result in global minima.

The problem of determining a meaningful direction-dependent local cost ψ from an image I will be tested in sections 4.4, 4.5 and 4.6.

[3] Anisotropic smoothing could be performed but it would have to be independent from the curve or surface. That would not be natural since some direction of space would be privileged and the overall segmentation scheme would cease to be rotation-invariant.

4.3.1 Gradient Descent

One way to minimize functionals such as (67) or (68) is to analyze the variation of the energy for an infinitesimal variation $\varepsilon\boldsymbol{\nu}$. For example, in the case of a curve, if the extremities of the curve are fixed, this can be written

$$\lim_{\varepsilon \to 0} \frac{\mathcal{E}(\boldsymbol{\Gamma} + \varepsilon\boldsymbol{\nu}) - \mathcal{E}(\boldsymbol{\Gamma})}{\varepsilon} = \int_{\boldsymbol{\Gamma}} \boldsymbol{\nu}.\boldsymbol{\alpha} \, \mathrm{d}s.$$

The optimal way to decrease the energy is therefore to perturb $\boldsymbol{\Gamma}$ by $-\boldsymbol{\alpha}$, i.e., to flow $\boldsymbol{\Gamma}_t = -\boldsymbol{\alpha}$ from some initial curve $\boldsymbol{\Gamma}(t = 0)$. At steady state, $\boldsymbol{\Gamma}$ will be a *local* minimum of $\mathcal{E}$. As in the non direction-dependent case, in order to limit the number of local minima of the functional $\mathcal{E}$, the local cost ψ has to be smooth. If that is not the case, the evolution might converge very rapidly to some undesired local minimum. In that case, some explicit smoothing of ψ needs to be introduced, for example by convolving it with some Gaussian kernel.

4.3.1.1 Curves

Consider the energy defined by Equation (67) as

$$\mathcal{E}(\boldsymbol{\Gamma}) \triangleq \int_{\boldsymbol{\Gamma}} \psi(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}_s) \, \mathrm{d}s,$$

where s denotes arc-length.

Choosing a parameterization p, of the curve, we compute the first variation for a perturbation $\boldsymbol{\Gamma}_t$.

$$\begin{aligned}
\frac{\partial \mathcal{E}}{\partial t} =& \frac{\partial}{\partial t} \int_0^1 \psi(\boldsymbol{\Gamma}, \frac{\boldsymbol{\Gamma}_p}{|\boldsymbol{\Gamma}_p|}) |\boldsymbol{\Gamma}_p| \, \mathrm{d}p \\
=& \int \underbrace{\nabla_{\mathbf{p}} \psi . \boldsymbol{\Gamma}_t |\boldsymbol{\Gamma}_p|}_{I_1} + \underbrace{\nabla_{\hat{\boldsymbol{d}}} \psi . (\frac{\partial}{\partial t} \frac{\boldsymbol{\Gamma}_p}{|\boldsymbol{\Gamma}_p|}) |\boldsymbol{\Gamma}_p|}_{I_2} + \underbrace{\psi \frac{\partial}{\partial t} |\boldsymbol{\Gamma}_p|}_{I_3} \, \mathrm{d}p
\end{aligned}$$

where $\nabla_{\mathbf{p}}\psi$ and $\nabla_{\hat{\boldsymbol{d}}}\psi$ are the gradients of ψ with respect to position and direction respectively. Note that $\nabla_{\hat{\boldsymbol{d}}}\psi$ is an element of the tangent space $T_{\boldsymbol{\Gamma}_s}\mathbb{S}^{n-1}$. We will consider that it is an element of $\mathbb{R}^n$ with zero component along the unit tangent $\mathbf{T} = \boldsymbol{\Gamma}_s$.

Our goal is to write the variation as $E_t = \int \boldsymbol{\Gamma}_t.\boldsymbol{\beta}|\boldsymbol{\Gamma}_p|\ \mathrm{d}p$ which will correspond to the optimal energy increasing evolution $\boldsymbol{\Gamma}_t = \boldsymbol{\beta}$.

The first term I_1 is already in the right form. We will now work independently on the second and third terms I_2 and I_3.

$$I_2 = \int \nabla_{\hat{\boldsymbol{d}}}\psi.(\frac{\partial}{\partial t}\frac{\boldsymbol{\Gamma}_p}{|\boldsymbol{\Gamma}_p|})|\boldsymbol{\Gamma}_p|\ \mathrm{d}p$$

using $\frac{\partial}{\partial t}\frac{\boldsymbol{\Gamma}_p}{|\boldsymbol{\Gamma}_p|} = (\mathrm{Id} - \boldsymbol{\Gamma}_s\boldsymbol{\Gamma}_s^t)\boldsymbol{\Gamma}_{p,t}/|\boldsymbol{\Gamma}_p|$ and omitting the projection since $\nabla_{\hat{\boldsymbol{d}}}\psi$ is normal to $\boldsymbol{\Gamma}_p$,

$$I_2 = \int \nabla_{\hat{\boldsymbol{d}}}\psi.\boldsymbol{\Gamma}_{p,t}\ \mathrm{d}p$$

integrating by parts, (the boundary terms $\nabla_{\hat{\boldsymbol{d}}}\psi(\boldsymbol{\Gamma}(0)).\boldsymbol{\Gamma}_t(0)$ and $\nabla_{\hat{\boldsymbol{d}}}\psi(\boldsymbol{\Gamma}(1)).\boldsymbol{\Gamma}_t(1)$ vanish since the extremities of the curve are fixed),

$$I_2 = -\int \boldsymbol{\Gamma}_t.\frac{\partial}{\partial p}\nabla_{\hat{\boldsymbol{d}}}\psi\ \mathrm{d}p$$

using $\frac{\partial}{\partial p} = |\boldsymbol{\Gamma}_p|\frac{\partial}{\partial s}$,

$$I_2 = -\int \boldsymbol{\Gamma}_t.|\boldsymbol{\Gamma}_p|\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi\ \mathrm{d}p$$

Similarly, the third term can be re-written:

$$\begin{aligned} I_3 &= \int \psi\frac{\partial}{\partial t}|\boldsymbol{\Gamma}_p|\ \mathrm{d}p \\ &= \int \psi\frac{\partial}{\partial t}\sqrt{\boldsymbol{\Gamma}_p.\boldsymbol{\Gamma}_p}\ \mathrm{d}p \\ &= \int \psi\frac{\boldsymbol{\Gamma}_{p,t}.\boldsymbol{\Gamma}_p + \boldsymbol{\Gamma}_p.\boldsymbol{\Gamma}_{p,t}}{2\sqrt{\boldsymbol{\Gamma}_p.\boldsymbol{\Gamma}_p}}\ \mathrm{d}p \\ &= \int \psi\boldsymbol{\Gamma}_{p,t}.\boldsymbol{\Gamma}_s\ \mathrm{d}p \end{aligned}$$

integrating by parts, (the boundary terms vanish as before because $\mathbf{\Gamma}(0) = \mathbf{\Gamma}(1) = \mathbf{0}$.)

$$\begin{aligned} I_3 &= -\int \mathbf{\Gamma}_t . \frac{\partial}{\partial p}(\psi \mathbf{\Gamma}_s)\ \mathrm{d}p \\ &= -\int \mathbf{\Gamma}_t . |\mathbf{\Gamma}_p| \frac{\partial}{\partial s}(\psi \mathbf{\Gamma}_s)\ \mathrm{d}p \\ &= -\int \mathbf{\Gamma}_t . |\mathbf{\Gamma}_p| \{\ (\nabla_{\mathbf{p}} \psi . \mathbf{\Gamma}_s)\mathbf{\Gamma}_s + (\nabla_{\hat{\boldsymbol{d}}} \psi . \mathbf{\Gamma}_{s,s})\mathbf{\Gamma}_s + \psi \mathbf{\Gamma}_{s,s}\ \}\ \mathrm{d}p \end{aligned}$$

Since $\nabla_{\hat{\boldsymbol{d}}} \psi$ is normal to $\mathbf{\Gamma}_s$, it is unchanged when projected on the subspace orthogonal to $\mathbf{\Gamma}_s$. This projection is denoted $P_{\mathbf{\Gamma}_s^\perp}$ and can be written in matrix form as $(\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t)$.

$$\begin{aligned} \nabla_{\hat{\boldsymbol{d}}} \psi =& (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t) \nabla_{\hat{\boldsymbol{d}}} \psi \\ \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi =& \frac{\partial}{\partial s} \{\ (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t) \nabla_{\hat{\boldsymbol{d}}} \psi\ \} \\ =& (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t) \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi - \mathbf{\Gamma}_{s,s} \mathbf{\Gamma}_s^t \nabla_{\hat{\boldsymbol{d}}} \psi - \mathbf{\Gamma}_s \mathbf{\Gamma}_{s,s}^t \nabla_{\hat{\boldsymbol{d}}} \psi \end{aligned}$$

since $\nabla_{\hat{\boldsymbol{d}}} \psi$ is normal to $\mathbf{\Gamma}_s$, only the first and last term remain,

$$= (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t) \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi - (\mathbf{\Gamma}_{s,s} . \nabla_{\hat{\boldsymbol{d}}} \psi) \mathbf{\Gamma}_s$$

Finally, the three terms can be combined to write,

$$\begin{aligned} \mathcal{E}_t(\mathbf{\Gamma}) =& I_1 + I_2 + I_3 \\ =& \int \mathbf{\Gamma}_t . \{\ (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t)(\nabla_{\mathbf{p}} \psi - \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi) - \psi \mathbf{\Gamma}_{s,s}\ \} |\mathbf{\Gamma}_p|\ \mathrm{d}p \\ =& \int_{\mathbf{\Gamma}} \mathbf{\Gamma}_t . \{\ (\mathrm{Id} - \mathbf{\Gamma}_s \mathbf{\Gamma}_s^t)(\nabla_{\mathbf{p}} \psi - \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi) - \psi \mathbf{\Gamma}_{s,s}\ \}\ \mathrm{d}s \end{aligned}$$

Using the Cauchy-Schwarz inequality, the geometric deformation $\mathbf{\Gamma}_t$ that optimally decreases the energy $\mathcal{E}(\mathbf{\Gamma})$ (in a L_2 sense) is therefore:

$$\mathbf{\Gamma}_t = -P_{\mathbf{\Gamma}_s^\perp}(\nabla_{\mathbf{p}} \psi - \frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi) + \psi \mathbf{\Gamma}_{s,s}. \tag{69}$$

This evolution is similar to that of the non direction-dependent conformal active contour with the addition of the term: $P_{\mathbf{\Gamma}_s^\perp}(\frac{\partial}{\partial s} \nabla_{\hat{\boldsymbol{d}}} \psi)$. See Section 4.3.1.5 for interpretation.

4.3.1.2 Implementation of the Curve Evolution

A particle-based approach can be used to evolve an open space curve according to Equation (69). Derivatives can be approximated by fitting a spline to these marker particles. If the parameterization of the spline is chosen such that $t_{i+1} - t_i = \|\mathbf{p}_{i+1} - \mathbf{p}_i\|$ then the parameterization t will approximate the arc-length parameterization s. The derivatives will therefore approximate the derivatives with respect to arc-length, which will alleviate the problem of uneven particle sampling on the curve.

In this work, the adaptive time-step RK45 numerical integration scheme [110] was used.

4.3.1.3 Surfaces

Consider the energy defined by Equation (68) as:

$$\mathcal{E}(\mathbf{\Sigma}) \triangleq \int_{\mathbf{\Sigma}} \psi(\mathbf{\Sigma}, \mathbf{N}) \, \mathrm{d}\mathbf{\Sigma}.$$

Choosing a parameterization $(p_1, \ldots, p_{n-1})$ of the surface, we want to determine the variation of $\mathcal{E}(\mathbf{\Sigma})$ for an evolution of the form $\mathbf{\Sigma}_t = \alpha \mathbf{N}$. Since the geometry of $\mathbf{\Sigma}$ is only affected by deformations along the normal $\mathbf{N}$, this does not result in loss of generality.

Using the notation $\Sigma_i \triangleq \frac{\partial \Sigma}{\partial p_i}$, let the vector $\mathbf{w}$ be the cross product of the partial derivatives $\mathbf{w} \triangleq \mathbf{\Sigma}_1 \wedge \cdots \wedge \mathbf{\Sigma}_{n-1}$ such that the unit normal is $\mathbf{N} = \mathbf{w}/|\mathbf{w}|$.

$$\begin{aligned}
\frac{\partial E}{\partial t} =& \frac{\partial}{\partial t} \int_p \psi(\mathbf{\Sigma}, \mathbf{N}) |\mathbf{w}| \, \mathrm{d}p \\
=& \int \underbrace{\nabla_{\mathbf{p}} \psi . \mathbf{\Sigma}_t |\mathbf{w}|}_{I_1} + \underbrace{\nabla_{\hat{\boldsymbol{d}}} \psi . \mathbf{N}_t |\mathbf{w}|}_{I_2} + \underbrace{\psi \frac{\partial}{\partial t} |\mathbf{w}|}_{I_3} \, \mathrm{d}p
\end{aligned}$$

where $\nabla_{\mathbf{p}} \psi$ and $\nabla_{\hat{\boldsymbol{d}}} \psi$ are the gradients of ψ with respect to position and direction respectively. Note that $\nabla_{\hat{\boldsymbol{d}}} \psi$ is an element of the tangent space $T_{\mathbf{N}} \mathbb{S}^{n-1}$. In order to avoid technicalities we will consider that it is an element of $\mathbb{R}^n$ with zero component along N.

For a geometric deformation of the form $\mathbf{\Sigma}_t = \alpha\mathbf{N}$, our goal is to write the variation as $E_t = \int_p \alpha\beta|\mathbf{w}| \, \mathrm{d}p$. The optimal energy-increasing evolution will then be obtained by taking $\alpha = \beta$.

The first term I_1 is already in the right form. We will now work independently on the second and third terms I_2 and I_3.

$$I_2 = \int \nabla_{\hat{\boldsymbol{d}}}\psi.\mathbf{N}_t|\mathbf{w}| \, \mathrm{d}p$$

using $\mathbf{N}_t = (\mathrm{Id} - \mathbf{N}\mathbf{N}^t)\mathbf{w}_t/|\mathbf{w}|$, and $\nabla_{\hat{\boldsymbol{d}}}\psi.\mathbf{N} = 0$ (i.e., as in the previous calculation, $\nabla_{\hat{\boldsymbol{d}}}\psi$ is an element of $T_{\mathbf{N}}\mathbb{S}^{n-1}$ but we will consider that it lives in $\mathbb{R}^n$ with zero component along $\mathbf{N}$), we get

$$I_2 = \int \nabla_{\hat{\boldsymbol{d}}}\psi.\mathbf{w}_t \, \mathrm{d}p$$

using the multi-linearity of the wedge product and introducing the notation $\mathbf{w}(k : \mathbf{v}) \triangleq \mathbf{\Sigma}_1 \wedge \ldots \wedge \mathbf{\Sigma}_{k-1} \wedge \mathbf{v} \wedge \mathbf{\Sigma}_{k+1} \wedge \ldots \wedge \mathbf{\Sigma}_{n-1}$ (this is a version of $\mathbf{w}$ where the k^{th} vector in the wedge product was replaced by $\mathbf{v}$; note that if $n = 2$ this reduces to $\mathbf{w}(1 : [x_1 \; x_2]^t) = \mathrm{perp}[x_1 \; x_2]^t = [-x_2 \; x_1]^t$.), we get,

$$I_2 = \int \sum_{k=1}^{n-1} \nabla_{\hat{\boldsymbol{d}}}\psi.\mathbf{w}(k : \mathbf{\Sigma}_{k,t}) \, \mathrm{d}p$$

the dot product of $\nabla_{\hat{\boldsymbol{d}}}\psi$ and the wedge product can be interpreted as a determinant, using the antisymmetry,

$$I_2 = -\int \sum_{k=1}^{n-1} \mathbf{\Sigma}_{k,t}.\mathbf{w}(k : \nabla_{\hat{\boldsymbol{d}}}\psi) \, \mathrm{d}p$$

integrating by parts (the surface being closed, the boundary conditions vanish),

$$I_2 = \int \sum_{k=1}^{n-1} \mathbf{\Sigma}_t.\frac{\partial}{\partial p_k}\mathbf{w}(k : \nabla_{\hat{\boldsymbol{d}}}\psi) \, \mathrm{d}p$$

when computing the partial derivative of the wedge products, all terms containing second derivatives $\mathbf{\Sigma}_{i,j}$ of the surface cancel out for $i \neq j$ and we are left with

$$I_2 = \int \sum_{k=1}^{n-1} \mathbf{\Sigma}_t.\mathbf{w}(k : \frac{\partial}{\partial p_k}\nabla_{\hat{\boldsymbol{d}}}\psi) \, \mathrm{d}p$$

if we write $\mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi = (\beta_{i,j})$ in the basis $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \cdots, \mathbf{N})$ then since $\frac{\partial}{\partial p_k}\nabla_{\hat{\boldsymbol{d}}}\psi$ is wedged with all $\boldsymbol{\Sigma}_l$ for $l \neq k$, then only the component $\beta_{k,k}$ will not cancel out.

$$I_2 = \int \sum_{k=1}^{n-1} \boldsymbol{\Sigma}_t.\mathbf{w}(k : \beta_{k,k}\boldsymbol{\Sigma}_k)\ \mathrm{d}p$$

then, using the multi-linearity of the wedge product and the fact that $\mathbf{w}(k : \boldsymbol{\Sigma}_k) = \mathbf{w}$,

$$I_2 = \int \sum_{k=1}^{n-1} \boldsymbol{\Sigma}_t.\beta_{k,k}\mathbf{w}\ \mathrm{d}p$$

finally, since $\boldsymbol{\Sigma}_t = \alpha\mathbf{N}$ and $\mathbf{N}.\mathbf{w} = |\mathbf{w}|$,

$$\begin{aligned} I_2 &= \int \sum_{k=1}^{n-1} \alpha\beta_{k,k}|\mathbf{w}|\ \mathrm{d}p \\ &= \int \alpha\ \mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi\ )|\mathbf{w}|\ \mathrm{d}p \end{aligned}$$

where trace($\mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi$) is the trace of the linear map $\mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi : T_{\mathbf{N}}\mathbb{S}^{n-1} \to T_{\mathbf{N}}\mathbb{S}^{n-1}$.

Similarly, the third term can be written

$$\begin{aligned} I_3 &= \int \psi\frac{\partial}{\partial t}|\mathbf{w}|\ \mathrm{d}p \\ &= \int \psi\frac{\partial}{\partial t}\sqrt{\mathbf{w}.\mathbf{w}}\ \mathrm{d}p \\ &= \int \psi\frac{\mathbf{w}_t.\mathbf{w} + \mathbf{w}.\mathbf{w}_t}{2\sqrt{\mathbf{w}.\mathbf{w}}}\ \mathrm{d}p \\ &= \int \psi\mathbf{w}_t.\mathbf{N}\ \mathrm{d}p \end{aligned}$$

as before,

$$I_3 = \int \psi \sum_{k=1}^{n-1} \mathbf{w}(k : \boldsymbol{\Sigma}_{k,t}).\mathbf{N}\ \mathrm{d}p$$

which is a determinant, and, by antisymmetry,

$$I_3 = -\int \psi \sum_{k=1}^{n-1} \mathbf{w}(k : \mathbf{N}).\boldsymbol{\Sigma}_{k,t}\ \mathrm{d}p$$

by parts (the boundary conditions vanish since the surface is closed)

$$I_3 = \int \boldsymbol{\Sigma}_t . \sum_{k=1}^{n-1} \frac{\partial}{\partial p_k} \psi \mathbf{w}(k : \mathbf{N}) \, \mathrm{d}p$$

since $\boldsymbol{\Sigma}_t = \alpha \mathbf{N}$, and the normal $\mathbf{N}$ appears in the wedge product, it will have to be derived in order to result in a non zero dot product,

$$I_3 = \int \alpha \mathbf{N} . \sum_{k=1}^{n-1} \psi \mathbf{w}(k : \frac{\partial}{\partial p_k} \mathbf{N}) \, \mathrm{d}p$$

as before, if we express $\mathrm{d}\mathbf{N} = (\gamma_{i,j})$ in the basis $(\boldsymbol{\Sigma}_1, \cdots \boldsymbol{\Sigma}_{n-1}, \mathbf{N})$ then

$$\begin{aligned} I_3 &= \int \alpha \mathbf{N} . \sum_{k=1}^{n-1} \psi \mathbf{w}(k : \gamma_{k,k} \boldsymbol{\Sigma}_k) \, \mathrm{d}p \\ &= \int \alpha \mathbf{N} . \sum_{k=1}^{n-1} \psi \gamma_{k,k} \mathbf{w} \, \mathrm{d}p \\ &= \int \alpha \psi \ \mathrm{trace}(\ \mathrm{d}\mathbf{N}\ ) |\mathbf{w}| \, \mathrm{d}p \end{aligned}$$

Note: $\mathrm{d}\mathbf{N}$ is the shape operator and its trace is $(n-1)H$ where H is the mean curvature of the surface (note that there is a sign difference with the definition in [111] due to the fact that in this work outward normals will be used).

Finally, the three terms can be combined to write, for an evolution of the form $\boldsymbol{\Sigma}_t = \alpha \mathbf{N}$,

$$\begin{aligned} \mathcal{E}_t(\boldsymbol{\Sigma}) &= \int_p \alpha \{\ \nabla_{\mathbf{p}} \psi . \mathbf{N} + \ \mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}} \psi\ ) + \psi \ \mathrm{trace}(\ \mathrm{d}\mathbf{N}\ )\ \} |\mathbf{w}| \, \mathrm{d}p \\ &= \int_{\boldsymbol{\Sigma}} \alpha \{\ (\nabla_{\mathbf{p}} \psi) . \mathbf{N} + \ \mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}} \psi\ ) + \psi \ \mathrm{trace}(\ \mathrm{d}\mathbf{N}\ )\ \} \, \mathrm{d}\boldsymbol{\Sigma} \end{aligned}$$

The geometric deformation $\boldsymbol{\Sigma}_t$ that optimally decreases the energy $\mathcal{E}(\boldsymbol{\Sigma})$ is therefore:

$$\boldsymbol{\Sigma}_t = -\{\ \nabla_{\mathbf{p}} \psi . \mathbf{N} + \ \mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}} \psi\ ) + (n-1)\psi H\ \} \mathbf{N}, \tag{70}$$

where n is the dimension of the spatial domain, H is the mean curvature and $\mathbf{N}$ is the outward normal.

As in the case of the curve, this is similar to the non direction-dependent case with the addition of the term $-\mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}} \psi\ ) \mathbf{N}$, which can be interpreted as a realigning term. For dimension $n = 2$, the argument is similar to that of the curve case and more insights are also available in Section 4.3.1.5.

4.3.1.4 Implementation of the Surface Evolution

Evolving closed hypersurfaces is numerically easier than evolving space curves thanks to the levelset technique developed by Osher [112] and Sethian [23]. An auxiliary *level set function* u is defined such that the surface Σ corresponds to the zero levelset of u:

$$\Sigma = \{ \mathbf{p},\ u(\mathbf{p}) = 0 \} \tag{71}$$

Therefore, by construction, $u(\Sigma) = 0$. Differentiating this relation with respect to the artificial time parameter t, we obtain

$$u_t(\Sigma) + \nabla u(\Sigma).\Sigma_t = 0. \tag{72}$$

The deformation Σ_t is given by Equation (70). The only difficulty is that (72) only defines the evolution of the level set function u on the surface Σ. This is due to the fact that Σ is the only location where u is properly defined by (71). The speed term Σ_t therefore has to be defined on the complete domain in a process known as *velocity extension*.

Geometric quantities such as the normal $\mathbf{N}$ or mean curvature H (that are needed to define Σ_t) can be computed on the curve from the level set function u.

Equation (72) together with Equation (70) can be written as a transport equation

$$u_t(\mathbf{p}) + \nabla u(\mathbf{p}).\tilde{\mathbf{v}}(\mathbf{p}) + (n-1)\tilde{\psi}H_u = 0. \tag{73}$$

where the tilde denotes quantities smoothly extended from Σ to the whole domain, for example by solving the transport equation

$$\begin{cases} \tilde{\psi} = \psi & \text{on } \Sigma \\ \nabla\tilde{\psi}.\nabla u = 0 & \text{everywhere else} \end{cases}$$

Note that the mean curvature H_u is not extended and that a proper upwinding scheme should be employed to handle shocks. A detailed discussion of these issues can be found in [112].

4.3.1.5 Synthetic two-dimensional examples

In order to gain some understanding of the new directional terms we will consider a simple synthetic example in dimension $n = 2$. In order to isolate the influence of directionality, we consider a cost that does not depend on position, for example:

$$\psi_{\mathbf{v}}(\mathbf{p}, \hat{\boldsymbol{d}}) = 1.0 - 0.5 * \exp(-0.5 * (\cos^{-1}(\hat{\boldsymbol{d}}.\mathbf{v}))^2/0.3^2) \tag{74}$$

The minimal value of this cost is 0.5, reached for $\hat{\boldsymbol{d}} = \mathbf{v}$. The value of the cost approaches 1 for other directions.

Note that in dimension $n = 2$, an hypersurface is a closed curve. Therefore Equations (70) and (69) will have exactly the same effect. The only subtelty is that the direction provided to the local cost is the normal in the surface case, Equation (70) and the tangent in the curve case, Equation (69). Therefore in order to obtain the same deformation, the two local cost have to be defined with a rotation of $\frac{\pi}{2}$, see Figures 11.a and 11.c.

As an example, consider the unit circle. Figures 11.b and 11.d represent the directional gradient of the local cost. Since this can be interpreted as the direction in which to rotate the tangent (resp. the normal) in order to *increase* the local cost, we represent the gradient at the tip of the unit tangent (resp. the normal). The norm of the gradient in the two case is the same (on the figures it has been scaled differently for visualization purposes). Let the function f be such that:

$$\nabla_{\hat{\boldsymbol{d}}}\psi_1(\Gamma, \mathbf{T}) = f\mathbf{N} \tag{75}$$

or equivalently

$$\nabla_{\hat{\boldsymbol{d}}}\psi_2(\Gamma, \mathbf{N}) = -f\mathbf{T} \tag{76}$$

The additional term resulting from the dependence on direction in Equation (69) is:

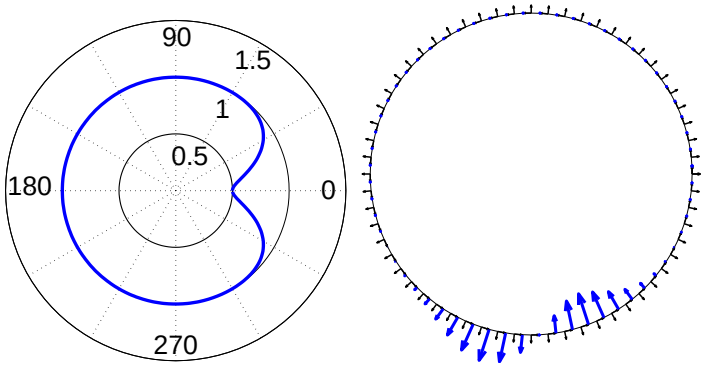
$$
\begin{aligned}
P_{\mathbf{\Gamma}_s^\perp}(\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi_1) &= P_{\mathbf{\Gamma}_s^\perp}(\frac{\partial}{\partial s}f\mathbf{N}) \\
&= P_{\mathbf{\Gamma}_s^\perp}(f'\mathbf{N} - f\kappa\mathbf{T}) \\
&= f'\mathbf{N}
\end{aligned}
$$

or equivalently, in the hypersurface case, $\mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi_2$ is the 1×1 matrix $[-f]$ and therefore the directional term in Equation (70) is also:

$$
\begin{aligned}
-\mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi_2\ )\mathbf{N} &= -\mathrm{trace}(\ \mathrm{d}(-f))\mathbf{N} \\
&= f'\mathbf{N}
\end{aligned}
$$

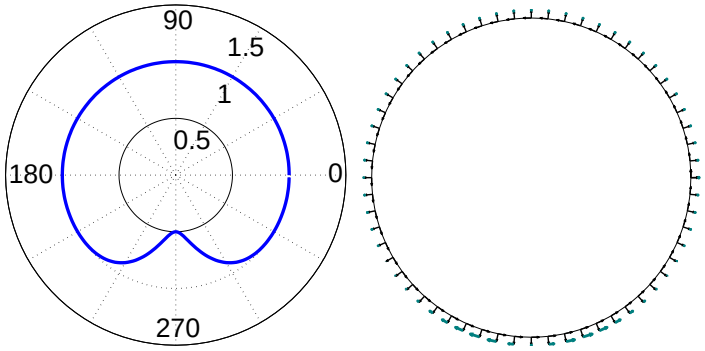
Function f and its derivative f' are represented on Figure 11.e. The deformation $P_{\mathbf{\Gamma}_s^\perp}(\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi_1) = -\mathrm{trace}(\ \mathrm{d}\nabla_{\hat{\boldsymbol{d}}}\psi_2\ )\mathbf{N} = f'\mathbf{N}$ is represented and applied to the initial curve. Its effect is to realign the curve locally the curve with the prefered direction. Note that the realigning occurs where the norm of the directional gradient is maximal. This is were sensitivity is maximal for the local cost and therefore were the optimal deformation in a L_2 sense focuses.

Figures 12 and 13 present the result on a smaller portion of the unit circle, as well as yet another piece of the unit circle for which the direction of the normal has been inverted (this would therefore correspond to an hpersurface patch with negative curvature). In all cases, the curve is realigned with the prefered horizontal direction.



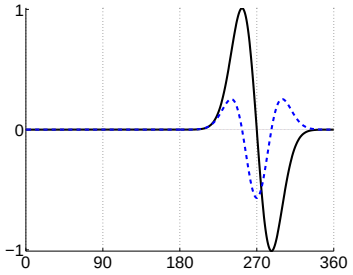
(a) Local cost ψ_1

(b) $\nabla_{\hat{\boldsymbol{d}}}\psi_1(\Gamma, \mathbf{T})$

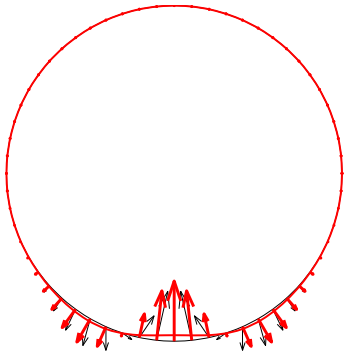


(c) Local cost ψ_2

(d) $\nabla_{\hat{\boldsymbol{d}}}\psi_2(\Sigma, \mathbf{N})$

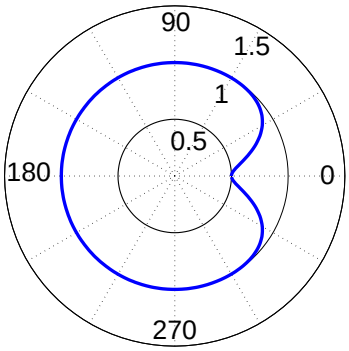


(e) f (solid) and f' (dashed)

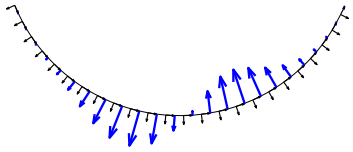


(f) Original curve and $\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi_1$ (thin), projection using $P_{\mathbf{\Gamma}_s^{\perp}}$ and resulting deformed curve (thick)

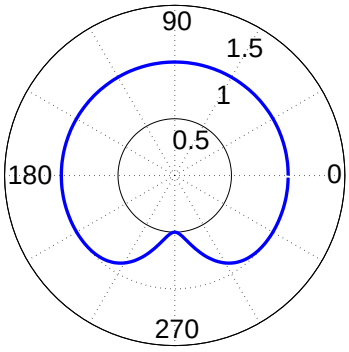
Figure 11: Synthetic two-dimensional example (circle). The effect of the new directional term is to realign the curve with the prefered direction (here right).



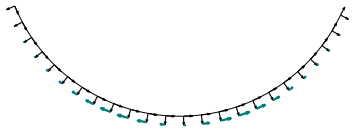
(a) Local cost ψ_1



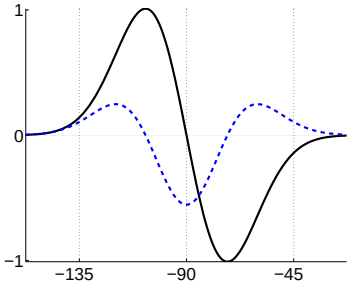
(b) $\nabla_{\hat{\boldsymbol{d}}}\psi_1(\Gamma, \mathbf{T})$



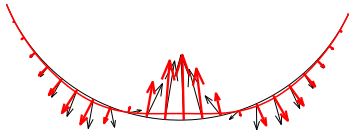
(c) Local cost ψ_2



(d) $\nabla_{\hat{\boldsymbol{d}}}\psi_2(\Sigma, \mathbf{N})$

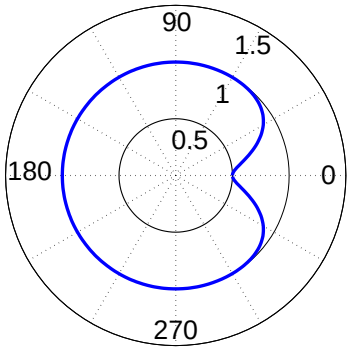


(e) f (solid) and f' (dashed)

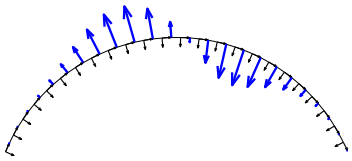


(f) Original curve and $\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi_1$ (thin), projection using $P_{\mathbf{\Gamma}_s^\perp}$ and resulting deformed curve (thick)

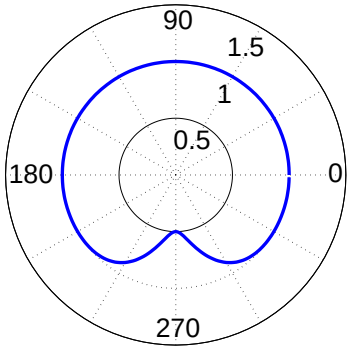
Figure 12: Synthetic two-dimensional example (portion of circle, compare to Figure 11). The effect of the new directional term is to realign the curve with the prefered direction (here right).



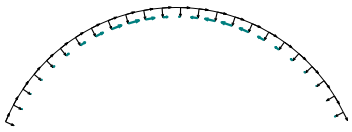
(a) Local cost ψ_1



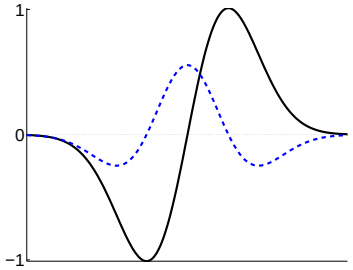
(b) $\nabla_{\hat{\boldsymbol{d}}}\psi_{\mathbf{T}}(\Gamma, 1)$



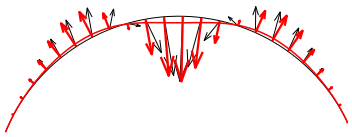
(c) Local cost ψ_2



(d) $\nabla_{\hat{\boldsymbol{d}}}\psi_{\mathbf{N}}(\Sigma, 2)$



(e) f (solid) and f' (dashed)



(f) Original curve and $\frac{\partial}{\partial s}\nabla_{\hat{\boldsymbol{d}}}\psi_1$ (thin), projection using $P_{\mathbf{\Gamma}_s^{\perp}}$ and resulting deformed curve (thick)

Figure 13: Synthetic two-dimensional example (portion of curve with negative curvature, compare to Figure 12). The effect of the new directional term is to realign the curve with the prefered direction (here right).

Another interesting experiment in order to compare the proposed direction-dependent framework to the non direction dependent framework we let two closed curves (initially a circle and a bean-shaped curve) evolve according to the minimizing flow derived in the previous sections.

Since the examples are for dimension $n = 2$, hyper-surfaces are closed curves and Equations (70) and (69) are equivalent.

In order to isolate the effect of direction information we study local costs that do not depend on position but only on the direction $\mathbf{N} = [n_1 n_2]^t$:

1. $\psi = 1$
2. $\psi = \max(\frac{1}{\sqrt{2}}|n_1 + n_2|, \frac{1}{\sqrt{2}}|n_2 - n_2|)^3/0.75$
3. $\psi = \max(|n_1|, |n_2|)^3/0.75$

The first cost is isotropic. In that case, the global cost of the curve is its Euclidean length and the minimizing flow is the classical Euclidean curvature flow [24, 25, 26, 27, 28, 29]. This flow shrinks any planar shape into a smaller and smaller circle. The vanishing point is the global minimum of any weighted-mean global cost because the Euclidean length of such a curves converges to 0. This is visible on the first column of Figures 14 and 15.

The second and third costs are defined using direction information. In particular, the second cost favors portions of the curve that are either horizontal or vertical. The third costs does exactly the opposite and favors portions of the curve that are diagonal. The corresponding evolutions can be observed on the second and third columns of Figures 14 and 15. The influence of direction information is very visible in that the curves no longer deform into a vanishing circle. The existence of preferred (less expensive) directions is very visible. For example the circle deforms into a rounded square. In effect, for that direction-dependent local cost, the rounded square is more economical than the sphere. Note that this rounded square will also shrink to a point which is still a global minimum of the global cost.

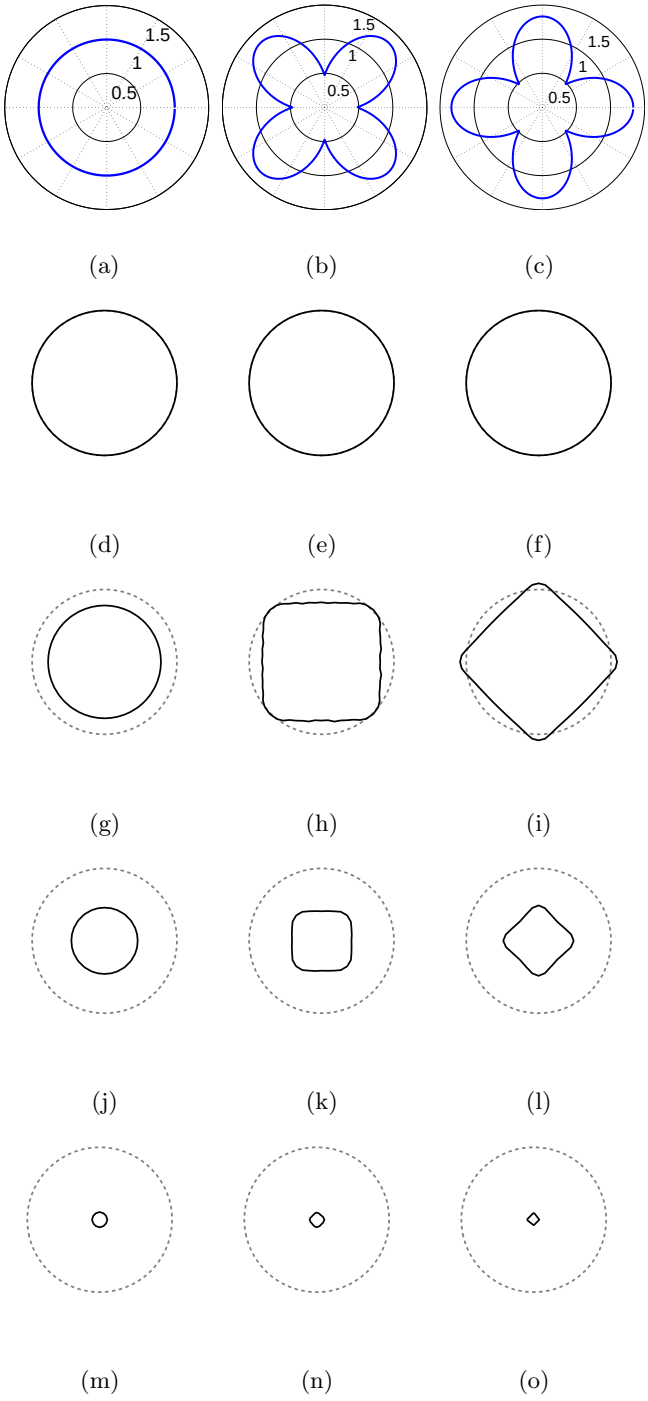


Figure 14: Synthetic two-dimensional example. These three different local cost depend only on direction. They are represented as polar plots (first row). The corresponding deforming shapes are presented on the following rows (black). The initial curve (gray dashed) is a circle. See text.

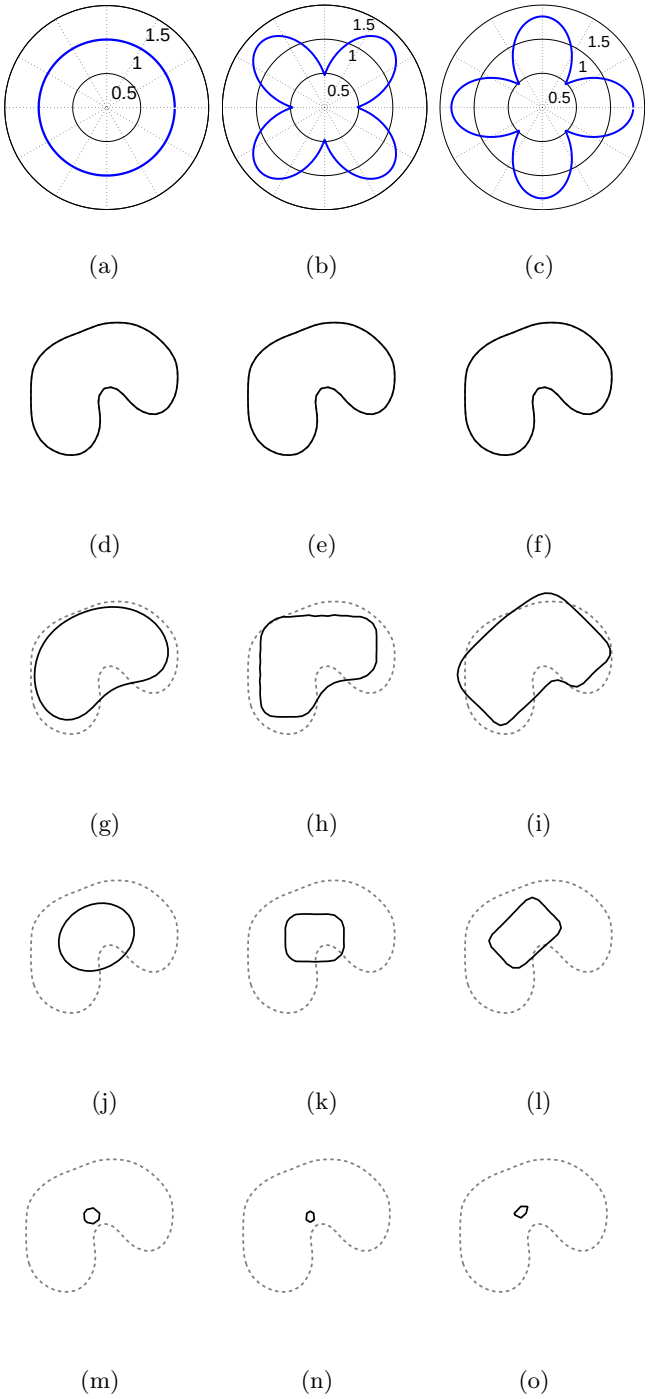


Figure 15: Synthetic two-dimensional example. These three different local cost depend only on direction. They are represented as polar plots (first row). The corresponding deforming shapes are presented on the following rows (black). The initial curve (gray dashed) is bean-shaped. See text.

4.3.2 Dynamic Programming

As in the isotropic case, given a seed region $S \subset \mathbb{R}^n$, the value function $\mathcal{E}^*(\mathbf{p})$ is defined to be the minimum cost over all curves between $\mathbf{p}$ and S.

If the local cost ψ depends only in position and not on direction, then, as discussed in Section 4.1.2, the value function $\mathcal{E}^*$ satisfies the Eikonal equation $|\nabla \mathcal{E}^*| = \psi$. If the local cost depends not only on position but also on direction, then the value function $\mathcal{E}^*$ satisfies the Hamilton-Jacobi-Bellman equation

$$\max_{\hat{\boldsymbol{d}} \in \mathbb{S}^{n-1}} \{ \nabla \mathcal{E}^*(\mathbf{p}).\hat{\boldsymbol{d}} - \psi(\mathbf{p}, \hat{\boldsymbol{d}}) \} = 0. \tag{77}$$

Consider the optimal control problem of determining a trajectory $\mathbf{p} : [0,1] \to \mathbb{R}^n$ that is optimal with respect to the functional

$$J(\mathbf{p}(\cdot), \mathbf{u}(\cdot)) = \int_0^1 L(\mathbf{p}(t), \mathbf{u}(t)) \, \mathrm{d}t.$$

The control $\mathbf{u}(\cdot)$ is defined by:

$$\frac{\partial \mathbf{p}}{\partial t}(t) = \mathbf{u}(t)$$

For any given starting point $\mathbf{p}_0$, define the value function as the minimum cost for reaching a seed region $S \subset \mathbb{R}^n$ from $\mathbf{p}_0$.

$$J^*(\mathbf{p}_0) = \inf_{\mathbf{u}(\cdot), \mathbf{p}(0)=\mathbf{p}_0, \mathbf{p}(1) \in S} J(\mathbf{p}(\cdot), \mathbf{u}(\cdot))$$

This problem satisfies Bellman's *principle of optimality* which states that if $\mathbf{p}^*(\cdot)$ is an optimal trajectory then all sub-path are also optimal. This can be expressed by the following relation:

$$J^*(\mathbf{p}_0) = \inf_{\mathbf{u}(\cdot), \mathbf{p}(0)=\mathbf{p}_0, \mathbf{p}(1) \in S} \{ \int_0^r L(\mathbf{p}(t), \mathbf{u}(t)) \, \mathrm{d}p + J^*(\mathbf{p}(r)) \}$$

This means that if an optimal trajectory $\mathbf{p}^*(\cdot)$ is found such that $\mathbf{p}^*(0) = \mathbf{p}_0$ and $\mathbf{p}^*(1) \in S$ then, for any $r \in]0,1[$ the sub-trajectories $\mathbf{p}^*_{|[0,r]}$ and $\mathbf{p}^*_{|[r,1]}$ are also optimal. See [113] for a detailed proof.

Assuming that all functions are sufficiently smooth and differentiating with respect to $r \to 0$

$$0 = \inf_{\mathbf{u(0)}} \{ \ L(\mathbf{p}(0), \mathbf{u}(0)). + \nabla J^*(\mathbf{p}(0)).\mathbf{p}_t(0) \ \}$$

Using $\mathbf{p}_t = \mathbf{u}$ and $\mathbf{p}(0) = \mathbf{p}_0$, the Hamilton-Jacobi-Bellman equation is obtained:

$$0 = \inf_{\mathbf{u(0)}} \{ \ L(\mathbf{p}_0, \mathbf{u}) + \nabla J^*(\mathbf{p}_0).\mathbf{u} \ \} \tag{78}$$

In general the value function may not be differentiable. In that case the differential equation (78) holds in the sense of *viscosity theory*. See [114].

This can be applied to cost functionals of the form

$$\begin{aligned} \mathcal{E}(\mathbf{\Gamma}) &= \int_0^{L(\mathbf{\Gamma})} \psi(\mathbf{\Gamma}(s), \mathbf{\Gamma}_s(s)) \ \mathrm{d}s \\ &= \int_0^1 \psi(\mathbf{\Gamma}(t), \mathbf{\Gamma}_t(t)/|\mathbf{\Gamma}_t(t)|)|\mathbf{\Gamma}_t(t)| \ \mathrm{d}t, \end{aligned} \tag{79}$$

where s is arclength, $L(\mathbf{\Gamma})$ is the complete length of the curve and t is some other parameterization.

The resulting Hamilton-Jacobi-Bellman equation is

$$0 = \inf_{\mathbf{\Gamma}_t(0)} \{ \ \psi(\mathbf{\Gamma}(0), \mathbf{\Gamma}_t(0)/|\mathbf{\Gamma}_t(0)|)|\mathbf{\Gamma}_t(0)| + \nabla \mathcal{E}^*(\mathbf{\Gamma}(0)).\mathbf{\Gamma}_t(0) \ \}$$

finally,

$$\begin{cases} 0 = \inf\limits_{\hat{\boldsymbol{d}} \in \mathbb{S}^{n-1}} \{ \ \psi(\mathbf{p}, \hat{\boldsymbol{d}}) + \nabla \mathcal{E}^*(\mathbf{p}).\hat{\boldsymbol{d}} \ \} \\ \mathcal{E}^*(s) = 0 \text{ for } s \in S. \end{cases} \tag{80}$$

This equation can be solved numerically [115, 116, 117]. From any point $\mathbf{p}_0 \in \mathbb{R}^n$, an optimal path in the sense of (79) can then be determined by following locally the vector $\hat{\boldsymbol{d}}^*$ for which the minimum is attained in (80).

The algorithm sweeps through all points $\mathbf{p}$ in search of the least expensive direction. The cumulated cost to reach $\mathbf{p}$ from direction $\hat{\boldsymbol{d}}$ is $f_{\mathcal{E}^*,\psi}(\mathbf{p}, \hat{\boldsymbol{d}}) \triangleq (\sum_{k=1}^n \alpha_k \mathcal{E}^*(\mathbf{p} + \boldsymbol{\delta}_k) +$

Algorithm 3: Sweeping algorithm to solve the Hamilton-Jacobi-Bellman equation (80), see [116]

Require: seed region S, direction-dependent local cost ψ

1: Initialize $\mathcal{E}^*(\cdot) \leftarrow +\infty$, except at starting points $s \in S$ where $\mathcal{E}^*(s) \leftarrow 0$
2: **repeat**
3: **sweep** through all voxels $\mathbf{p}$, in all possible grid directions
4: $\hat{\boldsymbol{d}}' \leftarrow \arg\min_{d \in \mathbb{S}^{n-1}} f_{\mathcal{E}^*,\psi}(\mathbf{p}, \hat{\boldsymbol{d}})$
5: **if** $f_{\mathcal{E}^*,\psi}(\mathbf{p}, \hat{\boldsymbol{d}}') < \mathcal{E}^*(\mathbf{p})$ **then** $\mathcal{E}^*(\mathbf{p}) \leftarrow f_{\mathcal{E}^*,\psi}(\mathbf{p}, \hat{\boldsymbol{d}}')$ and $\hat{\boldsymbol{d}}^*(\mathbf{p}) \leftarrow \hat{\boldsymbol{d}}'$ **end if**
6: **end sweep**
7: **until** convergence of $\mathcal{E}^*$
8: **return** value function $\mathcal{E}^*$, characteristics $\hat{\boldsymbol{d}}^*$

Note: as explained in [116] Step 3 can be written in C as:

```
for(w1=-1;w1<=+1;w1+=2)
   for(w2=-1;w2<=+1;w2+=2)
      for(x1=(w1<0?nx1:0);(w1<0?x1>=0:x1<=nx1);x1+=w1)
         for(x2=(w2<0?nx2:0);(w2<0?x2>=0:x2<=nx2);x2+=w2)
```

$\psi(\mathbf{p}, \hat{\boldsymbol{d}}))/(\sum_{k=1}^{n} \alpha_k)$, where the n neighbors[4] $\mathbf{p} + \boldsymbol{\delta}_1, \ldots, \mathbf{p} + \boldsymbol{\delta}_n$ of $\mathbf{p}$ in direction $\hat{\boldsymbol{d}}$ are interpolated using the components of the vector $\boldsymbol{\alpha} \triangleq [\boldsymbol{\delta}_1 \mid \ldots \mid \boldsymbol{\delta}_n]^{-1}\hat{\boldsymbol{d}}$. For example, if $\hat{\boldsymbol{d}} = \boldsymbol{\delta}_k/\|\boldsymbol{\delta}_k\|$ (i.e., $\hat{\boldsymbol{d}}$ points directly at one of the neighboring voxels) then $f = \mathcal{E}^*(\mathbf{p}+\boldsymbol{\delta}_k) + \psi(\mathbf{p}, \hat{\boldsymbol{d}})\|\boldsymbol{\delta}_k\|$. which is the cost for reaching voxel $\mathbf{p}$ from voxel $\mathbf{p} + \boldsymbol{\delta}_k$. This same quantity would be computed in Dijkstra's algorithm. Unlike Dijkstra's algorithm however, the search for the optimal direction is not restricted to discrete grid directions. and the minimization is performed continuously over the sphere $\mathbb{S}^{n-1}$. In our implementation the minimization is performed over 100 directions sampled uniformly on the sphere [5] and the coefficients $\boldsymbol{\alpha}(\hat{\boldsymbol{d}})$ are pre-computed.

4.4 Links to Biological Vision

4.4.1 Model of Direction-sensitive Neurons of the Visual Cortex

In a classical neuroscience experiment, Hubel and Wiesel [118, 119] showed that the spike rate of single neurons in particular areas of the visual cortex in response to drifting oriented

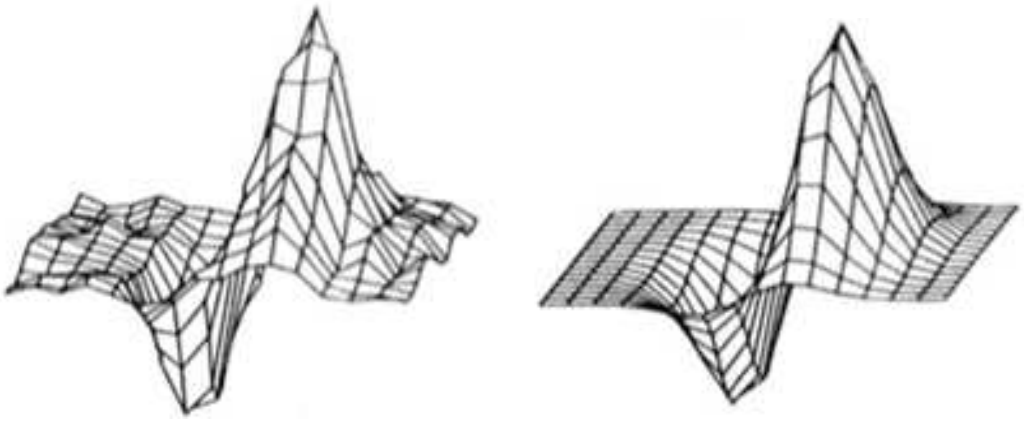
[4] in three dimensions, this is $n = 3$ neighbors among 26.
[5] for the algorithm to initialize properly, discrete grid directions have to be present

luminance bars and/or drifting luminance spots is strongly dependent on orientation. These results were later confirmed by many other researchers [120].

A linear model for the processing of such cells was proposed by Gaugman [121] that uses some sort of modulated Gabor function.

$$g(x,y) = \exp(\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \cos(2\pi \frac{x'}{\lambda} + \phi)$$

While it is clear to the neuroscience community that the processing of such cells is not purely linear, the model corresponds very well to experimental data [122], see Figure 16. This model has been used to detect edges and texture in image processing [123, 124].



(a) Experimental measure

(b) Mathematical model using Gabor windows

Figure 16: Measured receptive field of visual neuron and mathematical model, from [jones1987], see text.

In our case the pattern will be rotated in the direction of the outward normal **N**. A curve will therefore be optimum if it corresponds to the boundary between a dark region (inside) and a bright region (outside), see Figure 17.

4.4.2 Examples

This technique was applied to synthetic examples and real examples.

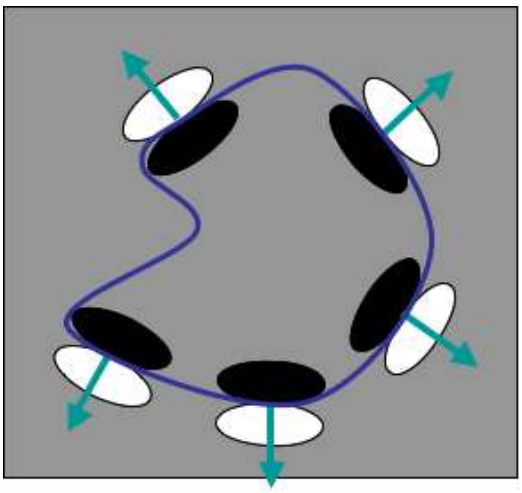


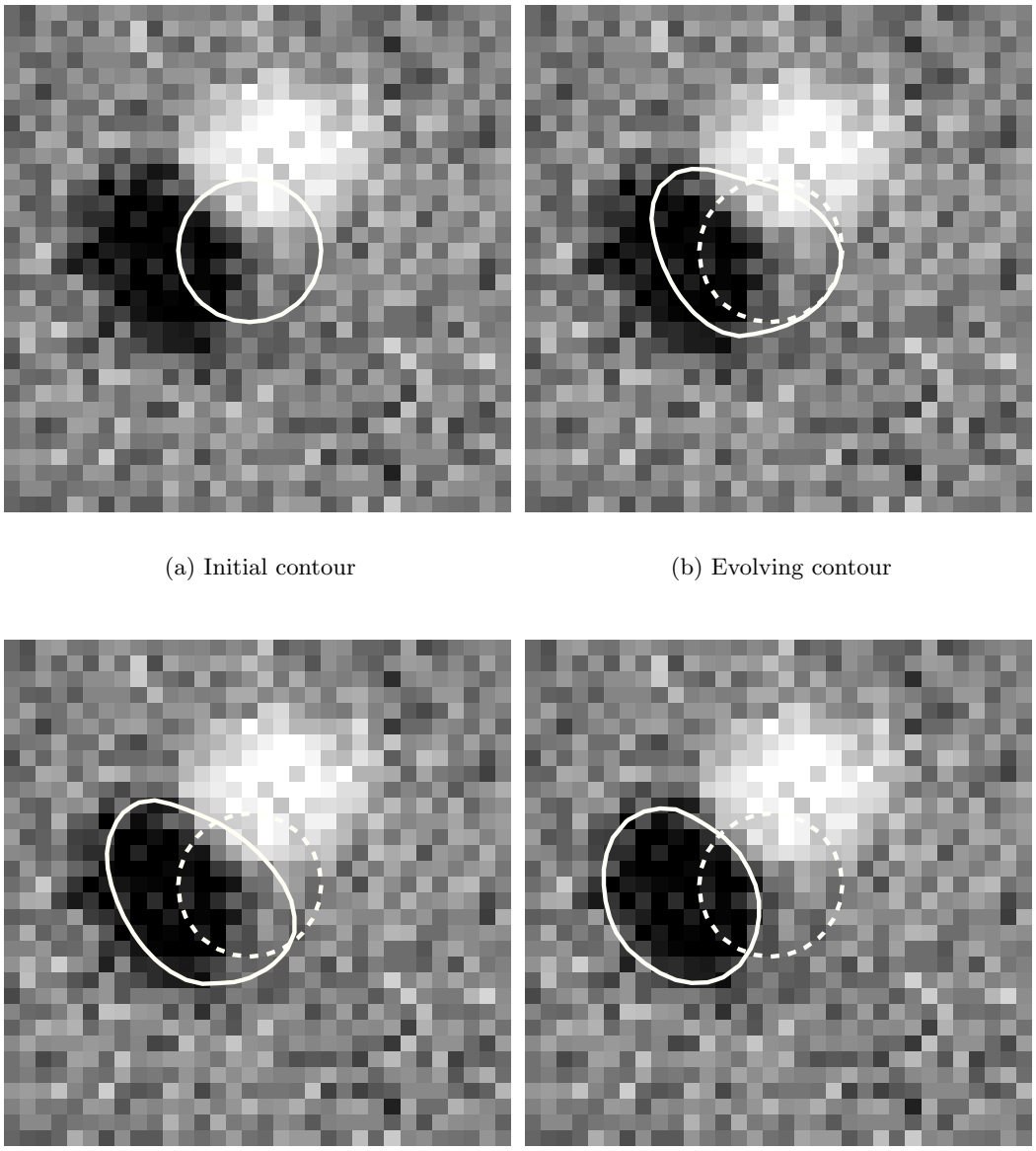
Figure 17: The total cost of a curve can be obtained by aligning the biologically-inspired pattern detector (see Figure 16, the positive and negative Gabor windows are represented here as bright and dark ellipses) with the normal of the curve and integrating along the curve. The cost is therefore low for the boundary of a dark region on a bright background.

4.4.2.1 Synthetic Images

In the synthetic case, if the closed surface Σ (here for $n = 2$, Σ is a closed planar curve) is initialized halfway between a dark and bright spot, it will, when deformed according to Equation (70) in order to minimize the total cost (68), be pushed away from the bright spot and match closely the boundaries of the dark spot that will be the steady state of the curve, see Figure 18.

If for the same image, intensities are reversed (such that the dark spot become the bright spot and vice-versa), then the steady state of the evolving curve corresponds to the boundaries of the other spot, see Figure 19.

Implementation with level sets (see Section 4.3.1.4) allows for changes in topology. For example, Figure 20 illustrates the breaking of a closed planar curve into two closed planar curves in order to capture the boundaries of two disconnected darks spots in another artificial image.



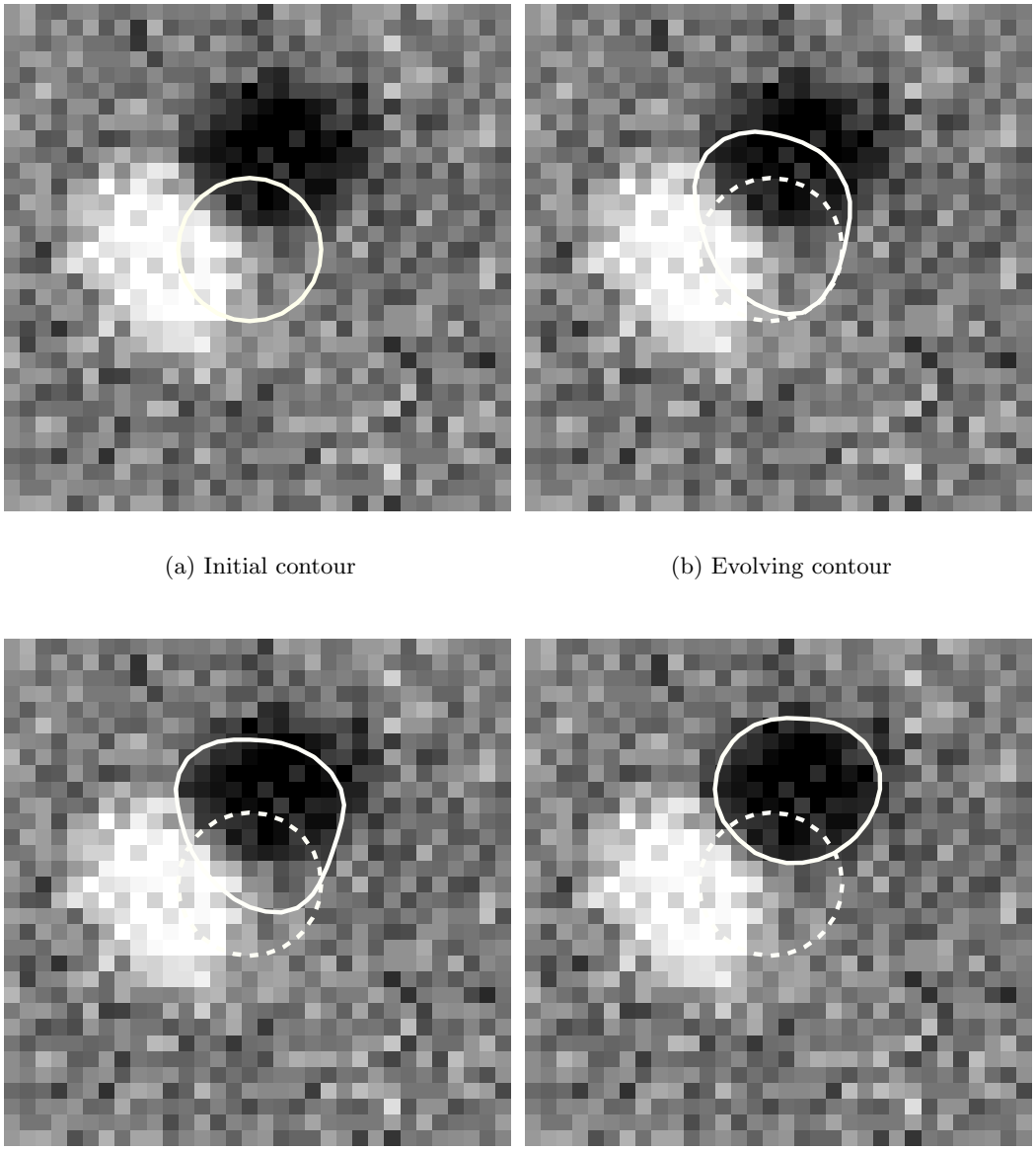
(a) Initial contour

(b) Evolving contour

(c) Evolving contour

(d) Steady state

Figure 18: Contour evolution on a synthetic example. The local cost is biologically inspired (see Section 4.4). The contour captures the dark region on the bottom left. Compare to Figure 19.



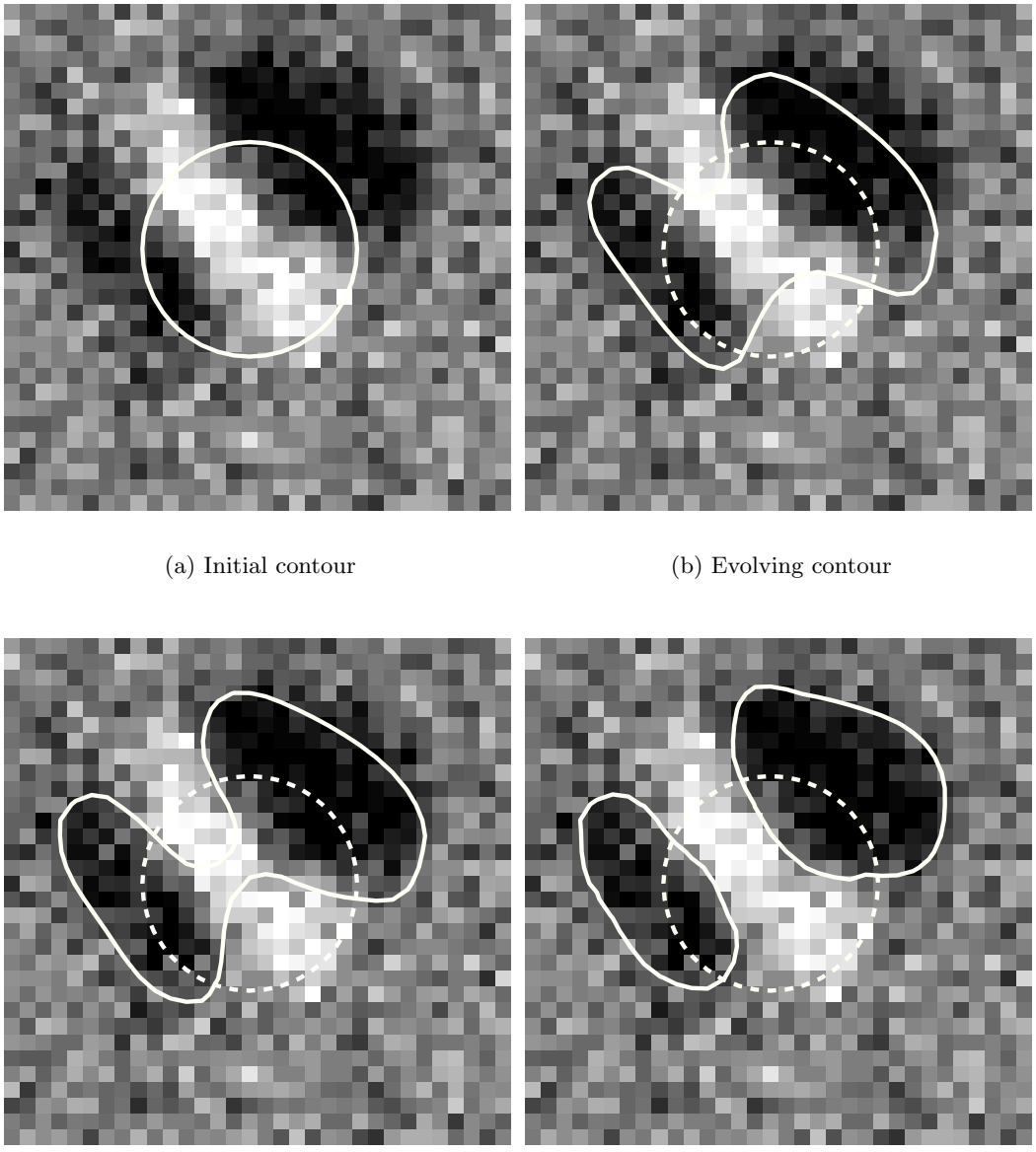
(a) Initial contour

(b) Evolving contour

(c) Evolving contour

(d) Steady state

Figure 19: Contour evolution on a synthetic example. The contour captures the dark region on the top right. Compare to Figure 18.



(a) Initial contour

(b) Evolving contour

(c) Evolving contour

(d) Steady state

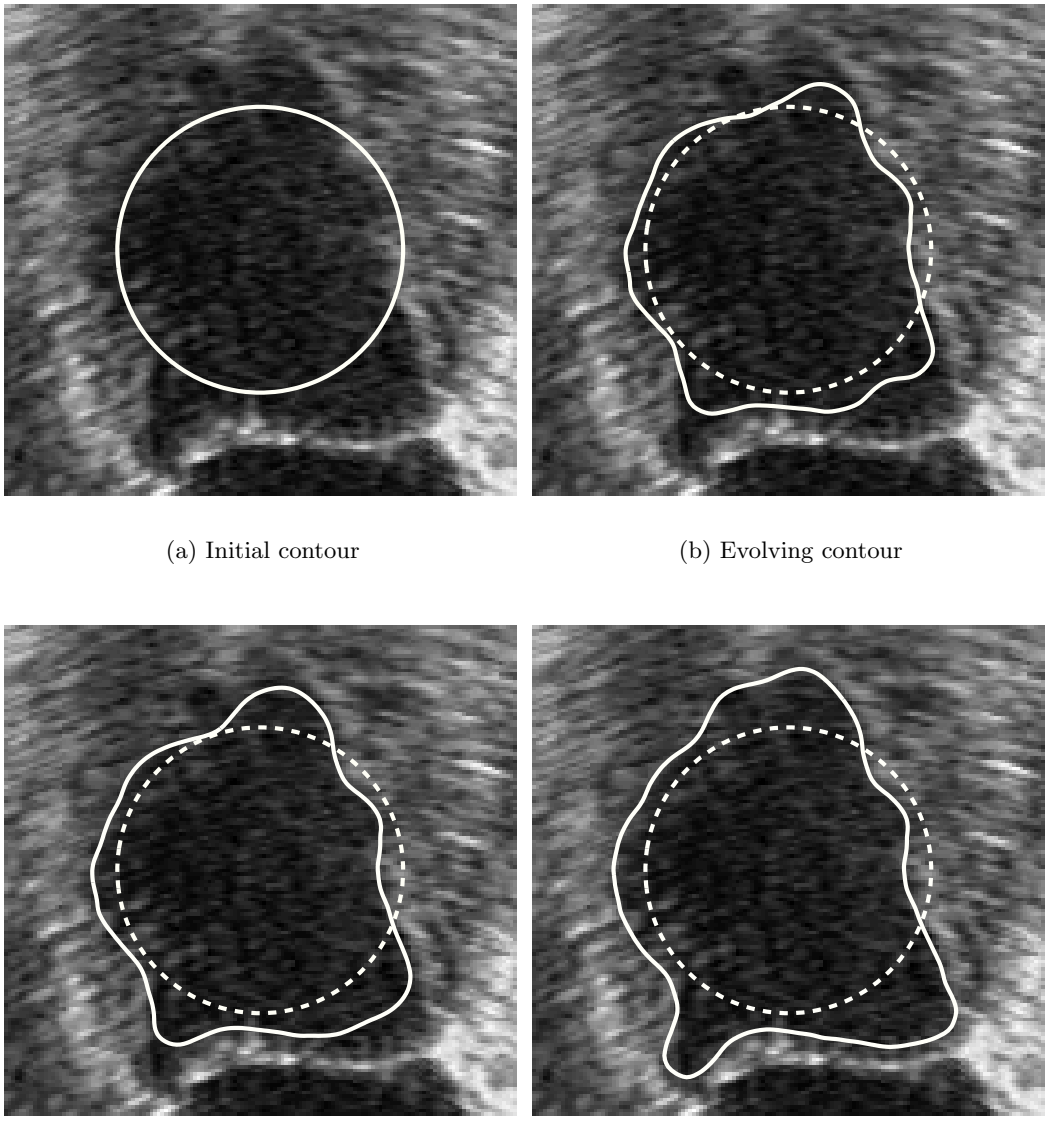
Figure 20: Contour evolution on a synthetic example. The evolving contour changes topology.

4.4.2.2 Real Images

When applied to real images, the proposed technique results in good segmentations in noisy images that would have been very challenging for previously proposed surface-based active contours. For example, see Figure 21 for the evolution of the contour on a real ultrasound image of the heart's left ventricle and Figure 22 on an ultrasound image of a breast cyst. The segmentation of these noisy ultrasound images would typically have necessitated a region-based technique in which all information inside the evolving contour is used. Instead in the proposed approach only information near the curve is used. Thanks to the presence of direction information, it is possible to discriminate between the inside and the outside of the curve and favor, for example here, a dark inside and a brighter outside. In previously proposed surface-based active contours, this information is not present and the immediate inside and outside of the curve must therefore be treated symmetrically which is very detrimental to the segmentation process.

While it is surface-based (the functional only involves integration on the surface Σ and not on the volume it encompasses) the proposed framework shares with region-based techniques the ability to discriminate between the inside and outside of the contour. Moreover, unlike previously proposed region-based techniques, the proposed framework can discriminate between information that is just on the outside of the curve or outside but relatively far away from the curve. This can prove to be very desirable on real images. For example, a typical region-based energy could be $\mathcal{E}(\Sigma) = H_{\text{inside}}(\Sigma) + H_{\text{outside}}(\Sigma)$ where H is some measure of the homogeneity of the image inside and outside Σ. While it might make sense to look for regions that are homogeneous inside, unless the image only contains one object on a uniform background, there is no reason to expect the outside to be also uniform. For example, when segmenting some brain structure such as the ventricle in an MRI dataset, the outside of that structure is itself composed of many different regions and it seems hardly justifiable that far away outside information such as say, gray levels in the region of the jaw, would play any role at all. In contrast, the proposed framework can use image information differently based not only on the binary fact that it is inside or outside but also on how far

away from the curve it is. In this way the processing can be limited to local information around the curve.



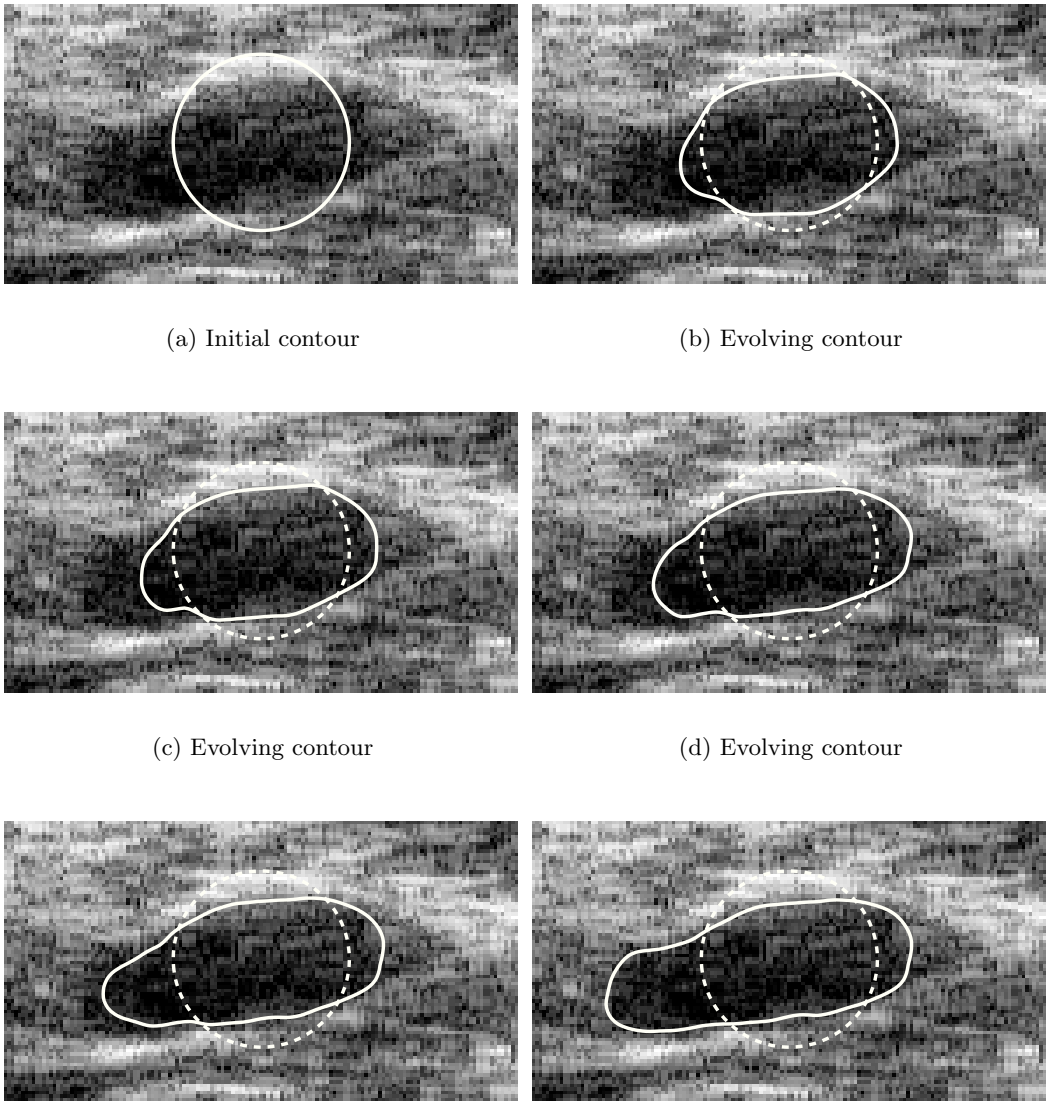
(a) Initial contour

(b) Evolving contour

(c) Evolving contour

(d) Steady state

Figure 21: Contour evolution on a real ultrasound image of the heart's left ventricle.



(a) Initial contour

(b) Evolving contour

(c) Evolving contour

(d) Evolving contour

(e) Evolving contour

(f) Steady state

Figure 22: Contour evolution on real ultrasound image of a breast cyst.

4.5 Links to Pattern Detection

In this section we propose to construct the conformal factor from some local pattern detector and propose a particular pattern detector inspired by image matching as well as a variational learning procedure.

4.5.1 Defining the Local Cost Using a Pattern Detector

The direction-dependent local cost $\psi(\mathbf{p}, \hat{\boldsymbol{d}})$ can be obtained by translating the image by $-\mathbf{p}$ and applying a direction-dependent pattern detector Φ.

More precisely, if a direction-dependent pattern detector is available of the form

$$\Phi : \mathbb{S}^{n-1} \times (\mathbb{R}^n \to \mathbb{R}^p) \to \mathbb{R}^{+*},$$

then the local cost can be obtained by

$$\psi(\mathbf{p}, \hat{\boldsymbol{d}}) \triangleq \Phi(\hat{\boldsymbol{d}}, I \circ T_{-\mathbf{p}}),$$

where $T_{-\mathbf{p}}$ is the translation by vector $-\mathbf{p}$ and the image is a map $I : \mathbb{R}^n \to p$.

Such a direction-dependent pattern can in turn be constructed by applying a regular pattern detector on a rotated image. This construction is illustrated on Figure 23.

For example, the local cost used in Section 4.4 could be constructed from the following pattern detector:

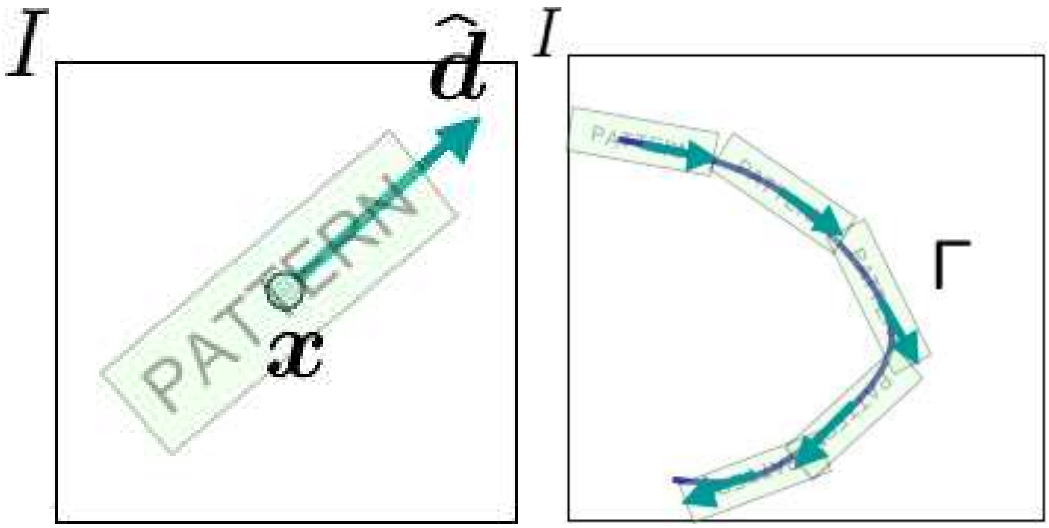
$$\Phi_{\hat{\boldsymbol{d}}}(I) = \exp \int_{\mathbb{R}^n} I(\mathbf{p})(G_{\hat{\boldsymbol{d}}}(\mathbf{p} + \hat{\boldsymbol{d}}) - G_{\hat{\boldsymbol{d}}}(\mathbf{p} - \hat{\boldsymbol{d}}))\ \mathrm{d}\mathbf{p},$$

where the exponential ensures that Φ is always positive and G is a Gabor function of the form

$$G_{\hat{\boldsymbol{d}}}(\mathbf{x}) = \alpha \mathbf{x}^t \hat{\boldsymbol{d}} \hat{\boldsymbol{d}}^t \mathbf{x} + \beta \mathbf{x}^t \mathbf{x}.$$

4.5.2 Properties

If the image and the pattern detector are smooth then the local cost is also smooth and the proposed construction verifies the following properties:



(a) The local cost $\psi(\mathbf{x}, \hat{\boldsymbol{d}})$ is the output of a pattern detector Φ that has been centered at position $\mathbf{x}$ and rotated in direction $\hat{\boldsymbol{d}}$.

(b) The global cost of a curve is then the integration of the output of the pattern detector along the curve.

Figure 23: The local cost ψ (and therefore also the global cost $\mathcal{E}$) can be constructed from a general pattern detector.

- Continuity with respect to the curve

$$(\boldsymbol{\Gamma}' \to \boldsymbol{\Gamma}) \Rightarrow (\mathcal{E}(\boldsymbol{\Gamma}', \psi_I) \to \mathcal{E}(\boldsymbol{\Gamma}, \psi_I)) \tag{P1}$$

  Human experts typically do not agree perfectly on the position of a certain feature. It is important to capture the inherently fuzzy nature of segmentation.

- Continuity with respect to the image

$$(I' \to I) \Rightarrow (\mathcal{E}(\boldsymbol{\Gamma}, \psi_{I'}) \to \mathcal{E}(\boldsymbol{\Gamma}, \psi_I)) \tag{P2}$$

  The presence of (limited) noise on the image should not impact drastically the segmentation result.

- Invariance with respect to Euclidean transformations

$$\forall T \text{ Euclidean } \mathbb{R}^n \to \mathbb{R}^n, \mathcal{E}(\boldsymbol{\Gamma}, \psi_{I \circ T}) = \mathcal{E}(T \circ \boldsymbol{\Gamma}, \psi_I) \tag{P3}$$

  The image is not assumed to be oriented in a meaningful way.

Moreover, if the pattern detector is local, i.e., if it only utilizes information within a finite radius[6] then the following additional property is fulfilled:

- Locality of information employed

$$\exists r_0 > 0, \forall \mathbf{\Gamma}, \forall I, I', (\forall \mathbf{x} \in \mathbb{R}^n, d(\mathbf{x}, \mathbf{\Gamma}) < r_0 \Rightarrow (I(\mathbf{x}) = I'(\mathbf{x})) \Rightarrow (\mathcal{E}(\mathbf{\Gamma}, \psi_{I'}) = \mathcal{E}(\mathbf{\Gamma}, \psi_I)) \tag{P4}$$

4.5.3 Learning

We propose to use a parametric model for the pattern detector Φ. If the image is expressed in cylindrical coordinates $I(l, r, \boldsymbol{\theta})$ where l is the abscissa along the axis $\hat{\boldsymbol{e_1}}$ we define

$$\Phi(I) \triangleq \int_l \int_r \int_{\boldsymbol{\theta}} [I(l, r, \boldsymbol{\theta}) - \mu(l, r)]^t \Sigma^{-1}(l, r) [I(l, r, \boldsymbol{\theta}) - \mu(l, r)] \, \mathrm{d}^{n-2}\boldsymbol{\theta} \; r^{n-2} \, \mathrm{d}r \, \mathrm{d}l.$$

This can be interpreted as a Mahalanobis distance with a mean vector field $\mu : \mathbb{R}^n \to \mathbb{R}^p$ and a positive semi-definite covariance matrix field $\Sigma : \mathbb{R}^n \to S^+(\mathbb{R}^p)$. In order to satisfy the locality property (P4), set $\Sigma^{-1}(l, r) = \mathbf{0}$ for l and r larger than some constant r_0. In order to satisfy the continuity properties, (P1) and (P2) the fields have to be continuous.

Given a sample image I for which the optimal curve $\mathbf{\Gamma}^*$ is given (for example by a human expert), define the optimal conformal factor:

$$\psi_I^* \triangleq \arg\min_{\psi_I} \mathcal{E}(\mathbf{\Gamma}^*, \psi_I) - \mathbb{E}_{\mathbf{\Gamma}}\{ \; \mathcal{E}(\mathbf{\Gamma}, \psi_I) \; \} + R(\psi_I), \tag{81}$$

where R is some regularizing term that ensures that ψ_I remains regular and the expectancy $\mathbb{E}_{\mathbf{\Gamma}}\{\cdot\}$ is taken on all curves of same length as $\mathbf{\Gamma}^*$, which is equivalent to considering the mean value of ψ along the curves. By applying the aligning transforms $T_{\mathbf{\Gamma}^*}(s)$ to I for different values of the arc-length parameter s, a set of positive examples $\{ \; I_p^+ \; \}_{p=1 \ldots n^+}$ is obtained. By taking random Euclidean transformations of I we obtain a set negative examples $\{ \; I_m^- \; \}_{m=1 \ldots n^-}$. If the number of examples is large enough a Monte-Carlo approximation of (81) is achieved. The corresponding optimal local pattern detector is then

[6] this is a bottom-up assumption. We exclude here higher forms of top-down reasoning, for example determining the location of an artery from that of the organs it is connected to.

$$\Phi^* = \arg\min_{\Phi} \frac{1}{n^+} \sum_p \Phi \circ I_p^+ - \frac{1}{n^-} \sum_m \Phi \circ I_m^- + R(\Phi).$$

This is a variational problem. Starting from some initial guess (for example the mean and covariance fields on the positive examples set), the fields $\Phi = (\mu, \Sigma)$ can be deformed to minimize the energy. After computing the first variation,[7]

$$\begin{aligned}
\frac{\partial \mu}{\partial t} =& \Delta\mu + S^t S \int_{\boldsymbol{\theta}} (I - \mu)\ \mathrm{d}^{n-2}\boldsymbol{\theta}\ r^{n-2} \\
\frac{\partial S}{\partial t} =& \Delta S - S[\int_{\boldsymbol{\theta}} (I - \mu)(I - \mu)^t\ \mathrm{d}^{n-2}\boldsymbol{\theta}\ r^{n-2}] \\
&+ [\Phi(I) / \iint \sum_{i,j} s_{i,j}^2 r^{n-2} \int\ \mathrm{d}^{n-2}\boldsymbol{\theta}\ \mathrm{d}r\ \mathrm{d}l] S,
\end{aligned}$$

where the evolution was constrained to $\|\Sigma^{-1}\|_2$ constant and, to ensure positive semi-definiteness it is S defined by $S^t S \triangleq \Sigma^{-1}$ that is deformed.

The learning procedure respects the extra property:

- Invariance with respect to invertible affine transformations of the color space

$$\boldsymbol{\Gamma}^*, I, T \text{ affine invertible } \mathbb{R}^p \to \mathbb{R}^p \Rightarrow \psi^*_{T \circ I} = \psi^*_I \tag{P5}$$

4.5.4 Examples

The proposed technique was applied to two two-dimensional images of blood vessels and roads.

4.5.4.1 Vessel Detection

The proposed framework was applied to an image obtained as a two-dimensional projection of an angiogram of blood vessels. A short section of the vessel was manually segmented (see Figure 24). This was used to learn a direction dependent pattern detector in the following way: for a given point on the curve, a rectangular window was extracted aligned with the tangent of of the curve, of width slightly more than the radius of the vessel and of length about the double of its width. This window was then made to slide along the curve an the

[7] this evolution is for only one positive example, the full case is obtained by linearity

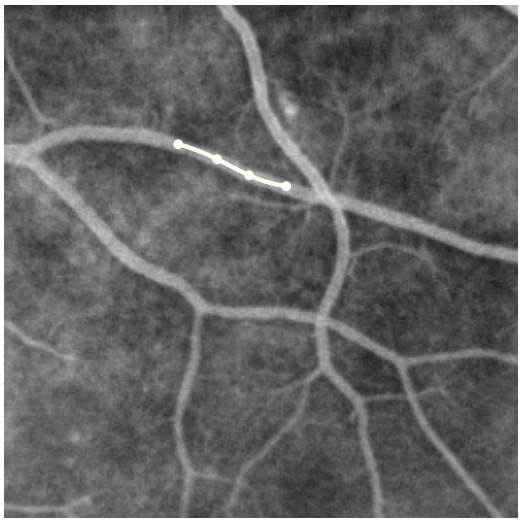


Figure 24: Vessel image and manually determined curve used for learning the pattern detector.

mean value of the intensity as well as the variance were computed. The pattern detector was then set to:

$$\Phi(\hat{\boldsymbol{d}}, I) = \exp \int_{\text{window}} (I_{\hat{\boldsymbol{d}}}(\mathbf{p}) - \mu(\mathbf{p}))^t \Sigma(\mathbf{p}) (I_{\hat{\boldsymbol{d}}}(\mathbf{p}) - \mu(\mathbf{p})) \, \mathrm{d}\mathbf{p},$$

where $I_{\hat{\boldsymbol{d}}}$ is the original image I rotated in direction $\hat{\boldsymbol{d}}$ and μ and Σ are the mean and variance of the intensities determined from the previously mentioned procedure. If the image intensities are very similar to the mean value μ (taking into account the variance Σ at that point) then the integration will be small and the output of the pattern detector will also small.

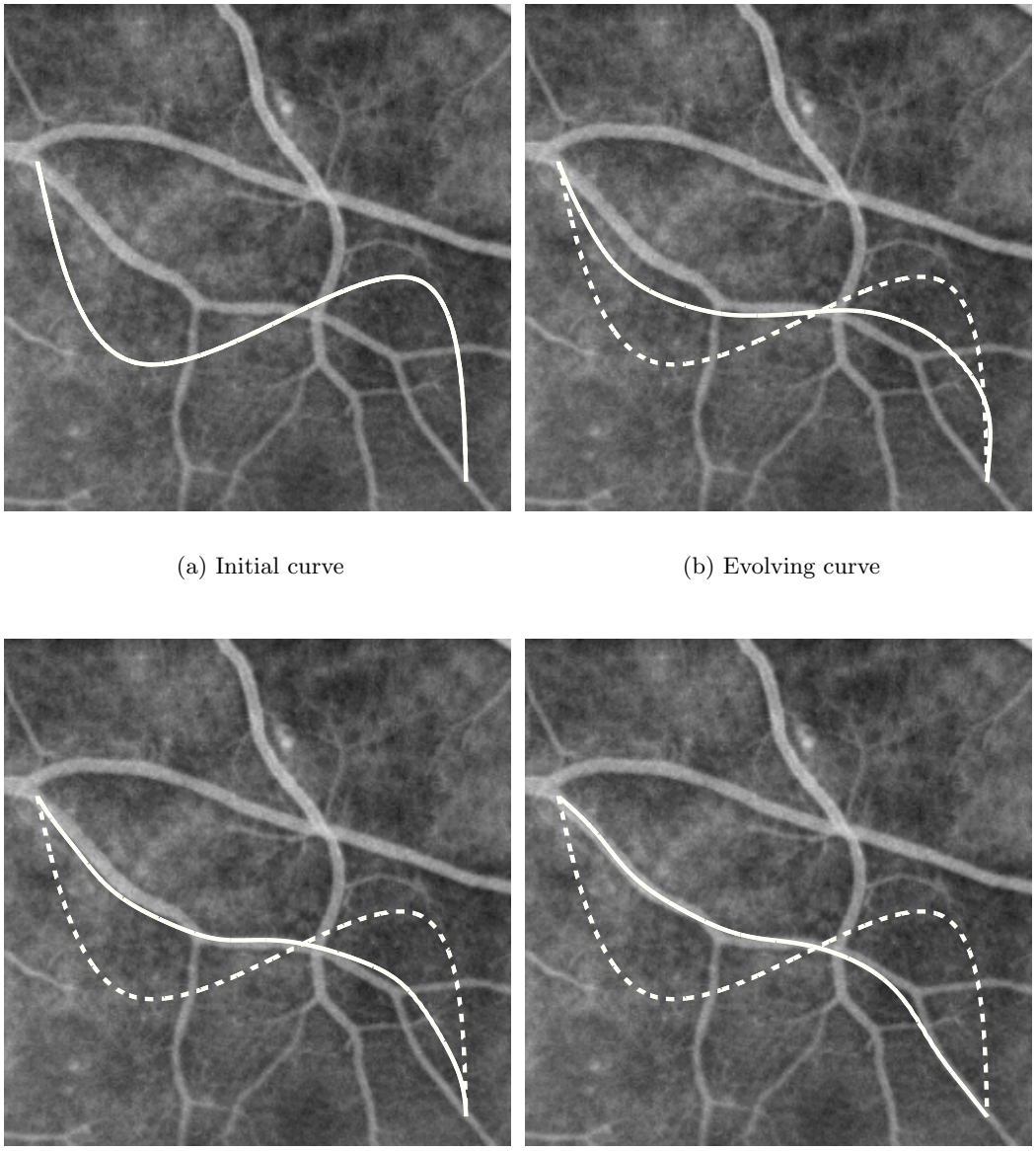
That pattern detector can then be applied to all points of the image to define the local cost.

Figure 25 shows the evolution of a curve as it is deformed according to the gradient flow (69) to minimize the global cost (67). This deformation optimally minimizes the integrated output of the learned pattern detector, i.e., it ensures that the curve locally resembles as much as possible the pattern learned from Figure 24. At steady state, the curve matches the road. It should be noted that since the global cost is a weighted length,

it can sometimes be advantageous to "cut corners" and incurring a slightly higher local cost (because the pattern is not matched perfectly) on a shorted trajectory. An alternative could be to minimize the average cost $\mathcal{E}(\Gamma)/L(cur)$ (where $L(\Gamma)$ is the length of the curve). However in such an energy the regularizing curvature term would drop out of the curve evolution equation.

Noise was then added to the image in order to test the robustness of the technique. The variance σ_I^2 of the image intensities was computed and Gaussian white noise was added with a variance σ_{noise}^2. The resulting image I_{noise} was then thresholded in order to maintain a dynamic range of $[0, 255]$ using the same sampling of the color space as I. The Signal to Noise Ratio was defined to be $\text{SNR} = \sigma_I^2/\sigma_{\text{noise}}^2$. Figures 26 and 27 illustrate the evolution of a user-defined initial curve to steady state. Even though the images are relatively noisy, the centerline of the vessel is recovered (except for a small section of the curve that got trapped into a local minimum on the left of Figure 27). This demonstrate the robustness of the technique due to the fact that the processing of the underlying pattern detector is not purely local. The use of information on a small window (rather than in a purely local way) allows for the averaging out of the noise and explains the good results of the algorithm even on low SNR images. This would not have been possible using the non direction-dependent in which local cost is defined directly by some function of the norm of the gradient.

Results using dynamic programming are illustrated for the original image on Figure 28 and for SNRs of 1.00 and 0.50 on Figure 29 and SNRs of 0.25 and 0.125 on Figure 30. The dynamic programming procedure was employed two times independently with different user-selected seed regions shown as one white and one black large discs.



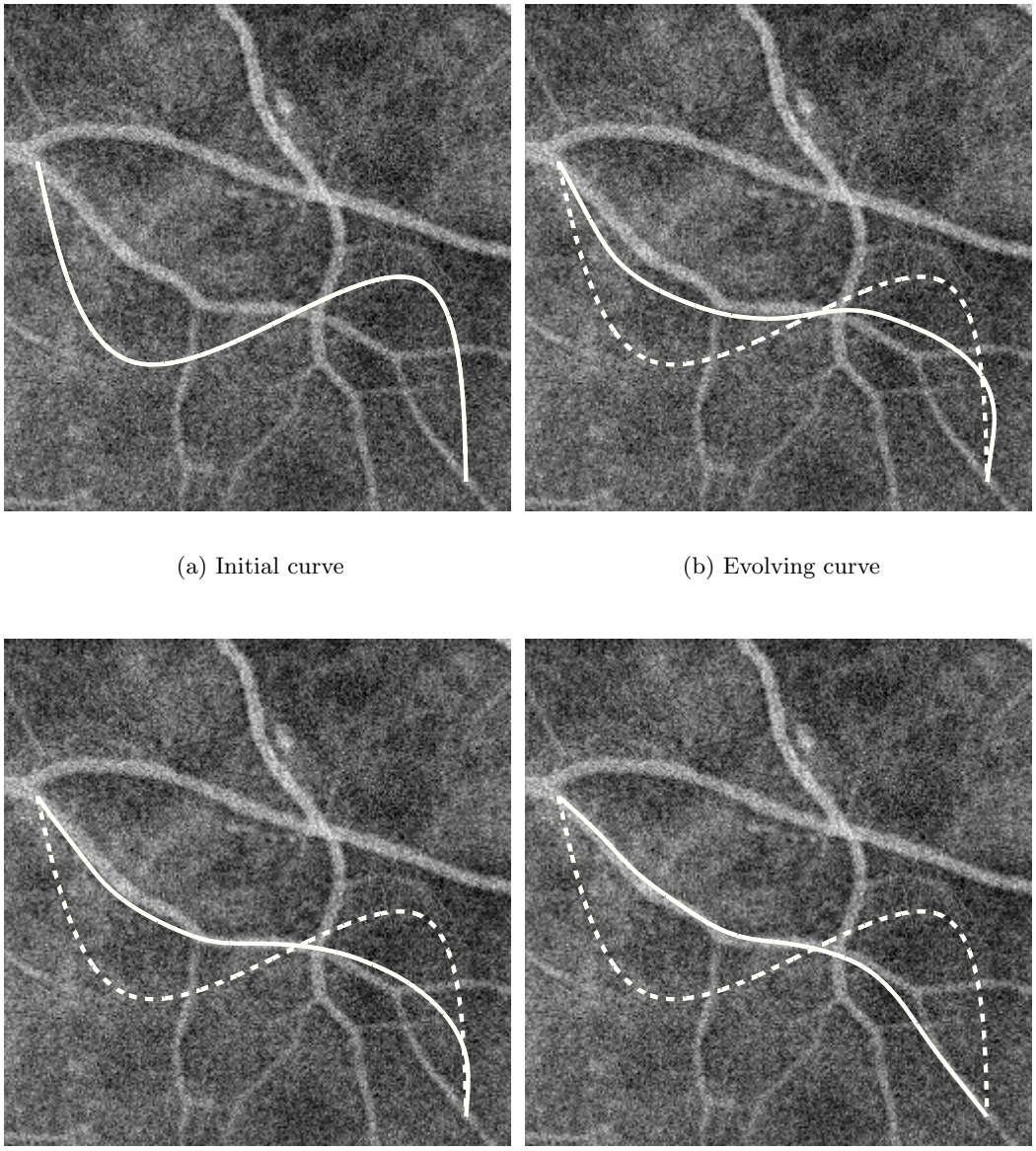
(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 25: Curve evolution on a real image. The local cost is determined using a pattern detector, see text and Figure 24.



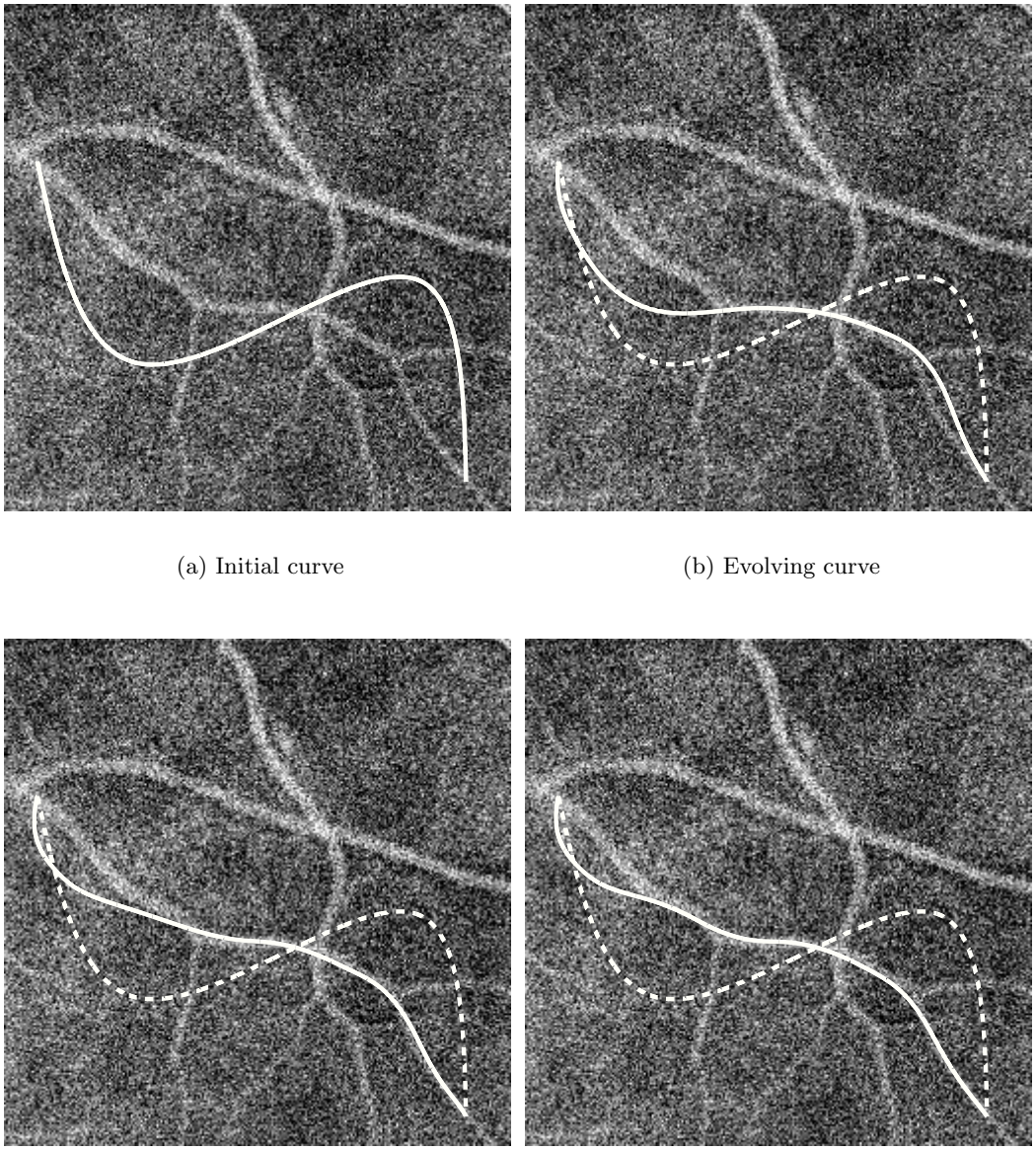
(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 26: Curve evolution on a real image. Signal-to-noise ratio is 1.00. Compare to Figure 25.



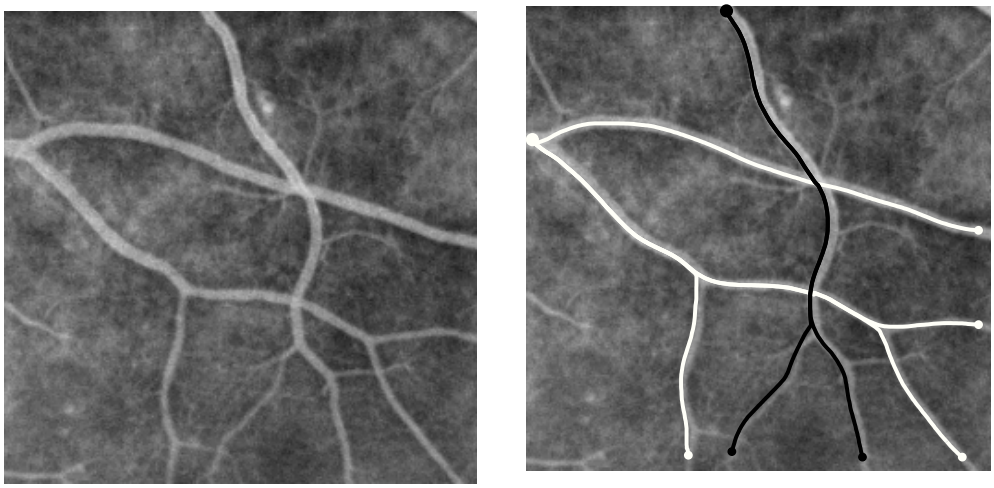
(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

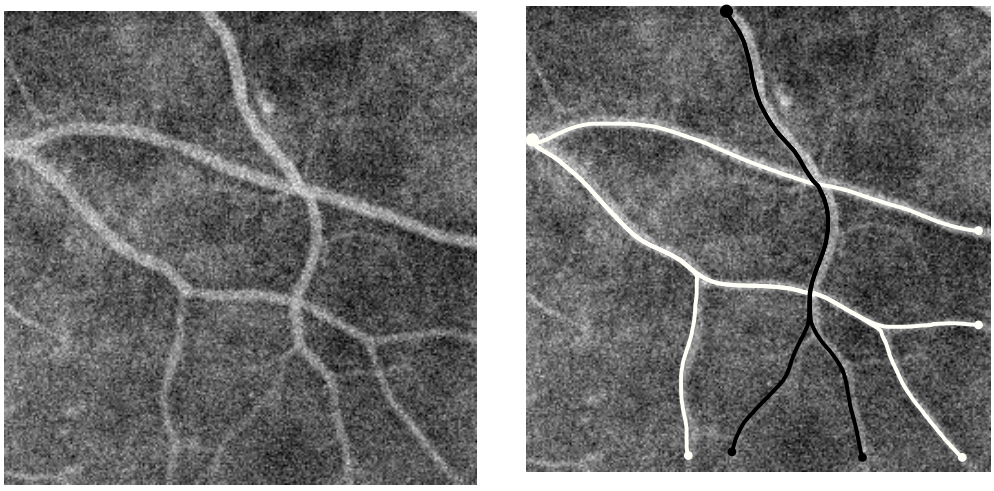
Figure 27: Curve evolution on a real image. Signal-to-noise ratio is 0.50. Compare to Figure 25.



(a) Original image.

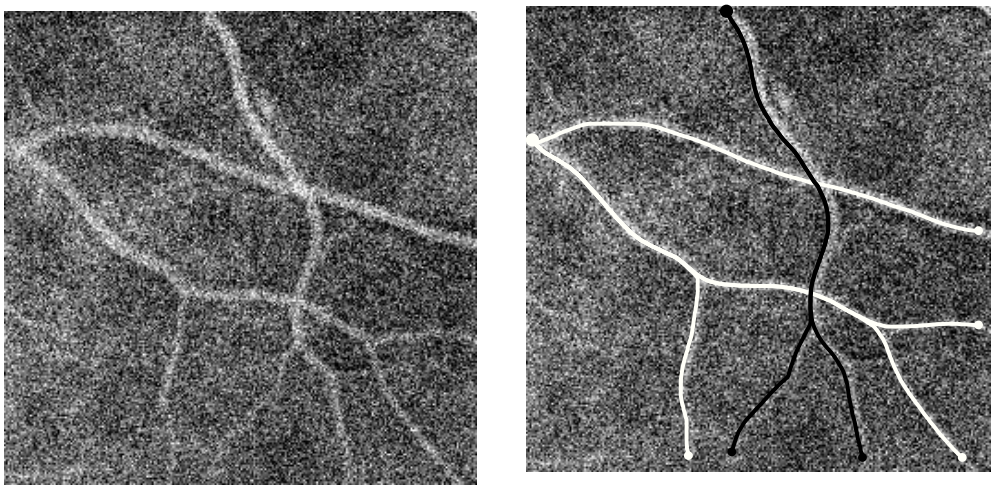
(b) Recovered curves.

Figure 28: Vessel detection using dynamic programming. The procedure was run independently for two seed points (large discs) and several target points (small discs).



(a) Original image, SNR=1.00.

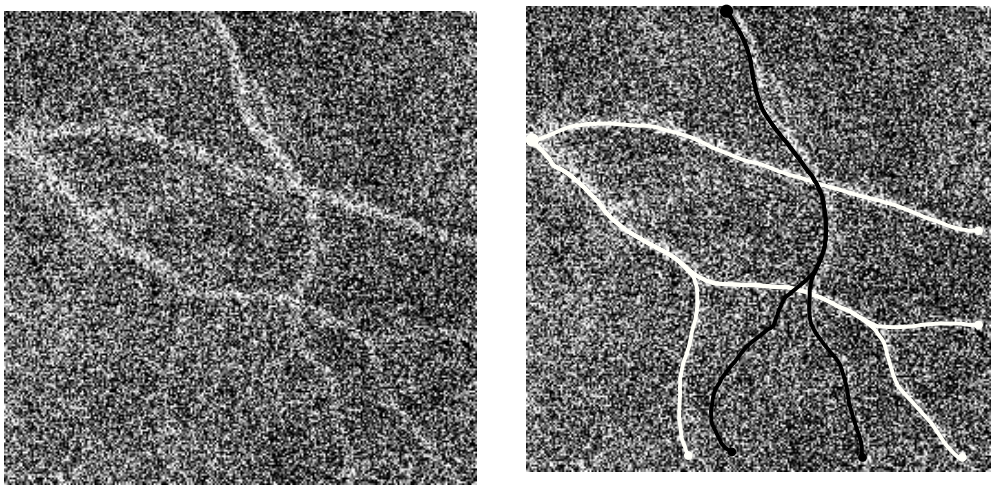
(b) Recovered curves.



(c) Original image, SNR=0.50.

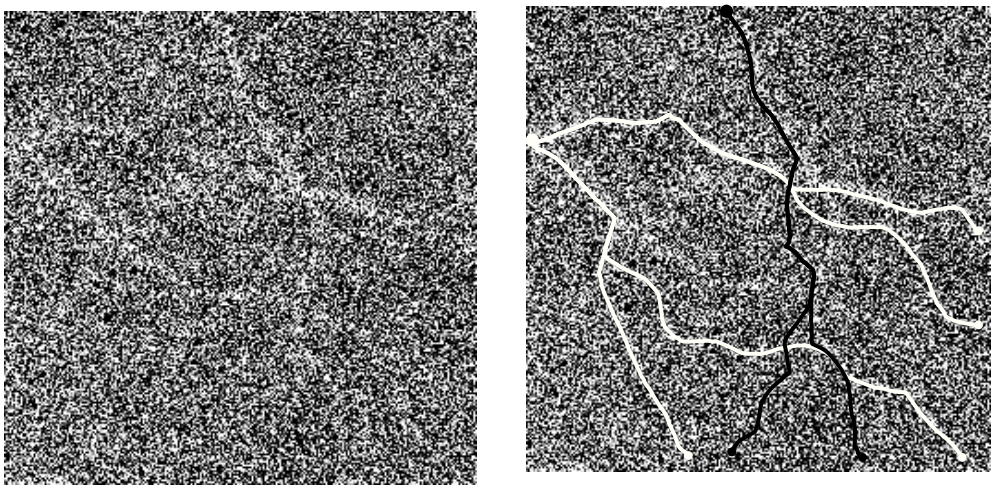
(d) Recovered curves.

Figure 29: Results of vessel detection using dynamic programming on noisy images. Compare to Figure 28.



(a) Original image, SNR=0.25.

(b) Recovered curves.



(c) Original image, SNR=0.125.

(d) Recovered curves.

Figure 30: Results of vessel detection using dynamic programming on noisy images. Compare to Figure 28.

Figure 31: Road image and manually determined curve used for learning the pattern detector.

4.5.4.2 Road Detection

The same protocol as in Section 4.5.4.1 was applied to a road detection task. Figure 31 shows the portion of the road that was used to learn the pattern detector. Figures 32 and 33 show two different initial curves converging to the same portion of the road. Figures 34, 35 illustrate results on images with artificially added noise.

Figures 36, 37 and 38 show the evolution of a self-intersecting initial curve for different levels of added noise.

Finally, Figures 39 and 40 illustrate the use of dynamic programming.

(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 32: Curve evolution on a real image. The local cost is determined using a pattern detector (see text and Figure 31). The initial curve is shown dashed. Compare to Figure 33.

(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 33: Curve evolution on a real image with a different initialization. Compare to Figure 32.

(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 34: Curve evolution on a real image. Signal-to-noise ratio is 1.00. Compare to Figure 33.

(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 35: Curve evolution on a real image. Signal-to-noise ratio is 0.50. Compare to Figure 33.

(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 36: Curve evolution on a real image for another initial curve.

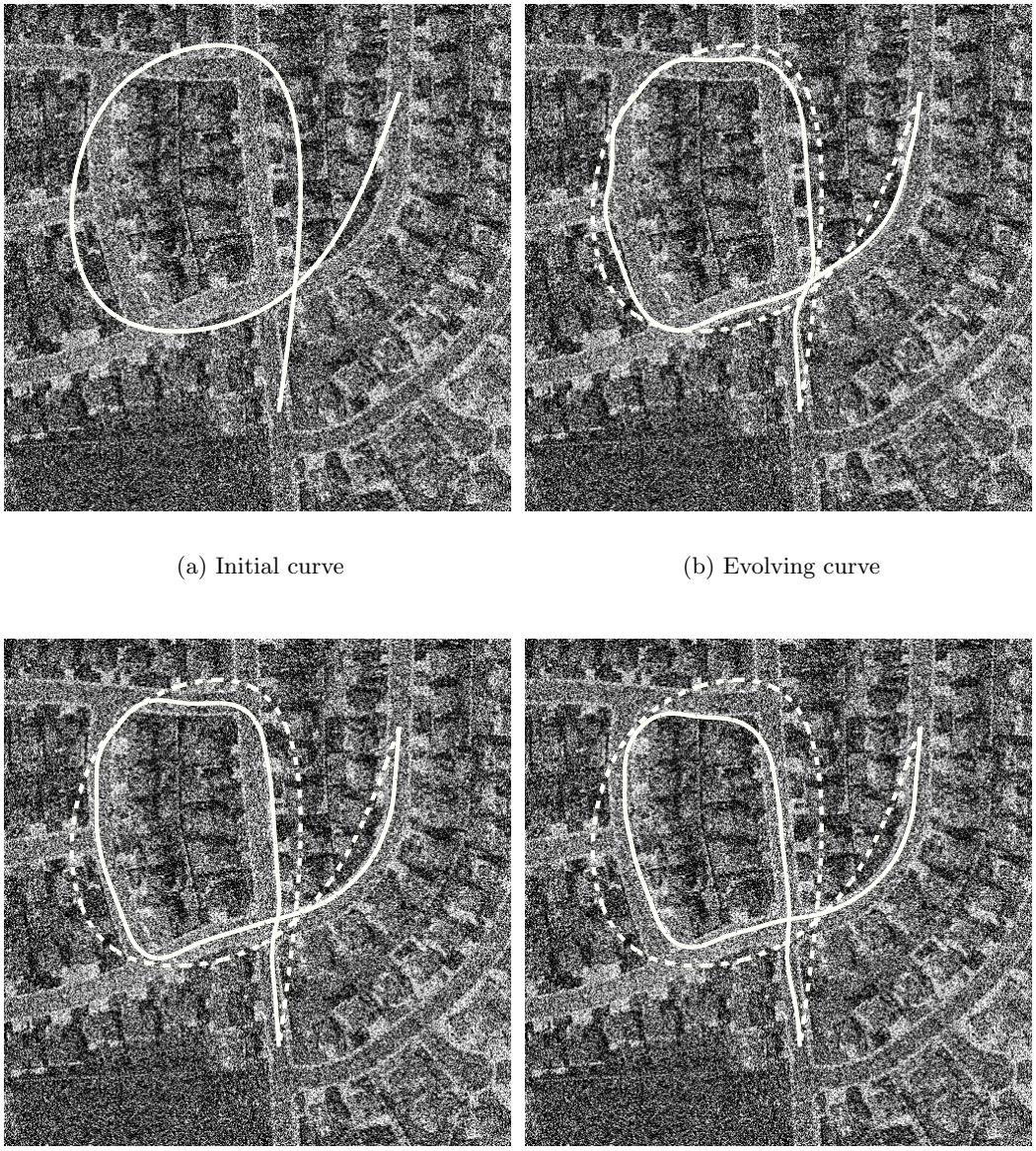
(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 37: Curve evolution on a real image. Signal-to-noise ratio is 1.00. Compare to Figure 36.



(a) Initial curve

(b) Evolving curve

(c) Evolving curve

(d) Steady state

Figure 38: Curve evolution on a real image. Signal-to-noise ratio is 0.50. Compare to Figure 36.

(a) Original image

(b) Recovered curve.

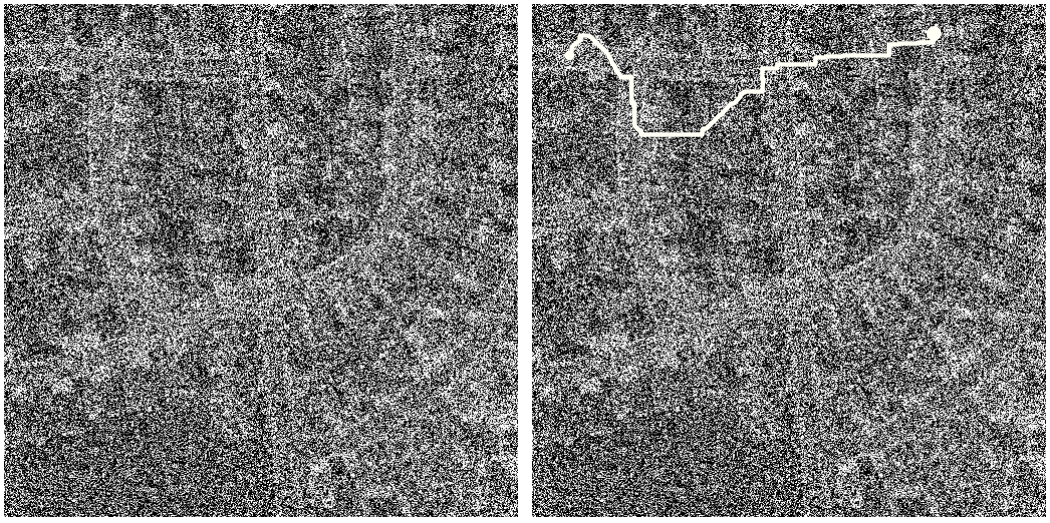
(c) Noisy image, SNR=1.00.

(d) Recovered curve.

Figure 39: Results of road detection using dynamic programming on noisy images. Compare to Figure 40.

(a) Noisy image, SNR=0.50.

(b) Recovered curve.



(c) Original image, SNR=0.25.

(d) Recovered curve.

Figure 40: Results of road detection using dynamic programming on noisy images. Compare to Figure 39.

4.6 Application to Neural Tracts Detection in Diffusion Magnetic Resonance Imagery

In this section we show how the proposed framework can be used for the extraction of neural fibers from diffusion magnetic resonance imagery (diffusion MRI). The main difference with previous applications of the technique is that no direction-dependent pattern detector needs to be applied to the image. Instead the image information at a given point can directly be treated as a direction-dependent pattern detector. This makes the proposed technique particularly natural for this type of imagery.

4.6.1 Diffusion MRI and Tractography

Water molecules in the biological tissues are in constant Brownian motion due to their thermal energy. This motion is influenced by the organization of the tissue. In particular, in the brain, membranes of axons and myelin sheaths tend to restrict movement in the direction of the neural fibers [125].

If during the acquisition of a Magnetic Resonance Imagery (MRI) dataset, electromagnetic gradient pulses are generated in a given direction $\hat{\boldsymbol{d}}$ of space, then an image $S(\hat{\boldsymbol{d}})$ is acquired. The intensity loss relative to an image $S(\mathbf{0})$ for which no gradient was applied is due to the diffusion of water molecules in direction $\hat{\boldsymbol{d}}$. This technique is called *Diffusion Weighted MRI*. If the diffusion is assumed to be governed by a Gaussian distribution, then all diffusion information can be captured by a tensor field D and the technique is then called *Diffusion Tensor MRI* [126, 127]. The development of Diffusion Tensor MRI has raised hopes in the neuroscience community for in vivo methods to track fiber paths in white matter.

A simple and effective method for tracking nerve fibers using DT-MRI is to follow the direction of maximum diffusion at each voxel, which is equivalent to the direction of the main eigenvector for each tensor [128, 129, 130, 131]. Although this method is widely spread and used in various ways the fiber trajectory is based solely on local information which makes it very sensitive to noise. Moreover the major direction of diffusion can become ill-defined for example at fiber crossings.

It has been proposed to shift from the Lagrangian, particle[8] based streamline approach described above to a Eulerian front propagation approach that can use full tensor information and is more robust to noise [132, 133]. This can be set in a Riemannian framework [134, 135].

The Gaussian assumption of diffusion tensor imaging does not hold for example if several fibers with different directions co-exist within the same voxel. Extensions to the tensor model, such as using multiple or higher order tensors, have been proposed [136].

High angular resolution diffusion modalities such as Q-Ball imaging [137] acquire diffusion information in potentially hundreds of directions thus measuring direction information in a non-parametric way.

Hagmann et al. [138] propose a streamline technique for such datasets. Anisotropy information is reduced to a multi-valued vector field corresponding to the detected directions and curves are sought that are locally aligned with one of the vectors of the set. Campbell [139] proposes a front evolution approach based on high angular resolution data.

Using the previously developed framework, tractography is set in a continuous minimum cost framework. This is different from [138, 139] who do not propose variational (cost minimizing) techniques. Local costs are defined for every direction on the unit sphere based on high angular resolution diffusion imagery. Equivalently this can be considered a minimum arrival time framework in which the speed of fictitious particles would be the inverse of the cost. The cost ψ can be interpreted as the inverse of the speed of a particle traveling along the curve. The value of function is dynamic programming, $\mathcal{E}^*(\mathbf{p})$, is then the minimum arrival time at $\mathbf{p}$. Alternatively, an infinite number of particle departing from S at the same time would propagate as a front geometrically evolving in the direction of its normal with a speed F. The corresponding equation is $\|\nabla\mathcal{E}^*(\mathbf{p})\|F(\mathbf{p}, \nabla\mathcal{E}^*(\mathbf{p})/\|\nabla\mathcal{E}^*(\mathbf{p})\|) = 1$. The two speeds F and $1/\psi$ are not identical because particles are not restricted to moving along the normal of the evolving front. Front speed and cost (or its inverse, particle speed)

[8] "particle" refers to the position of a fictitious evolving point

are related through the Legendre transformations [140]

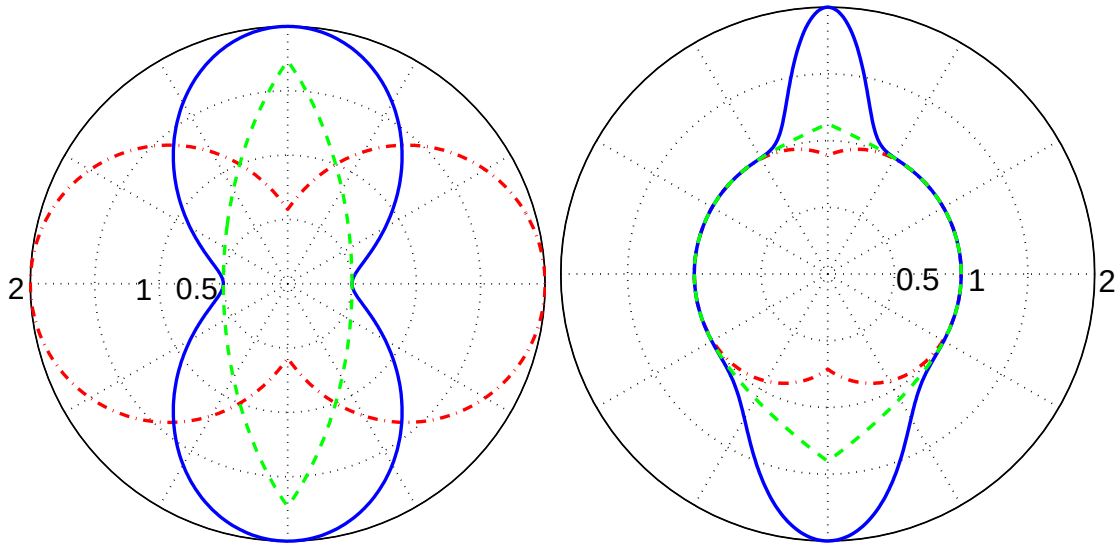
$$F(\mathbf{p},\hat{\boldsymbol{n}}) = 1/\min_{\hat{\boldsymbol{d}}.\hat{\boldsymbol{n}}>0}\{\ \frac{\psi(\mathbf{p},\hat{\boldsymbol{n}})}{\hat{\boldsymbol{d}}.\hat{\boldsymbol{n}}}\ \} \text{ and } \psi(\mathbf{p},\hat{\boldsymbol{d}}) = \max_{\hat{\boldsymbol{d}}.\hat{\boldsymbol{n}}>0}\{\ (\hat{\boldsymbol{d}}.\hat{\boldsymbol{n}})/F(\mathbf{p},\hat{\boldsymbol{n}})\ \}. \tag{82}$$

Previously proposed anisotropic front propagation techniques for diffusion MRI tractography are not set in a Hamilton-Jacobi-Bellman framework. Consequently it is the front speed F which is defined from the diffusion data. Curves will then be determined that are optimal for $\mathcal{E}$, which can only be interpreted as a cost (or an arrival time) in terms of ψ. The *max* operator present in the definition of ψ from F will tend to filter out the highest values of the front speed F that correspond to the preferred spatial directions (and therefore potential fibers) while preserving the slowest directions. Simulations and further analysis show that this distortion affects particularly speed functions with very localized direction information such as those encountered in high angular resolution diffusion imagery and can even result in loss of information. This problem can be avoided by setting directly the problem in a minimum cost framework (as proposed) or taking into account the Legendre transformation (82) when defining F. While the effects of this distortion are already noticeable for the very commonly used quadratic speed $F(\hat{\boldsymbol{d}}) = \hat{\boldsymbol{d}}^t D\hat{\boldsymbol{d}}$ (Figure 41(a)) the severity of the problem increases dramatically for more localized peaks that would correspond to fiber directions in high angular resolution diffusion datasets (Figure 41(b)). In summary, setting tractography in a minimum cost framework provides more control over the optimality criterion $\mathcal{E}$. Alternatively, this can also be achieved in a front propagation approach by using the appropriate Legendre transformation to define the front speed F. Note that when solving that after the Hamilton-Jacobi-Bellman equation has been numerically solved, $\mathcal{E}$-optimal curves can be obtained from any target point back to the seed points.

For these $\mathcal{E}$-optimal curves, the value of another global cost, K, corresponding to a different local cost Φ can be computed by solving the transport equation

$$\nabla K(\mathbf{p}).\hat{\boldsymbol{d}}^*(\mathbf{p}) = \Phi(\mathbf{p},\hat{\boldsymbol{d}}^*(\mathbf{p}))$$

with boundary condition $K = 0$ on the seed region S. In particular the length $L(\boldsymbol{\Gamma}^*(\mathbf{p})) = \int_{\boldsymbol{\Gamma}^*(\mathbf{p})} 1\ \mathrm{d}s$ of these optimal curves corresponds to $\Phi = 1$. The cost per unit length K/L can



(a) Front cost F (blue, solid) set to some quadratic form, equivalent particle cost and speed derived.

(b) Front cost F (blue, solid) models some synthetic non-Gaussian diffusion, equivalent particle cost and speed.

Figure 41: Whereas the particle speed $1/\psi$ (green, dashed) and the front speed F (blue, solid) can be derived from each other, the definition of the cost ψ (red, dot-dashed) from F does not preserve maxima (preferred directions) and therefore results in severe loss of anisotropy information especially for localized fast directions. (In both cases the preferred direction is vertical).

be used to define a validity index and rank curves that are optimal for one criterion using another criterion as in [132, 133].

4.6.2 Application to High Angular Diffusion MRI Tractography

Most front propagation techniques for diffusion tensor tractography use some ad hoc function f of the quadratic form $\hat{\boldsymbol{d}}^t D \hat{\boldsymbol{d}}$. If the Gaussian assumption holds, the diffusion weighted images follow $S(\mathbf{p}, \hat{\boldsymbol{d}}) \simeq S(\mathbf{p}, \mathbf{0}) \exp(-b\ \hat{\boldsymbol{d}}^t D(\mathbf{p}) \hat{\boldsymbol{d}})$. Tensor based techniques can formally be extended to high angular resolution diffusion datasets by setting

$$\psi(\mathbf{p}, \hat{\boldsymbol{d}}) \triangleq f(-\frac{1}{b} \log(\frac{S(\hat{\boldsymbol{d}})}{S(\mathbf{0})}))$$

Notice that Q-Ball datasets [137] and direction-dependent local costs ψ are both defined on the same space $\mathbb{R}^3 \times \mathbb{S}^2$.

However, due to the low signal to noise ratio of these datasets, it is desirable to consider more than one value at a time. The anisotropic cost can be defined by some decreasing function f' of the Funk-Radon transform[9] of the attenuation $S(\mathbf{p},\cdot)/S(\mathbf{p},\mathbf{0})$.

$$\psi'(\mathbf{p},\hat{\boldsymbol{d}}) \triangleq f'\left(\int_{\hat{\boldsymbol{v}}\perp\hat{\boldsymbol{d}}} \frac{S(\hat{\boldsymbol{v}})}{S(\mathbf{0})}\,\mathrm{d}\hat{\boldsymbol{v}}\right)$$

The cost $\psi'(\mathbf{p},\hat{\boldsymbol{d}})$ will therefore be small if and only if there is limited diffusion loss over the corresponding equator, i.e., if diffusion does not occur normal to $\hat{\boldsymbol{d}}$.

Mumford [81] showed that variational techniques, such as the one proposed here, can be set in an elegant and principled Bayesian framework by considering the cost $\psi = \psi_{\text{data}} + \psi_{\text{prior}}$. The extreme simplicity of this construction constitutes another advantage over non-variational front propagation approaches. Here, ψ_{data} would be as described above and ψ_{prior} could be obtained from an atlas of neural tracts.

The problems of generating such an atlas and registering it to the dataset at hand are well beyond the scope of this paper. Note however that masking off (with infinite cost values) the non white matter regions of the brain is a trivial and widely employed use of prior knowledge.

4.6.3 Results

We now show results obtained by applying the methodology described in the above sections to diffusion weighted data sets acquired using a single-shot diffusion-weighted EPI sequence, with 31 different gradient directions with b-values of 500, 1000, and 1500 s/mm2, on a 1.5 Tesla GE Echospeed system. The data was acquired with different b-values to enable comparisons of the results. Traditional eigenvector based tractography is normally carried out in data with b-values in the range of 700-1000 s/mm^2. Higher b-values give data with higher angular contrast but at the expense of more noise.

Cost per unit length, which can be interpreted as a validity index for the putative tracts was determined for all b-values Fig. 42.

All curves are optimal given their starting point. The cost per unit length is a measure of how good they are compared to each other. The best contrast (corresponding to the most

[9]Interestingly, the FRT is also central to the Q-Ball technique [137].

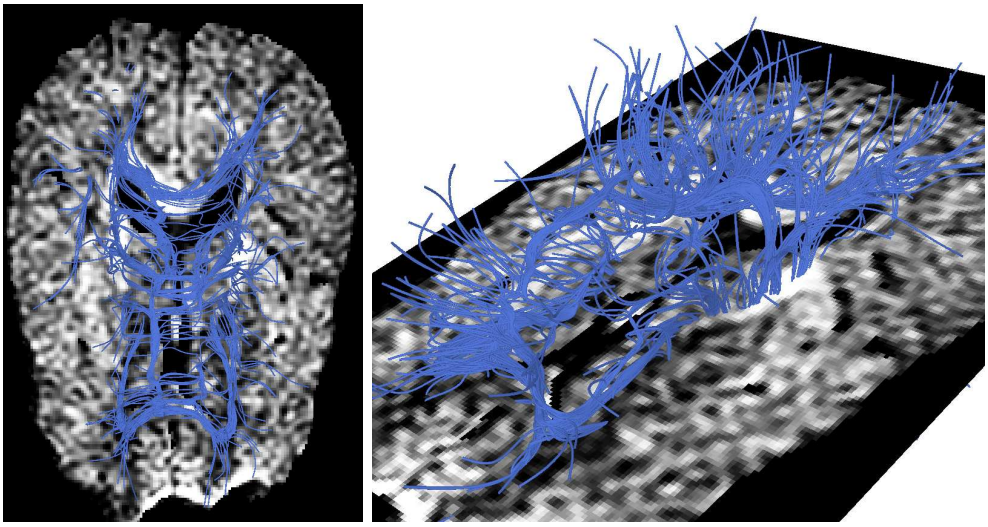
(a) $b = 500s/mm^2$ (b) $b = 1000s/mm^2$ (c) $b = 1500s/mm^2$

Figure 42: Cost per unit length of end points of optimal curves for different b-values (see text). Best results are achieved for the highest b-value.

coherent set of "super-optimal" tracts for a given seed point posterior of the corpus callosum) was obtained at the highest b-value available. This could indicate that the algorithm was able to take advantage of the higher angular contrast in spite of the lower SNR.

Tract results for several user defined seed points are presented on Fig. 43.

Finally the proposed technique was compared to a streamline technique for the needs of which the tensor field was computed (Fig. 44). While validation is a very challenging task due to the unavailability of ground truth, it can be noted that both algorithm give similar results even though their inputs are different. The tracts of the proposed technique tend to be more coherent as any noise in the data might set the streamline off course whereas the proposed technique is more global.



(a) (b)

Figure 43: Fiber tracking from high angular resolution dataset ($b = 1500s/mm^2$).

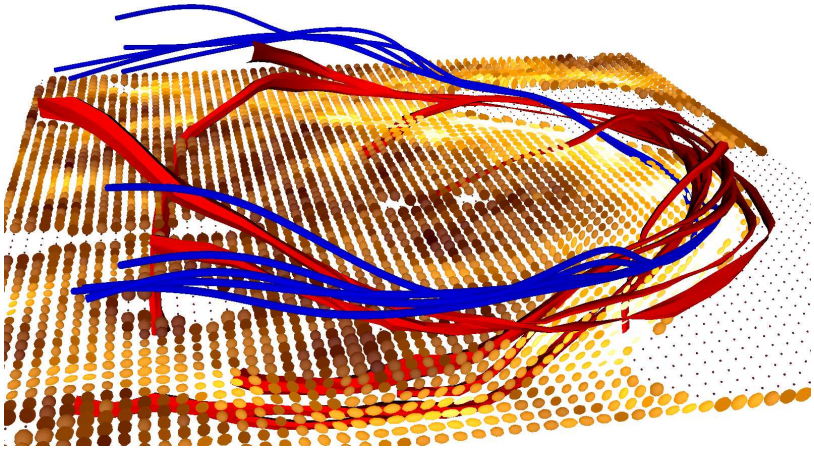


Figure 44: Proposed technique on high angular resolution data (blue) compared with streamline technique on tensor field (red) ($b = 1500s/mm^2$).

CHAPTER 5

SHAPE COMPARISON AND VALIDATION OF IMAGE SEGMENTATION

In this chapter we propose a technique for shape comparison based on the Laplace equation. By shape we mean some closed hypersurface that is the boundary of some region R of the spatial domain $\mathbb{R}^n$. This work is based on [141].

The ability to quantify both local and global differences between shapes is an important step in computer aided medical diagnosis. After imaging and segmentation have been performed, shape comparison techniques can be used to determine the deviation some structure or organ has to some standard template. Such differences in shape can be used for pathology diagnosis. For example, abnormal brain ventricle shape is a symptom of schizophrenia.

Another application of shape comparison is the validation of image segmentation techniques. The diversity of proposed segmentation algorithms raises the issue of performance evaluation. For a given organ, which algorithm results in the most accurate segmentation? How can its robustness be quantified? Using a measure of shape similarity, the performance of a particular algorithm can be quantitatively assessed by comparing its output to a known ground truth, a manual segmentation or the output of another segmentation algorithm.

In this chapter we describe a mathematical methodology to address the problem of shape comparison by determining one-to-one correspondences between two shapes and defining both global and local similarity statistics. To compare more than two shapes, statistics are computed with reference to a mean shape. Validating a complete medical procedure (only part of which would be a segmentation step) based on patient outcome is out of the scope of this work.

Section 5.1 motivates our approach. In Section 5.2 we introduce the Laplace equation and propose different methods to solve it. In Section 5.3 we show how this can be used to define corresponding points between two shapes. While these correspondences are very

well behaved it is possible that the surfaces would be locally so different that there is no very natural way to put them in correspondence. In particular, in that case a relatively large area of one of the surfaces might be mapped onto a relatively small area of the other surface. We show in Section 5.4 how these effects can be quantified. We also describe how it is possible to compare more than two shapes by using a mean shape. In Section 5.5 we illustrate our method for the purpose of validation of image segmentation. We show how to visualize locally the dissimilarity between the surfaces allowing the user to analyze the accuracy of a segmentation in an anatomically meaningful way. We also demonstrate how global metrics can be derived to quantify the accuracy of the segmentation. The complete methodology is illustrated on synthetic and real datasets.

5.1 Motivation and Related Work

This chapter is concerned with shape comparison, i.e., with quantifying similarities between shapes. Similarities may be defined on different levels of resolution, ranging from global shape metrics to local correspondences. A classical example for a global shape metric is the symmetric Hausdorff distance, measuring the maximal minimum Euclidean distance between two sets. While being able to provide global insight with regards to the "worst-case-deviation" between two shapes, the symmetric Hausdorff distance (as an example for a global shape metric) fails at resolving local shape variations. Global metrics may be computed from local ones but not vice versa. See [142] for an overview of shape matching techniques.

This chapter aims at calculating shape differences on a local level for descriptional flexibility. These local measures of shape differences may then be interpreted directly or may be used to compute more global shape similarity metrics.

We represent a shape by a closed hypersurface: e.g., a closed curve in the plane or a closed surface in three-dimensional space. Defining a local shape metric then hinges (in the most local case) on establishing point to point correspondences between surfaces. A method to obtain surface correspondence points should be computationally efficient, should generalize to different dimensions, and should yield a correspondence map that

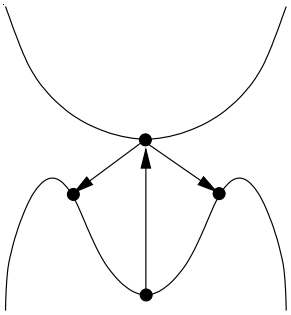
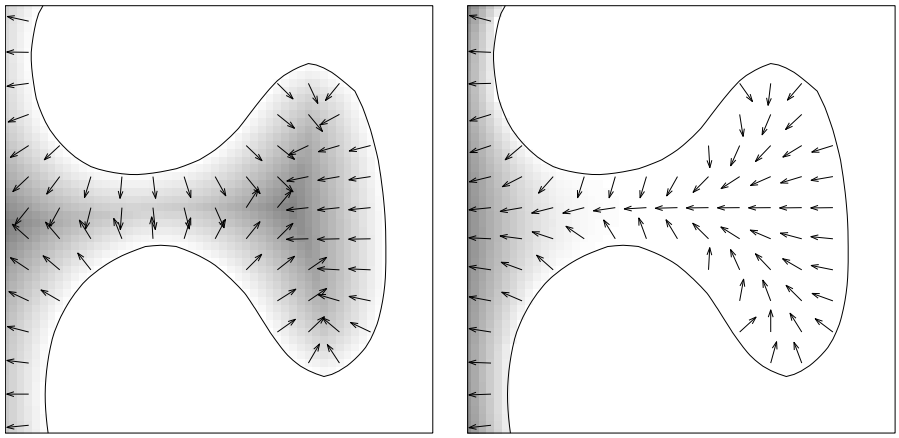


Figure 45: The minimal Euclidean distance between points may lead to correspondences that are not unique and not symmetric.

is symmetric, one-to-one and ideally also continuous (i.e., homeomorphic). Figure (45) shows some exemplary correspondences between two curves based on the smallest Euclidean distance from curve to curve and highlights the desirability of correspondences that are symmetric and one-to-one.

We advocate a Laplace equation based approach to find point correspondences between two surfaces with the aforementioned properties. This is a natural approach for the following reason: given two surfaces Σ_1 and Σ_2, envision a flow field transporting particles starting on surface Σ_1 to Σ_2. The conditions of (i) symmetry, (ii) one-to-one correspondence and (iii) continuity are in this setting equivalent to requiring that (i) the flow can be reversed, (ii) a particle does not get caught up in a vortex or at a stagnant point (unless it started there), and (iii) particles that start close to each other on Σ_1 arrive "close to each other" on Σ_2. Condition (ii) yields a potential flow, i.e., the flow field $\boldsymbol{v}$ can be expressed as the gradient of a scalar potential u, $\boldsymbol{v} = \nabla u$. If the fluid is assumed to be incompressible, $\boldsymbol{v}$ needs to be divergence free as well, i.e., $\nabla \cdot \boldsymbol{v} = 0$, but this implies $\nabla \cdot \nabla u = \Delta u = 0$, which is Laplace's equation. Solving Laplace's equation (with suitable boundary conditions) between Σ_1 and Σ_2 will induce point-to-point correspondences between Σ_1 and Σ_2 through the streamlines of $\boldsymbol{v}$. Figure (46) shows the vector fields induced by a distance function of a cavity and by the solution of Laplace's equation. All particles starting on the boundary of the cavity and moving along the vector field induced by the solution of Laplace's equation will eventually



(a) Vector field induced by the distance function.

(b) Vector field induced by the solution of Laplace's equation.

Figure 46: Vector fields induced by the distance function and by the solution of Laplace's equation. Dark values indicate larger solution values.

leave the cavity. This is not true for all particles for the vector field induced by the distance function.

Laplace equation approaches have previously been used in [143, 144] for colon surface flattening and centerline extraction and in [145, 146] for thickness measurements. This chapter introduces a framework to measure local shape variability. Given multiple surfaces Σ and a mean (or comparison) surface Σ_m point correspondences are used to measure local distance differences. The local distances between surfaces can then be used for visualization purposes (e.g., to color-code the mean shape) and to define local and global statistics. Statistics may be computed at a single point of a mean surface over all surfaces Σ to assess local shape variation. Alternatively one can compute statistical measures of distance variations over a single surface. The latter is particularly useful in the context of validation of surface segmentations: how good is a segmentation compared to a given ground truth.

5.2 Laplace Equation

Assume that we have two regions R_1 and R_2 whose boundaries Σ_1 and Σ_2 are surfaces implicitly represented as zero level sets of functions ϕ_1 and ϕ_2, i.e.,

$$\Sigma_i = \partial R_i = \{\mathbf{x} \ / \ \phi_i(\mathbf{x}) = 0\}, \ i \in \{1, 2\}. \tag{83}$$

Such a representation is natural for partial differential equations based segmentation algorithms. Regions defined by binary maps and triangulated surfaces can also be represented using level sets.

The Laplace equation

$$\begin{cases} \nabla^2 u(\mathbf{x}) = 0 & \mathbf{x} \in R_1 \ominus R_2 \\ u(\mathbf{x}) = 1 & \mathbf{x} \in \Sigma_1 \\ u(\mathbf{x}) = -1 & \mathbf{x} \in \Sigma_2 \end{cases} \tag{84}$$

underlies Fourier's law of heat conduction. In this context u can be interpreted as a temperature. Equation (84) describes the steady state temperature profile u in between the surfaces (here, $R_1 \ominus R_2 \triangleq (R_1 \setminus R_2) \cup (R_2 \setminus R_1)$ is the symmetric difference of the two regions) resulting from prescribed boundary conditions on Σ_1 and Σ_2. The field u is harmonic. It is smooth and its derivatives of all order exist [147]. Moreover the gradient does not vanish and therefore the vector field

$$\mathbf{v} = \frac{\nabla u}{\| \nabla u \|} \tag{85}$$

is also smooth. By following $\mathbf{v}$ (resp. $-\mathbf{v}$) starting at a point on Σ_1 (resp. Σ_2) we are assured to reach a unique point on Σ_2 (resp. Σ_1). This is a very desirable property for establishing one-to-one point correspondences between Σ_1 and Σ_2. A variety of schemes exist to solve (84) based on gradient descent (the Heat equation) or superposition of fundamental analytic solutions (Boundary Element Method).

Instead of solving the steady state problem (84) directly, we can use the heat equation

$$\frac{\partial u}{\partial t} = \nabla^2 u, \tag{86}$$

which with appropriate boundary conditions will converge to the solution of (84). In the simplest possible case, we discretize the time derivative using an Euler forward or backward approximation and the spatial derivatives using central differences.

The level set representation yields subpixel accuracy. Using subpixel boundary information is not straightforward with finite difference based schemes. The boundary element method (BEM) and the finite element method naturally handle intricately shaped boundaries. The boundary is approximated by N elements and the solution is constructed based on the superposition of the fundamental solutions of the Laplace equation for each boundary element [148]. For simple boundary parameterizations, analytical solutions exist. Furthermore, the boundary element method allows for analytical gradient computations which is very useful when calculating (85) [149, 150].

The BEM is a powerful and accurate method for solving the Laplace equation on the usually complicated domains resulting from the segmentation of medical datasets. A two-dimensional synthetic example is shown on Fig. 47 along with a zoom-in on the solution of the Laplace equation with the two proposed methods. The BEM is not restricted to the underlying grid and can take full advantage of the subpixel accuracy of the boundaries. It therefore outperforms the simple finite difference approach.

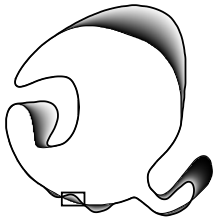
5.3 Defining a Distance Between Shapes

5.3.1 Local Dissimilarity

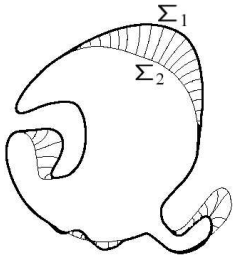
As explained in the previous section from any point s on surface Σ_1, (resp. Σ_2), a trajectory can be determined to Σ_2, (resp. Σ_1), by following the characteristics (i.e., the gradient) of the Laplace equation. A particle can be moved from one surface to the other according to

$$\begin{cases} s_t(t) = -\nabla u \circ s(t) & s \in \Sigma_1 \\ s_t(t) = +\nabla u \circ s(t) & s \in \Sigma_2 \\ s(0) = s \end{cases} \tag{87}$$

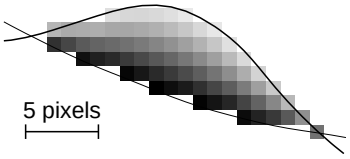
As explained previously such correspondence are well defined, one to one and symmetric.



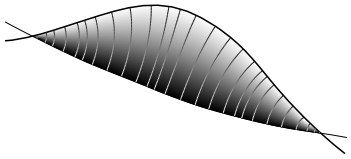
(a) Solution of the Laplace equation on $R_2 \ominus R_1$.



(b) Streamlines defined from the gradient of the solution of the Laplace equation.



(c) Finite difference method (Heat equation).



(d) Boundary element solution with streamlines of the gradient field.

Figure 47: Two-dimensional synthetic example and zoom in on rectangular region.

To each point s on one of the two surfaces we can therefore associate the length l to the corresponding point s' on the other surface. We propose to define the local error E as $-l$ where Σ_1 is inside Σ_2 and $+l$ otherwise. This scalar field is defined on both surfaces and describes locally the dissimilarity between the surfaces. See Figure 5.3.1 for an illustration on a synthetic example.

5.3.2 Visualization

The surfaces Σ_1 and Σ_2 can be colored with the local dissimilarity scalar field E (Fig. 51(a), 51(b), 51(c)). We use a perceptually linear color scale as proposed by Lefkowitz et al. [151].

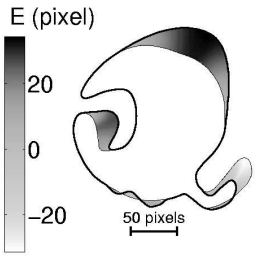


Figure 48: Local dissimilarity for a two-dimensional synthetic example.

5.3.3 Metrics

We define a probability space $(\Sigma_1 \cup \Sigma_2, P)$ on the union of the surfaces by $P(S) = \frac{\mathcal{A}(S)}{\mathcal{A}(\Sigma_1)+\mathcal{A}(\Sigma_2)}$ where $\mathcal{A}(S)$ is the area of an element of surface $S \subset \Sigma_1 \cup \Sigma_2$. The local dissimilarity field E can then be interpreted as a random variable.

The statistics of E can be used to define several metrics that will quantify the global dissimilarity between the surfaces.

- $\sup(|E|)$ is the largest distance between corresponding points of Σ_1 and Σ_2. This information is equivalent to the Hausdorff distance. Moreover, $\inf(E)$ and $\sup(E)$ are the largest under and over-segmentation errors. The quantiles of E convey the same information as the partial Hausdorff distance.

- $\text{mean}(|E|)$ is the average distance between the surfaces.

- $P(E \leq -3)$ is the probability that Σ_1 would be inside Σ_2 by 3 units.

- the probability density function p_E of E gives the full error distribution.

See Table 4 for the values of these statistics on the synthetic and brain datasets. Additional metrics can easily be defined. For example, the expectancy of $1/(1 + pE^2)$ would be equivalent to Pratt's Figure of Merit [152].

Table 4: Global statistics of the local dissimilarity E on the synthetic example.

	$\sup(E)$	99% percentile	$\inf(E)$	1% percentile	mean($\lvert E\rvert$)	$P(E \leq -3)$
Synthetic (pixel)	31.42	31.24	-28.21	-20.40	5.32	9.50%

5.3.4 Comparing Several Surfaces

In order to study the local variation in a set of n surfaces defined by the level sets function $\phi_1 ... \phi_n$, the level set function of the mean shape is defined to be

$$\phi_m \triangleq \sum_{i=1}^{n} \phi_i \tag{88}$$

and all n shapes are compared to the mean surface.

Figures 49 and 50 show the local dissimilarity to a mean shape for $n = 11$ prostates. Surfaces were segmented from pre-operative 1.5T MR images acquired using an endorectal coil integrated with a pelvic multi-coil (Signa LX, GE Medical systems, Milwaukee, WI). Patients were placed in a supine position in the closed-bore magnet. The slice thickness is of 3 mm with a slice gap of 0 mm, matrix of 256x192.

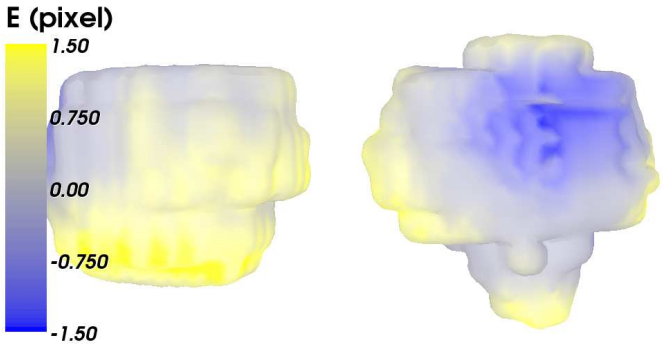
5.4 Quantifying Area Distortion

Consider two surfaces $\Sigma_0, \Sigma : \mathbb{S}^{n-1} \to \mathbb{R}^n$. Let u be a scalar field such that any point x on Σ can be put in correspondence with a point $x_0 = F_0(x)$ on Σ_0 by following the gradient lines of u.

If a probability space is constructed from Σ_0 by considering that all points of Σ_0 are equally likely, then the probability of an event $X \subset \Sigma_0$ is the ratio of the surfaces areas $\mathcal{A}(X)/\mathcal{A}(\Sigma_0)$.

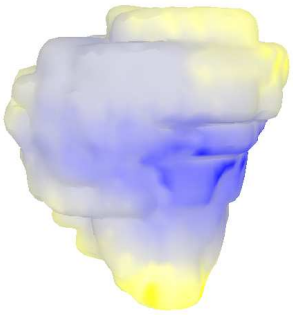
Some scalar field m_0 defined on Σ_0 can then be interpreted as a random variable. The integral

$$I_0 = \int_{\Sigma_0} m_0(\Sigma_0) \, \mathrm{d}\Sigma_0 \tag{89}$$

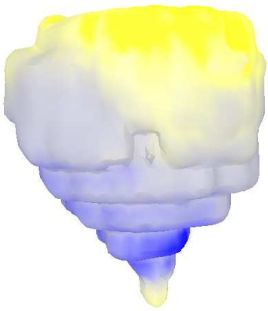


(a)

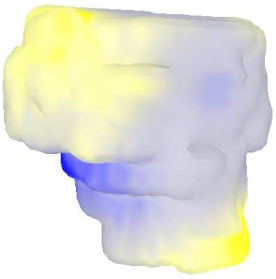
(b)



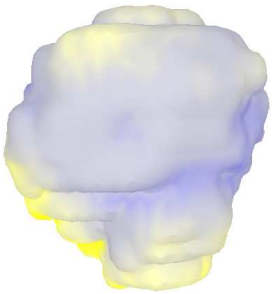
(c)



(d)

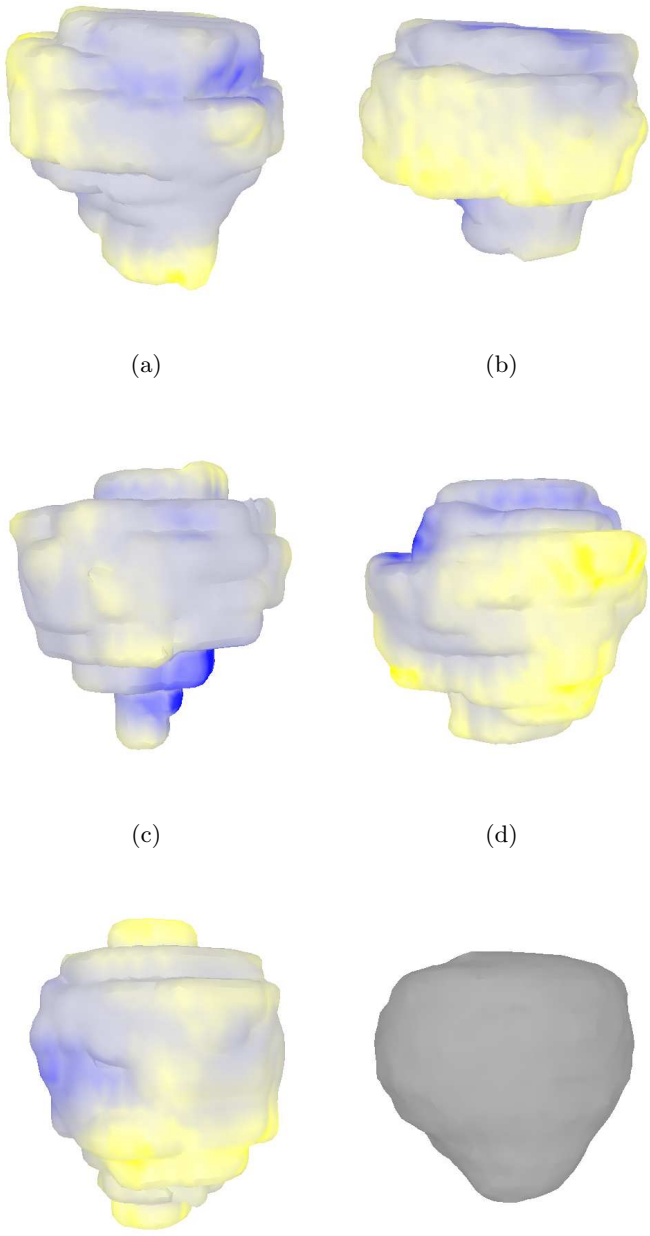


(e)



(f)

Figure 49: Comparison of $n = 11$ prostates to a mean surface. Prostates 1 to 6. These are color images. Continued on Figure 50.



(a) (b)

(c) (d)

(e) (f) Mean shape

Figure 50: Comparison of $n = 11$ prostates to a mean surface. Prostates 7 to 11 and corresponding mean shape. These are color images. Continued from Figure 49.

is then equivalent to the expectancy of m_0 multiplied by the total area of the surface Σ_0.

The probabilistic structure can be transported on Σ using the mapping F_0 from Σ to Σ_0. Specifically we consider, for $Y \subset \Sigma$, $P_\Sigma(Y) = P_{\Sigma_0}(F_0(Y)) = \mathcal{A}(F_0(Y))/\mathcal{A}(\Sigma_0)$.

The expectancy of m_0 can then be determined by integrating on Σ rather than Σ_0. In fact, this is equivalent to using the change of variable

$$I_0 = \int_\Sigma m_0(F_0(\Sigma))\ \mathrm{d}F_0\ \mathrm{d}\Sigma \tag{90}$$

It is not difficult to determine $m_0 \circ F_0$, i.e., transport m_0 onto Σ. Since F_0 puts points in correspondence along the gradient lines of u, this is done by solving the transport equation between the two surfaces:

$$\begin{cases} \nabla\tilde{m}.\nabla u = 0 \\ \tilde{m} = m_0 \text{ on } \Sigma_0 \end{cases} \tag{91}$$

then $m_0 \circ F_0 = \tilde{m}$ on Σ.

The computation of $\mathrm{d}F_0$ (i.e., the quantity by which areas are stretched by F_0) is a little more involved and will be the focus of the rest of this section.

Let $S(p,\theta) : \mathbb{S}^{n-1} \times [0,1] \to \mathbb{R}^n$ be the θ-levelset of u. By definition:

$$u(S(p,\theta)) = \theta$$
$$\text{therefore, } \frac{\partial}{\partial\theta}u(S(p,\theta)) = \nabla u.\frac{\partial}{\partial\theta}S = 1$$

Moreover since $S(p,\theta)$ is a levelset of u, its normal is $\nabla u/|\nabla u|$. The levelsets are therefore obtained by the evolution equation

$$\frac{\partial}{\partial\theta}S(\cdot,\theta) = \begin{cases} \mathbf{0} & \theta < u(S(\cdot,\theta)) \\ \mathbf{0} & \theta > u(S(\cdot,\theta)) \\ \nabla u(S(\cdot,\theta)/|\nabla u(S(\cdot,\theta)|^2 & otherwise \end{cases}$$

$$S(\cdot, \min_{\Sigma_0} u) = \Sigma_0(\cdot)$$

By construction, $S(\cdot,\theta) = \Sigma_0(\cdot)$ for all $\theta < \min_{\Sigma_0} u$ and $S(\cdot,\theta) = \Sigma(\cdot)$ for all $\theta > \max u$. Therefore we denote $S(\cdot,-\infty) = \Sigma_0(\cdot)$ and $S(\cdot,+\infty) = \Sigma_0(\cdot)$. Consider now the transform G_θ that maps $\Sigma_0(\cdot) = S(\cdot. - \infty)$ onto $S(\cdot.\theta)$ along the streamlines of u. For $\theta > \max u$, $G_\theta = F_0$ previously defined that maps Σ onto Σ_0.

$$\begin{aligned} G_\theta(\Sigma(p)) &= S(p,\theta) \\ (\det\ \mathrm{d}G_\theta)(\det\ \mathrm{d}_p\Sigma) &= \det\ \mathrm{d}_p S(p,\theta) \\ \frac{\partial \det\ \mathrm{d}G_\theta}{\partial\theta}(\det\ \mathrm{d}_p\Sigma) &= \frac{\partial}{\partial\theta}\det\ \mathrm{d}_p S(p,\theta) \\ &= (\det\ \mathrm{d}_p S(p,\theta))\ \mathrm{trace}\{\ (\ \mathrm{d}_p S)^{-1}\frac{\partial\ \mathrm{d}_p S}{\partial\theta}\ \} \end{aligned}$$

using $(\det A)' = (\det A)\ \mathrm{trace}\{\ A^{-1}A'\ \}$

$$\begin{aligned} &= (\det\ \mathrm{d}_p S(p,\theta))\ \mathrm{trace}\{\ (\ \mathrm{d}_p S)^{-1}\ \mathrm{d}_p\frac{\partial S}{\partial\theta}\ \} \\ &= (\det\ \mathrm{d}_p S(p,\theta))\ \mathrm{trace}\{\ (\ \mathrm{d}_p S)^{-1}\ \mathrm{d}_p\frac{\nabla u(S)}{|\nabla u(S)|^2}\ \} \end{aligned}$$

now, let $w = \nabla u/|\nabla u|^2$, then $\mathrm{d}_p w \circ S = \mathrm{d}_{TS} w \ \mathrm{d}_p S$ where $\mathrm{d}_{TS} w$ is the differential on the plane tangent to the surface S at that point. Here we will just consider that it is the differential $\mathrm{d}_{\mathbb{R}^n} w$ in $\mathbb{R}^n$ except that the last line and column that would correspond to the direction normal to the surface are omitted. Since the surfaces S are levelsets of u, this normal direction is the direction of ∇u. Simple algebral yields $\mathrm{d}_{\mathbb{R}^n} w = P_{\mathbf{v}^\perp} H_{\mathbb{R}^n}/|\nabla u|^2 w$ where $P_{\mathbf{v}^\perp}$ is the projection on the subspace orthogonal to $\mathbf{v} = \nabla u/|\nabla u|$ and $H_{\mathbb{R}^n}$ is the Hessian in $\mathbb{R}^n$. Finally, as far as $\mathrm{d}_{TS} w$ is concerned, that projection is superfluous and we get $\mathrm{d}_{TS} w = H_{\mathbf{v}^\perp}/|\nabla u|^2$.

$$
\begin{aligned}
&= (\det\ \mathrm{d}_p S(p,\theta))\ \mathrm{trace}\{\ (\ \mathrm{d}_p S)^{-1} \frac{\mathrm{H}_{\mathbf{v}^\perp} u}{|\nabla u(S)|^2}\ \mathrm{d}_p S\ \} \\
&= (\det\ \mathrm{d}_p S(p,\theta))\ \mathrm{trace}\{\ \frac{\mathrm{H}_{\mathbf{v}^\perp} u}{|\nabla u(S)|^2}\}
\end{aligned}
$$

using $\mathrm{trace}\{\ A^{-1}BA\ \} = \mathrm{trace}\{\ B\ \}$

$$
= (\det\ \mathrm{d}_p S(p,\theta)) \frac{\Delta_{\mathbf{v}^\perp} u}{|\nabla u|^2}
$$

Define the scalar field b between the surfaces by $b(S(p,\theta)) = \log(\det\ \mathrm{d}G_\theta)$. Differentiating with respect to θ we get

$$
\begin{aligned}
\nabla b . \frac{\partial}{\partial \theta} S(p,\theta) &= \frac{1}{\det\ \mathrm{d}G_\theta} \frac{\partial \det\ \mathrm{d}G_\theta}{\partial \theta} \\
\nabla b . \frac{\nabla u}{|\nabla u|^2} &= \frac{\Delta_{\mathbf{v}^\perp} u}{|\nabla u|^2} \\
\nabla b . \mathbf{v} &= \frac{\Delta_{\mathbf{v}^\perp} u}{|\nabla u|}
\end{aligned}
$$

the right hand side is proportional to the mean curvature h_u defined by $(n-1)h_u = \nabla . (\frac{\nabla u}{|\nabla u|}) = \frac{\Delta u}{|\nabla u|} + \nabla u . [-\frac{1}{2} \frac{2 H u \nabla u}{|\nabla u|^3}]$

$$
\nabla b . \mathbf{v} = (n-1) h_u
$$

Since by definition, $G_{-\infty} = \mathrm{Id}_{\Sigma_0}$ then $\det\ \mathrm{d}G_{-\infty} = 1$ and therefore $b = 0$ on $\Sigma_0(\cdot)$. This gives the boundary condition for the transport equation in b.

$$\begin{cases} \nabla b.\mathbf{v} = (n-1)h_u \\ b = 0 \text{ on } \Sigma_0 \end{cases} \tag{92}$$

And finally,

$$I_0 = \int_{\Sigma} m(G_0(\Sigma))\ \exp(b(G_0(\Sigma)))\ \mathrm{d}\Sigma \tag{93}$$

Note: $b(G_0(\Sigma))$ is obtained by integrating from Σ to Σ_0 and then transporting back the result on Σ along the same path. This is equivalent to integrating from Σ_0 to Σ.

If u is such that $\Delta u = 0$ in between the surfaces then the right hand side of the transport equation in b becomes $-u_{\mathbf{v},\mathbf{v}}/u_{\mathbf{v}}$. By integrating we will get $b = -\log u_{\mathbf{v}}|_{\Sigma} + \log u_{\mathbf{v}}|_{\Sigma_0}$ on Σ_0 and by taking the exponential, the area stretching factor is $\frac{u_{\mathbf{v}}|_{\Sigma}}{u_{\mathbf{v}}|_{\Sigma_0}}$ where $\mathbf{v}$ is normal to the surface.

This can also be found by using the fact that u being a laplacian field, the in and out fluxes are equal. Therefore $(\ \mathrm{d}\Sigma_0)N_0.\nabla u|_{\Sigma_0} = (\ \mathrm{d}\Sigma)N.\nabla u|_{\Sigma}$.

In the case where u is a Laplacian field it is therefore not necessary to solve the transport equation in b to determine the area stretching factor.

5.5 Validation Framework

The previously defined formalism can be applied directly to the quantitative validation of segmentation if the ground truth is known. In that case, one of the surfaces is set to be the segmentation result and the other the ground truth. The dissimilarity between the segmentation and the ground truth can then be compared both quantitatively and qualitatively. If the ground truth is not available, the results of segmentations obtained through different techniques can be compared (e.g. automatic and human expert).

In [153], Zhang reviews classical validation metrics. These discrepancy metrics are based either on the sole frequency of mis-segmented voxels (e.g, the Dice similarity coefficient) or on their position via some sort of distance information (e.g., the Hausdorff distance,

Yasnoff's Normalized Discrepancy etc.). Typically no distinction is made between under and over-segmentation and the Euclidean distance is employed and, as in the case of the Hausdorff distance, ill-defined underlying correspondences between the ground truth and the segmentation are implicitly used. Moreover many of these metrics are dimensionless scalars whose interpretation can be problematic.

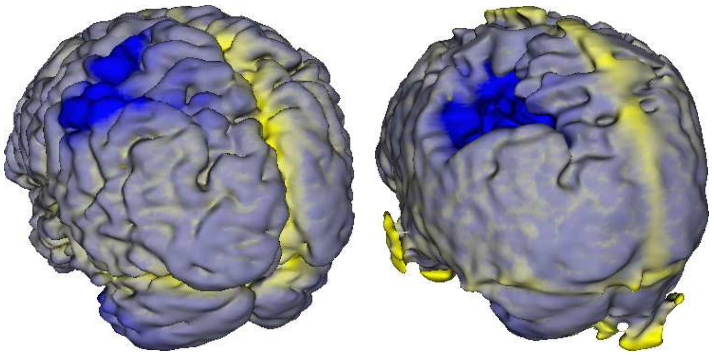
In contrast the error distance E discriminates between under and over-segmentation (negative values of the local dissimilarity scalar field E will correspond to under-segmentation and positive values will correspond to over-segmentation). Moreover it is based on well-defined, physically grounded correspondences and has the dimension of a distance.

Here, as an illustration we use, as the ground truth, the expert manual segmentations of a full brain from the Brain Tumor Database [106] and the output of a previously proposed algorithm [85] as the segmentation, see Figures 5.5 and 5.5 and Table 5.

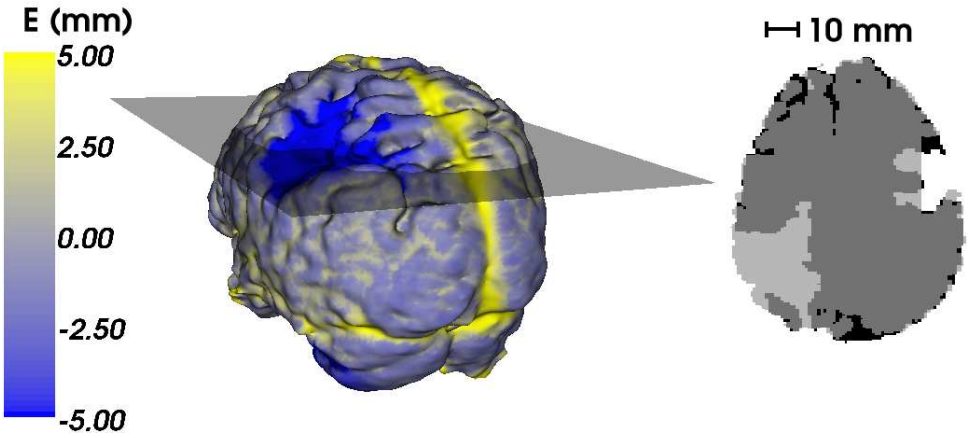
Table 5: Global statistics of the local dissimilarity E on the brain and tumor examples.

	$\sup(E)$	99% percentile	$\inf(E)$	1% percentile	mean($\lvert E\rvert$)	$P(E \leq -3)$
Brain (mm)	14.08	7.09	-86.10	-11.87	2.31	4.03%
Tumor (mm)	6.83	5.15	-4.75	-3.18	1.58	1.17%

In the proposed brain segmentation a large portion of the left hemisphere was left out (under-segmented) and the superior sagittal and transverse sinuses (venous channels in the back of the brain) were erroneously classified as brain (over-segmented). This visualization scheme allows researchers and clinicians to interpret the performance of a segmentation in an anatomically meaningful way.



(a) Ground Truth (b) Segmentation



(c) The underlying ground truth and segmentation on the gray slice are shown on the right. Over-segmented regions are in dark gray, under-segmented regions are in light gray.

Figure 51: Real dataset of a full brain The surfaces are colored with the local dissimilarity E. These are color images.

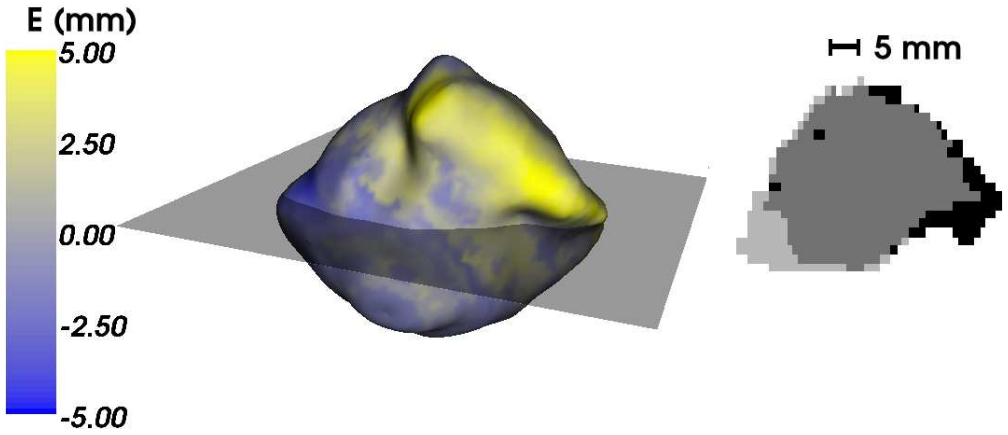


Figure 52: Real dataset of a brain tumor. The underlying ground truth and segmentation on the gray slice are shown on the right. Over-segmented regions are in dark gray, under-segmented regions are in light gray. The surfaces are colored with the local dissimilarity E. These are color images.

CHAPTER 6

CONCLUSION

Artificial vision is the problem of mimicking the human visual system. It is not possible currently to base algorithms closely on the working principles of the inner mechanisms of biological vision because those mechanisms are mostly unknown. The goal of the artificial vision researcher is therefore to devise algorithms that describe how artificial systems – in practice digital computers – can perform some processing from a digital image. Since they are not closely based on neurological processes, algorithms can be built from any mathematical framework. The value of a technique is to be assessed by the elegance of the mathematical formulation, the ease of implementation on a digital computer and ultimately, the results obtained for a particular[1] image analysis task.

This work is set in the framework of partial differential equations and differential geometry. Modern computers are limited to handling discrete information and it is not always straightforward to use them to simulate a continuous process such as the continuous deformation of a surface. However, ultimately, even if they are themselves discrete objects, digital images are the projection of a reality that is being, for now, better described by continuous mathematics and in particular partial differential equations. For this reason, very extensive mathematical theories have been devised to analyze PDEs. The problem of simulating continuous phenomena on digital computers is under vigorous study for applications in physics and engineering. Many of these techniques can be applied to the field of computer vision. The level set framework, which was initially created for the simulation of front propagation in physics, is a good example. In summary, the existence of a particular technology, such as digital computers, should not overly influence or restrict the field

[1]Some algorithms are more specific than others, but ultimately there is reason to believe that all segmentation problems are instances of a unique problem that could be solved by a unique algorithm. It is therefore not possible to define the efficiency of an algorithm independently of a given segmentation problem to which it is to be applied.

of artificial vision, which has much to benefit from all existing continuous mathematical techniques. This has been illustrated at great length in Chapter 2.

The problem of image segmentation is that of determining the contour of objects from their projection onto a digital image. While current understanding of the human high-level visual system is extremely fragmentary, it appears that at some point in the visual processing, objects are no longer perceived as the sum of independent dots but as a coherent whole. This leads the way for higher-level processing (e.g., reasoning based on the spatial relation between objects, detection of symmetries or occlusions etc.) that would not be possible otherwise. It has been observed that both bottom-up (the whole emerges from the parts) and top-down (the perception of the whole object is forced by some higher-level cognitive layer even though it is not necessarily consistent with lower-level perception) occur concurrently. The understanding and mimicking of the interplay between these two phenomena is one of the major difficulties of artificial vision. In a variational framework, a global cost functional is defined for structures to be detected. An optimal structure that minimizes this cost is then determined using local information. This framework offers a way to reproduce the interplay of top-down and bottom-up processes. Some global information is imposed (e.g., the structure has a given topology, it is an open curve, or a closed surface) and a global cost functional defines good and bad segmentations while, concurrently, local image information is used to deform the structure in order to minimize its cost.

The contribution of this thesis is threefold:

1. An image segmentation technique is proposed in which the global cost of a surface is defined by the spikiness of the distribution of intensities inside the region it encompasses (Chapter 3). Thus, if the surface corresponds to the boundary of a region of uniform intensities, the cost will be small. While these kinds of techniques had been proposed recently by other authors, the proposed functional is based on a non-parametric description of the distribution of intensities inside the region. It does not make use of the outside of the region. In particular bimodality of the image is not required. The functional leads to a unidirectional optimizing flow, which can be implemented in a very simple and fast way. The validation of image segmentation

techniques is a very important problem that is too often overlooked. Typically, when proposing a new technique, authors do not disclose source code, forcing others to re-implement the technique in order to compare it to their own. More often than not, the implementation of an algorithm from a published paper is a very arduous task requiring weeks of work for a skilled programmer. Because implementation details, such as the exact numerics, are very often omitted, there is no certainty that once re-implemented the algorithm will be perfectly equivalent to the one presented in the original publication. Moreover, publications typically demonstrate the efficiency of the algorithm with a figure showing the result of the algorithm on some test dataset.The reader can then assess the performance of the algorithm using his own visual system. While this procedure allows for some qualitative validation, it does not allow for quantitative validation. In fact, the dataset is usually not publicly available, which means that the exact simulation results demonstrated in the publication are not repeatable. Moreover, the "ground truth" corresponding to the proposed segmentation is often inaccessible. This is a general problem of image segmentation, which, being the mimicking of the processing of the human visual system, is inherently subjective. In contrast, the implementation of the proposed algorithm has been made publicly available through the 3D Slicer project. The algorithm has been quantitatively validated on several publicly available medical imaging datasets.

2. Directionality was added to the conformal active contour technique (Chapter 4). The cost of a curve or a surface is defined as the length of the curve (or area of the surface) weighted by some position and direction-dependent local cost based on image information. This allows for the asymmetric processing of information based on direction. For example, it is possible to define the local cost to be low not only if the curve is located on an edge of the image but also if the curve is actually oriented along that edge. In the case of surfaces, it is possible to treat image information inside or outside the surface asymmetrically. For example, it is possible to define a local cost that will be low if there is a transition from dark to bright from the inside of the surface to the outside of the surface. This will allow for the segmentation of dark regions whose

boundaries will correspond to a small cost. It was shown that the local cost can be defined from a direction-dependent pattern detector, which can be obtained after a learning step. A particular example of such a direction-dependent pattern detector is a model of edge detecting visual neurons that have been identified experimentally in the mammalian visual cortex.

3. A technique was proposed for the comparison of surfaces based on the Laplace equation (Chapter 5). In artificial vision, surfaces are usually obtained as the result of a segmentation step. Such a surface can be used for visualization by a human operator. Because human vision is inherently two-dimensional, surface rendering is much more natural than volume rendering, which attempts to represent a full three-dimensional image through a clever use of transparency. The technique can also be used for the extraction of quantitative shape information (e.g., position of center of gravity, area, volume). A more challenging and very important problem is that of classifying the obtained shape into for example "normal" and "abnormal" or more generally of determining its variation from some shape template. A technique based on the Laplace equation was proposed. One-to-one correspondences are determined between the two shapes and, based on that, local and global similarity metrics are defined. An important application is the comparison of the result of a segmentation algorithm with the ground truth (which could be defined as some manual segmentation by a human expert).

Further research could include a detailed study of the relation between the proposed direction-based segmentation technique and biological vision. The framework could be further extended and local costs defined by a pattern detector that would depend not only on direction but also on some parameter that would vary smoothly along the curve and would be recovered together with the curve (or surface) by minimization. In the surface comparison, the problem of determining a mean shape that would be optimal with respect to the metrics proposed could be investigated. The problem of extending the framework in order to determine correspondences between un-registered shapes could also be considered.

Finally, all the proposed techniques could be applied extensively to specific experimental problems in other research fields such as medicine or neuroscience.

The techniques proposed in this thesis are very general and could be used to extract information from many different types of imagery. They have been applied mostly to medical imaging datasets and in particular images of the brain. The development of new imaging modalities (such as structural, functional or diffusion magnetic resonance imagery) along with the necessary corresponding image analysis techniques is not only useful for medical applications but also for perfecting the fundamental understanding of the functioning of the human brain. Reciprocally, understanding the human brain could ultimately lead to the designing of better artificial vision and artificial intelligence techniques.

REFERENCES

[1] R. Gonzalez and R. Woods, *Digital Image Processing*. Prentice Hall, 2001.

[2] B. Horn, *Robot Vision*. MIT Press, 1986.

[3] R. Haralick and L. Shapiro, *Computer and Robot Vision*. Addison-Wesley, 1992.

[4] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing: Analysis and Machine Vision*. Brooks Cole, 2 ed., 1998.

[5] F. Chabat, D. Hansell, and G.-Z. Yang, "Computerized decision support in medical imaging," *IEEE Engineering in Medicine and Biology Magazine*, vol. 19, pp. 89 – 96, Sept.-Oct 2000.

[6] W. Köhler, "Gestalt psychology today," *American Psychologist*, vol. 14, pp. 727–734, 1959.

[7] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 12, no. 290, pp. 91–97, 1981.

[8] R. von de Heydt and E. Peterhans, "Illusory contours and cortical neuron responses," *Science*, vol. 224, no. 4654, pp. 1260–2, 1984.

[9] A. Turing, "Computing machinery and intelligence," *Mind*, vol. LIX, pp. 433–460, October 1950.

[10] S. Angenent, E. Pichon, and A. Tannenbaum, "Mathematical methods in medical image processing," *Bulletin of the American Mathematical Society*, 2005. Accepted.

[11] A. Hero and H. Krim, "Mathematical methods in imaging," *IEEE Signal Processing Magazine*, vol. 19, no. 5, pp. 13–14, 2002.

[12] T. Chan, J. Shen, and L. Vese, "Variational PDE models in image processing," *Notices of AMS*, vol. 50, pp. 14–26, Jan 2003.

[13] F. Guichard, L. Moisan, and J. Morel, "A review of PDE models in image processing and image analysis," *Journal de Physique IV*, no. 12, pp. 137–154, 2002.

[14] V. Caselles, J. Morel, G. Sapiro, and A. Tannenbaum, "Introduction to the special issue on partial differential equations and geometry-driven diffusion in image processing and analysis," *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 269–273, 1998.

[15] A. P. Witkin, "Scale-space filtering," *Int. Joint. Conf. Artificial Intelligence*, pp. 1019–1021, 1983.

[16] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.

[17] R. Hummel, "Representations based on zero-crossings in scale-space," in *IEEE Computer Vision and Pattern Recognition*, pp. 204–209, 1986.

[18] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel, "Axioms and fundamental equations of image processing," *Arch. Rat. Mech. Anal.*, vol. 123, no. 3, pp. 199–257, 1993.

[19] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 629–639, 1990.

[20] L. Alvarez, P. L. Lions, and J. M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Numer. Anal.*, vol. 29, pp. 845–866, 1992.

[21] S. J. Osher and J. A. Sethian, "Front propagation with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[22] S. Osher and R. P. Fedkiw, "Level set methods: An overview and some recent results," *Journal of Computational Physics*, vol. 169, pp. 463–502, 2001.

[23] J. Sethian, *LevelSet Methods and Fast Marching Methods.* Cambridge University Press, 1999.

[24] C. L. Epstein and M. Gage, "The curve shortening flow," in *Wave Motion: Theory, Modeling and Computation* (A. Chorin and A. Majda, eds.), Springer-Verlag, 1987.

[25] M. Gage and R. S. Hamilton, "The heat equation shrinking convex plane curves," *J. Differential Geometry*, vol. 23, pp. 69–96, 1986.

[26] M. Grayson, "The heat equation shrinks embedded plane curves to round points," *J. Differential Geometry*, vol. 26, pp. 285–314, 1987.

[27] G. Huisken, "Flow by mean curvature of convex surfaces into spheres," *J. Differential Geometry*, vol. 20, pp. 237–266, 1984.

[28] B. White, "Some recent developments in differential geometry," *Mathematical Intelligencer*, vol. 11, pp. 41–47, 1989.

[29] K.-S. Chou and X.-P. Zhu, *The Curve Shortening Problem.* Chapman & Hall/CRC, 2001.

[30] L. C. Evans and J. Spruck, "Motion of level sets by mean curvature," *J. Differential Geometry*, vol. 33, no. 3, pp. 635–681, 1991.

[31] Y. G. Chen, Y. Giga, and S. Goto, "Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations," *J. Differential Geometry*, vol. 33, pp. 749–786, 1991.

[32] G. Sapiro and A. Tannenbaum, "On affine plane curve evolution," *Journal of Functional Analysis*, vol. 119, no. 1, pp. 79–120, 1994.

[33] G. Sapiro and A. Tannenbaum, "Affine invariant scale-space," *International Journal of Computer Vision*, vol. 11, no. 1, pp. 25–44, 1993.

[34] G. Sapiro and A. Tannenbaum, "On invariant curve evolution and image analysis," *Indiana Univ. Math. J.*, vol. 42, no. 3, pp. 985–1009, 1993.

[35] S. Angenent, G. Sapiro, and A. Tannenbaum, "On the affine heat flow for non-convex curves," *J. Amer. Math. Soc.*, vol. 11, no. 3, pp. 601–634, 1998.

[36] L. Alvarez and J.-M. Morel, "Formalization and computational aspects of image analysis," *Acta Numerica*, vol. 3, pp. 1–59, 1994.

[37] S. O. L. I. Rudin and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[38] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser Verlag, 1984.

[39] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[40] A. Toga, *Brain Warping*. Academic Press, 1999.

[41] J. Hajnal, D. Hawkes, D. Hill, and J. Hajnal, eds., *Medical Image Registration*. CRC Press, 2001.

[42] J. Pluim and J. Fitzpatrick, "Special issue on image registration," *IEEE Transactions on Medical Imaging*, vol. 22, Nov 2003.

[43] G. Penney, J. Weese, J. J.A. Little, P. Desmedt, D. Hill, and D. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 586–595, Aug 1998.

[44] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187 – 198, 1997.

[45] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 986–1004, Aug 2003.

[46] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, "Optimal mass transport for registration and warping," *Int. Journal Computer Vision*, vol. 60, no. 3, pp. 225–240, 2004.

[47] D. Fry, *Shape Recognition Using Metrics on the Space of Shapes*. PhD thesis, Harvard University, 1993.

[48] S. Rachev and L. Rüschendorf, *Mass Transportation Problems*. Springer, 1998.

[49] L. V. Kantorovich, "On a problem of Monge," *Uspekhi Mat. Nauk.*, vol. 3, pp. 225–226, 1948.

[50] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions," *Comm. Pure Appl. Math.*, vol. 64, pp. 375–417, 1991.

[51] W. Gangbo and R. McCann, "The geometry of optimal transportation," *Acta Math.*, vol. 177, pp. 113–161, 1996.

[52] M. Knott and C. Smith, "On the optimal mapping of distributions," *J. Optim. Theory*, vol. 43, pp. 39–49, 1984.

[53] L. Ambrosio, "Lecture notes on optimal transport theory," in *Euro Summer School, Mathematical Aspects of Evolving Interfaces*, CIME Series of Springer Lecture Notes, Springer, July 2000.

[54] S. Angenent, S. Haker, and A. Tannenbaum, "Minimizing flows for the Monge-Kantorovich problem," *SIAM J. Math. Anal.*, vol. 35, no. 1, pp. 61–97 (electronic), 2003.

[55] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem," *Numerische Mathematik*, vol. 84, pp. 375–393, 2000.

[56] L. Zhu, *On Visualizing Branched Surfaces: An Angle/Area Preserving Approach*. PhD thesis, Department of Biomedical Engineering, Georgia Institute of Technology, 2004.

[57] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation*. Birkhäuser, 1994.

[58] D. Mumford and J. Shah, "Boundary detection by minimizing functionals," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 22–26, 1985.

[59] L. Roberts, *Optical and Electro-optical Information Processing*, ch. Machine perception of 3-D solids. MIT Press, 1965.

[60] I. Sobel, *Camera Models and Machine Perception*. PhD thesis, Stanford Univ., 1970.

[61] J. Canny, "Computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[62] J. Fram and E. Deutsch, "On the quantitative evaluation of edge detection schemes and their comparisions with human performance," *IEEE Transaction on Computers*, vol. 24, no. 6, pp. 616–627, 1975.

[63] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond.*, vol. B, no. 207, pp. 187–217, 1980.

[64] D. Marr, *Vision*. Freeman, 1982.

[65] S. Gunn, "On the discrete representation of the Laplacian of Gaussian," *Pattern Recognition*, vol. 32, no. 8, pp. 1463–1472, 1999.

[66] S. Ando, "Consistent gradient operators," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 252–265, March 2000.

[67] A. W. M. Kass and D. Terzopoulos, "Snakes: active contour models," *Int. Journal of Computer Vision*, vol. 1, pp. 321–331, 1987.

[68] L. D. Cohen, "On active contour models and balloons," *CVGIP: Image Understanding*, vol. 53, pp. 211–218, March 1991.

[69] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.

[70] T. McInerney and D. Terzopoulos, "Topologically adaptable snakes," in *Int. Conf. on Computer Vision*, pp. 840–845, June 1995.

[71] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik*, vol. 66, pp. 1–31, 1993.

[72] J. S. R. Malladi and B. Vemuri, "Shape modeling with front propagation: a level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, 1995.

[73] H. I. M. G. Crandall and P. L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *Bulletin of the American Mathematical Society*, vol. 27, pp. 1–67, 1992.

[74] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: from phase transitions to active vision," *Arch. Rational Mech. Anal.*, vol. 134, no. 3, pp. 275–301, 1996.

[75] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 11, pp. 61–79, 1997.

[76] M. Grayson, "Shortening embedded curves," *Annals of Mathematics*, vol. 129, pp. 71–111, 1989.

[77] L. Simon, "Lectures on geometric measure theory," in *Proceedings of the Centre for Mathematical Analysis, Australian National University, Canberra*, 1983.

[78] K. Siddiqi, Y. Lauziere, A. Tannenbaum, and S. Zucker, "Area and length minimizing flows for shape segmentation," *IEEE TMI*, vol. 7, pp. 433–443, 1998.

[79] E. Pichon, A. Tannenbaum, and R. Kikinis, "Statistically based flow for image segmentation," *Medical Imaging Analysis*, vol. 8, pp. 267–274, 2004.

[80] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.

[81] D. Mumford, *Geometry-driven Diffusion in Computer Vision*, ch. The Bayesian Rationale for Energy Functionals, pp. 141–153. Kluwer Academic Publisher, 1994.

[82] L. Ambrosio, "A compactness theorem for a special class of functions of bounded variation," *Boll. Un. Math. It.*, vol. 3-B, pp. 857–881, 1989.

[83] A. Tsai, A. Yezzi, and A. Willsky, "A curve evolution approach to smoothing and segmentation using the Mumford-Shah functional," *CVPR*, pp. 1119–1124, 2000.

[84] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[85] E. Pichon, A. Tannenbaum, and R. Kikinis, "A statistically based surface evolution method for medical image segmentation: presentation and validation," in *MICCAI*, vol. 2, pp. 711–720, 2003.

[86] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation*. Birkhäuser, Boston, 1994.

[87] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Computer Vision*, vol. 1, pp. 321–332, 1988.

[88] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Proc. ICCV*, pp. 694–699, 1995.

[89] R. Adams and L. Bischof, "Seeded region growing," *PAMI*, vol. 16, no. 6, pp. 641–647, 1994.

[90] R. Justice, E. Stokely, J. Strobel, R. Ideker, and W. Smith, "Medical image segmentation using 3-D seeded region growing," *Proc. SPIE Symposium on Medical Imaging Volume*, vol. 3034, pp. 900–910, 1997.

[91] R. Pohle and K. Toennies, "Segmentation of medical images using adaptive region growing," in *Proc. SPIE Medical Imaging*, vol. 4322, pp. 1337–1346, 2001.

[92] S. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *PAMI*, vol. 18, no. 9, pp. 884–900, 1996.

[93] N. Paragios and R. Deriche, "Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 249–268, 2002.

[94] A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 195–216, 2002.

[95] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. Wiley-Interscience, 2001.

[96] J. N. Tsitsiklis, "Efficient algorithms for globally optimal trajectories," *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1528–1538, 1995.

[97] Y. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.

[98] A. Zijdenbos, B. Dawant, and R. Margolin, "Morphometric analysis of white matter lesions in MR images: Method and validation," *IEEE TMI*, vol. 13, no. 4, pp. 716–724, 1994.

[99] W. Yasnoff, J. Miu, and J. Bacus, "Error measures for scene segmentation," *Pattern Recognition*, vol. 9, pp. 217–231, 1977.

[100] K. Strasters and J. Gerbrands, "Three-dimensional segmentation using a split, merge and group approach," *Pattern Recognition Letters*, vol. 12, pp. 307–325, 1991.

[101] D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing images using the Hausdorff distance," *PAMI*, vol. 15, no. 9, pp. 850–863, 1993.

[102] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Coupled multi-shape model and mutual information for medical image segmentation," in *Information Processing in Medical Imaging*, pp. 185–197, July 2003.

[103] X. Zeng, L. Staib, R. Schultz, and J. Duncan, "Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 927–937, 1999.

[104] D. Shattuck and R. Leahy, "BrainSuite: An automated cortical surface identification tool," *Medical Image Analysis*, vol. 6, pp. 129–142, June 2002.

[105] R. Kwan, A. Evans, and G. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE TMI*, vol. 18, no. 11, pp. 1085–1097, 1999.

[106] M. Kaus, S. Warfield, A. Nabavi, P. Black, F. Jolesz, and R. Kikinis, "Automated segmentation of MRI of brain tumors," *Radiology*, vol. 218, no. 2, pp. 586–591, 2001.

[107] E. Pichon and A. Tannenbaum, "Curve segmentation using directional information, relation to pattern detection," in *IEEE International Conference on Image Processing (ICIP)*, vol. 2, pp. 794–797, 2005.

[108] E. Pichon, C.-F. Westin, and A. Tannenbaum, "A Hamilton-Jacobi-Bellman approach to high angular resolution diffusion tractography," in *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2005. accepted.

[109] E. Mortensen, B. Morse, W. Barrett, and J. Udupa, "Adaptive boundary detection using live-wire two-dimensional dynamic programming," in *IEEE Proceedings of Computers in Cardiology*, pp. 635–638, 1992.

[110] W. H. Press, S. A. Teukolsky, W. T., and B. P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, 2 ed., 2002.

[111] M. P. D. Carmo, *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.

[112] S. J. Osher and R. P. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2002.

[113] W. Flemming and H. Soner, *Controlled Markov processes and viscosity solutions*. Springer-Verlag, 2003.

[114] H. Soner, "Dynamic programming and viscosity solutions," in *Lectures in the Annual AMS Meeting*, AMS, 1999.

[115] J. Sethian and A. Vladimirsky, "Ordered upwind methods for static Hamilton-Jacobi equations: Theory and applications," *SIAM J. on Numerical Analysis*, vol. 41, no. 1, pp. 325–363, 2003.

[116] C. Kao, S. Osher, and Y. Tsai, "Fast sweeping methods for static Hamilton-Jacobi equations," Tech. Rep. 03-75, UCLA CAM, 2003.

[117] C. Kao, S. Osher, and J. Qian, "Lax-Friedrichs sweeping scheme for static Hamilton-Jacobi equations," *Journal of Computational Physics*, vol. 196, pp. 367–391, May 2004.

[118] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J Physiol*, vol. 160, pp. 106–154, 1962.

[119] D. Hubel and T. Wiesel, "Receptive fields and functional architecture of the monkey striate cortex," *J Physiol*, vol. 195, pp. 215–243, 1968.

[120] R. DeValois, E. Yund, and N. Hepler, "The orientation and direction selectivity of cells in the macaque visual cortex," *Vis Res*, vol. 22, pp. 531–544, 1982.

[121] J. Daugman, "Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, vol. 2, pp. 1160–1169, 1985.

[122] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J Neurophysiol*, vol. 58, pp. 1233–1258, 1987.

[123] P. Kruizinga and N. Petkov, "Non-linear operator for oriented texture," *IEEE Trans. on Image Processing*, vol. 8, no. 10, pp. 1395–1407, 1999.

[124] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Contour detection based on nonclassical receptive field inhibition," *IEEE Trans. Image Processing*, vol. 12, no. 7, pp. 729–739, 2003.

[125] C. Beaulieu, "The basis of anisotropic water diffusion in the nervous system - a technical review," *NMR in Biomedicine*, vol. 15, pp. 435–455, 2002.

[126] P. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, pp. 259–267, 1994.

[127] D. L. Bihan, J. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: concepts and applications," *J Magn Reson Imaging*, vol. 13, pp. 534–46, Apr 2001.

[128] S. Mori, B. Crain, V. Chacko, and P. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann Neurol.*, vol. 45, pp. 265–269, Feb 1999.

[129] 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," in *Proc. Natl. Acad. Sci. USA*, pp. 10422–10427, August 1999.

[130] C.-F. Westin, S. E. Maier, B. Khidhir, P. Everett, F. A. Jolesz, and R. Kikinis, "Image processing for diffusion tensor magnetic resonance imaging," in *MICCAI*, pp. 441–452, 1999.

[131] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magnetic Resonance in Medicine*, vol. 44, pp. 625–632, 2000.

[132] G. Parker, C. Wheeler-Kingshott, and G. Barker, "Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 505–512, May 2002.

[133] M. Jackowski, C. Y. Kao, M. Qiu, R. T. Constable, and L. Staib, "Estimation of anatomical connectivity by anisotropic front propagation and diffusion tensor imaging," in *MICCAI*, pp. 663–671, 2004.

[134] L. O'Donnell, S. Haker, and C.-F. Westin, "New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space," in *MICCAI*, 2002.

[135] C. Lenglet, R. Deriche, and O. Faugeras, "Inferring white matter geometry from diffusion tensor MRI: Application to connectivity mapping," in *ECCV*, 2004.

[136] C. Liu, R. Bammer, B. Acar, and M. Moseley, "Characterizing non-Gaussian diffusion by using generalized diffusion tensors," *Magnetic Resonance in Medicine*, vol. 51, no. 5, pp. 924–937, 2004.

[137] D. Tuch, "Q-ball imaging," *Magn Reson Med.*, vol. 52, pp. 1358–1372, 2004.

[138] 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 *ISMRM*, 2004.

[139] J. S. Campbell, *Diffusion Imaging of White Matter Fibre Tracts*. PhD thesis, McGill University, 2004.

[140] G. Strang, *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, 1986.

[141] E. Pichon, D. Nain, and M. Niethammer, "A Laplace equation approach for shape comparison," in *SPIE Medical Imaging*, 2006. Submitted.

[142] R. C. Veltkamp, "Shape matching: similarity measures and algorithms," Tech. Rep. UU-CS-2001-03, Utrecht University, 2001.

[143] S. Haker, S. Angenent, A. Tannenbaum, and R. Kikinis, "Nondistorting flattening maps and the 3D visualization of colon CT images," *IEEE Trans. on Medical Imaging*, vol. 19, pp. 665–670., 2000.

[144] D. Nain, S. Haker, W. Eric L. Grimson, E. Cosman Jr, W.W. Wells, H. Ji, R. Kikinis and C.F. Westin, "Intra-patient prone to supine colon registration for synchronized virtual colonoscopy," in *MICCAI*, 2002.

[145] S. Jones, B. Buchbinder, and I. Aharon, "Three-dimensional mapping of cortical thickness using Laplace's equation," *Human Brain Mapping*, vol. 11, pp. 12–32, 2000.

[146] A. Yezzi and J. L. Prince, "A PDE approach for measuring tissue thickness," in *CVPR*, vol. 1, pp. 87–92, 2001.

[147] L. Evans, *Partial Differential Equations*, vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.

[148] G. Gipson, *Boundary Element Fundamentals - Basic Concepts and Recent Developments in the Poisson Equation*. Computational Mechanics Publications, 1987.

[149] T. Kuwabara and T. Takeda, "Boundary element method using analytical integration for three-dimensional Laplace problem," *Transactions of the Institute of Electrical Engineers of Japan*, vol. 106, no. 6, pp. 25–31, 1986.

[150] M. Fratantonio and J. J. Rencis, "Exact boundary element integrations for two-dimensional Laplace equation," *Engineering Analysis with Boundary Elements*, vol. 24, pp. 325–342, 2000.

[151] H. Lefkowitz and G. Herman, "Color scales for image data," *IEEE Computer Graphics and Applications*, vol. 12, no. 1, pp. 72–80, 1992.

[152] W. Pratt, *Digital Image Processing*. Wiley-Interscience, 1978.

[153] Y. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.

VITA

Eric Pichon was born in Meudon, near Paris in France on March 8th, 1978. He received a "Diplôme d'ingénieur" from the Ecole Supérieure d'Electricité (Supélec) in France in 2001 and a Master of Science in Electrical and Computer Engineering from the Georgia Institute of Technology in 2002. In 2005, he was graduated from the Georgia Institute of Technology with a Ph.D. in Electrical and Computer Engineering. His research focused on computer vision and its applications to the automatic analysis of medical imaging datasets.

Journal Publications

[1] S. Angenent, E. Pichon, and A. Tannenbaum, "Mathematical methods in medical image processing," *Bulletin of the American Mathematical Society*, 2005. Accepted.

[2] E. Pichon, A. Tannenbaum, and R. Kikinis, "A statistically based flow for image segmentation," *Medical Image Analysis*, vol. 8, pp. 267–274, September 2004.

[3] E. Pichon, G. Sapiro, and A. Tannenbaum, *Lecture Notes in Control and Information Sciences*, vol. 286 of *Directions in Mathematical Systems Theory and Optimization*, ch. Segmentation of Diffusion Tensor Imagery, pp. 239–247. Springer-Verlag Heidelberg, January 2003.

Conference Publications

[1] E. Pichon, D. Nain, and M. Niethammer, "A Laplace equation approach for shape comparison," in *SPIE Medical Imaging*, 2006. Accepted.

[2] E. Pichon, C.-F. Westin, and A. Tannenbaum, "A Hamilton-Jacobi-Bellman approach to high angular resolution diffusion tractography," in *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 180–187, 2005.

[3] E. Pichon and A. Tannenbaum, "Curve segmentation using directional information, relation to pattern detection," in *IEEE International Conference on Image Processing (ICIP)*, vol. 2, pp. 794–797, 2005.

[4] E. Pichon, C. Novak, A. Kiraly, and D. Naidich, "A novel method for pulmonary emboli visualization from high-resolution CT images," in *SPIE Conference on Medical Imaging*, vol. 5367, pp. 161–170, 2004.

[5] A. Kiraly, E. Pichon, D. Naidich, and C. Novak, "Analysis of arterial subtrees affected by pulmonary emboli," in *SPIE Conference on Medical Imaging*, vol. 5370, pp. 1720–1729, 2004.

[6] E. Pichon, A. Tannenbaum, and R. Kikinis, "A statistically based surface evolution method for medical image segmentation: presentation and validation," in *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 2, pp. 711–720, 2003. best student presentation in image segmentation award.

[7] E. Pichon, M. Niethammer, and G. Sapiro, "Color histogram equalization through mesh deformation," in *IEEE International Conference on Image Processing (ICIP)*, vol. 2, pp. 117–120, 2003.

[8] M. Niethammer, E. Pichon, A. Tannenbaum, and P. J. Mucha, "A Stokes flow boundary integral measurement of tubular structure cross sections in two dimensions," in *Proceedings of the International Conference on Image Processing*, pp. 825–828, IEEE, 2003.

[9] T. N. Mundhenk, C. Ackerman, D. Chung, N. Dhavale, B. Hudson, R. Hirata, E. Pichon, Z. Shi, A. Tsui, and L. Itti, "Low-cost high-performance mobile robot design utilizing off-the-shelf parts and the Beowulf concept: the Beobot project," in *SPIE Conference on Intelligent Robots and Computer Vision*, vol. 5267, pp. 293–303, Oct 2003.

[10] E. Pichon and L. Itti, "Real-time high-performance attention focusing for outdoors mobile Beobots," in *AAAI Spring Symposium*, (Stanford, CA), p. 63, March 2002.

[11] D. Chung, R. Hirata, T. N. Mundhenk, J. Ng, R. J. Peters, E. Pichon, A. Tsui, T. Ventrice, D. Walther, P. Williams, and L. Itti, "A new robotics platform for neuromorphic vision: Beobots," in *Workshop on Biologically Motivated Computer Vision (BMCV)*, 2002.

[12] J. Ng, R. Hirata, N. Mundhenk, E. Pichon, A. Tsui, T. Ventrice, P. Williams, and L. Itti, "Towards visually-guided neuromorphic robots: Beobots," in *Joint Symposium on Neural Computation (JSNC)*, 2002.

Patents

[1] E. Pichon, C. Novak, and A. Kiraly, "System and method for visualization of pulmonary emboli from high-resolution computed tomography images." US Patent, filed on 04/11/05, 2005. Application Serial No. 11/103,108.