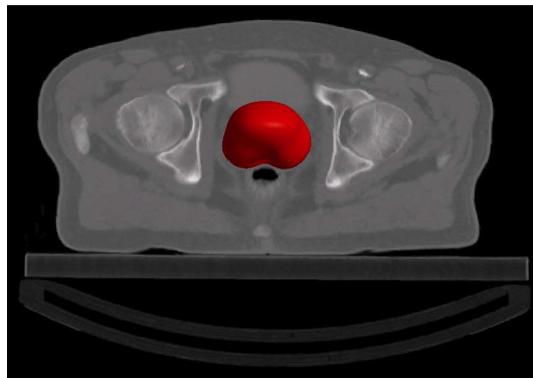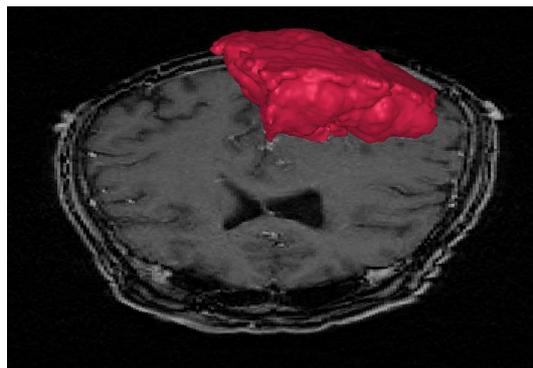# Committing Medical Image Analysis

Erik B. Dam

August, 2003

Preface

This constitutes Erik Dam's Ph.D. dissertation. The dissertation is submitted at the IT University of Copenhagen in partial fulfillment of the requirements for the degree of doctor of philosophy.

The work behind the dissertation was delivered in the period from July 2000 to June 2003 under the supervision of professor Mads Nielsen and co-supervision of assistant professor Ole Fogh Olsen from the IT University of Copenhagen. The work was carried out at the IT University of Copenhagen with the exception of a six month visit followed by a one month visit at the University of North Carolina, Chapel Hill, hosted by professor Stephen M. Pizer.

Front page, top image: A segmentation of a brain tumor. The brain scan is used in chapter 4 courtesy of the University Medical Center, Utrecht, the Netherlands. The segmentation was performed using the ∇*Vision* segmentation program [Dam et al., 2003a].

Front page, bottom image: A segmentation of a prostate. The abdominal CT scan is used in chapter 9 courtesy of the Radiation Oncology department, University of North Carolina, Chapel Hill, US. The segmentation was performed using the *Pablo* segmentation program [Fletcher et al., 2002].

Abstract

The topic of this dissertation is medical image analysis with focus on multi-scale methods and the task of segmentation. The main theme is how to move from explicitly non-committed approaches to methods that incorporate prior knowledge and thereby become specialized to the given task. The presented work explores a path from a non-committed multi-scale segmentation method towards more committed methods for segmentation and shape modeling.

The multi-scale watershed segmentation method by Fogh Olsen is the starting point for much of the work in this dissertation. In this non-committed method, the segmentation problem is addressed with the only assumption that there should be some contrast across the boundaries of the desired objects.

The first body of work is focused on specializing the multi-scale segmentation method through the use of non-linear diffusion. We present a *Generalized Anisotropic Non-linear* diffusion scheme with methods for optimizing the parameters and evaluating the performance. The results apply for segmentation of both 2D and 3D objects and show considerable performance improvements through the use of GAN. Upper limits for this performance gain are empirically established through experiments on artificial ideal objects. The methodology is evaluated on segmentation of brain structures in both 2D and 3D and shows significant improvements in segmentation efficiency.

Non-linear diffusion is then analyzed in more detail via the introduction of the *Diffusion Echo* that allows explicit analysis of the local filters in non-linear diffusion schemes. We use this approach to investigate the connection between linear and non-linear diffusion given by scale selection.

While non-linear diffusion allows a large degree of incorporation of prior knowledge and thereby facilitates specialization towards a specific task, the work shows that for the multi-scale segmentation program, the possible performance improvement due to non-linear diffusion is not unlimited. Specifically, near-automatic segmentation seems unfeasible without further commitment of the method. The most promising way of incorporating additional prior knowledge is through the use of a shape model.

The final body of work is focused on shape modeling via the medial shape representation known as the m-rep. We present an essentially automatic method for generating a statistical shape model from a training collection. We demonstrate the method for constructing a prostate shape model.

List of Publications

The majority of this dissertation has been published in one of the publications below (the full references are available in the bibliography). They have all been subject to peer review or are currently in the review process.

Non-linear Diffusion for Interactive Multi-scale Watershed Segmentation
   E. Dam & M. Nielsen.
   Medical Image Computing and Computer-Assisted Intervention — MICCAI 2000.

Exploring Non-linear Diffusion: The Diffusion Echo
   E. Dam & M. Nielsen.
   Scale-Space Theories in Computer Vision, 3rd international conference, 2001.

Approximating Non-linear Diffusion
   E. Dam, O.F. Olsen, & M. Nielsen.
   Scale-space Methods in Computer Vision, 4th international conference, 2003.

Non-linear Diffusion in 3D for Interactive Segmentation of Brain Tissue
   E. Dam & M. Letteboer.
   Submitted to MICCAI, 2003.

Prostate Shape Modeling based on Principal Geodesic Analysis Bootstrapping
   E. Dam, P.T. Fletcher, S.M. Pizer, G. Tracton, & J. Rosenman.
   Submitted to the Variational, Geometric and Level Set Methods workshop of ICCV, 2003.

In addition, the following supporting publications form the basis for some of the presented work but are not directly included in the dissertation. First, the supporting publications that have been subject to peer review:

Interactive Multi-Scale Segmentation in Clinical Use
   E. Dam, P. Johansen, O.F. Olsen, A. Thomsen, M. Lillholm, T. Darvann, A. Dobrzeniecki, N. Hermann, N. Kitai, S. Kreiborg, P. Larsen, & M. Nielsen.
   European Congress of Radiology, 2000.

Interactive multi-scale watershed segmentation of tumors in MR brain images
   M. Letteboer, W. Niessen, P. Willems, E. B. Dam, & M. Viergever.
   The Interactive Medical Image Visualization and Analysis workshop of MICCAI, 2001.

Segmentation of Tumors in MR Brain Images using an Interactive Multi-scale Watershed Algorithm
   M. Letteboer, O. F. Olsen, E. B. Dam, P. Willems, M. Viergever, & W. Niessen.
   Submitted to IEEE Transactions on Medical Imaging, 2003.

Finally, the supporting publications that have been subject to limited or no peer review:

Nabla Vision
A. Thomsen, E. Dam, O.F. Olsen, P. Johansen, M. Nielsen, & M. Lillholm.
Segmentation program developed 1999–2003.
http://www.itu.dk/image/nablavision.

Evaluation of Diffusion Schemes for Watershed Segmentation
E. Dam.
Master's thesis, University of Copenhagen, 2000.

Method and System for Multi-dimensional Segmentation
O. F. Olsen, E. B. Dam, M. Lillholm, A. Thomsen, P. Johansen, & M. Nielsen.
Danish patent application filed October 2000.

M-Rep Based Classification of Organ Conformations for Radiation Treatment Planning of Prostate Cancer
E. Dam, S.M. Pizer, J. Rosenman, & G. Tracton.
The 17th Annual UNC Radiology Research Symposium, 2002.

Billedanalyse og Strålebehandling
E. Dam.
(in english: "Image Analysis and Radiation Treatment")
At the science web portal of Danmarks Radio (national danish public service broadcasting company), 2003.

Deep Structure I & II
E. Dam & B. ter Haar Romeny.
Chapters 13 and 14 in *Front End Vision and Multi-Scale Image Analysis*, 2003.

Deep Structure III
B. ter Haar Romeny & E. Dam.
Chapter 15 in *Front End Vision and Multi-Scale Image Analysis*, 2003.

Acknowledgements

A large numbers of persons have been valuable to me during my Ph.D. years. I hope you all know who you are. I will only mention a few persons that have been particularly helpful in the work that lies behind this dissertation:

Professor Mads Nielsen (IT University of Copenhagen)
    For supervision and inspiration.

Associate Professor Ole Fogh Olsen (IT University of Copenhagen)
    For co-supervision, inspiration, and for answering silly questions.

Professor Stephen M. Pizer (University of North Carolina, Chapel Hill)
    For inspiration during the last one and a half year.

Martin Lillholm (IT University of Copenhagen)
    For answering silly questions and coming up with brand new ones of your own.

Contact Information

The author of this dissertation can currently be contacted at the following address:

Erik Dam
IT University of Copenhagen
Glentevej 67
DK-2400 Copenhagen NV
Denmark
Email: `erikdam@itu.dk`

Contents

Chapter 1

Introduction

The topic of this dissertation is medical image analysis with focus on multi-scale methods and the task of segmentation. However, most of the results are generally applicable to other computer vision areas and other tasks than segmentation.

The main theme is how to move from explicitly non-committed approaches to methods that incorporate prior knowledge and thereby become specialized to the given task.

1.1 The Big Picture

A complete understanding of the field of medical image analysis requires perfect models of the anatomy, the image formation process, and especially the variation within these depending on inter and intra patient variability, imaging machinery, and choices made by the radiologist. Explicit modeling of all these factors is not feasible. Even though some of the variations can be modeled with a sufficient accuracy, others can still only be modeled through simplifying approaches. This is in particular the case for the modeling of patient variability.

Even so, impressive results have been achieved through models that ignore the underlying cause for the image or shape formation and simply model the data. One obvious example is the *Active Shape Model* [Cootes et al., 1995].

However, these models are still somewhat naive. Admittedly, in many cases the active shape model provides a model of shape variation that allows generalization from training shapes to other organs of the same type. However, these models often suffer from a lack of specificity in the sense that they allow the formation of illegal shapes. Either these shapes are illegal in the sense that they are not realistic examples for the given object — or alternatively mathematically illegal (i.e. objects with unsuitable topology or self-intersecting surfaces).

More realistic models are possible through modeling of the actual physical properties of the anatomical structures. Using this approach, nice results have been achieved, for instance through the use of *Finite Element Models* [Crouch et al., 2003]. However, research is needed before these methods are fully matured.

Currently a full model that explains the entire data formation process including patient variability and image aquisitioning is not realistic. Models that approximate the observed image data directly and ignore the underlying image formation process do in many cases provide sufficiently accurate modeling and help solve the task at hand. The work in this dissertation is focused on this last approach.

1.1.1 Non-Committed Methods

When facing a specific task, a full model would allow incorporation of this understanding of the patient variability and image aquisition process as prior knowledge in a probabilistic framework. The lack of a model suggests non-committed methods. Ad hoc methods typically have implicit assumptions about the data, and therefore commit the methods to specific problem classes. A non-committed method explicitly enforces a minimal set of assumptions.

Perhaps surprisingly, making a method non-committed is not a trivial task — most methods are committed in implicit and undesired ways. *Scale-space theory* [Koenderink, 1984, Witkin, 1983, Ijima, 1962] provides one viewpoint on how to observe and analyze data in the absence of an explicit model of the image formation process.

An example of a method that follows this paradigm is the *Multi-scale Watershed Segmentation* method [Olsen, 1997] that is the starting point for much of the work is this dissertation. In this method, the segmentation problem is addressed with the only assumption that there should be some contrast across the boundaries of the desired objects. Size, shape, texture, and other features are assumed unknown. Without knowledge of the desired objects, automatic segmentation is obviously impossible. The method uses the approach where the user is presented with 3D building blocks that can then be used to interactively sculpt the desired objects. Even with this few assumptions on the observed data and the desired objects, the method allows good performance on complicated tasks such as segmentation of the mandible [Dam et al., 2000] and brain tumors [Letteboer et al., 2001] (see figure 1.1).

While following an appealing theoretical paradigm, non-committed methods also suffer from this very principle. They will not perform worse than average for any problem, but unfortunately they will not offer superior performance on many problems either. Even though the full understanding of most problems is not feasible, often even a limited modeling of the observed data allows a certain degree of specialization of a method.

The work in this dissertation investigates the path from non-committed towards more specialized methods.

1.1.2 Implicit Commitment via Evaluation

Incorporation of prior knowledge allows commitment of a method to the given task. Simple as this sounds, it is often problematic due to lack of ability to quantitatively model the prior knowledge. In cases where there is ground truth available, that describes the desired output of the method, this problem can be overcome. If a method has a number of free parameters to allow specialization, these can be optimized through evaluation against the ground truth.
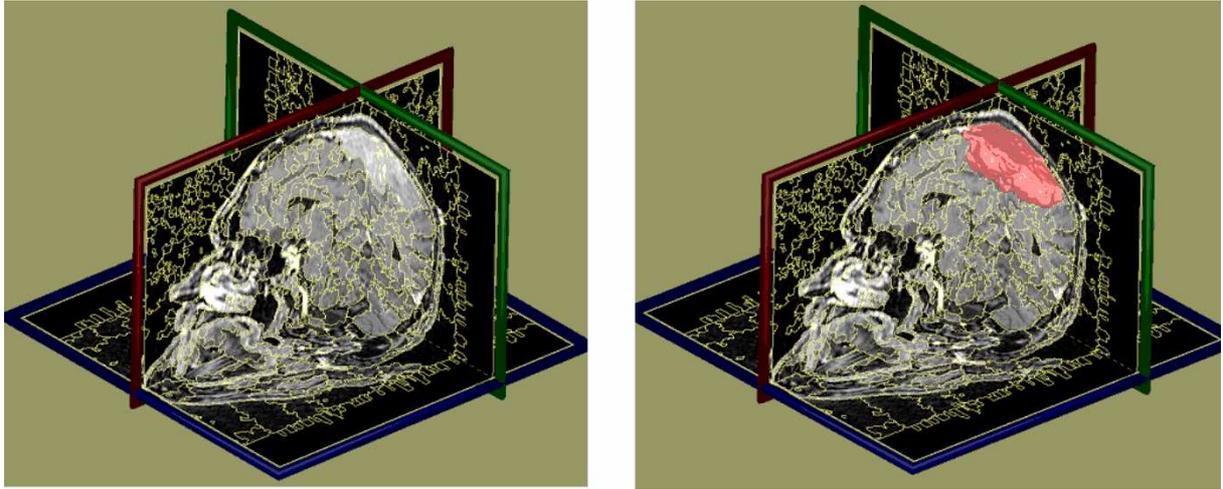
Figure 1.1: *Segmentation of a brain tumor using the ∇Vision segmentation program based on multi-scale watershed segmentation [Dam et al., 2003a, Olsen, 1997].* **Left**: *The building blocks that the user can select and deselect to sculpt the desired object. In this case the tumor at the top of the brain.* **Right**: *The segmented tumor. The illustration is from [Letteboer et al., 2003].*

We use this approach to design the *Generalized Anisotropic Non-linear* diffusion schemes for use in the multi-scale watershed segmentation method. Committing the building blocks through optimization of the parameters adds prior knowledge of the ground truth objects and thereby allow construction of building blocks better suited for the given task [Dam & Nielsen, 2000, Dam & Letteboer, 2003].

The analysis and optimization of non-linear diffusion schemes inspires the development of the *Diffusion Echo* [Dam & Nielsen, 2001]. This leads to better understanding of the non-linear diffusion schemes and in particular to the connection between non-linear and linear diffusion. The practical outcome of the work could be efficient implementation of simpler approximations of non-linear schemes [Dam et al., 2003c].

These approaches have little explicit representation of the actual prior knowledge that allows them to be specialized towards a specific task. The non-linear diffusion schemes have parameters that are optimized through evaluation. These optimal parameters do reveal some information about the anatomical objects. For instance, the soft edge threshold in the *Edge Enhancing diffusion* scheme [Weickert, 1998a] gives an explicit measure for the object boundary contrast, that the method is specialized towards. The specialization implied by the other optimal parameters is less obvious. For instance, boundary curvature and overall object elongation are implicitly adressed, but this specialization is quite intangible.

1.1.3   Explicit Commitment via Shape Modeling

The work on committing the multi-scale watershed segmentation method through the use of non-linear diffusion schemes shows that the performance can be significantly improved. However, achieving essentially automatic segmentation is not possible for most tasks. The implicit specialization does not allow sufficient commitment to the given segmentation task.

3

Specialization through a shape model of the desired anatomical objects is more explicit. In this approach, the shape model is trained on a database of example organs. For parametric shape models with probabilisticly defined variations (such as the active shape model or the *m-rep* [Pizer et al., 2003a]) this is actually very similar to commitment through optimization of parameters via evaluation against ground truth data. However, a major difference is the ability to sample from parametric shape models. This allows visualization of the shape class that the method is committed to. Thereby the commitment becomes more explicit.

Extending the multi-scale watershed segmentation method with an explicit shape model is a non-trivial problem. The segmentations are achieved through adding and removing regions at different scales. This corresponds to selecting and deselecing nodes in the multi-scale linking structure. Thereby the shape of the resulting object is implicitly represented as an attributed graph. For similar but different objects, the size and topology of this graph will differ. Even though promising preliminary results exist for related shape representations [Shokoufandeh et al., 2002, Demirci et al., 2003, Siddiqi et al., 1999], no shape model has yet been presented for an implicit shape representation of this kind. The main problem compared to the cited approaches is that the multi-scale watershed linking tree represents the entire image and not only the desired object.

Another shape model, that is also based on an underlying graph representation with varying topology, is the m-rep. A sampled medial axis will in general differ in topology and size for different objects. The approach chosen in the m-rep is to make it a generative model with fixed topology and sampling for a given object class. A similar approach could be relevant in the multi-scale watershed method.

The concluding work presented in this dissertation presents a shape modeling method that automatically generates an m-rep shape model with corresponding probabilistic variations from a training collection of shapes. The method is used to create a prostate shape model.

1.2 The Contributions in this Dissertation

The contributions of this dissertation fall in three main categories:

Non-linear Diffusion in Multi-Scale Watershed Segmentation in 2D & 3D
 The generalized anistropic non-linear diffusion scheme is introduced and evaluated for the purpose of feeding prior knowledge of the desired anatomical structures into the interactive multi-scale watershed segmentation method. This work is concentrated on segmentation of brain structures in 2D and 3D.

The Diffusion Echo
 The diffusion echo is introduced and offers intuitive visualization and analysis of non-linear diffusion schemes. In particular, the analysis is focused on the feasibility of approximating non-linear diffusion schemes with simpler, local Gaussian filters.

Shape Modeling via a Medial Representation
 A bootstrapping method for automatic construction of a medial shape model from a training set of anatomical structures is introduced. The method is applied for building a prostate shape model.

1.3   Reading this Dissertation

The following chapters assume basic knowledge of standard medical image analysis methods. The presentation is kept short and concise. A more elaborate introduction to the specific concepts required to read this dissertation can be found in [Dam, 2000].

The majority of the following chapters are based on edited versions of the publications listed in the preface. Each publication is assigned a separate chapter. These chapters have been slightly re-written in order to present the work coherently. However, the main contents and results are left untouched. Where later work improves earlier published results, the original results are therefore included in order to faithfully present the published work. The improved results are then added in a following chapter.

Chapter 2 introduces the multi-scale watershed segmentation method. The work on the use of non-linear diffusion in the multi-scale watershed segmentation is presented in chapters 3, 4, and 5. The diffusion echo and its uses are presented in chapter 6 and 7. The work on shape modeling is presented in chapters 8, 9, and 10. Finally, the concluding summary is in chapter 11.

Chapter 2

Multi-scale Watershed Segmentation

As announced in the introductory chapter, the multi-scale watershed segmentation (MSWS) method acts as the starting point for much of the work in this dissertation. In this chapter, the method and an implementation of it is presented.

2.1   Introduction

In many medical segmentation tasks, the images do not sufficiently clearly outline the relevant anatomical structures for making simple automated segmentations. The counter-example is simple thresholding of bone structures in CT images [Lorensen & Cline, 1987]. However, such techniques do not work for most anatomical structures such as soft tissue in CT (due to varying and indistinguishable attenuation [Webb, 1988]), most structures in MR (due to the image inhomogeneities [Webb, 1988]), and most structures in PET/SPECT (due to noise [Webb, 1988]). Here, the alternatives are either performing a tedious manual outline slice per slice, or creating specialised algorithms heavily supported by prior information [Cootes et al., 1995].

The interactive 3D multi-scale watershed segmentation tool, $\nabla Vision$, may successfully be applied in these situations [Dam et al., 2000]. The image scale-space is created by Gaussian convolution [Ijima, 1962, Witkin, 1983, Koenderink, 1984]. The watersheds of the gradient magnitude are computed independently at all scales. A linking procedure gives the simpler large scale watersheds the small scale localization [Olsen, 1996, Lindeberg, 1994, Gauch, 1999, Lifshitz & Pizer, 1990, Griffin & Colchester, 1995]. The linked watershed regions constitute a multi-scale partitioning of the images.

Ideally, a given anatomical structure may be outlined by a single region. However in most situations, the linked watersheds do not directly compare to the anatomical structures. $\nabla Vision$ lets the user arbitrarily change scale and select and deselect regions, and thereby sculpt the anatomical structure. All interaction is geometrical and thereby intuitive to the clinician. The speed-up compared to manual segmentation depends on the interactions required to outline the anatomical structure.

The Gaussian scale-space is a least committed scale-space. A number of other scale-space qualify as being least committed since they posess similar axiomatic foundations. Among these are alpha scale-space [Duits, 2003], total variation minimizing diffusion [Rudin et al., 1992], mean curvature motion [Gage, 1983], and affine morphological scale-space. A certain measure of personal preference guides our choice of Gaussian scale-space as the elevated least committed scale-space. Furthermore, the seperability property of Gaussian scale-space allows simple and fast implementations. Though the use of a least committed scale-space the resulting segmentation method is considered general with no specific task in mind.

2.2   MSWS Basics

During rain the drops gather in pools. The topology of a landscape defines the regions of support for each pool — the *catchment basins*. The boundaries between the catchment basins are termed *watersheds*. On large scale, the watersheds of a landscape are a particular kind of ridges and the catchment basins are the dales. The geographical concept *watershed* was introduced to mathematicians in [Maxwell, 1870].

The watersheds allow a simple partitioning of an image. However, for segmentation purposes the regions border should be defined as the watersheds of a dissimilarity measure instead of the original image. A simple, non-committed dissimilarity measure is the gradient magnitude.

The structures that are outlined by this partitioning are defined with respect to the scale at which the gradient is calculated. Different scales are therefore needed to locate objects of different sizes. The theory of scale-space suggests looking at the *deep structure* [Koenderink, 1984, Witkin, 1983, Sporring et al., 1997] — how the catchment basins develop over scale.

Each catchment basin corresponds to a local minimum for the gradient magnitude. The generic events for the gradient magnitude minima have been derived with the conclusion that *fold annihilation*, *fold creation*, *cusp annihilation*, and *cusp creation* catastrophes are stable and therefore to be expected for typical images. For the catchment basins, this corresponds to the *annihilation*, *creation*, *merge*, and *split* events [Olsen, 1996, Olsen & Nielsen, 1997]. These are illustrated in figure 2.1.

Linking of the catchment basins across scale combines the simplification at the detection scale with the fine scale precision at the localization scale (see figure 2.2). The segmentation method presented in [Olsen, 1996] uses these localized basins as building blocks for the segmentation. The user can shift the detection scale and thereby select building blocks appropriate for sculpting the desired objects. As scale increases, regions merge with no movement of the boundaries due to the linking to localization scale.

Approaches related to the multi-scale watershed segmentation of [Olsen, 1996] are presented in [Lifshitz & Pizer, 1990], [Gauch & Pizer, 1993], [Jackway, 1996], and [Gauch, 1999]. However, important contributions of [Olsen, 1996] are the robust linking of regions using maximum overlap and that an intuitive interface is presented that allows the user to interact directly with the three-dimensional building blocks. This forms the basis for an implementation that is presented in the following section.

No change   Annihilation   Creation   Merge   Split
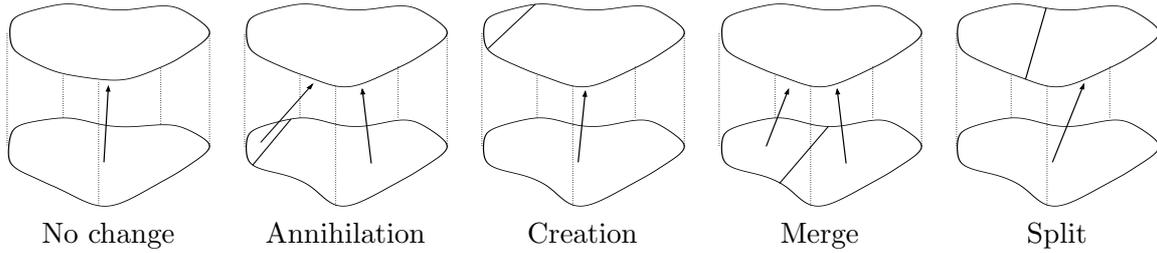
Figure 2.1: *Linking of catchment basins as scale increases. The five possible linking types of watershed regions between two adjacent scale levels [Olsen, 1996]. From low scale to high scale, each region is linked to the region at high scale with maximum area/volume overlap. This creates a linking tree (as opposed to a graph) since creations and split are implicitly ignored.*

Original example image   Watersheds at localization scale   Watersheds at detection scale   Linked to localization scale
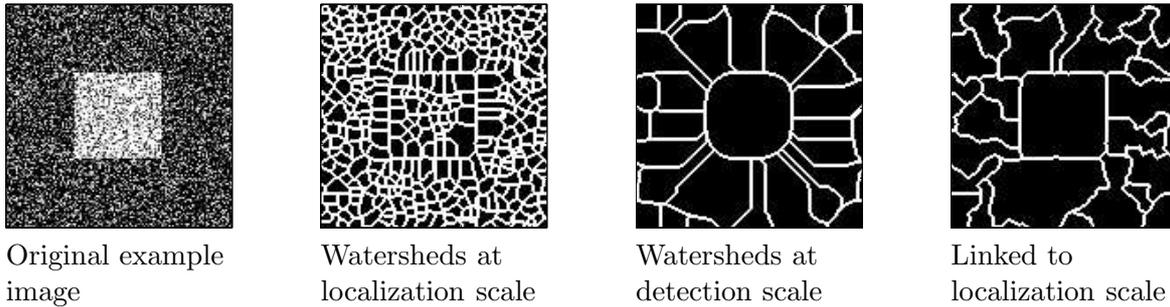
Figure 2.2: *Linking of watershed regions down to localization scale. The catchment basins at detection scale is linked down to the localization scale and thereby get fine scale precision.*

2.3 ∇ *Vision*

The segmentation program does not attempt automatic segmentation — rather the goal is to provide the expert user with an intuitive interaction that aids the segmentation process and thereby reduce interaction time and increase reproducability.

The philosophy of the segmentation methods of ∇ *Vision* is to provide effective means for simplifying the segmentation task — in a non-committed fashion that allows the expert user to reach the desired segmentation.

Segmentation using Building Blocks

The user selects and deselects among the generated building blocks in order to sculpt the desired anatomical objects. The segmentation process is illustrated in figure 2.3.

The choice of scale is essential for the selection of building blocks. If the scale is too small, many building blocks must be added in order to reach the desired object. A typical process starts with the selection of some large building blocks followed by refinement at a smaller scale through addition and removal of smaller building blocks. It is possible to reach any desired segmentation since the smallest building blocks are voxel-sized.
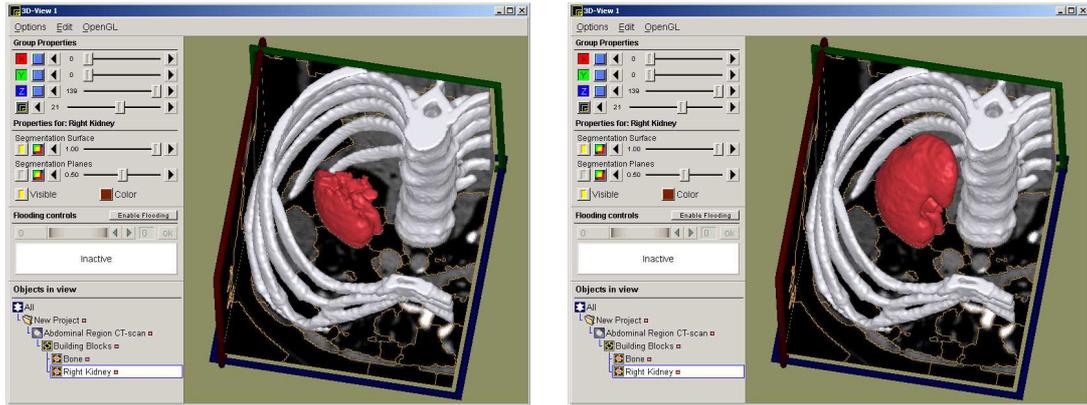
Figure 2.3: *The building blocks generated in the preprocessing step are used for segmenting the kidney.* **Left**: *A single building block have been selected. The size of the building blocks is controlled by the scale/resolution slider (at 21 in the figure).* **Right**: *An additional building block is selected by clicking on the partial kidney where the top is missing. The object is a segmentation of the kidney.*

A key point is that the actions are 3D actions. Even though the building blocks are only displayed in terms of the boundaries on the planes, an action will add/remove a 3D shape.

Furthermore, actions are done in 3D. A selection on top of an object will add the neighboring building block — even though the boundaries for this building block are not displayed on any plane. An example of this effect is illustrated in figure 2.3 where the top half of the kidney is added without moving the planes to the slices that contain that part.

Preprocessing using Intensity Transformation

The preprocessing step generating the building blocks can be customized using an intensity transformation. This changes the contrast between the intensities that corresponds to specific anatomical structures. Since the construction of the building blocks are based on the contrast, a suitable intensity transformation can ensure that the building blocks are better suited for segmentation of a given object. This is a simple, interactive commitment of the method to a specific segmentation task.

Segmentation using Histrogram Thresholding

Some structures are easily distinguishable from their intensity values in a scan. The generic example is bone structure in CT scans. For easy segmentation ∇ *Vision* has a simple thresholding segmentation method.

This method is illustrated in figure 2.4. Note that a threshold segmentation will generally not offer a perfect segmentation due to noise and other artifacts (contrast enhancement agents, for example). Morphological operations can sometimes effectively be used to clean the small noisy segments but are less useful for the larger segments due to artifacts. In both cases, deselection of building blocks offers an effective way of cleaning up the segmentation.
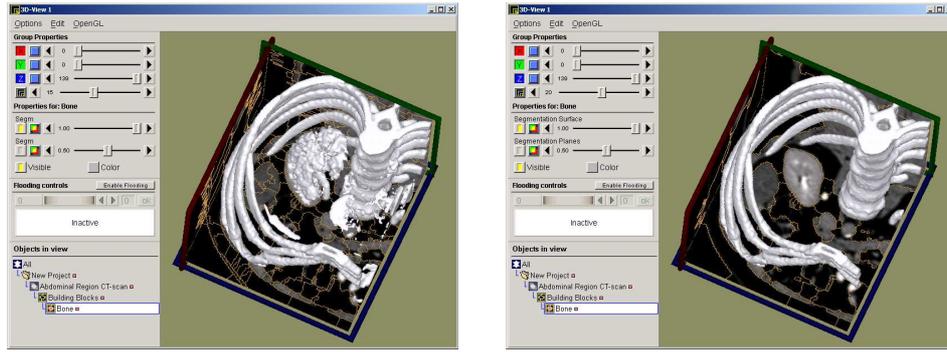
Figure 2.4: *Segmentation using histogram thresholding.* **Left:** *An intensity interval has been specified (from a histogram of the intensities) that selects the higher intensities. As a result the bone structure is selected. However, a spongy kidney and some blood vessel pieces are selected as well (their high intensity values are due to a contrast agent) together with small sections with noisy values. Relatively large building blocks can now be applied to deselect the undesired structure in approximately 20 mouse clicks.* **Right:** *The final segmentation result is achieved.*

Segmentation using Flooding

Some anatomical structures appear to be "easy" to segment due to a high degree of contrast. An example could be a rib from the abdominal CT scan shown previously. However, range segmentation is not applicable since the entire bone structure will be selected. While it is quite easy to segment the rib by simply selected all the corresponding building blocks, this is somewhat tedious. In order to simplify this the program supports "flooding" of the building blocks.

The flooding process helps select the neighboring building blocks that are similar to the first selected. The process is illustrated in figure 2.5. The procedure is defined by a region merging process where neighboring building blocks are sorted by difference between mean grey value from the each building block to the starting block.

2.4   Applications of ∇ *Vision*

The program ∇ *Vision* has been applied for segmentation of the masseter (chewing muscle) in MR. This is a highly complicated task. The evaluation shows a speed up factor of two compared to manual segmentation, which is the only available alternative [Dam et al., 2000, Murakami, 1998].

The program has also been evaluated for segmentation of brain tumors [Letteboer et al., 2001, Letteboer et al., 2003]. The evaluation compared the segmentation program with manual outlining with respect to accuracy and reproducibility (measured as inter- and intra-observer variability). The results showed that the two segmentation methods are interchangeable in terms of accuracy. Furthermore, ∇ *Vision* had higher reproducibility than manual segmentation.

Finally, the interaction time for the MSWS program was on average one third of the time used for manual outlining. The evaluation data set is illustrated in figure 2.6 and an example segmentation is shown in figure 1.1 (chapter 1, page 3).
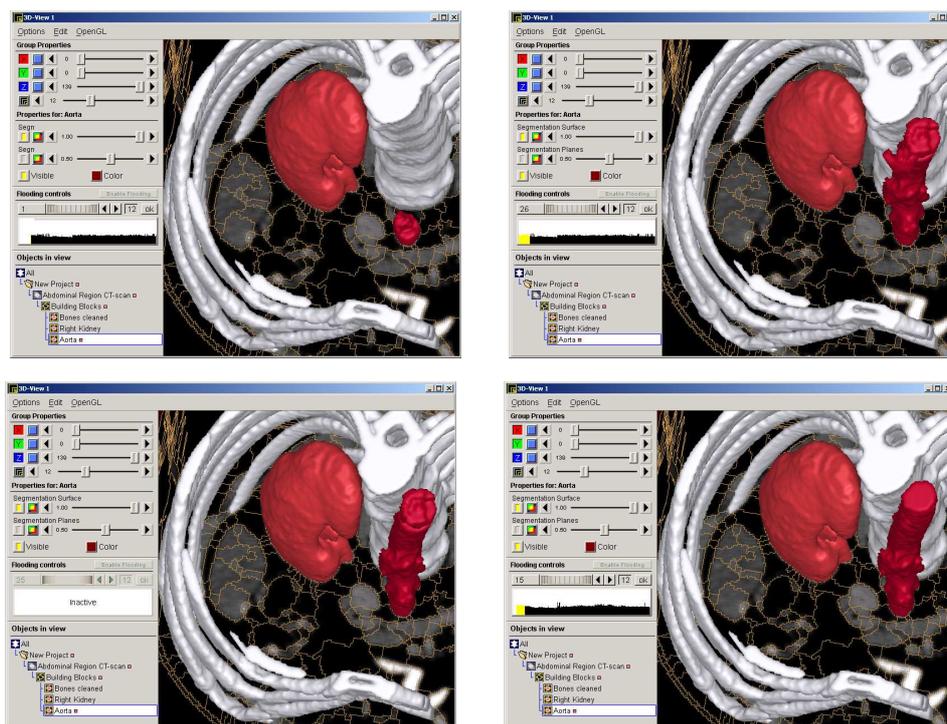
Figure 2.5: *Segmentation using flooding.* **Top left:** *Flooding is initiated on a building block at the lower part of the Aorta. This also selects the scale used for the entire flooding process.* **Top right:** *By scrolling the horizontal "wheel" in the Flooding controls window more building blocks are added until the flooding spills into undesirable segments.* **Bottom left:** *Rewinding the wheel removes the spilled segments. The partially segmented aorta is the result of one flooding operation (using 25 building blocks).* **Bottom right:** *A few selections add the remaining part at the top. Note that the bumps on the middle of the aorta are where smaller blood vessels branch.*
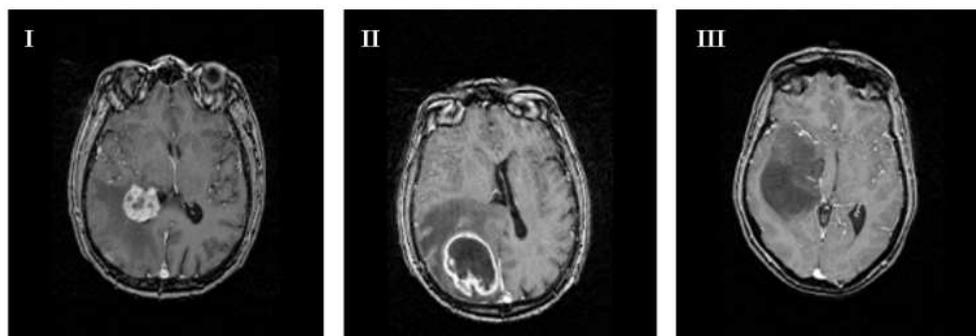
Figure 2.6: *The data set consists of 20 scans with brain tumors. These are divided into 7 full-enhancing tumors (type I), 6 ring-enhancing tumors (type II), and 7 non-enhancing tumors (type III). Each tumor has been manually segmented twice by three operators by slice-wise outlining.*

Recycling in this Chapter

The introduction and section 2.2 are edited parts of [Dam & Nielsen, 2000] and [Dam, 2000].

The segmentation program ∇*Vision* was developed by Andreas Thomsen, Erik Dam, and Ole Fogh Olsen [Johansen et al., 1999, Dam et al., 2003a, Olsen et al., 2000].

The brain segmentation figures in section 2.4 are from [Letteboer et al., 2003].

Chapter 3

Evaluating Non-linear Diffusion in MSWS

In this chapter we evaluate a broad class of non-linear diffusion schemes in the setting of the multi-scale watershed method presented in the previous chapter. Through a new scheme GAN, we show that diffusion similar to the non-linear Perona-Malik scheme is superior to the other evaluated diffusion schemes. This specialization of the segmentation method provides a speed up factor of two for the task of interactively segmenting gray and white matter of the brain.

3.1   Introduction

The Gaussian scale-space is a least committed scale-space. A non-linear scale-space commits itself to certain intensity variations through the non-linear function and to certain local edge shapes through the diffusion structure. This is formalised through the connection between energy minimization methods [Mumford & Shah, 1985] and non-linear diffusion in the biased non-linear diffusion [Nordstrom, 1990]. In this light, one may argue that the use of the non-linear diffusion schemes is the first step in commitment towards using prior shape and intensity knowledge as in the active contour and core-based segmentation methods [Cootes et al., 1995, Pizer et al., 1994, Cremers et al., 2002b].

In this work, we evaluate a number of non-linear diffusion schemes for the multi-scale watershed segmentation method in 2D: non-linear isotropic Perona-Malik [Perona & Malik, 1990], Weickert's non-linear image enhancing anisotropic schemes [Weickert, 1998a], and a generalization of these. In section 3.2.4, we argue that these schemes in a natural way span a space of diffusion schemes supporting segmentation.

The flavor of our work is close to the comparison of diffusion schemes for segmentation performed on the *hyper-stack* [Vincken, 1995, Koster, 1995, Niessen et al., 1997]. The major differences are that the hyper-stack is based on isophote linking and that it constitutes an automated segmentation algorithm.

Another related work is the evaluation of the use of non-linear diffusion as preprocessing before watershed segmentation [Weickert, 1998b]. In this work we specifically evaluate how the *deep structure* of the various scale-spaces support the segmentation.

In section 3.2 the evaluated diffusion schemes are presented. The evaluation method is outlined in section 3.3 with results in section 3.4.

3.2 Diffusion Schemes for Multi-scale Watershed Segmentation

The original multi-scale watershed segmentation method relies on the linear Gaussian scale-space to simplify the image. This simplification determines how the catchment basins group into gradually larger building blocks corresponding to image structures at a given scale.

The linear scale-space for an image $I(\vec{x})$ is described by the PDE

$$\frac{\partial L(\vec{x};t)}{\partial t} = \Delta L(\vec{x};t) = L_{ii}(\vec{x};t)$$

with the initial condition: $L(\vec{x};0) = I(\vec{x})$. The Laplace operator Δ is written using the Einsteins summation convention: $L_{ii} = L_{xx} + L_{yy}\dots$ The Gaussian convolution kernel with standard deviation $\sigma = \sqrt{2t}$ is the Green's function for the PDE.

The Linear Gaussian diffusion scheme (here denoted LG) has extremely nice theoretical properties [Lindeberg, 1994, Weickert, 1998a]. In particular, the causality property, the average gray level invariance property, and the fact that the image gets uniform intensity for scale tending to infinity ensures that the linear scale-space is applicable for the multi-scale watershed segmentation method. However, these properties do not ensure that it is an optimal diffusion scheme for the method. Specifically the isotropy property will tend to favor roundish objects.

In [Olsen, 1996, Olsen & Nielsen, 1997] the generic events for the gradient magnitude minima are derived. When the diffusion scheme is replaced by non-linear schemes, the analysis of the generic events for the watershed regions is no longer applicable. Some of these non-linear schemes have been analyzed [Damon, 1997]. However, from a practical viewpoint, the linking of the discrete scale levels can handle nearly any diffusion scheme with suitable simplification properties due to robust matching of regions [Dam, 2000]. The linking of watershed regions are done by linking a region a low scale to the region at the next, higher scale level with the largest area overlap. Thereby, the catastrophes are not detected and no information regarding the specific catastrophe type is applied in the linking procedure. Furthermore, no extrapolation and matching of singularity strings in scale-space in necessary.

3.2.1 Regularised Perona-Malik

The classical Perona-Malik diffusion scheme is designed to preserve edges during the diffusion [Perona & Malik, 1990]. The regularisation due to [Catté et al., 1992] is denoted RPM:

$$\frac{\partial L(\vec{x};t)}{\partial t} = div(\ p(|\nabla L_\sigma|^2)\ \nabla L\ ) \quad \text{where} \quad p(|\nabla L_\sigma|^2) = \frac{1}{1 + \frac{|\nabla L_\sigma|^2}{\lambda^2}} \tag{3.1}$$

The parameter λ is a threshold for the gradient magnitude required to make the scheme preserve an area (an edge). The σ determines the Gaussian regularisation scale at which the gradient ∇L_σ is evaluated. The divergence operator is defined $div = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \dots$

3.2.2 Anisotropic Non-linear Diffusion

Weickert [Weickert, 1998a] defines the *anisotropic non-linear diffusion* equation for a 2D image I by the PDE:

$$\frac{\partial L(\vec{x};t)}{\partial t} = div(\ D(J_\rho(\nabla L_\sigma))\ \nabla L\ ) \quad \text{where} \quad L(\vec{x};0) = I(\vec{x}) \tag{3.2}$$

The diffusion tensor $D \in C^\infty(R^{2\times2}, R^{2\times2})$ is assumed to be symmetric and uniform positive definite. The *structure tensor* J_ρ is evaluated at *integration scale* ρ, and the gradient ∇L_σ at *sampling scale* σ.

Throughout this dissertation we will ignore the structure tensor (implicitly setting it to identity) in the explored diffusion schemes.

The diffusion equation possesses the same simplification properties mentioned for linear Gaussian diffusion above (obviously not including the isotropy property) [Weickert, 1998a] that ensures that the diffusion schemes are applicable for the segmentation method.

For the following diffusion schemes, the diffusion tensor is defined in terms of the eigenvectors $\bar{v}_1 \parallel \nabla L_\sigma$, $\bar{v}_2 \perp \nabla L_\sigma$ and the corresponding eigenvalues λ_1 and λ_2. Furthermore, Weickert presents a diffusivity function w_m designed to preserve edges more aggressively than the Perona-Malik diffusivity function p.

$$w_m(|\nabla L_\sigma|^2) \quad = \quad \begin{cases} 1 & |\nabla L_\sigma| = 0 \\ 1 - exp\left(\frac{-C_m}{\left(\frac{|\nabla L_\sigma|^2}{\lambda}\right)^m}\right) & |\nabla L_\sigma| > 0 \end{cases} \tag{3.3}$$

Here m determines the aggressiveness of the diffusivity function, and C_m is derived from m such that the flux magnitude function $|\nabla L|\ w_m(|\nabla L_\sigma|^2)$ is increasing for $|\nabla L|^2 < \lambda$ and decreasing for $|\nabla L|^2 > \lambda$ [Weickert, 1998a, Dam, 2000]. In the following the diffusivity function w_m is used to define the eigenvalues λ_1 and λ_2.

Isotropic Non-linear Diffusion

Weickert [Weickert, 1998a] designs an isotropic non-linear diffusion scheme (here denoted IND) by the following eigenvalues $\lambda_1 = \lambda_2 = w_m(|\nabla L_\sigma|^2)$. For $m = 0.75$ this is qualitatively equivalent to Perona-Malik diffusion (this can be seen by comparing the respective diffusivity functions). Intuitively, this is an increasingly aggressive version of the Perona-Malik scheme for $m > 0.75$.

Edge Enhancing Diffusion

The anisotropic version is termed *edge enhanced diffusion* and defined by the eigenvalues $\lambda_1 = w_m(|\nabla L_\sigma|^2)$ and $\lambda_2 = 1$. The choice $m = 4$ (which implies $C_m = 3.31488$) is used with visually appealing results in [Weickert, 1998a]. Here, we also exploit $m = 2$ ($C_m = 2.33666$) and $m = 3$ ($C_m = 2.9183$) for the edge enhancing diffusion scheme.

The isotropic non-linear scheme enhances edges so aggressively that noise is preserved around edges for a long scale interval. The anisotropic schemes (here denoted EE2, EE3, and EE4 depending on the choice of m) remedy this by smoothing along the edges.

3.2.3 Maximally Anisotropic Diffusion

Inspired by Weickert's anisotropic diffusion equation we define the *maximally anisotropic diffusion* scheme by defining the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$. Thereby the diffusion is completely restricted across the potential edge while being allowed to flow along the isophotes. The diffusion is similar in spirit to *Mean curvature motion* — however the schemes are not identical.

Since one eigenvalue is zero, the diffusion tensor is not positive definite. Therefore we cannot assume that this scheme will fulfill the properties proven for the anisotropic non-linear diffusion equation (equation 3.2).

3.2.4 Generalized Anisotropic Non-linear Diffusion

The diffusion schemes previously presented are defined by the diffusivity functions in the gradient direction and the isophote direction. Figure 3.1 illustrates this "space of diffusion schemes". This inspires the new *Generalized Anisotropic Non-linear* diffusion scheme (denoted GAN) defined by the following diffusivity functions λ_1 and λ_2:

$$\begin{aligned} \lambda_1 &= w(m, \lambda, |\nabla L_\sigma|^2) \\ \lambda_2 &= \theta + (1 - \theta)\, \lambda_1 \end{aligned} \qquad (3.4)$$

The Weickert diffusivity function (equation 3.3) is written $w(m, \lambda, s^2)$ instead of $w_m(\lambda, s^2)$ since m is to be perceived as a regular parameter of the diffusivity function. The parameter θ determines the degree of anisotropy.

The GAN scheme is named *Generalized Anisotropic Non-linear diffusion* since it offers a straighforward generalization of the previously presented diffusion schemes. The choice of the parameters, in particular the aggressiveness parameter m and the anisotropy parameter θ, allows the scheme to cover the white area in figure 3.1. The specific schemes are realised by the following:

LG	Linear Gaussian diffusion is defined by $\lambda \to \infty$.
IND	Isotropic Non-linear Diffusion is achived for $\theta = 0$.
RPM	Regularised Perona-Malik scheme is approximated by $\theta = 0$ and $m = 0.75$.
EEx	The Edge Enhancing schemes EE2, EE3, and EE4 are defined by $\theta = 1$ and the corresponding m.
MAD	Maximal Anisotropic Diffusion is achieved for $\theta = 1$ and $\lambda \to 0$.

Figure 3.1: *Space of diffusion schemes. The gradient direction and isophote direction diffusivity functions determine the positions on the horizontal and vertical axes, respectively. The diffusion schemes with aggressive diffusivity functions are mapped closest to the lower, left corner. The gray area is populated by diffusion schemes that are not suited for segmentation-like purposes — they diffuse more across edges than along them. Thereby the catchment basins merge across possible object borders before merging inside the regions likely to correspond to the desired objects.*
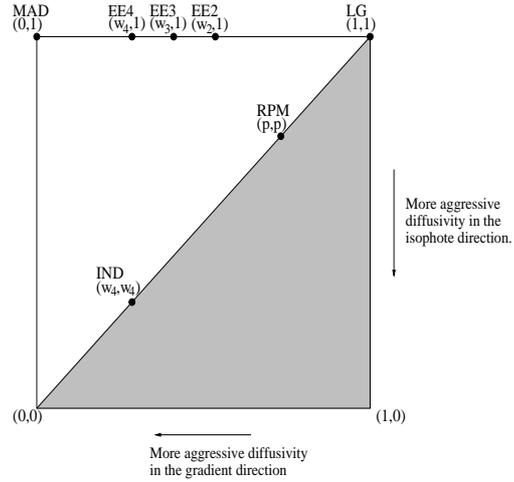
3.3   Evaluation

An obvious evaluation is to let clinicians test the segmentation method on real segmentation tasks with the different diffusion schemes. However, this is not objective and requires extensive work by the clinicians. The alternative is to measure the quality of the segmentations with respect to a "correct segmentation" — the *ground truth*.

The quality measure should be general, objective, and quantitative. For specific segmentation tasks, the quality measure could be defined in terms of specific features of the desired segmentations (shape, topology, etc.) or even ultimately by the ability to facilitate a correct diagnosis. However, for a general evaluation method, the measure must be simple and geometric. This allows area or boundary oriented measures such as the area overlap or the mean minimal boundary distance between the ground truth segmentation and the segmentation being evaluated.

Volume overlap is the simplest measure to compute for segmentations represented by pixels. However, even the normalized volume overlap is not an optimal measure since the measure is increasing with the compactness of the objects (the area to boundary ratio). Therefore the normalized volume overlap does not offer a good objective absolute evaluation of a segmentation. It does however allow fair relative evaluation of competing segmentations on the same object. Evaluation on a collection of objects can also impose problems if the compactness vary greatly across the collection. This would bias the importance of each object in the overall evaluation. If the objects in the ground truth collection are similar in size and compactness, this is not a problem. Therefore we choose the simpler quality measure based on normalized area overlap. The discussion is equivalent for 3D segmentation using volume, surface, and voxels as the basic representations.

In this chapter, ground truth segmentations of white and gray matter for both real and simulated MRI brain scans are used (figures 3.2 and 3.3). The real data is from the *Internet Brain Segmentation Repository* [ibs, 1999]. The simulated data is from the *BrainWeb* [Cocosco et al., 1997, Kwan et al., 1996, Collins et al., 1998].

17
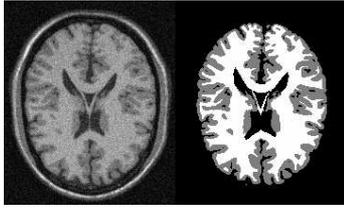
Figure 3.2: *Simulated T1 MR brain scan with ground truth white and gray matter. The volume is 181x217x181 with coronal slice thickness 1 mm, intensity nonuniformity level 20%, noise level 9%. From http://www.bic.mni.mcgill.ca/brainweb*
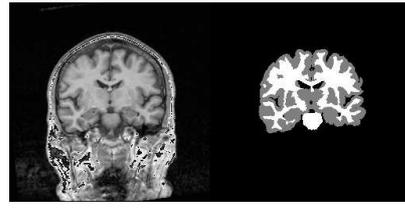
Figure 3.3: *Real T1 MR brain scan of a 55 year old male. The set contains 60 256x256 slices with slice thickness 3.0mm. The slices are coronal with a flip angle of 40 degrees. From http://neuro-www.mgh.harvard.edu/cma/ibsr*

The quality measure is defined via the relative error (misclassified pixels relative to number of pixels in ground truth object, where a pixel is uncorrectly segmented if it is included only in the segmentation or only in the ground truth). From each data source 5 slices are used for training and another 5 slices are used for evaluation.

The pixel-sized building blocks allow the user to reach an arbitrary segmentation — therefore any diffusion scheme allows perfect segmentation. The evaluation of the semi-automatic method measures the user effort required to reach a specific *quality threshold*. In an evaluation with clinicians, the user effort could be measured as the time needed. For this evaluation, the effort is naturally measured as the minimal number of basic user actions required. The canonical actions are *selection* and *deselection* of the building blocks.

In order to use the minimal number of user actions needed as a measure of the user effort, this minimal count must be established. This is done efficiently by algorithm 1 described in section 3.3.1. Once the user effort is quantifiable, this measure can be used to optimize the parameters for the non-linear schemes to minimize the user effort. This optimization method is described in section 3.3.2.

3.3.1 Optimal User Actions

The linking tree determines how the building blocks merge as scale increases. Each node corresponds to a single region at the given scale — and corresponds to a set of regions at the localization scale. The leaves in the tree are the single voxel/pixel building blocks. A simple heuristic for reaching the optimal segmentation is to start at the largest scale, and then visit all nodes in the linking tree in a top down order. At each node, the subtree is either marked as selected or deselected depending on whether the corresponding regions are mostly inside or outside the ground truth segmentation. During this traversal, actual user actions are only needed where the attribute changes for a node (if a region is selected as part of a large scale region, it need not be selected again at low scale). This simple top-down heuristic yields reasonable user actions that yield the ground truth segmentation — but it does not offer the optimal actions.

18

In algorithm 1 (on page 20) we define the set of possible actions as *Select*, *Deselect*, and *Ignore* in order to be able to perform an action at each node in the tree. Obviously only *Select* and *Deselect* correspond to user actions. Furthermore, we define the node states *In* and *Out* depending on whether the node is selected or deselected by its predecessors. Finally, we add an artificial root node that corresponds to the entire image. This node has state *Out* by definition.

The algorithm is a bottom up traversal of a linking tree. At each node the minimal action count is determined both if the node is assumed to have state *In* and *Out*. This is trivial at the leaf node, since a single action is needed only if the state and the actual position compared to the ground truth segmentation differs. Moving up through the linking both minimal actions counts are determined — assuming the node have *In* or *Out*. When the root node is reached the state is *Out* by definition, and the overall minimal action count is determined.

Note that we do not have to create the leaves explicitly. For each leaf-parent (corresponding to the regions at the localization scale level) the *In*-count and *Out*-count can simply be calculated from a count of the number of pixels inside and outside of the ground truth for each region.

Algorithm 1 determines the minimal count of actions it requires to achieve the perfect segmentation according to the ground truth segmentation. This number can obviously be established by exhaustive search of all combinations of the possible actions *Select*, *Deselect*, and *Ignore* at each node. However, the computational complexity of this approach is not appealing. It is quite simple to see that algorithm 1 provides the minimal processing cost by structural induction. At the leaf nodes the minimal processing is trivially performed. At an inner node, the minimal processing cost is obtained since the minimal of the two possibilities (corresponding to *Ignore* or *Select*/*Deselect*) is chosen.

Let p be the number of pixels/voxels and n be the number of inner nodes in the tree. The algorithm has a complexity of $O(p+n)$. Each inner node is inspected once from its parent. This is optimal since an algorithm has to visit every node in the tree (or convert the tree into some other data structure — which would require a visit at each node anyway).

In [Dam, 2000] an extended algorithm is presented that finds the minimal error given a limited number of user actions.

3.3.2   Optimization of Parameters

The diffusion schemes have a number of parameters that determine the performance for the given segmentation task in the multi-scale segmentation method. The quantitative measure for user effort allows automatic optimization of the parameters.

The optimization method is treated in more detail in chapter 5. Here we simply present the very naive double-scale gradient descent method in algorithm 2.

For additional details on the evaluation method, the algorithms for establishing the optimal combinations of building blocks, the optimization method, and the optimal parameter sets, see the technical report [Dam, 2000].

Algorithm 1 *Optimal User Actions*

Input: The localization scale watershed regions, the multi-scale watershed linking tree, and the ground truth segmentation.

Output: The minimal count of user action required to reach the ground truth segmentation.

1. For each leaf:

   - Determine whether the leaf pixel is inside or outside the ground truth.

   - Determine *In*-count (cost of reaching optimal segmentation if parent has state *In*):
     If the leaf is outside, it is necessary to *Deselect* it and the action count is one.
     If the leaf is inside, no action is necessary and the action count is zero.

   - Determine *Out*-count (cost of reaching optimal segmentation if parent has state *Out*):
     If the leaf is outside, no action is necessary and the action count is zero.
     If the leaf is inside, it is necessary to *Select* it and the action count is one.

   - Keep both of these conditional counts — denoted *In*-count and *Out*-count — as attributes for the leaf.

2. Visit each inner node in the tree in a bottom up order:
   Assume the node has m children. Let the *In*-counts for the children be denoted $I_1, \ldots, I_m$. Let the *Out*-counts be denoted $O_1, \ldots, O_m$.

   - Determine *In*-count (assuming the parent has state *In*):
     We can either choose to keep the default state *In* or choose to *Deselect* the node.
     If we keep the state, we can simply add the *In*-counts of the children in order to get an *In*-count I^I for the node. If we *Deselect* the node we get an *In*-count I^A by adding one to the sum of the *Out*-counts of the children. In order to get the optimal count *In*-count I we choose a smallest of these two possible counts:

   $$I = min\{I^I, I^A\} \quad \text{where} \quad I^I = \sum_{i=1}^{m} I_i \quad \text{and} \quad I^A = 1 + \sum_{i=1}^{m} O_i$$

     The superscripts in I^I and I^A are short for *Ignore* and *Action*.

   - Determine *Out*-count (assuming the parent has state *Out*):
     The *Out*-count O is achieved in the same manner:

   $$O = min\{O^I, O^A\} \quad \text{where} \quad O^A = 1 + \sum_{i=1}^{m} I_i \quad \text{and} \quad O^I = \sum_{i=1}^{m} O_i$$

3. At the root:

   - Since the root state is *Out* by definition, the optimal action count is the *Out*-count for the root node.

Algorithm 2 *Optimization of Diffusion Scheme Parameters*

Input: The collection of image data with corresponding ground truth segmentations.
Output: The optimal parameters for the given diffusion scheme.

1. Define inital value and suitable step values for each parameter.

2. Evaluate the performance of the initial parameter set as the average for the entire collection of a quantifiable measure. This could be the minimal number of user action required to reach the ground truth segmentation or the minimal error given 50 user actions.

3. Repeat until no more improvement is found:

   - Inspect each parameter in turn.
   - Evaluate the performance of the parameter sets where the current parameter value is added and subtracted three times the corresponding parameter step (as defined in step 1).
   - If either of the added or subtracted parameter values provide improvement, use this parameter set as the current.
   - Try next parameter.

4. Repeat the previous step where single parameter steps are attempted.

5. The resulting parameters are denoted the optimal parameters.

3.4   Results

The performance is illustrated by the error as a function of the number of actions for each scheme (with parameters optimized to the specific data set). This is put into perspective by the performance of a *Quad tree* linking scheme [Samet, 1984] (here denoted QT). The number of quad tree blocks required has a close relation to the *box-counting dimension* of the ground truth objects — also denoted the *Hausdorff dimension* [Ott, 1993]. Furthermore, the building blocks of the quad tree are not adapted to the geometry of the image. Thereby the performance of the quad tree linking gives a frame of reference for the performances of the diffusion schemes defined in terms of the complexity of the ground truth objects.

Figures 3.4 and 3.5 display the performance on the simulated and the real data, respectively. The best of the existing schemes is the regularised Perona-Malik scheme. The new GAN scheme is slightly better than this. However, the graphs do not give a clear notion of the quantitative differences in performance. Figure 3.6 delivers the desired relative performance indicator on average for all data sets. The relative performance is determined both for the training set used for optimization of the parameters for the diffusion schemes and for an independent data set.
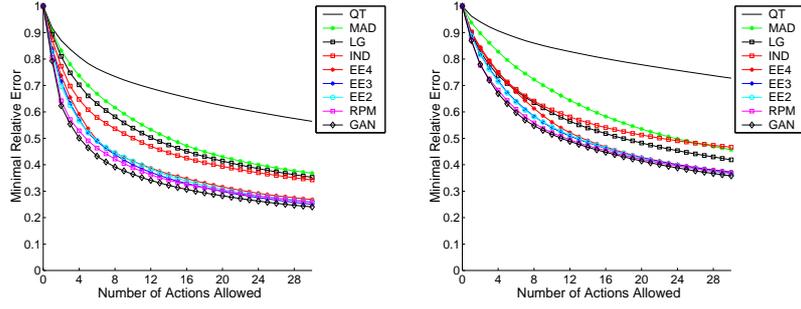
Figure 3.4: *Evaluation on simulated data from figure 3.2. **Left**: white matter. **Right**: gray matter. Slices 60, 80, 100, 120, and 140 from the data set are used.*
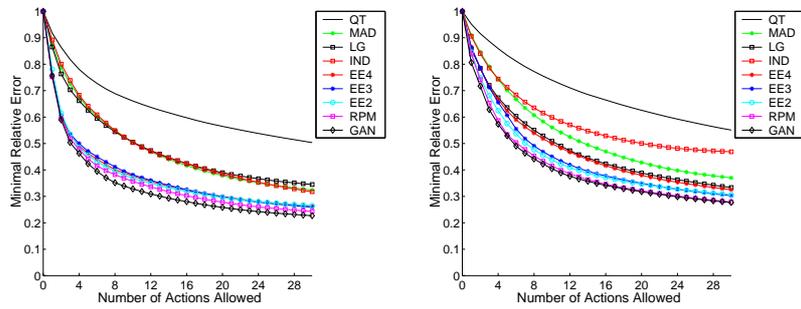
Figure 3.5: *Evaluation on real data from figure 3.3. **Left**: white matter. **Right**: gray matter. Slices 10, 20, 30, 40, and 50 from the data set are used.*
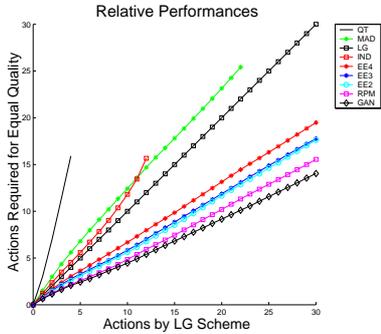
Scheme	Performance	
	Training	Evaluation
QT	3.83	3.34
MAD	1.18	1.19
LG	1.00	1.00
IND	1.20	1.17
EE4	0.66	0.66
EE3	0.59	0.63
EE2	0.58	0.62
RPM	0.51	0.51
GAN	0.46	0.49

Figure 3.6: *Average performances with the Gaussian scheme as reference. For a given number of actions the performance of the LG scheme is noted — for each scheme is measured the actions required to get equal quality. The incline of a curve determines the number of actions required to obtain a given segmentation quality relative to the performance of the Gaussian scheme. This performance indicator is displayed in the table for each diffusion scheme. The new GAN scheme requires less than half as many actions compared to the Gaussian scheme for both the training set and an independent data set.*

The results show that Perona-Malik is superior among the existing diffusion schemes — somewhat surprising, the anisotropic schemes show inferior performance. Furthermore, the optimal parameter sets for the GAN scheme reveal the optimal degrees of anisotropy and edge preservation aggressiveness. The anisotropy parameter θ is close to zero for all data sets (0.06, 0.12, 0.0, and 0.0). The improved performance compared to RPM is due to slightly increased aggressiveness in the diffusivity function (the parameter m is 1.4, 1.0, 1.1, and 1.0 compared to the approximate value of 0.75 for RPM).

In the technical report [Dam, 2000], a number of other results are documented as well:

- With a tolerance area 1 pixel wide around the borders of the ground truth segments, 97% of the ground truth pixels can be segmented without the use of pixel-sized building blocks.

- The simple top-down heuristic for user actions presented in section 3.3.1 requires around one third more user actions than the optimal actions used by the evaluation method for all diffusion schemes.

- The diffusion schemes require up to 30 scale levels for the discrete linking to be sufficiently closely discretised.

- Similar results are measured for higher number of actions, from simulated data with less noise and from real data of a schizophrenic brain.

It is worth noting that the improved performance between RPM and GAN is achieved at the expense of introducing two additional parameters (m and θ). Additional parameters increase the complexity of the optimization process and thereby require a larger training data set in order to attain a performance that generalizes to use on an evaluation data set (or eventually real data). Therefore, in situations with limited training data it can actually be advantageous to stick with the simpler RPM scheme.

3.5   Conclusion

We present a generalized anisotropic diffusion scheme GAN capturing many known diffusion schemes. Specializing this for interactive multi-scale watershed segmentation of white/gray matter in T1-weighted MR slices of the brain shows that diffusion similar to regularised Perona-Malik is superior to the other diffusion schemes. Furthermore, the aggressiveness in the diffusion cut-off is more important than the degree of anisotropy.

The best among the tested diffusion schemes yields a decrease in interaction time with more than a factor two compared to linear Gaussian scale-space. Our expectation is that the gain is even higher in a 3D implementation — this is explored further in chapters 4 and 5. The conclusion is linked to the segmentation task of white/gray matter in the brain. For cases like vessels or abdominal organs, other diffusion schemes may be optimal.

Acknowledgements

Recycling in this Chapter

This main results in this chapter are published in [Dam & Nielsen, 2000] and in part also presented in [Dam & ter Haar Romeny, 2003].

Chapter 4

Evaluating Non-linear Diffusion in MSWS in 3D

The previous chapter on segmentation of brain tissue from 2D slices shows that the watershed building blocks that forms the basis for the interaction method improve significantly through non-linear diffusion.

Furthermore, the results from [Letteboer et al., 2001, Letteboer et al., 2003] briefly introduced in section 2.4 show that the $\nabla Vision$ implementation of the MSWS method is effective for segmenting brain tumors in 3D.

These two results inspire the obvious attempt of extending the non-linear scheme GAN to 3D in order to improve the segmentation of brain tumors.

In addition to the direct implications for segmentation of brain tumors, it is certainly also theoretically interesting whether the results on specialization through the use of non-linear diffusion carry over from 2D to 3D.

4.1 Non-linear Diffusion in 3D for Interactive Segmentation of Brain Tissue

In this chapter we address the task of segmentation for anatomical structures where the use of statistical shape models is unfeasible due to the large variability in shape and appearance of the desired objects.

Often pathologies lead to this situation — here we investigate segmentation of brain tumors. For some types of brain tumors, automatic methods show great potential (one such example combines tissue statistics with an asymmetry measure to detect the tumors [Lorenzen et al., 2001]).

However, in many cases automatic segmentation is unattainable and interactive methods must be used. We investigate the interactive segmentation method based on multi-scale watersheds.

The method has shown good performance on segmentation of the mandible [Dam et al., 2000] and of brain tumors [Letteboer et al., 2001]. In both cases the quality of the segmentations are as good as manual outlining while the interaction time is greatly reduced.

In demonstrated in chapter 3, the performance of this method can be improved by using non-linear diffusion as basis for the multi-scale method [Dam & Nielsen, 2000]. However, this previous work only address segmentation in 2D.

Here we investigate the use of non-linear diffusion in 3D for the multi-scale watershed method. Furthermore, we qualitatively categorize the anatomical structures where non-linear diffusion can be expected to improve the performance of the segmentation method.

The performance is evaluated on three types of brain tumors (non-enhancing, ring-enhancing, and full-enhancing) and on white matter brain tissue. The results confirm the categories of anatomical structures that can benefit from the use on non-linear diffusion. Unfortunately, this implies that the performance is only moderately improved for segmentation of the brain tumors. Surprisingly, the evaluation shows that a more geometrically complicated anatomical objects such as white matter brain tissue also only allow moderate improvement.

We briefly re-introduce the multi-scale watershed segmentation method using segmentation of a brain tumors used for the evaluation as example. Then the use of non-linear diffusion in the method is described with the definition of the class of the evaluated non-linear diffusion schemes in 3D. Finally we show results of the evaluation on brain tumors and white matter tissue.

4.2 Interactive Segmentation of Brain Tumors

The multi-scale watershed segmentation method described above has been implemented in the program $\nabla Vision$ [Dam et al., 2003a]. Evaluation of this implementation has shown good performance for segmentation of brain tumors [Letteboer et al., 2001].

The evaluation compared the segmentation program with manual outlining with respect to accuracy and reproducibility (measured as inter- and intra-observer variability). The results showed that the two segmentation methods are interchangeable in terms of accuracy. Furthermore, the interactive MSWS program had higher reproducibility than manual segmentation.

Finally, the interaction time for the MSWS program was on average one third of the time used for manual outlining. The evaluation data set is illustrated in figure 4.1 and an example segmentation in figure 4.2.

In the following, we evaluate whether the use of non-linear diffusion in the MSWS method improves these results.
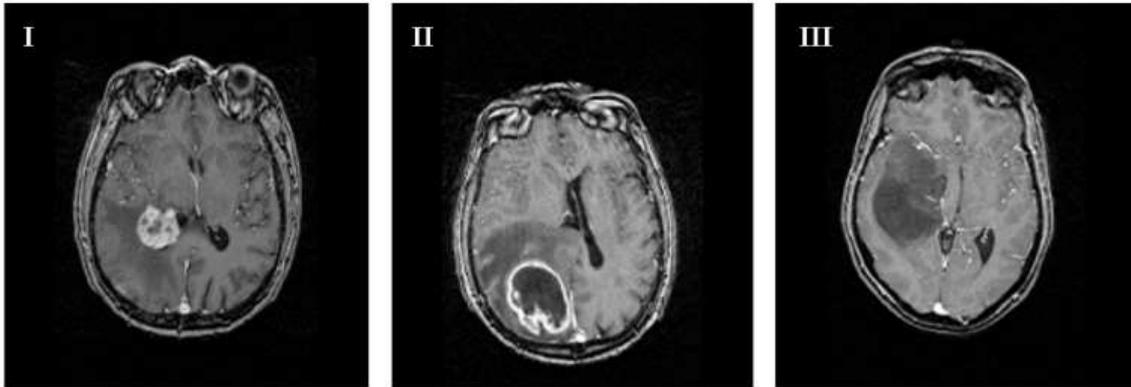
Figure 4.1: *The data set consists of 20 scans with brain tumors. These are divided into 7 full-enhancing tumors (type I), 6 ring-enhancing tumors (type II), and 7 non-enhancing tumors (type III). Each tumor has been manually segmented twice by three operators by slice-wise out-lining. The MR scans are aquired with T1-weighted post-contrast acquisition, scanned with a slice thickness of 2.2 mm and reconstructed at 1.1 mm. The datasets consist of 120 to 150 slices of 256 x 256 voxels with voxels size is 1.0 x 1.0 x 1.1 mm.*
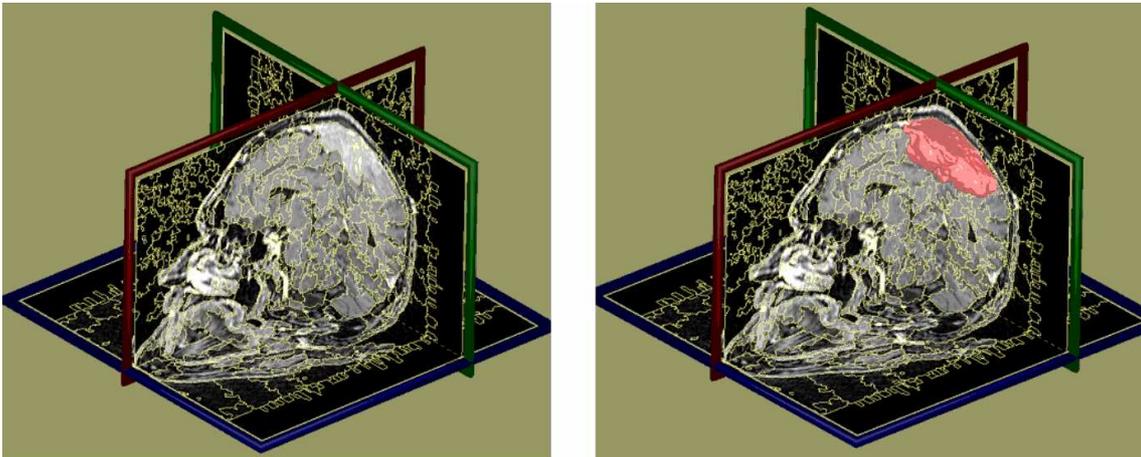
Figure 4.2: *Segmentation of a brain tumor using the program ∇Vision [Dam et al., 2003a].* **Left**: *A visualization of the scan.* **Right**: *The segmented tumor. Repeated from figure 1.1.*

Slice from	Groundtruth	Segments from	Segments from
BrainWeb scan	white matter	linear diffusion	GAN diffusion
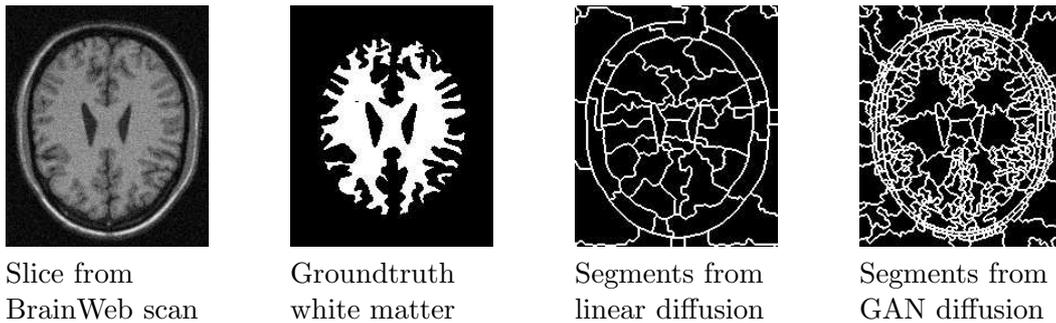
Figure 4.3: *Segmentation of white matter in 2D slice: Linear diffusion makes the building blocks merge across brain structures before reasonably large segments are formed. The use of non-linear diffusion allows the building blocks to grow within the tissue boundaries. The implication is that where 18 action are required to select 70% of the white matter using building block created using linear diffusion, only 5 are necessary with GAN diffusion. For 80% the action counts are 39 and 11, respectively.*

4.3   Non-linear Diffusion in MSWS

The original MSWS method relies on linear Gaussian scale-space to simplify the image. This simplification determines how the watershed regions group into gradually larger building blocks corresponding to image structures at a given scale.

4.3.1   Revisiting Segmentation of White Matter in 2D Slices

In chapter 3 (and [Dam & Nielsen, 2000]) the use of non-linear diffusion in MSWS is explored for the task of segmenting grey and white matter from 2D slices from the BrainWeb collection [Collins et al., 1998]. Figure 4.3 illustrates how the building blocks resulting from non-linear diffusion are better suited to the task at hand.

The use of non-linear diffusion is evaluated based on a count of the minimal number of selections and deselections of building blocks in the segmentation. The parameters for the non-linear schemes are determined such that this count is minimized. The evaluation results are presented by normalizing the performance (of the building blocks resulting from a non-linear scheme) with respect to the performance of linear diffusion. A good descriptor is then the ratio of actions required compared to linear diffusion.

In chapter 3 the evaluation is based on both real and artificial brain scans and with the task of segmenting both white and grey matter. Here we include only segmentation of white matter and only use data from the BrainWeb collection. The evaluation data set is 9 slices evenly distributed from this artificial brain scan. For RPM and GAN are we use the parameter sets resulting from the optimization in chapter 3.

The table below shows that for each action used on building blocks resulting from linear diffusion, it is on average only necessary to use 0.42 or 0.32 when the building blocks resulted from Perona-Malik or GAN (where the optimal parameters give very little anisotropy, 0.06, and relatively low aggressiveness at 1.2 — see section 4.3.2 for definition of these terms).

Diffusion Scheme	Ratio	Std. Dev.
Linear Gaussian	1.00	0.00
Regularized Perona-Malik	0.42	0.15
Generalized Anisotropic Non-linear	0.32	0.07

These number are relative and does not state whether any of the schemes produce buildings blocks that are actually usable for the task at hand. Furthermore, they only compare the performance of the first 50 actions for each diffusion scheme. So the evaluation from [Dam & Nielsen, 2000] is only to be considered a feasibility study on the use of non-linear diffusion in the method.

4.3.2   Non-linear Diffusion in 3D

The *Generalized Anistropic Non-linear* diffusion scheme (GAN) is presented in chapter 3 (and introduced in [Dam & Nielsen, 2000]) as a generalization of a number of prominent 2D diffusion schemes — among these linear Gaussian diffusion, the classical Perona-Malik scheme [Perona & Malik, 1990], and Weickert's edge enhancing diffusion scheme [Weickert, 1998a].

Here we define the extension of GAN to 3D. It is based on the anisotropic diffusion equation [Weickert, 1998a] in equation 4.1. The diffusion on the image U is defined by the eigenvalues λ_1, λ_2, and λ_3 for the diffusion tensor D — where the corresponding eigenvectors $\bar{v}_i$ are defined from ∇U_σ (the gradient at scale σ) such that $\bar{v}_1 \parallel \nabla U_\sigma$ and $\bar{v}_2, \bar{v}_3$ are chosen as two orthogonal vectors both orthogonal to v_1. The diffusivity function w gives the eigenvalues.

$$\frac{\partial U(t, \vec{x})}{\partial t} = div(\ D(\nabla U_\sigma)\ \nabla U\ ) \quad \text{where} \quad U(0, \vec{x}) = I(\vec{x}) \tag{4.1}$$

$$\lambda_1 = w(m, \lambda, |\nabla U_\sigma|^2) \tag{4.2}$$

$$\lambda_2 = \theta + (1 - \theta)\ \lambda_1$$

$$\lambda_3 = \theta + (1 - \theta)\ \lambda_1$$

$$w(m, \lambda, |\nabla U_\sigma|^2) = \begin{cases} 1 & |\nabla U_\sigma| = 0 \\ 1 - exp\left(\frac{-C_m}{\left(\frac{|\nabla U_\sigma|^2}{\lambda}\right)^m}\right) & |\nabla U_\sigma| > 0 \end{cases}$$

The parameter θ determines the degree of anisotropy and m the aggressiveness with which the edges are preserved (where edges are defined by the soft threshold λ, and C_m is calculated from m such that w is increasing for $|\nabla U_\sigma|^2 < \lambda$ and decreasing for $|\nabla U_\sigma|^2 > \lambda$). Weickert's edge enhancing diffusion scheme is a special case ($\theta = 1$ and $m = 4$), the regularized Perona-Malik scheme can be approximated ($\theta = 0$ and $m = 0.75$) [Dam & Nielsen, 2000], and for $\lambda \to \infty$ GAN becomes linear Gaussian diffusion.

In equation 4.2, the eigenvalues λ_2 and λ_3 are defined as being equal. This implies equivalent local diffusion in all directions perpendicular to the gradient. Another possibility would be to choose the two corresponding eigenvectors according to the main isophote curvature directions and then limit diffusion in the direction with large curvature. Preliminary experiments with this approach show high sensitivity to noise (similar to the behavior for the CED scheme presented in section 6.2.2 as mentioned in chapter 8).

It should further be noted that this choice does not imply isotropic diffusion perpendicular to the gradient — just like the Perona-Malik scheme does not offer isotropic diffusion even though the diffusion tensor has two equal eigenvalues (see section 6.3.2 for an illustration of this).

The diffusion scheme is implemented using a simple explicit numerical discretization scheme. This imposes severe time step restrictions in order to ensure stability. In the 2D version the time step is required to be below 0.25 [Weickert, 1998a]. In the 3D version above the time steps should be below 0.16 (in 2D $\frac{1}{4}$, in 3D $\frac{1}{6}$).

4.3.3 When does Non-linear Diffusion Improve Performance?

The potential gain via the use of non-linear diffusion is due to the ability to incorporate task-specific prior knowledge into the diffusion process — typically by optimizing parameters to the task at hand. In the multi-scale watershed method this changes the rate of region merging as scale increases in different area of the data. Thereby regions inside the desired anatomical structures can be allowed to grow larger before the boundaries are blurred away and the regions merge with regions outside the object.

We can qualitatively describe the categories of objects that will benefit from the use of non-linear diffusion in this framework. The categories are cascading such that objects have higher potential the more requirements they fulfill:

1. The boundaries of the object are quantifiable. This allows prevention of early diffusion across the boundary.

2. The interior of the object has few places which have the same properties that define the boundaries. This allows full diffusion inside the object and thereby large building blocks can be formed.

3. The geometry of the object is non-sphere-like. Linear diffusion favours sphere-like object — potatoes — due to the isotropy. Non-linear diffusion can allow regions to merge and form more complicated shapes.

4.4 Evaluating GAN for Segmentation of Brain Tumors

The three types of tumors (see figure 4.1) fall into two of the categories above:

Ring-enhanced: The boundaries are quantifiable. However, due to ring effect, the object has the inner edge of the ring inside the object, so this tumor class fall in category 1.

Non-enhanced & Full-enhanced: The boundaries are quantifiable. The inside has no edges, and the tumor class fall in category 2.

The tumors are generally "potato-shaped" so none of them qualify for category 3. Therefore we can expect little improvement for ring-enhanced while non- and full-enhanced should allow some improvement.

The evaluation for the white matter tissue in 2D reviewed above measures the efficiency of the first 50 actions — normalized with respect to the performance for linear diffusion building blocks. Here we evaluate both accuracy and efficiency of the optimally selected building blocks compared to the observers segmentation. The accuracy measure is the error measured by the normalized volume overlap also used in section 3.3 (so high accuracy is a low number). The efficiency is the count of actions needed to reach the optimal accuracy. The algorithm for determining the optimal building block actions is described in section 3.3.1.

The evaluation results are as summarized in the table.

Tumor class	Accuracy		Efficiency	
	Linear	GAN	Linear	GAN
Full-enhancing	0.12 ± 0.03	0.12 ± 0.04	45 ± 31	37 ± 26
Non-enhancing	0.16 ± 0.05	0.16 ± 0.05	76 ± 67	62 ± 59
Ring-enhancing	0.14 ± 0.05	0.14 ± 0.04	57 ± 57	61 ± 57

The evaluation number for each class is average and standard deviation of the results for all tumors and all observers (i.e. full-enhancing covers 6 manual segmentations on each of 7 tumors). The results show that no improvement is gained for ring-enhancing while the reduction is approximately 18% for non- and full-enhancing.

The problem with the ring-enhancing tumors is the inner edge of the ring boundary profile. Therefore it could be advantageous to simply segment the outside instead. However, this is exactly what the optimal actions selected by algorithm 1 will do — if it is actually advantageous.

It should be stressed that the same 20 tumor data sets are used for training and evaluation — thereby potentially allowing overfitting to the data. Since only 6 or 7 data sets from each class are avaliable, we decided not to split the data set into a training and an evaluation collection. The parameters used for linear and GAN diffusion are obtained through the heuristic gradient descent method described in section 3.3.2. Due to the computation time needed to generate scale-spaces for each of the tumors in order to evaluate a parameter set, a proper optimization method requires more computational resources than were available. Thereby the optimization is fragile with respect to local minima — as seen by the fact that GAN performs worse than linear diffusion for ring-enhancing tumors even though linear diffusion is a special case for GAN.

4.4.1 Majority Standard Segmentations

The results above are simply averages of the individual results for evaluation versus each observer on each tumor. An evaluation method that better captures the variability in the observers segmentation is by evaluating versus a "majority standard" segmentation.

The majority standard segmentation is obtained by letting the observers "vote" on the result for each voxel. The voxel is then considered inside or outside depending on the majority. In the case where the vote is a tie the voxel is allowed to be either. Evaluating against the majority standard gives:

Tumor class	Accuracy		Efficiency	
	Linear	GAN	Linear	GAN
Full-enhancing	0.08 ± 0.03	0.08 ± 0.04	35 ± 24	29 ± 20
Non-enhancing	0.10 ± 0.03	0.10 ± 0.03	64 ± 63	54 ± 60
Ring-enhancing	0.09 ± 0.03	0.09 ± 0.04	48 ± 56	52 ± 56

The errors and the actions needed are reduced for all tumor classes when evaluating against the majority standard. The overall picture of no improvement for ring-enhancing tumor, and moderate improvement for the other tumor classes is the same as above.

4.5 Evaluating GAN for White Matter Segmentation

The results above are in line with the categories given in section 4.4: ring-enhancing tumors allow no improvement while the other allow moderate improvement. In order to confirm category 3, we evaluate the use of non-linear diffusion for segmentation of white matter brain tissue. The geometry is highly complicated, and thereby the potential for improvement should be larger.

For the evaluation we have used the same brain scan from BrainWeb [Collins et al., 1998] as seen in figure 4.3. In order to reduce computation resources required, the evaluation is performed on half the volume only (split along the center sagittal slice). The results are:

Tissue	Accuracy		Efficiency	
	Linear	GAN	Linear	GAN
White matter	0.18	0.17	1510	1118

The action count can be reduced by 26% through the use of non-linear diffusion. The error is even reduced slightly. This improvement in performance is larger than for the brain tumors but admittedly not quite as large as expected.

4.6 Conclusion

We evaluate the use of GAN diffusion in 3D for multi-scale watershed segmentation. The performance of GAN compared to linear diffusion depends on the appearance of the anatomical structures to be segmented. We propose a categorisation and evaluate a number of significantly different anatomical structures in order to investigate each category.

The results confirm the categorisation. The ring-enhancing tumors show no improvement since the property that defines the boundary of the tumors is also widely present inside the tumors. The non-enhancing and full-enhancing tumors allow moderate improvement. These objects have quantifiable boundaries and interiors different from their boundary. The reduction in actions needed is 18%. White matter brain tissue is an example of the last category with complicated geometry. The reduction in actions needed is here 26%.

While confirming the categorisation, the results are not nearly as good as expected from the previous work in 2D. While reducing the need for the time-consuming 3D implementation of GAN in the preprocessing step this is disappointing.

In the next chapter we address whether there is a fundamental problem with exploring GAN (or similar schemes) in 3D, whether the brain tumors are simply a special case, or whether the method is flawed.

As stated in the preface, the reevaluation is kept in a separate chapter in order to keep this chapter aligned with the publication.

Recycling in this Chapter

This chapter is an edited version of [Dam & Letteboer, 2003].

Chapter 5

Reevaluating Non-linear Diffusion in MSWS in 3D

Chapter 3 shows how the use on non-linear diffusion in 2D multi-scale watershed segmentation allows a significant reduction of the user interaction needed. This is evaluated for segmentation of white matter and grey matter from slices of brain scans. In chapter 4 this approach is extended to 3D. However, the results are less encouraging since the reduction in user interactions apparently is less significant in 3D.

The explanation for this is to be found in or more of the following possible causes:

- The object boundaries are harder to quantify consistently in 3D than in 2D.

- The geometry is more complicated in 3D making it harder for the non-linear schemes to excel compared to linear diffusion.

- The specific segmentation tasks evaluated simply does not offer more improvement through the use of non-linear diffusion.

- The optimization process that seeks the optimal parameters for the non-linear scheme reaches a local minimum.

- There is some unknown error in the implementation.

In the following the optimization method is scrutinized in order to reveal whether it actually determines the desired optimal parameters. This results in the presentation of a more advanced optimization method.

With the new optimization method, the basic difference between 2D and 3D in this MSWS setting is then investigated.

5.1 Improving the Optimization Method

Any optimization method that seeks the optimal parameters for some objective function will need to evaluate the function many times. Either directly or indirectly through the evaluation of the derivatives of this function. With discrete functions the derivatives are approximated using finite differences — again requiring evaluations of the function.

The non-linear diffusion scheme GAN (equation 4.2, page 29) has 7 parameters: starting scale, end scale, number of scale levels, regularization scale, soft gradient threshold with corresponding aggressivenes, and degree of anisotropy.

The evaluation objective functions that need to be optimized are first the *Accuracy* and then the *Efficiency*:

Accuracy:
> The best possible segmentation accuracy given a set of multi-scale building blocks. This is defined to be without the use of the articifical level with single-voxel building blocks. This is defined as the relative volume overlap between the best segmentation and the ground truth. Only the lowest scale level (the localization level) has influence on the accuracy. Therefore the accuracy for a given parameters set can be relatively quickly evaluated.

Efficiency:
> The minimal number of actions needed to obtain a segmentation with a given accuracy. Different requirements for the sufficient accuracy will lead to slightly different optimal parameter sets. However, the parameters are always optimized to give optimal performance for obtaining the optimal accuracy. In order to evaluate a given parameter set the entire scale-space of building blocks needs to be calculated making this quite time-consuming. This is particularly problematic when the evaluation is based on multiple images and the performance is given by the average of the individual cases.

In order to obtain optimal accuracy and efficiency simultaneously the optimzation is performed in two steps. First the parameters are optimized for accuracy. This optimal parameter set is then fixed for the first level in the gradient magnitude watershed linking scale-space. A new parameter set is then used for the remaining levels in the watershed scale-space. This new parameter set is then optimized for efficiency. Thereby the scheme actually has 14 parameters, but they are optimized independently in groups of 7.

In the following, the two alternative methods for obtaining optimal parameter sets are presented. Since optimization methods are not a central part of this dissertation the presentation focuses on the overview.

5.1.1 Naive Descent Algorithm

The rationale behind the original optimization method used is to avoid too many evaluations of the efficiency function due to the high cost of generating the multi-scale building blocks.

This lead to the *naive descent* algorithm. Starting from a suitable inital parameter set each parameter is tested in turn for values $\pm 3 step_i$ where $step_i$ is some suitable epsilon step for the i'th parameters. The best among the alternative values for the i'th parameter is then chosen as the new value. When none of the parameters offer improvement a new round is performed with test values $\pm step_i$. This simple double-scale descent algorithm is used to generate the optimal parameter sets in 2D in chapter 3 and in 3D in chapter 4.

5.1.2 Multi-scale Conjugate Gradient Descent Algorithm

Due to the very local nature of the optimization steps in the naive descent algorithm it is likely to be sensitive to local minima in the objective function.

As an alternative to this naive algorithm a more advanced algorithm is used as well. This is a multi-scale version of the *conjugate gradient descent* algorithm (from Numerical Recipies [Press et al., 1999]). The main difference between the algorithm used and the original is the addition of an outer multi-scale layer. This applies the algorithm in a coarse to fine approach where the gradient that defines the search for the minimum are evaluated at large scale first and then at smaller scales as the optimization ceases to improve[1]. To be specific, 4 levels are used where the scale is divided by 4 between each level. This appears to be suitable.

5.1.3 Regularization and Illegal Parameter Values

Figure 5.1 illustrates how the naive optimization algorithm fails to reach optimal parameters for the GAN scheme for segmenting white matter in 3D in chapter 4. The plots show that the large scale steps work reasonably well, but the small scale steps are simply a mess.

The parameter plots show exactly the behavior that is problematic for the naive optimization algorithm, namely the presence of a plethora of local minima.

The more advanced multi-scale conjugate descent algorithm is less sensitive to local minima. However, the non-smooth objective function potentially causes problems for this algorithm. Furthermore, the illegal parameter values (i.e. negative scale values or the edge preservation aggressiveness m in equation 4.2 below 0.5) that are simply implemented as giving a very high value in the objective function, are also problematic. Gradients that are evaluated using these artificially high values cause the gradient to push away from the bad values instead of pulling towards the good.

---

[1]The multi-scale extension of the conjugate gradient descent algorithm was developed in cooperation with Thomas Fletcher at UNC, Chapel Hill. The original purpose was for optimization in the m-rep shape modeling setting in chapter 9.
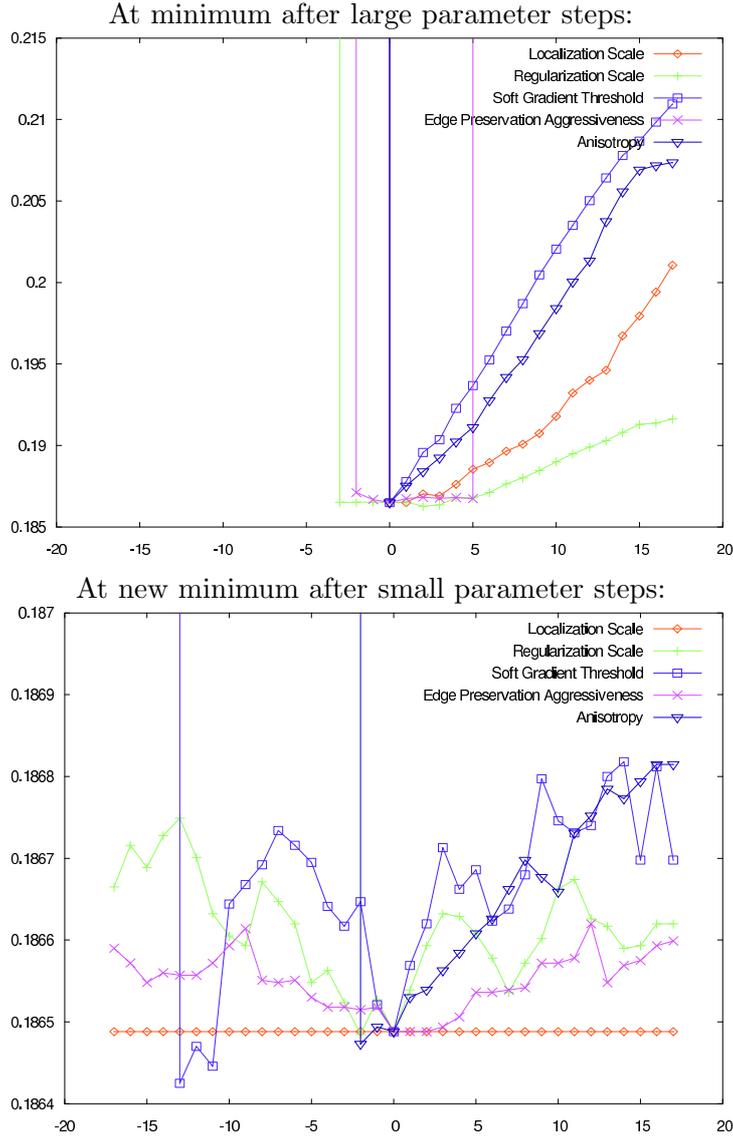
Figure 5.1: *Optimization of parameters for GAN using the naive double-scale optimization algo-rithm. The aim here is to minimize error (maximizing accuracy) when segmenting white matter tissue in a single half brain scan. A minimal error around 0.1865 is reached. The plots illustrate the effect on the error of varying the parameters individually from a reached minimum — each curve is the effect of changing a single paremeter independently of the others. The number of parameter steps away from the minimum are on the abscissa axis.* **Top***: The situation at the minimum reached after the large scale steps. From this minimum, the errors values are plotted for up to ± 17 large scale steps for each parameter. The parameters are respectively scale, regu-larization scale, soft gradient threshold, gradient threshold aggressiveness, and global anisotropy (see section 4.3.2). Most curves behave nicely except the regularization scale that reveals that only a local minima is reached. Only a single data set is used.* **Bottom***: From the large scale minimum, the optimization is continued using smaller steps (one third the large scale steps). The plots show the situation at the minimum that the small scale steps reach. All curves for the individual parameter are littered with local minima.*

The problem with the non-smooth objective function can be handled through regularization of the function. Evaluating the gradient at high scale offers a simple means of regularization but does not effectively remove the many local minima. Another simple regularization method is to average the performance over a number of training data sets. Since this is needed for robust determination of the optimal parameter set anyway, this is the simplest and most obvious regularization.

A reparameterization could be used to get rid of the illegal parameter values. For instance the positive-only scale values could be mapped to cover the entire real axis through an exponential mapping.

This reparameterization approach is not pursued in this dissertation. In the following, regularization through averaging is investigated. This simple regularization could very well explain why the work in chapter 3 on segmentation of white matter in 2D slices from brain scan gives good results whereas the work on 3D scans in chapter 4 shows poor performance. Instead of being related to dimensionality, the explanation can simply be that the work in 2D is performed on many slices whereas the work on white matter segmentation in 3D is optimized on a single brain scan (actually only half a brain split along the center sagittal slice).

5.2   Segmentation of Elongated Objects

In order to investigate this further, an evaluation on artificial shapes with simple geometry is presented. The non-linear diffusion schemes are expected to offer better performance on elongated objects. This is evaluated in 2D by optimizing the parameters for performing segmentation of a very elongated rectangle in 2D and a very elongated box in 3D. Both objects are added random uniform noise as illustrated in figure 5.2.

The collections of images used is simply produced by repeating the image generation — the random noise then differs between the example images. For completeness, objects with different orientation with respect to the coordinate axes are also included.

Three collections are evaluated. The first, denoted *Std*, has objects with widths 8 and length 480 pixels/voxels and consist of 10 images. The second collection with 5 images, denoted *Air*, has equal objects but in a larger image with more space around. The third collection with 5 images, denoted *Fat*, has objects with double width. All three collections exist in both 2D and 3D versions. It should be noted that the same collections are used for parameter optimization and for evaluation. This is not proper evaluation methodology, but since this is just a "proof of concept" evaluation on synthetic objects, this sloppy approach is allowed.

The building blocks that the multi-scale watershed method using linear diffusion produces are approximately potato-shape due to the isotropy of the diffusion. This means that no main orientation is expected to be drastically larger than the others. Non-linear diffusion has the potential for making the building blocks more elongated and thereby more effective for the segmentation task.
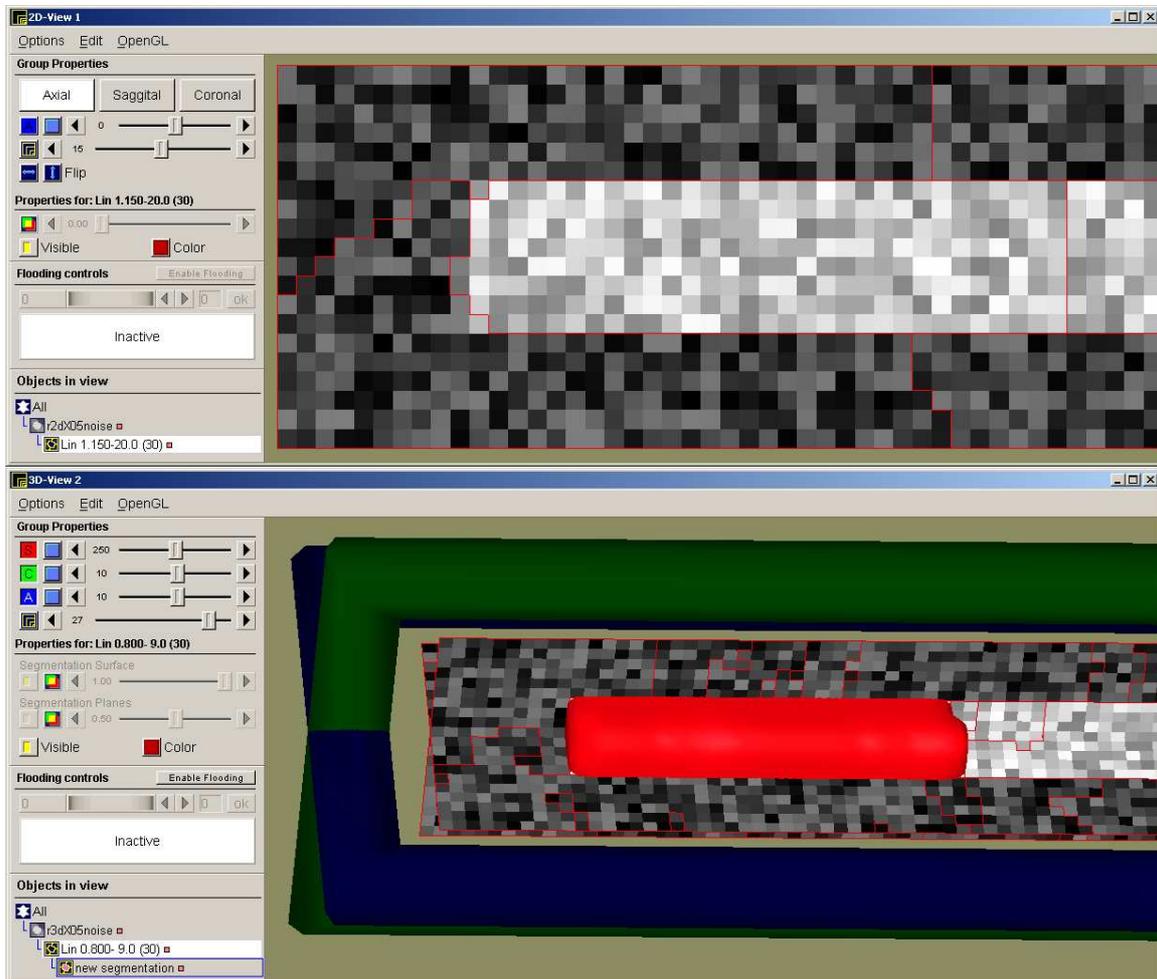
Figure 5.2: *Elongated objects in 2D and 3D.* **Top**: *Elongated rectangle in 2D. The ground truth rectangle is 8 by 480 pixels — only one end is shown. The images are generated by adding uniform random noise with a maximum intensity equal to the intensity given the ground truth object. The red lines show boundaries between building blocks generated by the multi-scale watershed method using linear diffusion.* **Bottom**: *Elongated box in 3D measuring 8 by 8 by 480 voxels. A single linear diffusion building block at the end of the box has been selected. The 2D rectangles and 3D boxes with widths 8 and lengths 480 are considered the standard collection (denoted Std). This collection has 10 2D and 10 3D images where the only difference is the random noise and the orientation of the object. Alternative 2D and 3D objects with more space around (collection denoted Air with 5 2D and 5 3D images) and with double width (collection denoted Fat with 5 2D and 5 3D images) are also tested.*

39

Std 480 × 8	2D			3D		
	Accuracy	Efficiency	Elongation	Accuracy	Efficiency	Elongation
Linear	0.00260	19.9	3.0	0.02070	16.8	3.6
GAN	0.00029	7.9	7.6	0.00095	8.0	7.5

Air 480 × 8	2D			3D		
	Accuracy	Efficiency	Elongation	Accuracy	Efficiency	Elongation
Linear	0.00234	9.4	6.4	0.02072	17.2	3.5
GAN	0.00021	3.6	16.7	0.00132	5.6	10.7

Fat 480 × 16	2D			3D		
	Accuracy	Efficiency	Elongation	Accuracy	Efficiency	Elongation
lin	0.00102	10.4	2.9	0.00525	8.8	3.4
GAN	0.00005	6.8	4.4	0.00025	3.0	10.0

Figure 5.3: *Evaluation on synthetic, elongated objects in 2D and 3D. The non-linear GAN scheme achieves building blocks that are on average 9.5 times longer than their width. The elongation of the linear diffusion building blocks is only 3.8 on average. The results are consistent for both 2D and 3D.*

A simple measure of the effectiveness of the building blocks is how elongated they are on average. In 2D, the ground truth rectangle has width w and length l. In 3D, the ground truth box has to sides equal to w and one equal to l. When the object can be segmented in a actions (up to the optimal accuracy), the average elongation of the building blocks is then $\frac{l}{wa}$. This is an alternative efficiency measure, denoted *elongation* that can be used to compare the performances of linear and non-linear diffusion for producing elongated building blocks.

The overall conclusions from the tables in figure 5.3 are that the performances in 2D and 3D are quite similar. For all collections, the accuracy obtained through the GAN scheme is significantly better compared to linear diffusion — the error is approximately an order of magnitude smaller.

Also the efficiency measured as the elongation of the building blocks shows consistent perfomance in 2D and 3D. In 2D linear diffusion has an average elongation score of 4.1 versus 9.6 for GAN. In 3D the average elongation score for linear diffusion is 3.5 versus 9.4 for GAN. These specific elongation numbers obviously depend on the level of noise in the images, but the relative performance when comparing linear diffusion and GAN is not expected to be affected by this (within reasonable limits).

The rule of thumb is that for each action needed using linear diffusion only approximately 0.4 actions are required for GAN diffusion building blocks. This rule of thumb holds for both 2D and 3D.

This invariance with respect to dimension is exactly the desired and expected behavior. Both the 2D rectangles and the 3D boxes are essentially thin one-dimensional shapes — therefore the corresponding results are to be expected.

The nice consequence is that the disappointing performance in chapter 4 is then most likely due to the naive optimization algorithms and/or the missing regularization — and not due to a fundamental problem with using non-linear diffusion in 3D.

5.3   Segmentation of Flat Objects

The behavior on the thin, elongated objects is expected to be similar in 2D and 3D. The diffusion process that merges the regions into larger building blocks, has in both 2D and 3D one free direction where the flow is desirable. The building blocks will merge in this direction and become more elongated until some scale is reached where the diffusion causes the regions to merge across the boundaries of the object and make the building blocks unsuitable for the task. The contrast across the boundary survives up to some scale and thereby allows the building blocks to become elongated. Non-linear diffusion allow this contrast to survive even longer in the diffusion process and make the building blocks even more elongated.

In 3D there is an extra direction, but for thin, elongated objects, this direction is limited by an extra object boundary with the same width and contrast. Therefore this extra dimension allows not extra elongation of the building blocks and no additional advantage due to non-linear diffusion.

However, in 3D flat objects are also possible. This type of object has two open dimensions where the diffusion can flow while being limited by the single narrow dimension. Thereby the non-linear schemes — that excels in allowing extra flow in the open dimensions before running across the narrow dimension — can possibly be even better compared to linear diffusion. As demonstrated on the elongated objects, with one open dimension, non-linear diffusion only needs approximately 0.4 actions for each action needed through linear diffusion. This compares to an elongation 2.5 times the elongation for linear diffusion building blocks.

With two open dimensions, the non-linear diffusion building blocks can then be expected to be extended by a factor of 2.5 in each direction — and thereby only needing approximately 0.16 times the actions required through linear diffusion building blocks.

This expectation is evaluated through another simple experiment. Collections of flat object are constructed in the exact same manner as the elongated objects. One collection (denoted *Pancake*) has objects with thickness 8 and side lengths 480. Another collection has smaller objects with thickness 8 and side lengths 180 (denoted *Small*). Both collections have 5 images. Again, the same collections are used for training and evaluation.

For thin, elongated objects, the elongation is a natural descriptor of the building blocks. For the flat objects we keep the elongation for comparison, but keep in mind that relevant efficiency measures should be volume oriented (and thereby the elongation squared).

We see the expected qualitative behavior in the results in figure 5.4. The elongation of the building blocks are comparable to those observed on the elongated objects (slightly worse for linear, slightly better for GAN). The consequence is that for flat objects, non-linear diffusion can reduce the building blocks needed by more than a factor of ten.

Pancake $480^2 \times 8$	3D		
	Accuracy	Efficiency	Elongation
Linear	0.00067	436.2	2.9
GAN	0.00002	≤ 28.4	≥ 11.3

Small $180^2 \times 8$	3D		
	Accuracy	Efficiency	Elongation
Linear	0.00177	52.6	3.1
GAN	0.00009	3.8	11.5

Figure 5.4: *Evaluation on synthetic, flat objects in 3D. The elongation numbers are comparable to those from figure 5.3 on elongated objects (slightly worse for linear, slightly better for GAN). In the tables, the use of $\leq$ and $\geq$ indicates that, unfortunately, the parameter optimization method has not yet reached the minimum at the time of writing. Possibly the final values will therefore be even better for GAN. For flat objects in 3D, the use of GAN can reduce the number of actions needed by more than a factor of ten.*

5.4 Reevalution of White Matter Segmentation in 3D

Above, the results on synthetic, elongated and flat objects in 3D demonstrate the qualitative behavior that is to be expected from the experiments in 2D. So the natural conclusion is that the disappointing results in chapter 4 are then not due to an inherent problem with the use of non-linear diffusion in 3D.

An appealing explanation is then the use of the naive optimization algorithm and/or the fact that the optimization is performed on a single data set. Alternatively, the problem could be that the boundaries in the brain scan can not be quantified satisfyingly through the parameters in the GAN diffusion scheme.

In order to investigate this, the experiment is repeated with a collection of BrainWeb data volumes. The BrainWeb site allows generation of data volumes with varying noise and varying intensity in-homogeneity. Combining noise levels 20% and 40% with intensity non-uniformity levels 5%, 7%, 9% gives a collection of six brain scans. Again, we use the same data collection for training and evaluation.

It is extremely difficult to predict the reduction in actions that the GAN scheme can optimally provide compared to linear diffusion. The geometry of the white matter tissue is a mixture of small, thin, flat lobes and larger "potato-shaped" regions. The performance on the synthetic, flat objects above are definitely not realistic due to the highly curved geometry of the white matter lobes.

The results shows that now, GAN allows a reduction of at least 44% in the number of building blocks used compared to linear diffusion — again $\leq$ means that the optimization process has not ended at the time of writing. This is a significant improvement compared to the results from the previous chapter. The informal rule of thumb seems to hold in 3D as well as 2D, that for segmentation of white matter tissue, the number of actions needed can at least be halved through the use of GAN diffusion.

	Accuracy		Efficiency		
	Linear	GAN	Linear	GAN	Reduction
Naive optimization, Single volume, Half brain only, Chapter 4	0.18	0.17	1510	1118	26%
Advanced optimization, Six volumes	0.18	0.18	2742	≤ 1548	$\leq 44\%$

Figure 5.5: *Reevaluation of GAN for white matter segmentation. The repeated evaluation using the improved optimization algorithm and a larger data collection allows GAN the performs significantly better compared to linear diffusion. Since the data collection now includes the entire brain, the absolute number of building blocks needed is higher than for the half brain used in the previous experiment in chapter 4. Again, the use of $\leq$ means that the parameter optimization is not concluded at the time of writing. Therefore GAN possibly performs even better than shown.*

For completeness, the improved parameter optimization method should be used to reevaluate the results from chapters 3 and 4. This would reveal whether it is the regularization due to larger data collections, or whether it is the actual optimization algorithm that make the improved results in 3D possible. Furthermore, it could actually improve the results for segmentation of brain tumors in chapter 4 significantly — possibly to a degree that would make the use of non-linear diffusion applicable for this segmentation task. This reevaluation of the previous results is left for future work.

5.5 Summary: GAN for MSWS in 2D and 3D

This dissertation presents a number of results on the use of non-linear diffusion in the multi-scale watershed segmentation method. The main results are summarized here.

The objects that will benefit from the use of non-linear diffusion can be categorized with increasing potential as follows (with accumulating categories, such that category 3 assumes the properties of the previous two):

1. The boundaries of the object are quantifiable.

2. The interior of the object has limited areas that share the properties that define the boundaries.

3. The geometry of the object is non-sphere-like.

The possible improvements in performance trough using GAN compared to linear diffusion can be loosely estimated:

- For thin, elongated objects in 2D and 3D, the efficiency for GAN is up to 2.5 higher than for linear diffusion.

- For flat objects in 3D, the efficiency for GAN is up to 10 times higher than for linear diffusion.

- For anatomical objects with more complicated geometry, these upper bounds for efficiency improvements are un-attainable. For segmentation of white matter tissue in 2D and 3D, the efficiency for GAN is approximately 2 times the efficiency for linear diffusion.

Recycling in this Chapter

This contents of this chapter are previously unpublished.

Chapter 6

Exploring Non-linear Diffusion: The Diffusion Echo

Chapters 3, 4, and 5 show that designing non-linear diffusion schemes to replace linear diffusion in multi-scale watershed segmentation is no trivial matter. While the desired behavior of the non-linear schemes can be elegantly expressed in terms of the desired diffusion across and along boundaries, is it not obvious what a specific diffusion scheme is actually "doing".

This is a fundamental problem with non-linear diffusion processes. For the linear diffusion equation, the Gaussian serves as Green's function and as a source for intuitive understanding of the linear diffusion process. Non-linear diffusion equations have in general no known closed form solutions and thereby no equally simple description.

This chapter describes a simple, intuitive description of these processes in terms of the *Diffusion Echo*. The focus of the presentation is on the ability of the diffusion echo to offer intuitive visualizations for non-linear diffusion processes. The methodology is general for arbitrary dimension, however for practical purposes we return to 2D.

In the next chapter we investigate further how the diffusion echo can be used to describe and analyze the diffusion.

6.1   Introduction

Linear scale-space [Koenderink, 1984, Witkin, 1983, Lindeberg, 1994] is a least committed scale-space with appealing theoretical properties. Among these are the existence of a Green's function for the PDE (partial differential equation) in terms of the Gaussian. Besides providing a closed form solution to the PDE, the Gaussian yields a clear, intuitive understanding of the local filtering process.

Non-linear scale-spaces are appropriate for enhancement of desired features and for extraction of certain deep structure features (in for instance edge detection [Perona & Malik, 1990] and segmentation). These diffusion schemes can typically be formulated as PDE's where a diffusion tensor determines the non-linear nature [Weickert, 1998a]. In general, most PDE's have no known closed form solutions (exceptions exist such as the analytic solution to TV diffusion [Brox et al., 2003]). This necessitates iterative numerical approximation schemes which offer less intuition.

Section 6.2 contains a presentation of the diffusion schemes used. Some of the equations are repetitions from previous chapters in order to show the simple extension to the corner enhancing scheme. The diffusion echo is introduced in section 6.3 with examples of how the diffusion echo can be used for visualization of the diffusion schemes. Finally, potential applications of the diffusion echo are presented:

- Grouping of features, for instance used for segmentation (section 6.4).

- As a deep structure summary that can serve as an alternative to multi-scale linking or flooding techniques (section 6.5).

6.2 Diffusion Schemes

A number of diffusion schemes are explored. The notation follows the use in chapter 3. All schemes use a PDE to define a scale-space $L(\vec{x};t)$, where $\vec{x}$ are spatial coordinates and t the scale parameter. The PDE's have an image I as initial condition: $L(\vec{x};0) = I(\vec{x})$.

Linear diffusion [Koenderink, 1984, Witkin, 1983] can be defined by: $L_t(\vec{x};t) = \Delta L(\vec{x};t)$, the heat diffusion equation. The Gaussian with standard deviation $\sigma = \sqrt{2t}$ is Green's function for the PDE.

The non-linear Perona-Malik scheme [Perona & Malik, 1990] attempts to preserve edges during the diffusion: $L_t(\vec{x};t) = div(\ p(|\nabla L_\sigma|^2)\ \nabla L\ )$ where $p(|\nabla L_\sigma|^2) = 1/(1 + \frac{|\nabla L_\sigma|^2}{\lambda^2})$. The regularisation parameter σ is due to [Catté et al., 1992]. The notation ∇L_σ means the gradient evaluated at scale σ. The parameter λ is a soft threshold for the gradient magnitude required to locally slow the diffusion and preserve an edge. Following the terminology of Weickert [Weickert, 1998a], the scheme is termed "isotropic" since the diffusivity function p is scalar-valued.

6.2.1 Generalized Anisotropic Non-linear Diffusion

Weickert [Weickert, 1998a] defines the *anisotropic non-linear diffusion equation*:

$$L_t(\vec{x};t) = div(\ D(J_\rho(\nabla L_\sigma))\ \nabla L\ ) \qquad (6.1)$$

The diffusion tensor $D \in C^\infty(R^{2\times2}, R^{2\times2})$ is assumed to be symmetric and uniform positive definite. The *structure tensor* J_ρ is evaluated at *integration scale* ρ, and the gradient ∇L_σ at *sampling scale* σ. As stated previously we set the structure tensor to identity.

Diffusion schemes can be defined in terms of the eigenvalues λ_1 and λ_2 for the corresponding eigenvectors $\bar{v}_1 \parallel \nabla L_\sigma$, $\bar{v}_2 \perp \nabla L_\sigma$ for the diffusion tensor D.

A large class of diffusion schemes (including the previous) are generalized by the *Generalized Anisotropic Non-linear* scheme (GAN) [Dam, 2000, Dam & Nielsen, 2000], where the diffusion tensor eigenvalues are defined:

$$w(m,\lambda,s) = \begin{cases} 1 & |\nabla L_\sigma| = 0 \\ 1 - exp\left(\frac{-C_m}{\left(\frac{s^2}{\lambda}\right)^m}\right) & |\nabla L_\sigma| > 0 \end{cases} \quad (6.2)$$

$$\lambda_1 = w(m,\lambda,|\nabla L_\sigma|)$$
$$\lambda_2 = \theta + (1-\theta)\,\lambda_1 \quad (6.3)$$

The scheme is *anisotropic* when the eigenvalues are not equal (then D can not simply be replaced by a scalar-valued function). The global parameter θ determines the degree of anisotropy (0 is isotropic diffusion and 1 is full anisotropic), λ is the soft edge threshold, and m is the "aggressiveness" that the edges are preserved with.

The GAN scheme has the following schemes as special cases:

- Linear Gaussian diffusion is defined by $\lambda \to \infty$.

- The regularised Perona-Malik scheme is approximated by $\theta = 0$ and $m = 0.75$ (which implies $C_m = 0.762689$ [Dam, 2000]).

- Weickert's *Edge Enhancing diffusion* (EED) is defined by $\theta = 1$ and $m = 4$ (which implies $C_m = 3.31488$ [Weickert, 1998a]).

6.2.2 Corner Enhancing Diffusion

Near "edges" the Perona-Malik scheme slows diffusion in all directions. For image enhancement, EED is appropriate since diffusion is full along edges.

However, the EED scheme tends to round corners due to the full diffusion along the edge. Therefore, while full anisotropic diffusion is desirable at edge-like structures, a diffusion scheme with a milder degree of anisotropy is desired at corners. A local steering of the degree of anisotropy therefore seems sensible.

The following *Corner Enhancing Diffusion* scheme (CED) is similar to GAN but steers the anisotropy locally using a corner measure: the isophote curvature κ times the gradient to a power k.

$$\lambda_1 = w(m_g, \lambda_g, |\nabla L_\sigma|)$$
$$\theta = w(m_i, \lambda_i, |\kappa_\sigma|\,|\nabla L_\sigma|^k)$$
$$\lambda_2 = \theta + (1-\theta)\,\lambda_1 \quad (6.4)$$

The Corner Enhancing scheme is similar to the CID scheme from [Dam, 2000].

The use of the curvature in the definition of the eigenvalue function makes the equation for corner enhancing diffusion a third-order PDE. Thereby previous results on well-posedness and scale-space properties are no longer valid.

6.3 Visualizations

It takes a strong mathematician to get intuition about the differences between the diffusion schemes above. A standard way of illustrating the schemes is to visualize the local diffusion at key points in an image like in the following.

The non-linear diffusion processes are implemented using iterative numerical approximation schemes. For each iteration a diffusion tensor is determined for each point in the scale-space image. This diffusion tensor can be visualized by an ellipse where the orientation and the size are determined by the eigenvectors and corresponding eigenvalues.

In figure 6.1 EED is illustated like this. Isolated, the third image seems to offer an understanding of the intensions of the diffusion scheme. However, the illustrated diffusion tensors are deceiving since they evolve during the diffusion. Furthermore, they fail to capture the interaction with the surrounding area.
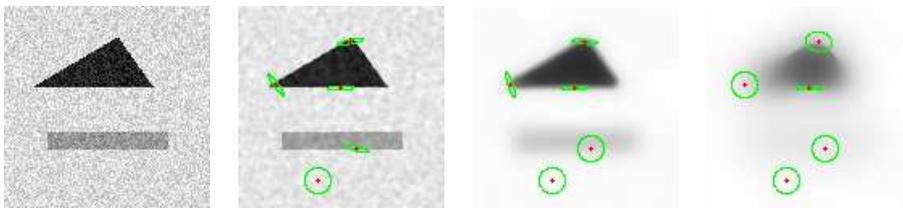
Figure 6.1: *Visualizations of diffusion tensors for EED scheme.* **Left:** *test image (64x64 pixels, intensities 0-255, SN ratio 2.5).* **Right three images:** *the local diffusion tensors illustrated as ellipses at five points for three different iterations (t = 0.4, 20, 100). An explicit approximation scheme with a nonnegativity discretisation is used [Weickert, 1998a].*

6.3.1 The Diffusion Echo

The diffusion echo is inspired by the Gaussian that defines the local filtering in linear diffusion. The equivalent is obtained for non-linear schemes in two steps:

Diffusion Echo: Source

For a fiducial point p, construct an auxiliary image with the value 1 at the point p and zero otherwise: the discrete impulse function.

For each iteration in a diffusion process for an image I, the values are computed by assigning each pixel a weighted average of a neighborhood of pixels.

The auxiliary image is treated with the same weighting as the image I. The result is a distribution that records the flux that propagates from the source pixel p. This is the *diffusion echo source distribution* and is denoted $S_p(\cdot)$.

48

Diffusion Echo: Drain

The *diffusion echo drain distribution* is the opposite of the source. For a point q, the value for the drain distribution at a given point p is defined in terms of the source at p. Specifically, the drain distribution $D_q(\cdot)$ is $D_q(p) \equiv S_p(q)$. Note that the drain for a point requires the sources for the entire image.

The diffusion echo drain distribution is the local filter kernel for the diffusion process equivalent to the Gaussian filter for the linear diffusion process.

The algorithm for computing the diffusion echo distributions is included as algorithm 3 — just to show how simple it is. The algorithm applies to any diffusion scheme that can be defined in terms of the anisotropic non-linear diffusion equation (equation 6.1) — and most other diffusion schemes with few modifications. Note that during the diffusion, only step 4e is added to the basic iterative diffusion algorithm. For more detail on calculating the local weights corresponding to a diffusion tensor, see [Weickert, 1998a].

Algorithm 3 *Diffusion Echo Distributions*

Input: Original image I, diffusion scheme parameters.
Output: Diffusion echo source $S_p(\cdot)$ and drain distributions $D_p(\cdot)$ for all image points p.

1. Initialize diffused image to $I^0 = I$.

2. Initialize echo source to zeros for all image points p, q: $S_p^0(q) = 0$.
   In 2D this is a 4D matrix, in 3D a 6D matrix.

3. Set $S_p^0(p) = 1$ for all image points p. Each distribution is then a discrete impulse function.

4. For $i = 1$ to desired diffusion iterations T do:

   For each image point p do:

   (a) Calculate local image geometry as needed using Gaussian derivatives at desired regularization scale.
   (b) Calculate diffusion tensor eigenvectors and eigenvalues as defined by the diffusion scheme..
   (c) Calculate local weights $w(n)$ for points n in the neighborhood N of p.
   (d) Diffuse image: $I^i(p) = \sum_{n \in N} w(n) \, I^{i-1}(n)$
   (e) Calculate echo source distribution for each image point e:
   $S_e^i(p) = \sum_{n \in N} w(n) \, S_e^{i-1}(n)$

5. Diffused image is now $I^T(\cdot)$.

6. Diffusion echo source distribution is $S_p^T(\cdot)$ for all image points p.

7. Assign echo drain distributions for each combination of image points p, q: $D_p(q) = S_q^T(p)$.

Diffusion Echo Properties

For linear diffusion the source and the drain distributions are identical Gaussian distributions. However, in general the distributions are not equal.

The diffusion echo drain distribution is the local convolution filter kernel for the diffusion process:
$L(\vec{x}; t) = \int L(q; 0) \, D_{\vec{x}}(q) \, dq = \int I(q) \, D_{\vec{x}}(q) \, dq$

The distributions can be interpreted as affinity measures. However, note that in general both $S_p(q) \neq S_q(p)$ and $D_q(p) \neq D_p(q)$.

Since both source and drain are unity distributions they can also be interpreted as probabilistic distributions. The source distribution $S_p(q)$ (or the drain distribution $D_q(p)$) states the probability for an "atom" originating at point p to end at point q as a result of the diffusion.

The maximum for both source and drain distributions remain at the origin for the distribution for most diffusion schemes. Mean and higher order moments are in general not located at the origin and can be used to characterise the distributions. The definitions are applicable for images of arbitrary dimensions.

6.3.2   Diffusion Echo Visualizations

The diffusion echo is a summary of the diffusion process up to a certain time/scale. In the following we show that illustrations using this principle offer significantly more information than the illustrations in the previous section.

Basic Comparison

In figure 6.2 we illustrate this for the four diffusion schemes presented in section 6.2. The figure displays the diffusion echo drain distributions for the five selected points from figure 6.1. These are equivalent to the local convolution filter kernels that would yield the diffusion directly.

The ellipses in figure 6.2 highlight the properties of the diffusion schemes. Linear diffusion uses the same diffusion tensor at all points. The non-linear Perona-Malik reduces the diffusion gradually determined by the gradient magnitude compared to a soft threshold value. The anisotropic EED scheme reduces diffusion across edges quite agressively but maintains full diffusion along the edges.

Finally, the CED scheme reduces diffusion perpendicular to the gradient as well at corner-like structures. However, the diffusion echo distributions reveal that differences between the schemes are not quite as characteristic. Apparently, there is more like a smooth transitition between the schemes — like the existence of the GAN scheme implies. Even though the Perona-Malik scheme is termed isotropic it has a preferred diffusion direction along the edge. This is much more pronounced for the anisotropic EED scheme but not qualitatively different.
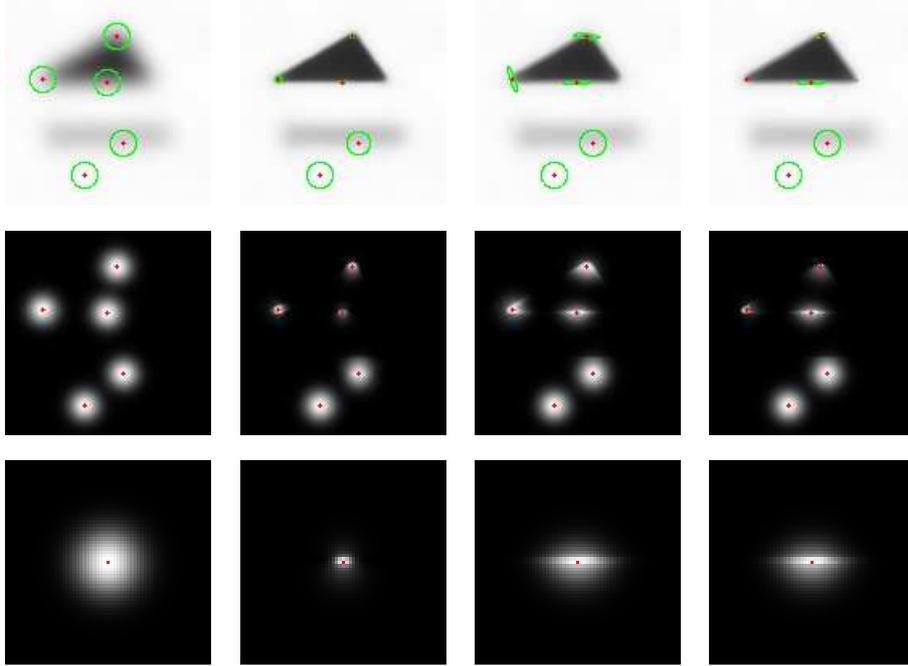
Figure 6.2: *Diffusion Echo drain distributions.* **Top row**: *Ellipse-illustrations for the four diffusion schemes (linear, Perona-Malik, EED, and CED).* **Center row**: *The corresponding drain distributions. Note that the distribution is computed separately at each point — the illustrations are mosaics of these separate illustrations. By definition, each distribution has the same total energy (they are unity filters), but they are scaled individually for better visual appearance.* **Bottom row**: *Close-ups of the distributions for the point right below the center of the triangle.*
An explicit approximation scheme with a nonnegativity discretisation was used with $t = 20$ [Weickert, 1998a]. The edge threshold parameter is set to 220 for all schemes. This corresponds to characterising the contour around the triangle as edge and the contour around the rectangle as non-edge. The regularisation scale is 1.2.

Effects of Discretisation Scheme

Apart from the relative differences between the schemes, it appears that the Edge Enhancing scheme is not quite able to enhance the straight edges as well as in previous publications [Weickert, 1998a]. This is simply because of the numerical approximation scheme. For the previous illustrations we use the implementation that ensures non-negative weights in the local diffusion stencil ([Weickert, 1998a] page 95). This restricts the spectral condition number of the diffusion tensor to be below 5.8284 — meaning that the local degree of anisotropy is limited. The eigenvectors are correspondingly limited such that $\lambda_1 < 5.8284\ \lambda_2$. For the edge enhancing scheme where $\lambda_1 \equiv 1$ this sets a lower limit on λ_2 and thereby some diffusion across the edges is allowed.

In figure 6.3 the same diffusion processes are repeated using the non-restricted, standard approximation scheme [Weickert, 1998a, Scharr & Weickert, 2000]. It is apparent that a more effective preservation of the edges is possible. The shapes of the distributions are especially interesting at the corners. The different abilities of the schemes with respect to supporting diffusion along the edge through the corner is evident.

The illustration clearly reveals that the change of discretisation scheme has a major effect on the diffusion for some of the schemes.
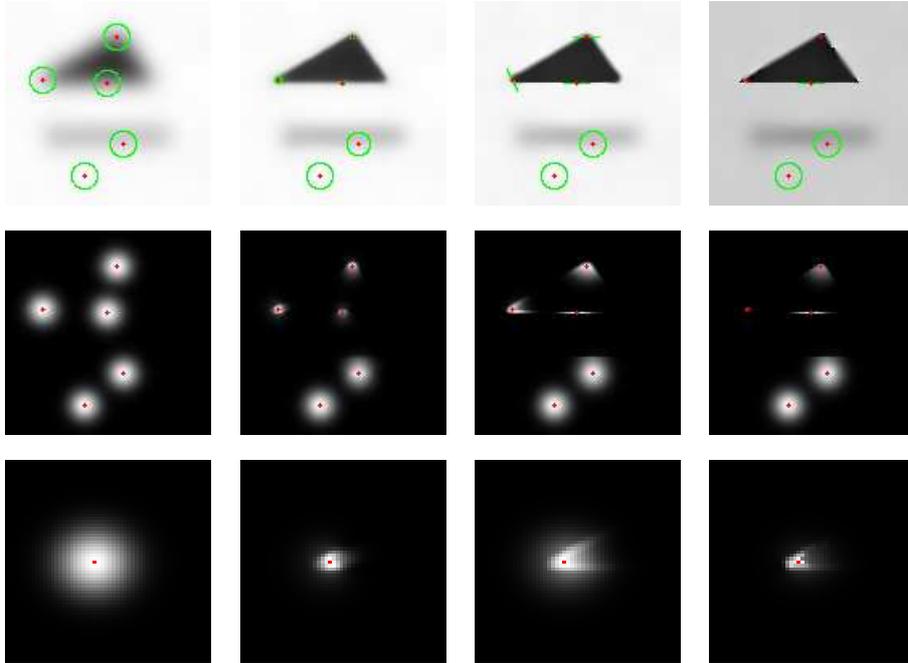
Figure 6.3: *Diffusion Echo drain distributions for the four diffusion schemes (linear, Perona-Malik, Edge Enhancing, and Corner Enhancing) where the nonnegativity approximation scheme used in figure 6.2 is replaced with the simpler standard approximation scheme. The standard scheme allows more pronounced anisotropy.* **Bottom row**: *Close-ups for the point inside the left corner of the triangle.*

An obvious discretization scheme is to use the larger nonnegativity stencil — but not enforce the upper limit on the condition number. This would offer a better numerical approximation than the simpler standard discretization scheme while still allowing arbitray local anistropy. This is not investigated in this work.

The Visualization Ability

The previous illustrations show that the diffusion echo is able to visualize properties of the diffusion schemes that are not otherwise apparent. Extended experiments with the diffusion schemes allows similar intuition but the diffusion echo illustrations offer this understanding directly.

In the following sections we offer a few appetisers indicating that the diffusion echo can be used for more than just illustrations.

6.4 Diffusion Echo Application: Grouping

The diffusion echo expresses affinity between points in an image. This affinity measure can be used for grouping of pixels into regions or grouping of feature points in general.

Since the affinity measure is defined in terms of the underlying diffusion, the diffusion scheme needs to be appropriate for the specific grouping task.

Edge detectors often produce edge pieces with small gaps in-between. The diffusion echo for the Edge Enhancing diffusion scheme would be a very appropriate measure for determining the connectivity of the edge pieces.

Attributes of the diffusion echo can be used for determining grouping as well. An example could be using the shape of the distribution to guide grouping of pixels into regions. This is illustrated by the distributions in figure 6.4.

The differences in diffusivity from the low diffusivity near-edge areas to the high diffusivity areas away from the edges causes a flux away from the edges. Thereby the means of the distributions move away from their origins in a direction away from the edges (if any edges are within "striking distance" depending on diffusion parameters, especially regularisation and diffusion time). This can be used to create a drift field that can be used to group the pixels. The field will create a *sink* in each region that the pixels drift towards.
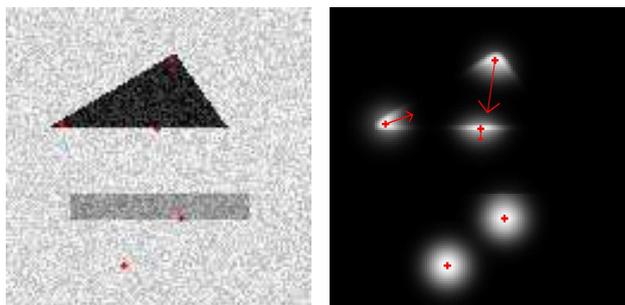
Figure 6.4: *Diffusion Echo grouping drift fields. For the five points a vector from the point to the mean for the drain distribution gives a drift field that can be used for grouping. For visualization purposes, the vectors have been scaled to 10 times their actual lengths. The vectors from the points just inside the corners aim towards the center of the triangle. The vector for the point just below the triangle (it is a single pixel outside) aims away from the triangle. The remaining vectors are practically zero-vectors since the parameters for the Edge Enhancing scheme dictate that there are no edges near them — their distributions are approximately Gaussian with means at their origins.*

6.5 Diffusion Echo Application: Deep Structure Summary

Instead of grouping the pixels individually, the diffusion echo can also be used for grouping image regions. This could be an alternative to existing multi-scale linking schemes. An example of this is shown in figure 6.5 where grouping based on the diffusion echo is compared to multi-scale watershed segmentation.

The diffusion echo grouping uses a simple threshold to determine whether two neighboring regions are merged into one region. This threshold is compared to the average affinity measure between pairs of pixels in the two regions using the affinity measure directly from the diffusion echo source distribution. The result is a simple flooding-like algorithm.

Even though both examples are quite simple, they illustrate that the diffusion echo distributions capture what can be considered a deep structure summary to such an extent that even simple attributes offer powerful grouping abilities.
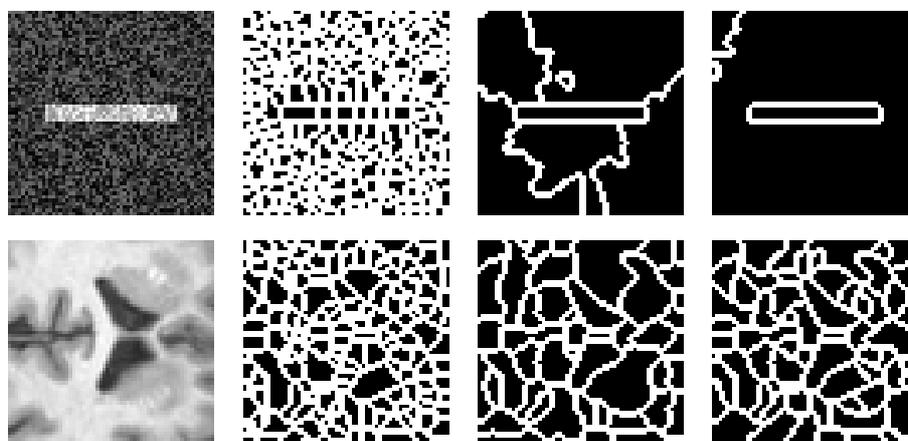
Figure 6.5: *Segmentation by grouping of watershed regions using the Diffusion Echo.*
Top row*: A simple example where the segmentation task is to capture the rectangle. First the test image followed by a watershed segmentation at low scale. Third image shows the multi-scale linking of the regions where the underlying diffusion is 30 levels of the edge enhancing scheme from $t = 1$ to $t = 600$ [Dam, 2000]. The rightmost image is the result of flooding with the threshold 0.0025 using the average diffusion echo source affinity between neighboring regions.*
Bottom row*: Equivalent where the segmentation task is to capture the ventricles from a data set from the Internet Brain Segmentation Repository [ibs, 1999]. Here, the linking uses 10 levels from $t = 0.6$ to $t = 80$ in 10 levels. The flooding threshold is 0.0059.*
The diffusion echo flooding method groups the desired regions for both images.

6.6   Conclusion

We introduce the diffusion echo: the distributions that equip non-linear diffusion schemes with what corresponds to the Gaussian for linear diffusion.

The diffusion echo offers illustrations of non-linear diffusion schemes that reveal the true local diffusion in an intuitive manner.

Furthermore, we argue that the affinity nature of the distributions can be used for grouping. This is demonstrated through two examples. First, the diffusion echo distributions are used to generate a drift field, where each pixel is equipped with a vector stating the preferred direction of grouping. Secondly, we group regions by the average affinity between pixels in the regions. This proves to be a simple but effective region grouping scheme.

Computation time and memory requirement for the computation of the diffusion echoes are quadratic in the number of image pixels — this obviously is problematic for larger images. However, actual applications will not use the basic definition directly. Where only linear attributes of the diffusion echo distributions are needed, these can be computed directly during the diffusion iterations (in linear time and memory) as shown in the following chapter.

Diffusion echo based methods imply a shift from the diffusion scheme being an underlying information-simplifying step to being the central information-collecting process. The basic principle behind Diffusion Echo based methods is *the diffusion knows*. Future work will reveal the application tasks where the proper non-linear diffusion scheme really is omnipotent.

Recycling in this Chapter

This chapter is an edited version of [Dam & Nielsen, 2001].

Chapter 7

Approximating Non-linear Diffusion

Chapters 3, 4, and 5 demonstrate how non-linear diffusion offers superior performance compared to linear diffusion as a simplifying proces that retains the relevant information during diffusion. This is due to the ability to specialize non-linear diffusion.

However, as illustrated in sections 5.2 and 6.5 the isotropic linear diffusion actually allows clearly elongated building blocks in the multi-scale setting and thereby exhibits some anisotropic behavior. This suggests that linear and non-linear diffusion might not be fundamentally different in multi-scale methods.

The results in this chapter are general for any use of the non-linear schemes, but as the evaluation method reveals we have the multi-scale watershed segmentation method in mind.

We assess the feasibility of approximating non-linear diffusion processes with simple local Gaussian filters. Thereby the approach has a certain flavor of replacing non-linear diffusion with scale-selection in linear scale-space. The purpose of doing this is twofold. Firstly, the theoretical implications are by themselves interesting. Secondly, a successful method would reduce the need for computationally expensive implementations of non-linear diffusion schemes.

We evaluate using isotropic and affine Gaussian filters for the task of approximating the local diffusion for a number of non-linear diffusion schemes. The approximations are first explored using an information theoretical approach and secondly evaluated based on their performance in the multi-scale watershed segmentation method.

The results show that while the approximations do not perform quite as well as the original non-linear schemes, the decrease in performance is acceptable for the evaluated task. Furthermore, the affine approximations perform significantly better than the isotropic, as expected.

7.1 Introduction

Non-linear diffusion have proven extremely useful in numerous applications: noise-reduction, enhancement, restoration, and multi-scale segmentation [ter Haar Romeny, 1994, Weickert, 1998a, Sapiro, 2001, Dam & Nielsen, 2000]. This success is due to the ability to incorporate prior task-specific knowledge into the diffusion process — typically by tuning parameters to the task at hand.

However, non-linear diffusion does also introduce a couple of basic problems. Firstly, the parameters for the non-linear schemes have to be determined in some more or less well-founded manner. This is in itself not a trivial matter as explored in chapter 5. We do not address this issue in this chapter. Secondly, the non-linear schemes are expensive in terms of computational complexity. For applications that use non-linear diffusion for noise-reduction in a pre-processing step, this is not problematic due to relatively short diffusion times. For more demanding applications the use of non-linear diffusion is often un-feasible — e.g. high resolution medical 3D scans will often impose severe computational time problems.

One way to attack this problem of computational complexity is to introduce sophisticated numerical implementations. One such example is the AOS scheme for anisotropic diffusion [Weickert, 1998a].

The work in this chapter aims at an alternative solution where the non-linear diffusion schemes such as anisotropic diffusion schemes are replaced by simpler schemes based on local Gaussian filters. Since the goal is to determine the "best" local Gaussian filter this has a certain scale-selection flavour and could be inspired by the methods that do this by maximization of scale-invariant expressions [Lindeberg, 1994] or MDL minimization [Gomez et al., 2000]. However, we aim at replacing a given non-linear diffusion scheme and therefore the scale-selection mechanism must have the desired diffusion time as a parameter.

We do not present such a scale-selection method — this is a fundamental feasibility study. However, we illustrate how such a method can optimally perform. But the main focus is on the evaluation of the performance of local Gaussian filters that approximate a given non-linear diffusion scheme optimally. So we answer the question: *If we design the perfect scale-selection method for approximating non-linear diffusion, how good will it be?* An actual scale-selection method could be inspired by related work on the subject [Nitzberg & Shiota, 1992, Fischl & Schwartz, 1997].

The approximation method relies heavily on the use of the diffusion echo presented in the previous chapter. The local filter kernels extracted in this manner is then approximated by isotropic and affine Gaussian filters satisfying maximal entropy constraints. As a secondary result, these approximations allow us to quantify how non-linear the schemes are (i.e. how different from linear Gaussian diffusion they are).

Having defined the approximating filter kernels we turn to the evaluation that we perform in two steps. First we explore the performance of the approximating schemes by evaluating them on random points in natural images from the Van Hateren collection [van Hateren & van der Schaaf, 1998].

Secondly we evaluate the approximating filter kernels in the multi-scale segmentation setting where non-linear diffusion schemes have been shown to offer superior performance compared to linear diffusion.

7.2 Approximating Non-linear Diffusion

In order to approximate the non-linear diffusion schemes we first perform the desired non-linear scheme implemented through a simple explicit iterative scheme. This is done with an augmented implementation that records the actual local filter kernels that are implicitly used in the diffusion as explained in the previous chapter.

These recorded local non-parametric filter kernels are approximated with simple, parametric Gaussian filters based on information-theoretic criteria.

7.2.1 The Diffusion Echo Moments

Most non-linear diffusion schemes have no explicit expression for (or representation of) the local filter kernel that determines the diffusion in each point (i.e. pixel or voxel, the following works for arbitrary spatial dimension).

However, a simple method exists for obtaining these local filter kernels. For each pixel an auxiliary image is created with the value 1 in this pixel and zero elsewhere. This is the discrete equivalent of the impulse function. The diffusion that is performed on the actual image is then performed in parallel on each auxiliary image. Thereby the impulse responses for the diffusion process are acquired. They determine where the "mass" in a pixel flows to during diffusion. The local diffusion filter kernel for a pixel is then collected by picking the flow from each impulse response that flows to the specific pixel. The recorded impulse responses and local filter kernels are called the *Diffusion Echo* [Dam & Nielsen, 2001].

The downside of this simple method is the computational complexity and the memory requirements. Independent of the efficiency of the underlying implementation of the diffusion scheme, the augmented method becomes at least $O(P^2)$ (where P is the number of pixels in the image, assuming a fixed diffusion time) in the straightforward implementation. This can be lowered to $O(P \times F)$ (where F is an upper limit of the number of pixels in a local filter kernel not being zero) if the extent of the impulse response can be limited. Nevertheless, this is still a quite restrictive complexity.

However, for some purposes only the moments of the filter kernels are necessary and not the actual filters. Many implementations of non-linear diffusion schemes are based on iterative schemes where a local stencil is used to form a weighted average of some neighborhood for each pixel/voxel. The specific scheme defines the local stencil.

In this case it is simple to record the moments (specifically mean and variance) of the local filter kernel directly without recording the local impulse response. This is done by using the local stencil to average the moments from the previous iteration for each iteration step. This process adds no computational complexity to the diffusion method (in the $O(\cdot)$ sense).

7.2.2 Maximum Entropy Approximation Filters

As described above we summarize the diffusion process with a local diffusion filter kernel in each point of the domain. The convolution with the filter kernel and the original image in that specific point will give exactly the same result as the diffusion process in the point. The filter values depend on the chosen diffusion scheme (including choice of parameters — e.g. iterations) and the original image.

We want to approximate this filter kernel. A straightforward approach is to approximate the first few moments, say the mean and the variance. In order to select among the filters with the same mean and the variance we choose the filter with maximum entropy when the filter is viewed as a distribution.

A maximum entropy solution is a least committed appoach in the sense that it treats all locations as equally as possible under the given restrictions (here a specified mean and variance). In the case of no restrictions, a maximum entropy solution results in a uniform distribution. In this way we avoid to introduce a bias towards a specific (unknown) purpose in our approximation. For a distribution $P(x)$ on the domain D, the entropy $H(P)$ is defined:

$$H(P) = \int_{x \in D} -P(x) \log P(x) \ dx$$

The maximum entropy solution given a specific mean and variance is in the continuous case a Gaussian distribution. To stay in the continuous domain we calculate the variance of the diffusion echo filter kernel by modelling it as a piecewise constant function on a continuous domain. When we apply the resulting approximating filter kernel we also model the image as a piecewise constant function on a continuous domain.

Another possible measure for defining the best approximating filter kernel is to limit the *Kullback-Leibler divergence* (or the *relative entropy*, [Kullback & Leibler, 1951]). For a distribution $P(x)$ and an approximating distribution $A(x)$ on the domain D, the Kullback-Leibler divergence d is defined:

$$d(P, A) = \int_{x \in D} P(x) \log \frac{P(x)}{A(x)} \ dx$$

By following Jaynes' maximum entropy principle [Jaynes, 2003] we limit the possible approximating filter kernel to be Gaussian. It is fairly easy to show that a Gaussian $G(\mu, \sigma)$ with the same mean μ and variance σ^2 as the original filter kernel P also minimizes the Kullback-Leibler divergence d between the Gaussian and the filter kernel (done in 1D for simplicity):

$$
\begin{aligned}
\arg \min_{\mu, \sigma} d(P, G) &= \arg \min_{\mu, \sigma} \int_{-\infty}^{\inf} P \log \frac{P}{G} \ dx = \arg \min_{\mu, \sigma} \int_{-\infty}^{\inf} -P \log G \ dx \\
&= \arg \min_{\mu, \sigma} \int_{-\infty}^{\inf} P \log \sqrt{2\pi\sigma^2} \ dx + \int_{-\infty}^{\inf} P \frac{(x-\mu)^2}{2\sigma^2} \ dx \\
&= \arg \min_{\mu, \sigma} \log \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \int_{-\infty}^{\inf} P (x-\mu)^2 \ dx
\end{aligned}
$$

The last integral minimizes to the variance of P for μ equal to the mean of P. Then, the remaining expression is easily shown to minimize for σ^2 equal to the variance of P.

Thus, by following Jaynes' maximum entropy principle we minimize the Kullback-Leibler divergence between the diffusion echo filter kernel and the approximating filter.

Using an information theoretic approach is relevant since the filters can be naturally perceived as probability distributions. Specifically they measure the probability of flow between two points in the image (the two directions of the flow being in source and drain echo respectively).

7.2.3 Illustrating Diffusion Approximations

The non-linear diffusion schemes used in the evaluation are presented in previous chapters. We investigate the use of linear diffusion, Perona-Malik diffusion, Weickert's edge enhancing diffusion, and finally generalized anisotropic non-linear diffusion. These schemes are described in the previous chapter (section 6.2).

The approximation method is illustrated by comparing the result with the use of the original non-linear schemes.

For the illustrations we use the simple image in figure 7.1 where the point of interest is the center point. From a diffusion approximation point of view, this is a relatively challenging point since it is located just inside a corner.

Figures 7.1 and 7.2 show that the approximating filters do a reasonably good job for this example point — the quantitative differences are obvious, but the qualitative appearance of the diffused images are quite similar.

7.3 Information Theoretical Evaluation

The diffusion echo filter kernels from the diffusion processes are approximated by a Gaussian filter. We evaluate both the use of isotropic and affine Gaussian filters for the approximations. The expectation is that in some cases the affine can give a better approximation than the isotropic. More advanced filters could be chosen, but that would defeat the purpose of making simple, approximating schemes.

The diffusion echo depends on the diffusion process as well as the original image. The influence of the original image is addressed by 10,000 repetitions of randomly selecting a point in a randomly selected image from the Van Hateren database of natural images (actually approximately half of the Van Hateren images are discarded due to poor image quality [Griffin et al., 2003]). Around each selected point, a neigborhood of $N \times N$ pixels is selected and the diffusion is done on this subimage. The neighborhood size N is defined such that a linear Gaussian filter at the center has three standard deviations inside the subimage. The diffusion time t gives the standard deviation by $2t = \sigma^2$ — we use 100 iterations with time step 0.2 so $N = 39$.
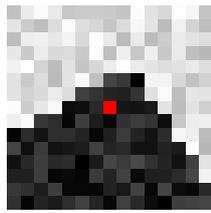
Image Dimensions: 15x15
Intensities (min/mean/max): 0.0 / 132.5 / 255.0
Iterations: 15, Time step: 0.2
Echo Approximation Criterium: MaxEntropy

LIN — Echo — Affine — Isotropic

Entropy: 4.618
Entropy: 4.637 / KL div: 0.0008
Entropy: 4.637 / KL div: 0.0008

RPM — Echo — Affine — Isotropic

Entropy: 3.669
Entropy: 3.785 / KL div: 0.0845
Entropy: 3.797 / KL div: 0.0971

EED — Echo — Affine — Isotropic

Entropy: 3.479
Entropy: 3.774 / KL div: 0.2633
Entropy: 3.777 / KL div: 0.2673
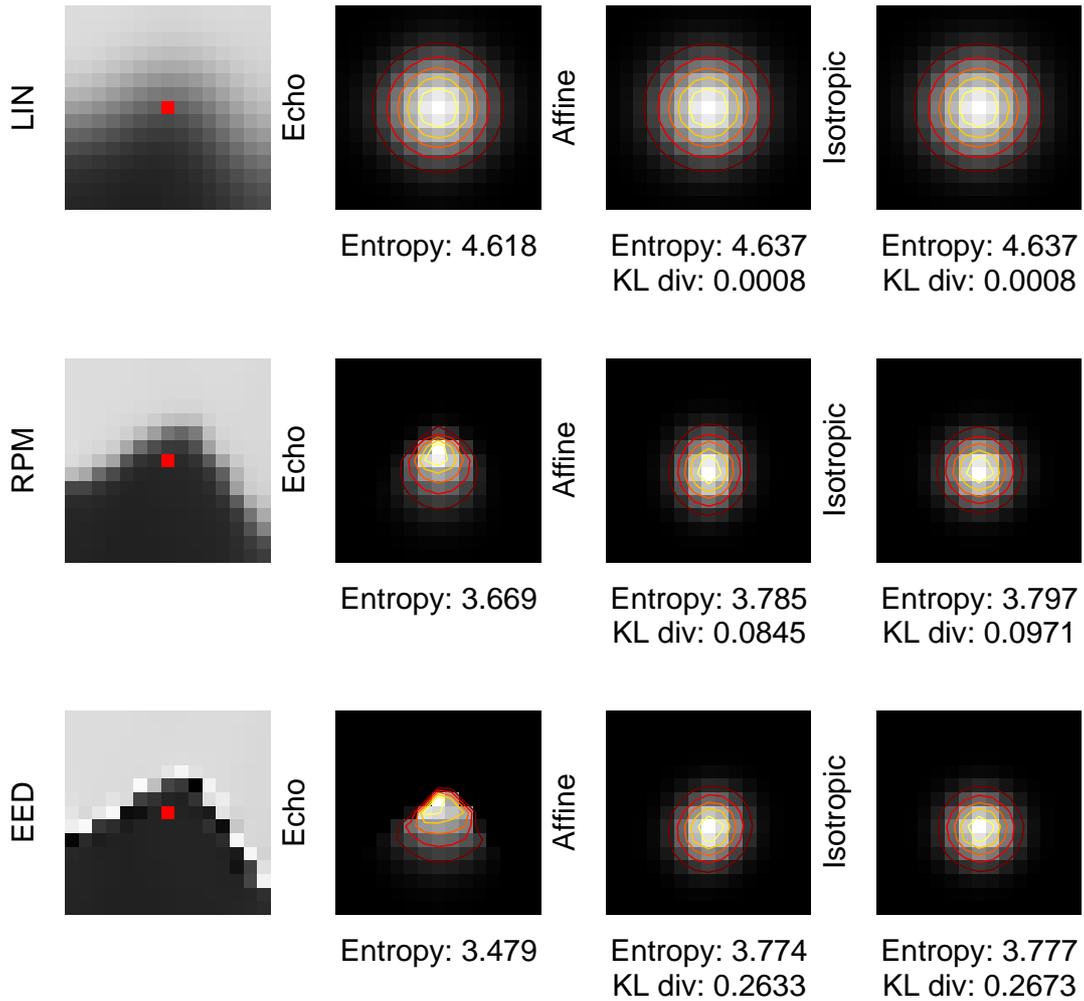
Figure 7.1: *Comparison of the actual diffusion echo filter kernels with the approximated filters.* **Left column**: *The original image and the result of applying linear Gaussian, Perona-Malik, and edge enhancing diffusion. The red dot is the center point where the filters are approximated.* **Second column**: *The diffusion echo filter kernels for each diffusion scheme.* **Third and fourth columns**: *Affine and isotropic approximating filters, respectively. The overlaid contours in warm colors are iso-curves. The entropy and KL measures show that EED is clearly more non-Gaussian than RPM, and that there is a surprisingly small difference between the isotropic and affine approximations.*
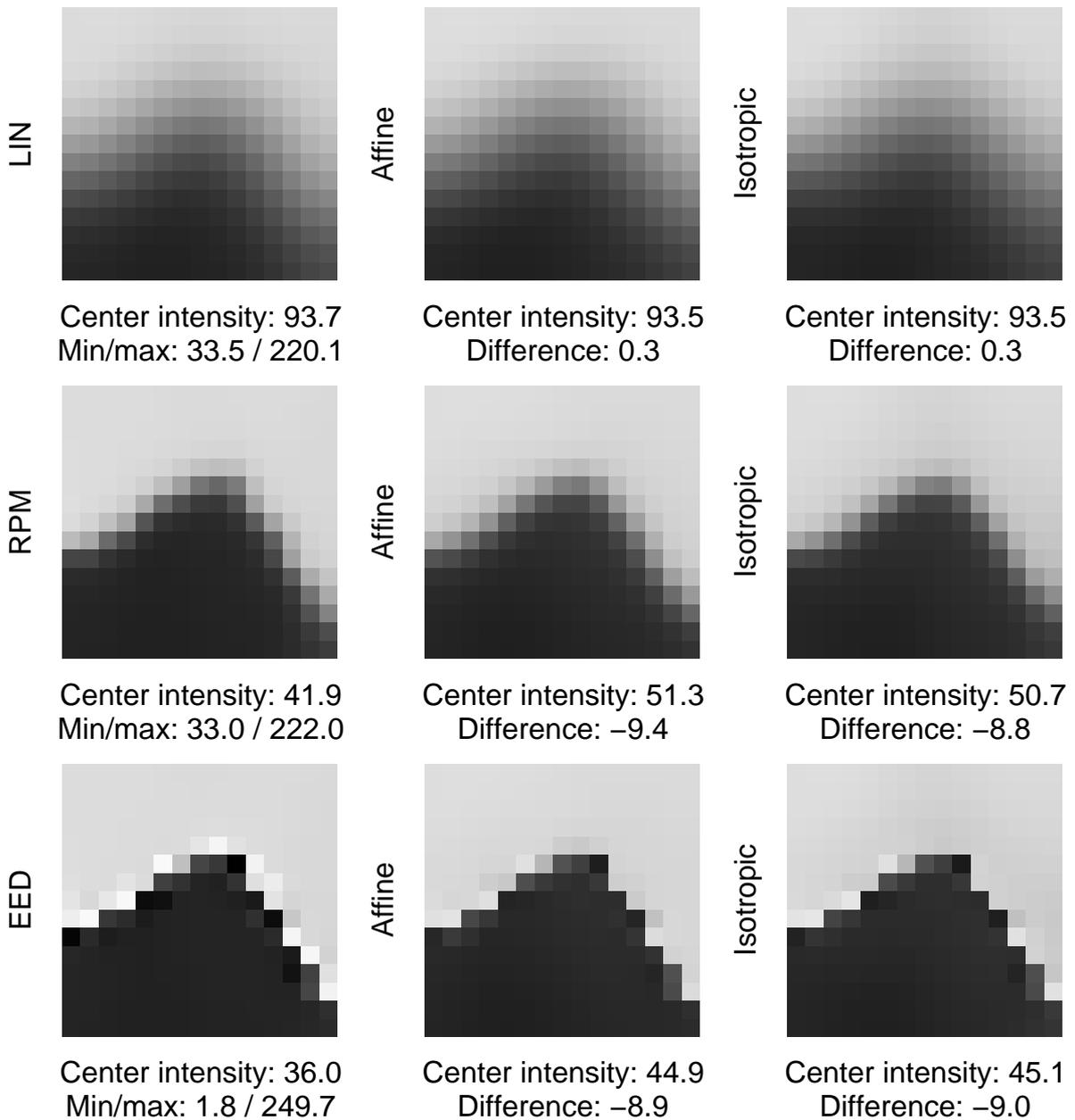
Figure 7.2: *The effect of applying the diffusion echo filter kernel compared to the result of applying the approximating filters. The rows are linear Gaussian, Perona-Malik, and edge enhancing diffusion. First column is the diffused images, second and third columns are the approximated diffused images. Even though the quantitative differences between applying the non-linear filter kernels and the approximated filters are evident, the qualitative appearence of the resulting images are quite similar.*

In order to assess the correspondence between the diffusion echo filter kernel and the approximating filter we compare both the filters themselves and their effect on the image. The filters can be perceived as distributions that govern the flow of mass in the image. Thereby a natural measure of difference is the Kullback-Leibler divergence. When measuring the difference in the effect of the diffusion on the images we measure the resulting intensity difference. The local intensity difference is the difference between intensities in the point of interest of the diffused image and the intensity resulting from convolving the original image with the approximating filter in the center of the subimage.

An alternative measure of the effect of approximating the filter kernels could be the shift of the isophotes in the resulting images. In some cases this will be a more informative measure — for instance indicating the expected feature detection imprecision resulting from the approximation.

The figures 7.3 and 7.4 display the Kullback-Leibler divergence and the local intensity difference respectively with histograms over the 10,000 samples. Each subfigure consists of the histogram and a smaller figure with the same data but different scaling and bin distribution.

All the large plots in the subfigures have the same scaling and the y-axis (the counts) has been scaled logarithmic. The smaller inserted are scaled according to the range of the data for that specific histogram and with linear y-axis.

In Figure 7.3 the histograms show mean Kullback-Leibler divergences between the approximation and the diffusion process. Pairs of mean KL divergence for (affine, isotropic) approximations for linear, RPM, and EED are respectively (0.0006, 0.0006), (0.0542, 0.0663), and (0.2898, 0.3566).

First, we see that linear and Perona-Malik diffusion can be approximated quite well. In the linear case that is trivial. If the linear diffusion process was a perfect approximation to convolving with a Gaussian the difference would be zero but due to the numeric limitation in discretisation that is not the case. The KL divergence of 0.0006 can been interpreted as the level of precision. In the case of edge enhancing diffusion the difference is quite high but as can be seen from the histogram a lot of cases can be approximated well but some approximations are very poor which on average gives a large difference.

Secondly, the numbers offer an ordering of the diffusion processes with EED as having behavior furthest away from linear. As expected, the aggressive anisotropic EED scheme is more non-Gaussian than Perona-Malik. Finally the affine approximations are evidently significantly better than the isotropic.

In order to put the measured Kullback-Leibler divergences into perspective, they can be compared to a simple example. Assume that we want to approximate a Gaussian distribution with another Gaussian distribution. Then the KL divergence is a function of the shift in mean and the change in standard deviation in the approximating distribution. For fixed, equal standard deviations σ, the KL divergence d as a function of the shift of the mean $\Delta\mu$ is $d = \frac{\Delta\mu^2}{2\sigma^2}$ (this is fairly easy to derive). For a given KL divergence d, the mean shift is then $\sqrt{2d}\ \sigma$. The KL divergences around 0.05 and 0.3, as measured for the approximations for RPM and EED above, then correspond to a mean shifts of respectively $\sqrt{0.1}\ \sigma$ and $\sqrt{0.6}\ \sigma$. Obviously the measured KL divergences are due to differences in shape not in shift of the mean (by definition this shift is zero). Even so, the implication is that KL divergences in this range are definitely measuring significant differences between the original and approximating distributions.
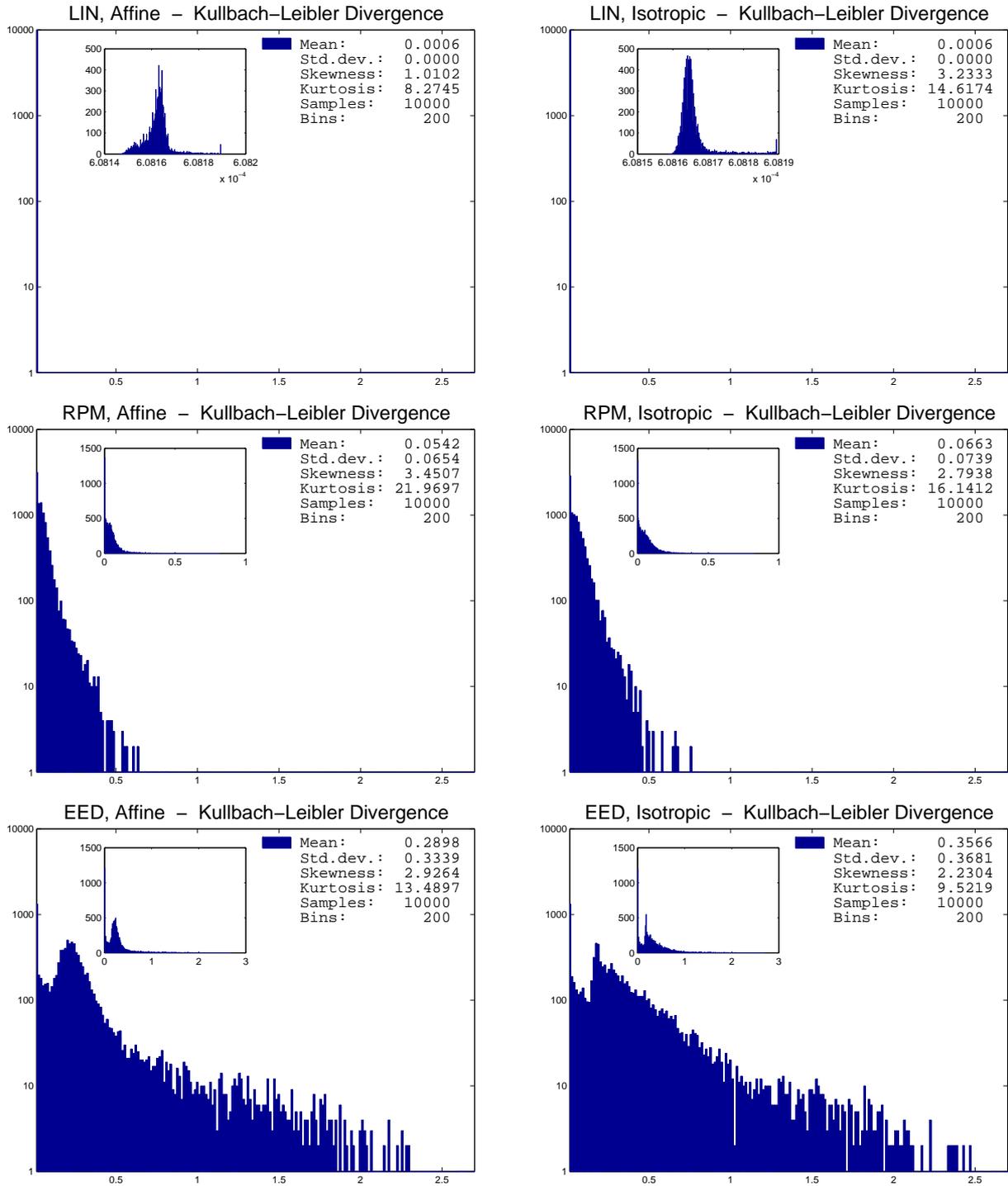
Figure 7.3: *Measuring Kullback-Leibler divergence between diffusion echo filter kernels and approximations. The columns show the affine and isotropic approximations respectively. The rows show Linear, RPM and EED. The small plots show same data — with different axis, bins and scaling. Note that the large plots are scaled logarithmically.*

In figure 7.4 the local intensity differences are presented, hence this plot assesses the actual outcome of applying the approximations compared to the original diffusion. The local intensity differences should be compared to the range of the image (which is zero to one due to normalization).

Due to the symmetry in the formulation a mean of zero is expected and confirmed within the precision.

Pairs of standard deviations for (affine, isotropic) approximations for linear, RPM, and EED are respectively (0.0002, 0.0002), (0.0067, 0.0096), and (0.0087, 0.0141) This supports the trend established in figure 7.3 which again indicates that the good approximations in the information theoretical sense actually gives good approximations of the diffusion processes. The standard deviations again support that the affine approximations give significantly better results than the isotropic as expected.

It should be noted that the evaluation above is specific to the chosen diffusion schemes and the choice of parameters for these. Especially the choice of diffusion time could influence the results since non-linear diffusion can become "arbitrarily non-linear" with increasing diffusion times. However, the following application evaluation adresses this issue indirectly since the multi-scale method requires both short and long diffusion times.

7.4   Application Evaluation

The evaluation on natural images above show that the approximations perform quite well — especially for the less non-Gaussian Perona-Malik scheme. However, the lack of a specific task makes the interpretation of the results slightly vague. As a counter-part to that we evaluate the performance of the approximated non-linear diffusion schemes in the interactive multi-scale watershed segmentation method used in the previous chapters.

The chosen task is segmentation of white matter tissue from brain scans. This task is chosen since the geometry of white matter tissue is quite complicated and therefore challenging for a non-committed segmentation method that uses no prior knowledge on the intensity distributions of brain matter.

The scans and corresponding ground truth segmentations are obtained from the *BrainWeb* site [Collins et al., 1998] (we use 9 slices with 10 slices between each from a simulated T1 MR brain scan, intensity non-uniformity level 20%, noise level 9%).
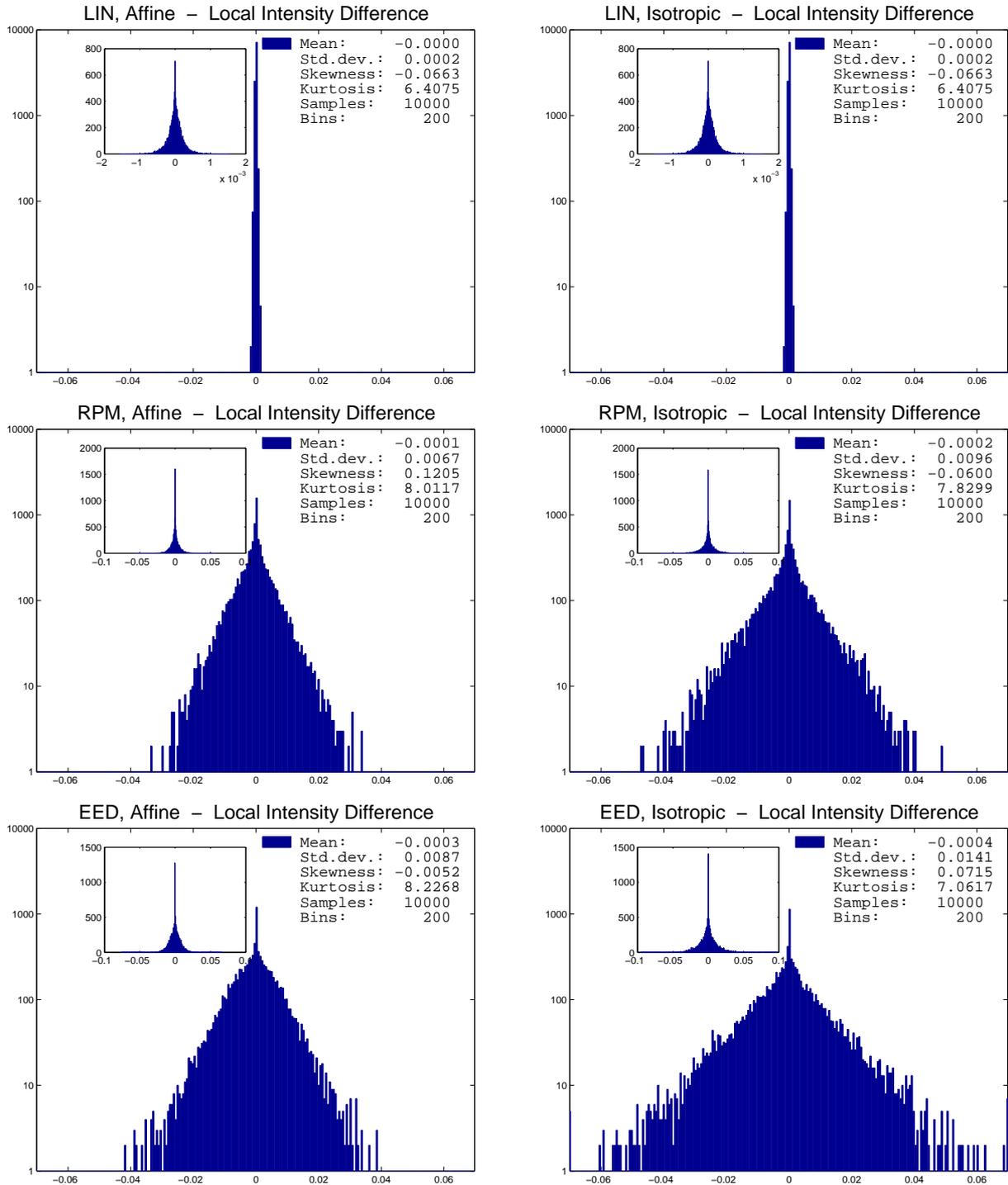
Figure 7.4: *Comparing local intensity difference between the diffusion process and the approximated diffusion process. Like figure 7.3 the columns are isotropic and affine approximations and the rows are Linear, RPM and EED. The intensities in the images are zero to one due to normalization.*

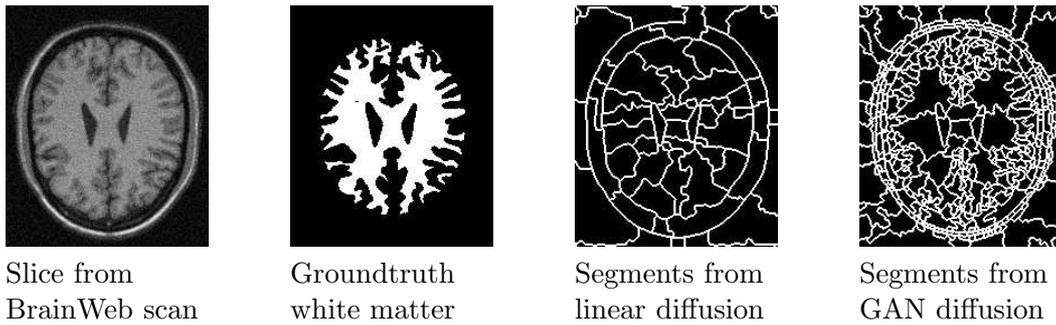| Slice from | Groundtruth | Segments from | Segments from |
| BrainWeb scan | white matter | linear diffusion | GAN diffusion |

Figure 7.5: *Linear diffusion makes the building blocks merge across brain structures before reasonably large segments are formed. Therefore smaller building blocks must be selected in order to segment the white matter tissue. The use of non-linear diffusion allows the building blocks to grow within the tissue boundaries. The practical implication is that where 18 action are required to select 70% of the white matter using building block created using linear diffusion, only 5 are necessary with GAN diffusion. For 80% the action counts are 39 and 11, respectively.*

7.4.1 Non-linear Diffusion in MSWS

The original MSWS (multi-scale watershed segmentation) method relies on linear Gaussian scale-space to simplify the image. This simplification determines how the watershed segments group into gradually larger building blocks corresponding to image structures at a given scale. In [Dam & Nielsen, 2000] the use of non-linear diffusion in MSWS is explored and evaluated (see chapter 3 for details). Figure 7.5 illustrates how the building blocks are better suited to the application at hand.

The evaluation is based on a count of the minimal number of selections and deselections of building blocks in the segmentation. The parameters for the non-linear schemes are determined such that this count is minimized.

A simple way to present the evaluation results is to normalize the performance (of the building blocks resulting from a non-linear scheme) with respect to the performance of linear diffusion. The descriptor is then the ratio of actions required compared to linear diffusion. Table 7.1 shows that for each action used on building blocks resulting from linear diffusion, it is on average only necessary to use 0.42 or 0.32 when the building blocks are from Perona-Malik or GAN.

We have chosen to disregard the edge enhancing diffusion scheme EED for this evaluation since it does not perform well (chapter 3 and [Dam & Nielsen, 2000]), and therefore is less interesting to approximate. The parameters for GAN given above produce a diffusion scheme that intuitively is in-between RPM and EED with a qualitative behavior closer to RPM.
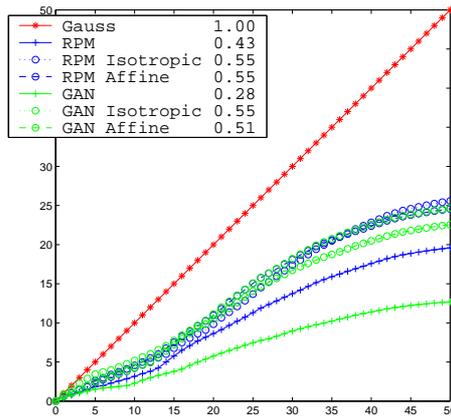
Figure 7.6: *The diffusion schemes are evaluated on the image in figure 7.5 on the number of actions needed to reach a given segmentation quality where the performance of linear diffusion building blocks is used as normalization. The graph shows the actions needed to reach the quality reached by a given number of actions using linear diffusion.*

7.4.2 Approximating Non-linear Diffusion in MSWS

In section 7.2.1 we present how to approximate the non-linear diffusion scheme using local isotropic and affine Gaussian filters. The approximation method can be used directly to generate filters approximating the non-linear diffusion in the segmentation method. For each scale we approximate the diffusion using the mean and variances recorded for each pixel. The approximating Gaussian filters are then applied and the gradient magnitude calculated. From there the MSWS method is exactly as described above.

The performance of the approximating filters can then be evaluated in the same manner as the non-linear schemes are evaluated in chapter 3 [Dam & Nielsen, 2000]. Figure 7.6 shows the performance of the approximating filters on the image in figure 7.5.

Table 7.1 shows how the filters perform on average on the entire test set consisting of 9 brain scan slices. As stated above, for each action used on building blocks resulting from linear diffusion, it is on average only necessary to use 0.42 or 0.32 when the building blocks resulted from Perona-Malik or GAN.

The approximating schemes perform significantly worse than the original non-linear schemes (they use between 27% and 73% more actions) but still maintain the main part of the advantage compared to linear diffusion (the ratios are between 0.50 and 0.65). As expected, the affine approximations perform significantly better than the isotropic. As a rule of thumb, it can be stated that for each 6 actions necessary to reach a given quality using linear diffusion, then only 2 actions are required with GAN diffusion, and 3 actions with the approximations of GAN.

Scheme	Ratio	Std. Dev.	Ratio compared to approx. scheme
Linear Gaussian	1.00	0.00	
Regularized Perona-Malik	0.42	0.15	
RPM approximated Isotropic	0.65	0.14	1.55
RPM approximated Affine	0.53	0.13	1.27
Generalized Anisotropic Non-linear	0.32	0.07	
GAN approximated Isotropic	0.56	0.08	1.73
GAN approximated Affine	0.50	0.10	1.55

Table 7.1: *The performance of the approximated non-linear diffusion scheme evaluated by the usability of the resulting multi-scale watershed segmentation building blocks.* **First column**: *All schemes are compared using the performance of linear diffusion building blocks as yardstick.* **Last column**: *The approximating schemes are compared to the schemes they are approximating — so the approximating schemes use between 27% and 73% more actions than ideally.*

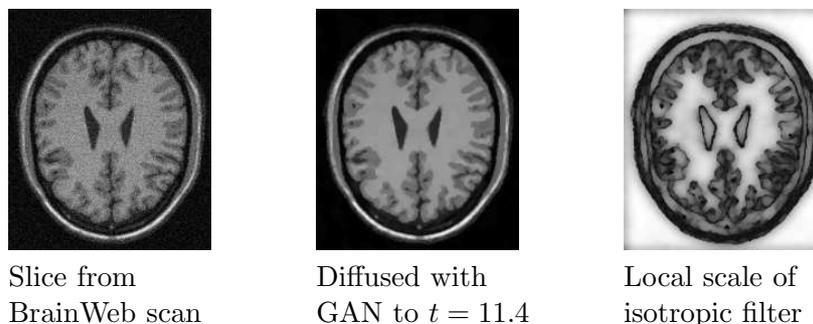| Slice from	Diffused with	Local scale of
BrainWeb scan	GAN to $t = 11.4$	isotropic filter

Figure 7.7: *The local scale as determined by the approximating isotropic filter. In areas far from a edge (as defined by the parameters for GAN) full diffusion is allowed. Around edges the diffusion is restricted, thus giving a lower local scale for the diffusion.*

7.4.3 Multi-scale Scale Selection

In the following we investigate the use of approximating filters in multi-scale watershed segmentation a bit further. Due to the simpler parameterisation of isotropic Gaussian filters we restrict ourselves to those for the analysis.

Implicit Scale Selection

At a given scale, in a given pixel the isotropic approximation filter implicitly determines a local scale. This local scale is illustrated in figure 7.7. It should be noted that we will implicitly treat the isotropic local Gaussian filters as if they only have a single parameter in each pixel. This is a simplification since the mean of the filter kernel is generally not centered at the given pixel and therefore needs to be represented as well.
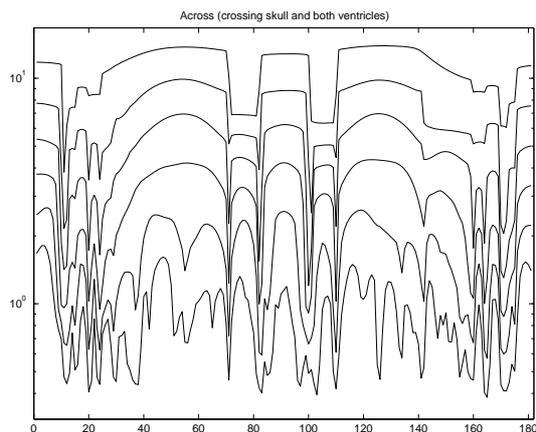
69

Figure 7.8: *The local scale selected on a line across the ventricles in the image in figure 7.5 for 6 scale levels. The crossings of the ventricles are evident: at low scale levels we see drops in local scale at each boundary — at high scale levels there is a single nose-dive across the entire ridge.*

Implicit Multi-scale Scale Selection

In a standard multi-scale setting based on linear diffusion, each scale level is a hyper-plane in scale-space with a constant scale value. For the approximating isotropic filters, the local scale is selected at each point — thereby the scale levels become hyper-surfaces. These surfaces are bounded upwards in scale by the maximal scale the non-linear diffusion can reach with the given diffusion time (e. g. GAN diffusion becomes linear diffusion when used on a constant starting image — it then reaches this maximal scale in all points). At points where the diffusion is restricted (e.g. due to edges in the image) the hyper-surface will drop down to lower scales.

This local scale selection is illustrated in the multi-scale setting in figure 7.8. Again, the illustration is a simplification since the local isotropic Gaussian filters can not be parameterized by a single parameter only. It is this behavior that a scale selection method would have to approximate in order to implement the approximating diffusion schemes directly. Note that the scale level surfaces do not intersect. The local scale is monotonically increasing as a function of the diffusion time — provided the diffusion scheme meets a suitable causality principle.

70

7.5 Conclusion

We evaluate the feasibility of approximating non-linear diffusion schemes with isotropic and affine Gaussian filters. The basis are the local diffusion filter kernels from the non-linear schemes that are extracted from the diffusion echo. These filter kernels are compared to the approximating Gaussian filters.

The approximating filters are evaluated on natural images. Both the measures that compare the filters directly (Kullback-Leibler divergence) and the measure that compare the resulting diffused images (intensity difference) show that especially Perona-Malik diffusion can be approximated quite well. Furthermore, the affine approximation shows significantly better performance than the simpler isotropic Gaussian filters.

Secondly, the approximating filters are evaluated on their performance in the multi-scale segmentation method for the task of segmenting white matter brain tissue. This evaluation offers more concrete results due to the specific task. The results show that for each 6 actions necessary to reach a given segmentation quality using linear diffusion, then only 2 actions are required with GAN diffusion, and 3 actions when the approximations of GAN are used. While the approximations perform significantly worse than the original non-linear scheme, the major part of the advantage compared to linear diffusion is retained. Again, this evaluation show that the affine filters approximate the non-linear schemes significantly better.

It should be noted that the approximating processes have not been optimized for solving the task but optimized to mimic the diffusion. Hence it is possible that another choice from the same class of filters would give a better performance in the segmentation task.

So, do we need the original non-linear diffusion schemes? The somewhat predictable answer is: If optimal performance is needed then yes — but if a relatively small decrease in performance is acceptable then no.

The obvious and necessary direction for future work is to establish methods for determining the parameters for the approximating filters directly from the desired diffusion process parameters and the local image structure.

Furthermore, if the approximating filters are to be used as fast approximations compared to running the original non-linear schemes, much work is needed in comparing efficient implementations of both.

Recycling in this Chapter

This chapter is an edited version of [Dam et al., 2003c].

Chapter 8

Exploring Shape Models

As stated in the introduction (chapter 1), the interactive multi-scale watershed segmentation method has provided inspiration for a large part of the work in this dissertation.

The segmentation program $\nabla Vision$ shows good performance already in the original formulation based on linear diffusion (see chapter 2). Furthermore, the work in chapters 3 and 5 shows that this performance can be significantly improved through the use of non-linear diffusion. The use of generalized anistropic non-linear diffusion improves the performance by specializing the building blocks towards a specific segmentation task.

The work on the diffusion echo started as an attempt to visualize and understand non-linear diffusion schemes better (chapter 6) — with the implied purpose of designing non-linear schemes for improving these building blocks even further. This inspired the corner enhancing diffusion scheme (presented in chapter 6). However, experiments show that there is an implicit limit for how much the diffusion schemes can be specialized towards a specific segmentation task. Larger specialization implies more terms and parameters that determines where the local diffusion is lowered in order to preserve local structures. This introduces two fundamental constraints. The introduction of more terms that lower the diffusion locally makes the schemes more sensitive to noise. And the introduction of more parameters complicates the optimization of the scheme for the specific segmentation task.

The conclusion is therefore that we cannot expect to improve the building blocks in the segmentation method drastically further through better use of non-linear diffusion along the lines projected in the previous chapters.

In particular, it will not be possible to reach near-automatic segmentation for most segmentation tasks. Experience shows that the level of automation is a critical parameter when considering using the method for actual clinical applications. Therefore higher specialization, committing the method even further to the given task, is desirable.

Now is a good time to put on our thinking caps.

8.1   Improving User Interaction

One possible path for improving the segmentation program is to optimize the interaction. It is not a problem that 50 building blocks are needed to sculpt the aorta if these can be selected in a single interaction. A simple example of this approach is illustrated in section 2.3 (see illustration page 11). This approach is not explored in this dissertation.

8.2   Adding a Shape Model

The use of non-linear diffusion in the multi-scale watershed approach described in the previous chapters is based on a local definition of object homogeneity. This allows specification of preferred local boundary behavior (such as a soft gradient threshold and curvature information). Global features of the objects are only indirectly addressed through the multi-scale approach.

A way to achieve a higher level of automation of the segmentation process would be to incorporate this higher level information. The natural type of information to add to the object model is shape information that incorporates prior information on the expected shape of the objects of the given segmentation task.

A simple analogy is the step from a *Snake* (or *Active Contour Model* [Kass & Terzopoulos, 1988]) to an *Active Shape Model* [Cootes et al., 1995]. The snake allows specification of expected local boundary behavior in terms of image gradients and boundary curvature — the ASM approach adds terms specifying the likelihood of the global shape and the preferred deformations in the optimization process.

The shape information could be incorporated in the non-linear diffusion process, as done in the *Diffusion-Snake* [Cremers et al., 2002b, Cremers et al., 2002a]. This approach is not pursued here. Alternatively, the shape information can be extracted from the watershed linking tree.

Augmenting the multi-scale watershed segmentation method with a shape model is no simple task. The multi-scale linking graph models the entire image with the desired object being represented by selected nodes in this graph. The way we handle creations of singularities ensures that the graph is actually a tree. The object is then implicitly represented through the selected and deselected nodes in the tree.

Unfortunately the object is in general not a single sub-tree — this would mean that it could be segmented using a single action in the interactive method. But is also makes the representation of the object more complicated. Specifically it means that the representation is a combined representation of both the object and the background.

Another complicating factor is that the topology of the tree will vary for different, but very similar objects. This makes the concept of correspondence difficult to define.

In the following, a few related shape matching methods are briefly introduced.

8.3 Related Shape Representation Methods

The most mature graph-based matching schemes are the methods developed for matching of shock graphs [Siddiqi et al., 1999]. The shock graph is generated by shrinking the boundary of the shape recording the collapses where opposing boundary points meet. A categorization of these points allows the process to be represented as a graph. This method has been used for shape matching and indexing with promising results.

Another promising approach is more related to classical scale-space methods using scale-space hierarchies of features [Demirci et al., 2003]. Here classical scale-space features such as blobs and ridges are linking and grouped into shape representations for hands.

Worth mentioning is also the curvature scale-space [Abbasi et al., 1999] that is similar to the shock graph but generates the graph representation by following zero-crossing for the boundary curvature as the curve evolves.

Also, promising research on using shape representations based on scale-space structures is currently underway (such as matching of top-points [Kanters et al., 2003] for face recognition).

Common for these shape matching methods is that they have severe limitations that restrict them from being the ideal choice for inspiring a shape model in the multi-scale watershed segmentation method. First, they define a representation of a shape — not an image (except for the representation based on top points in scale-space). Furthermore, the matching schemes do not provide a measure of probability for the query shape (this problem could be solved). Furthermore, the methods have no obvious way of providing a shape mean and corresponding deformations that can steer a segmentation process. Finally, the lack of an explicit statistical interpretation means that it is problematic to generate and sample new shapes.

The last two requirements, being able to model shape variations and being able to statistically sample new shapes, are extremely useful properties for a shape model.

Some preliminary work exists on achieving this for shape representations based on trees. This statistic method aims at providing the analogy of principal components analysis for trees, and is termed *functional data analysis* [Wang, 2003]. However, this approach imposes severe rescriptions since the trees are required to be binary and have correspondence based on the level-order position in the tree.

8.4 Generative Representation with Fixed Topology

A central problem in basing a shape representation on the multi-scale watershed linking tree is that varying topology for similar images of the same object class.

A way to attack this problem is to move from a shape representation derived from the image to a generative model that is deformed to fit the image data as well as possible. Using this approach, the topology is determined by the generative model and not the image.

This approach has been used successfully for a representation that is similar to the watershed linking tree. This medial representation is based on the Blum medial axis [Blum & Nagel, 1978]. The medial axis also offers highly varying topology for similar shapes [Pizer et al., 2003b], but in the *m-rep* setting the generative approach is used and the topology is fixed [Pizer et al., 2003a].

The topology can be fixed to the appropriate mean topology using a somewhat complicated method, where each shape in converted into *spherical harmonics* and then *Voronoi skeletons* representations [Styner & Gerig, 2001]. However, empirical experiments show that in many cases a very simple topology is adequate.

The desire to reach a shape modeling method for the multi-scale watershed method inspires the work on automatic generation of a medial shape model that is presented in the next chapter.

Recycling in this Chapter

This contents of this chapter are previously unpublished.

Chapter 9

Prostate Shape Modeling based on Principal Geodesic Analysis Bootstrapping

The use of statistical shape models in medical image analysis is growing due to the ability to incorporate robustly prior organ shape knowledge for tasks such as segmentation, registration, and classification.

Shape models are constructed from collections of segmented organs. Though interaction can ensure correspondence, it also introduces bias and ruins reproducibility — so a high degree of automation is desirable in the training process.

We present a novel shape model construction method via a medial shape representation. The essentially automatic iterative bootstrap method is based on an iterative bootstrap method that alternates between shape representation optimization and analysis of shape mean and variations.

The method is used to create a model from 46 segmented prostates with quantitatively and intutitively good results.

9.1   Introduction

Methods based on analysis of shape variation are becoming widespread in medical imaging. These methods allow incorporation of statistical prior shape knowledge in tasks where the image information alone often is not strong enough to solve the task automatically. The obvious example is the use of deformable models in segmentation, in which the preferred deformations are determined by a statistical shape model. Another important task is shape analysis and classification, in which a statistical shape model offers information for diagnostic methods.

Most statistical shape models consists of a mean shape with deformations. The mean and the corresponding deformations are constructed through statistical analysis of shapes from a collection of training data. Each shape in the training set is represented partially by the chosen shape representation, and analysis of the parameters for the representation gives the mean and variations [Cootes et al., 1995, Cremers et al., 2002a].

The best known model from this class is the *Active Shape Model* (ASM) [Cootes et al., 1995]. Here, the shapes are represented by a *point distribution model* (PDM) with given point-wise correspondence. The mean model is achieved through Procrustes alignment of the shapes followed by mean computation of each point in the model. *Principal component analysis* (PCA) is used to provide the variations.

This work pursues the medial shape representation known as the *m-rep* [Pizer et al., 1996]. The m-rep offers an intuitive representation of the shape by means of the sheet of sampled medial atoms. Compared to PDMs this representation is less simple since the parameter space is not Euclidean but consists of a combination of position, scaling, and orientation parameters. Standard PCA is therefore not applicable. However, the analogue of PCA has been developed for a more applicable space of shape representations. This is the *Principal Geodesic Analysis* (PGA) that applies to shape representations that form Lie groups [Fletcher et al., 2003a, Fletcher et al., 2003b].

A key element in constructing shape models is the representation of the shapes in the training collection. This should be done in a manner that defines/preserves correspondence across the population. For PDMs the simplest method is manual selection of the boundary points by an expert of the specific anatomical structure. In 2D this is a time-consuming and tedious process — in 3D it is even worse. However, this process can be automated. The approach by Davies [Davies et al., 2002] starts by generating boundary points from a spherical harmonics shape representation. This set of boundary points and their correspondences are then optimized through a *Minimum Description Length* (MDL) approach.

This work presents an essentially automatic shape modeling method. The core is a fully automatic bootstrap process that iteratively optimizes the shape model on a training collection and then derives the PGA mean and modes of deformation. Through the bootstrap iterations, the PGA mean and variations are optimized to allow automatic fitting of all shapes in the training collection.

The flavor of this work resembles the MDL method in [Davies et al., 2002]. The main difference is that the MDL approach starts the optimization process from representations with good training shape fit and poor correspondence. The MDL process then keeps the shape fits while optimizing the correspondence. The PGA bootstrap starts from a generative model with explicit correspondence but with poor fit to the individual training shapes. The bootstrap process then keeps the correspondence while optimizing the fit to the training shapes.

There exists another method for generating an m-rep mean model from a set of training shapes [Styner & Gerig, 2001] that uses a spherical harmonics representation followed by generation of the mean medial sheet from pruned Voronoi skeletons. Our approach is intuitively cleaner since m-rep is the only shape representation in play. Furthermore, our approach provides modes of variation as well as the shape mean.

We evaluate the presented PGA bootstrap method for the task of constructing a shape model for a population of prostates. The training collection consists of 46 cases where the prostates were segmented in the course of prostate cancer external-beam radiation treatment. Especially in CT scans with slice thickness 2mm or larger, the boundaries of the prostate have low contrast — therefore, prior knowledge in a statistical shape model is essential to making automatic segmentation possible. This prostate shape model is a key step towards a pelvis multi-object shape model that hopefully will achieve this goal.

Furthermore, current research aims at using the shape model to analyze the prostates in order to make a shape classification used for the radiation treatment planning. This is concentrated on the problematic *saddle-back* cases where the prostate reaches around the rectum — that complicates giving radiation to the prostate without hitting the rectum.

The contributions of this work are twofold: a) The presented PGA bootstrap method that allows essentially automatic generation of a shape model with mean and corresponding main modes of variation. b) The resulting prostate model that will be central in segmentation and analysis of prostates and eventually allow better radiation treatment planning.

9.2   The UNC Pelvis Collection

In radiation treatment, accurate segmentation of the prostate and surrounding organs is vital. Low image contrast across the prostate boundary makes this a difficult task.

The segmentation programs, *MASK* [Tracton et al., 1994] and *anastruct_editor*, from the PLan-UNC suite of radiotherapy treatment tools developed at UNC-CH Radiation Oncology, have slice-based contour drawing tools and visualization of reconstructed sagittal and coronal views. Both programs have interactive 12-bit intensity-windowing, which is required to find and draw both the prostate boundaries across from the bladder and the prostate's apex (the superior tip). The contours are scan-converted to labelled images, which introduces less than one pixel of error not significant to this shape study. Prostatic fat is included in the prostate's shape, as is seen in clinical practise, both because of the difficulty of finding the border between these and the prostate and the chance that these will contain significant counts of cancer cells. Seminal vesicles are excluded from the prostate.

Clinical contours are used but adjusted when obvious errors were found, such as missing contours, overlapping contours (eg. between the rectum and prostate), or just sloppy contouring – all of these shortcuts are due to clinical time constraints and are not perceived to affect clinical care but can affect shape studies.

The ungated CT scans are acquired from non-immobilized supine patients at UNC Healthcare (Chapel Hill, NC, USA) and Western Wake Radiology (Cary, NC, USA) on Siemens Somotom 4+ scanners without administering contrast agents. Note that while the prostate is quite hard, the multi-object statistics eventually produced will be sensitive to prostate shifts based on the patient's position and the CT couch shape (flat vs. rounded) because of the surrounding tissues' malleability. These effects should not affect this shift-invariant prostate shape analysis.

Figure 9.1: *Sagittal slices of the manual segmentations of rectum, prostate, and bladder from two cases in the UNC pelvis collection.*

Retrospective patient images are selected from the patient archives based on technical criteria, such as adequate image quality and anatomical coverage (the entire bladder down through the prostate apex), as well as shape and anatomical considerations such as very large bladders, prosthetic hips, or surgical procedures proximal to the prostate, yielding "normal cancerous" prostates.

The collection has 46 sets with manual segmentations for prostate, bladder, and rectum. All cases are diagnosed with prostate cancer so the resulting shape model will not necessarily model prostates in general. For instance, an increase of the size of the prostate is common for prostate cancer patients. However, since the shape model is to be used for segmentation and analysis of patients diagnosed with prostate cancer, this bias towards cancerous prostates is desirable.

The volumes of the prostates varies from $12cm^3$ to $144cm^3$. Figure 9.1 illustrates the large variation in shape.

9.3   Medial Shape Representation: m-rep

We use a medial representation, m-rep, to model shape. Here, we briefly review the geometry of m-reps and the deformable m-rep framework for image segmentation [Pizer et al., 2003a, Joshi et al., 2002].

9.3.1   m-rep Geometry Overview

The shape representation we use is based on the medial axis of Blum [Blum & Nagel, 1978]. In this framework, a 3D geometric object is represented as a set of connected continuous medial sheets, which are formed by the centers of all spheres that are interior to the object and tangent to the object's boundary at two or more points. Here we focus on 3D objects that can be represented by a single medial figure.
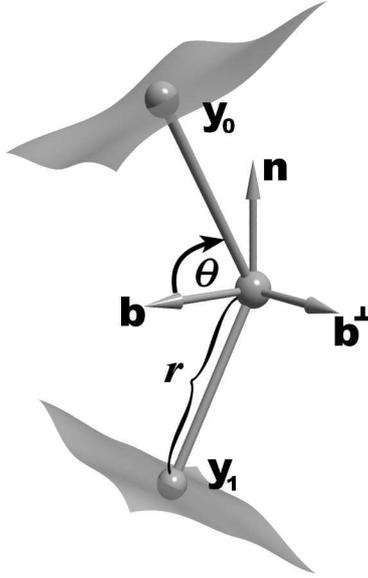
Figure 9.2: *Medial atom with a cross-section of the boundary surface it implies.*

We sample the medial sheet $\mathcal{M}$ over a spatially regular lattice. Each sample point also includes first derivative information of the medial position and radius. The elements of this lattice are called *medial atoms*. A medial atom (figure 9.2) is defined as a 4-tuple $\mathbf{m} = \{\mathbf{x}, r, \mathbf{F}, \theta\}$, consisting of: $\mathbf{x} \in \mathbb{R}^3$, the center of the inscribed sphere, $r \in \mathbb{R}^+$, the local width defined as the radius of the sphere, $\mathbf{F} \in \mathbf{SO}(3)$ an orthonormal local frame parameterized by $(\mathbf{b}, \mathbf{b}^\perp, \mathbf{n})$, where $\mathbf{n}$ is the normal to the medial sheet, $\mathbf{b}$ is the direction in the tangent plane of the fastest narrowing of the implied boundary sections, and $\theta \in [0, \pi)$ the object angle determining the angulation of the implied sections of boundary relative to $\mathbf{b}$. The medial atom implies two opposing boundary points, $\mathbf{y}_0, \mathbf{y}_1$, with respective boundary normals, $\mathbf{n}_0, \mathbf{n}_1$, which are given by

$$
\begin{aligned}
\mathbf{n}_0 &= \cos(\theta)\mathbf{b} - \sin(\theta)\mathbf{n}, & \mathbf{n}_1 &= \cos(\theta)\mathbf{b} + \sin(\theta)\mathbf{n}, \\
\mathbf{y}_0 &= \mathbf{x} + r\mathbf{n}_0, & \mathbf{y}_1 &= \mathbf{x} + r\mathbf{n}_1.
\end{aligned}
$$

Given an m-rep figure, we fit a smooth boundary surface to the model. We use a subdivision surface method [Thall, 2002] that interpolates the boundary positions and normals implied by each atom.

9.3.2 Segmentation using m-reps

Following the deformable models paradigm, an m-rep model $\mathbf{M}$ is deformed into an image I by optimizing an objective function, which we define as

$$ F(\mathbf{M}, I) = L(\mathbf{M}, I) + \alpha\, G(\mathbf{M}). $$

The function L, the *image match*, measures how well the model matches the image information, while G, the *geometric typicality*, gives a prior on the possible variation of the geometry of the model. The relative importance of the two terms is weighted by $\alpha \geq 0$.

This objective function is optimized in a multiscale fashion. That is, it is optimized over a sequence of transformations that are successively finer in scale. Here we will only be concerned with two levels of scale: the figural level, and the medial atom level. At the figural level the transformation we use is a similarity transformation plus an elongation of the entire figure. At the atom level each medial atom is independently transformed by a similarity plus a rotation of the object angle.

m-rep models are fit to binary segmentation images of the prostates. These binary images are blurred slightly to smooth the objective function, which is optimized with a conjugate gradient method. The image match term of the objective function is computed as a correlation with a Gaussian derivative kernel in the normal direction to the object boundary:

$$L(\mathbf{M}, I) = \int_{\mathcal{B}(\mathbf{M})} \int_{-\epsilon}^{\epsilon} \partial_t G(t) \, I \left(\mathbf{s} + (t/r)\mathbf{n}\right) \, dt \, d\mathbf{s},$$

where $\mathbf{s}$ is a parameterization of the boundary $\mathcal{B}(\mathbf{M})$, $\partial_t G$ is the Gaussian derivative kernel, r is the radius function, and $\mathbf{n}$ is the boundary normal.

The geometric typicality term is defined as

$$G(\mathbf{M}) = (1 - \beta) \, P(\mathbf{M}) + \beta \, N(\mathbf{M}), \tag{9.1}$$

where $\beta \in [0, 1]$ is a weighting term. The function P measures the change in the boundary from the previous level of scale:

$$P(\mathbf{M}) = -\int_{\mathcal{B}(\mathbf{M})} \frac{||\mathbf{s} - \mathbf{s}_0||^2}{r^2} d\mathbf{s},$$

where $\mathbf{s}_0$ is the initial position of the boundary at this scale level. The function N seeks to keep medial atoms in the same relationship with their neighboring atoms. It is defined as

$$N(\mathbf{M}) = -\int_{\mathcal{B}(\mathbf{M})} \frac{||\mathbf{s} - \mathbf{s}'||^2}{r^2} d\mathbf{s},$$

where now $\mathbf{s}'$ is the boundary surface of the model in which the current medial atom is in the position predicted by its neighbors. The neighbor term is only used at the atom scale level, i.e., $\beta = 0$ during the figural level. The role of the neigbor penalty term is to keep the shape nice locally — comparable to the curvature term in active contour models.

9.4   Shape Modeling

The goal is to arrive at a shape model that describes the shape variation within the population of prostates. A shape model is a parametric shape representation with predefined rules for deformation that allow the model to represent the class of shapes encountered. In medical image analysis the shape classes are typically defined by shapes of organs or other anatomical structures. This introduces some basic modeling trade-offs:

Compactness vs Accuracy

The model should be a more compact representation than the basic representation of the organ (typically a binary image). This introduces a trade-off between compactness and precision in the representation. The desire for compactness suggests coarse sampling of the basic elements of both the underlying shape representation (boundary points, medial points, control points in splines, spherical harmonic coefficients etc.), and the possible deformations (similarity transformation, PCA modes etc.). The desire for accuracy suggests finer sampling.

Generality vs Specificity

The deformations should be flexible enough to allow the shape model to fit the organs up to the precision possible with a specific compactness/sampling. This ensures the generality of the model. The opposing property is specificity, that states that the model should not deform into shapes not encountered in the organs. This trade-off inspires a statistical approach that allows a soft transition between likely and un-likely shapes.

Correspondence vs Accuracy

Another key property of shape models is the ability to give an explicit or implicit coordinate system on the shapes that offers correspondence of locations among the shapes. This introduces yet another requirement on the allowed deformations of the shape model. Being able to fit the individual organs shape is not sufficient — the model must also ensure that anatomically corresponding locations on the organs are equipped with corresponding locations given by the coordinate system of the shape model. This suggest that the shape model should restrict large deformations that violate correspondence — and thereby the attainable accuracy is limited.

9.4.1   First Attempt: The Potato Model

The segmentation program *Pablo* provides a user interface that allows construction of m-rep models and optimization of the parameters such that the constructed model is fitted to a specific training case [Pizer et al., 2003a].

The Potato in figure 9.3 is such a handcrafted m-rep model based on inspection of a subset of the prostate collection and some experimentation. If we can automatically deform this model into all the prostates in the training collection, we actually have the desired shape model.

Batch Optimization of m-reps

For this shape modeling we developed a batch non-interactive back-end to Pablo. It is of little theoretical interest, but without this automatic fitting program, our results would not be realistically possible.

The fitting program reads m-rep optimization constraints from an options file and performs the corresponding m-rep optimization steps. The process is completely free from interaction. The initial hand placement of the model is performed automatically by translating to the center of gravity and scaling to the volume of the relevant training case (represented as a binary image).

To fit the Potato in all the training prostates, the batch optimizer is run with $\alpha = 0.1$ and $\beta = 0.75$ — high confidence in the segmentations, and low importance of geometric typicality.
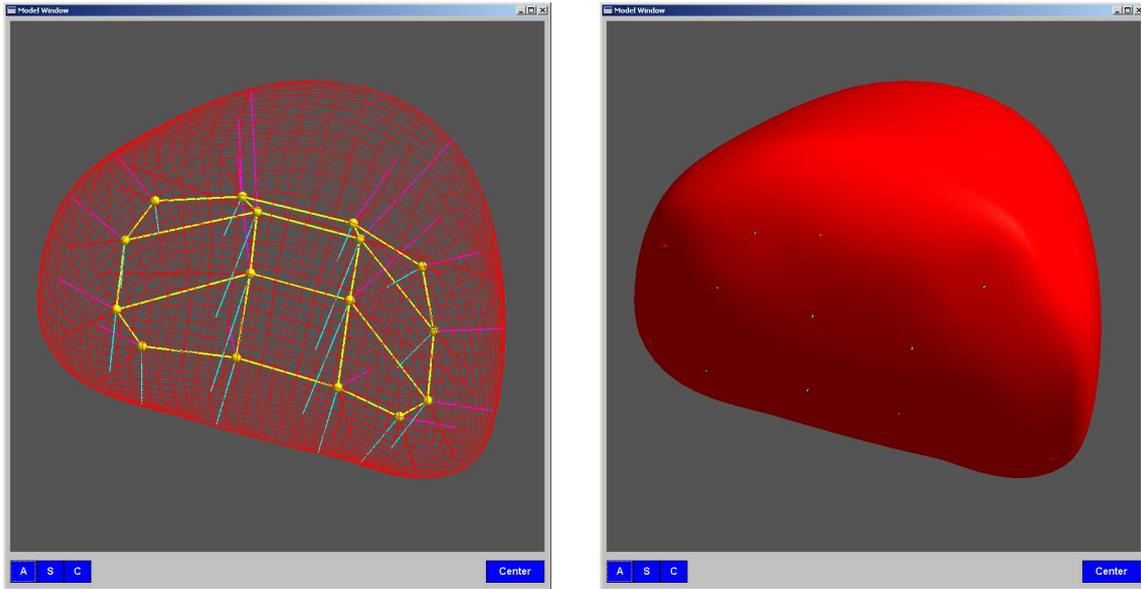
Figure 9.3: *The Potato m-rep model.* **Left**: *The medial grid with the implied boundary as wireframe.* **Right**: *The model as a solid surface. The viewpoint is from above, behind the prostate.*

Rating the Potato

In order to determine whether the Potato can be deformed satisfactorily into the training prostates we look at the image match values for the 46 cases. Recall that the match value is a normalized correlation measure in the range -1 to 1 with approximately 0.95 as the practical maximal value. Furthermore, heuristic experience from other organs show that values above 0.80 are quite good.

The Potato match values for the 46 cases are in the range 0.60 – 0.84 with a mean of 0.75. This is not acceptable for our task. Figure 9.4 illustrates the worst and best case.

The geometric penalty prevents the model from deforming enough to fit the worst cases satisfactorily. Therefore the model certainly is not general enough, and the Potato cannot be considered a good prostate prototype.

However, even though the batch optimization of the Potato does not give satisfying shape representations of all the prostates in the training collection, it does ensure a rough fit in most cases. In the following we present the theory that allows us to use these rough fits to improve the Potato.
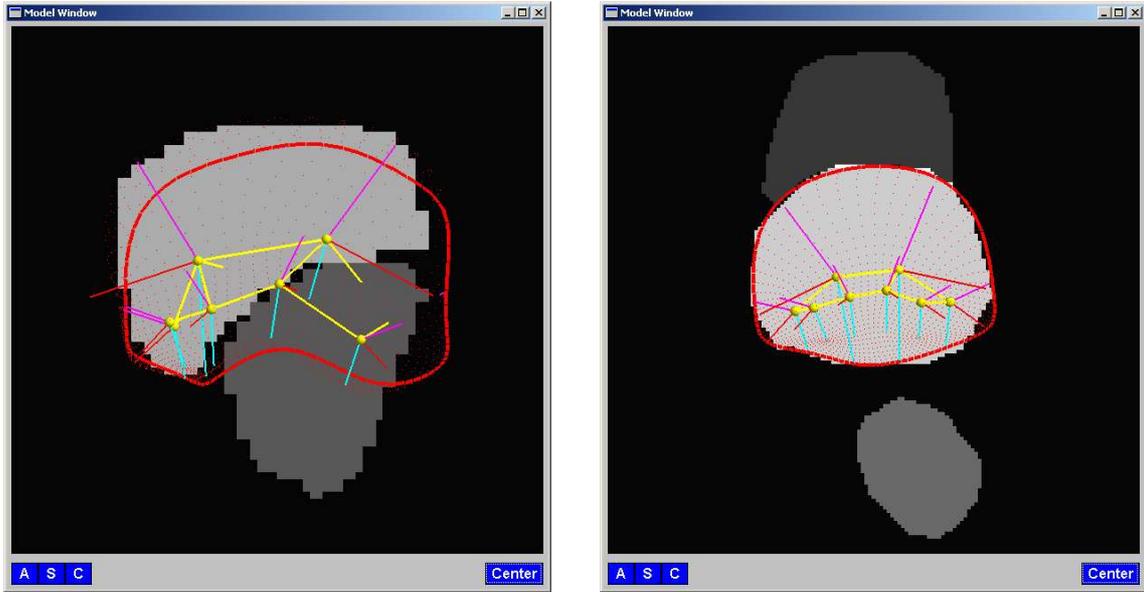
Figure 9.4: *Axial slices from the worst (left, image match 0.60, volume 27 cm^3) and best (right, image match 0.84, volume 127 cm^3) fitting cases for the deformed Potato model.*

9.5 Principal Geodesic Analysis

Principal geodesic analysis (PGA), introduced in [Fletcher et al., 2003b], is a generalization of principal component analysis (PCA) to curved manifolds. It is shown in [Fletcher et al., 2003a, Fletcher et al., 2003b] that m-rep models form a Lie group, and the necessary algorithms for computing the mean and PGA of a collection of m-rep models are given. We review these results briefly here.

9.5.1 Lie Groups

First we present a brief overview of Lie groups (see [Duistermaat & Kolk, 2000] for a detailed treatment). A Lie group G is a differentiable manifold that also forms an algebraic group, where the two group operations,

$$\mu : (x, y) \mapsto xy \qquad : \quad G \times G \to G \qquad \qquad \textit{Multiplication,}$$
$$\iota : x \mapsto x^{-1} \qquad : \quad G \to G \qquad \qquad \textit{Inverse,}$$

are differentiable mappings.

Let e denote the identity element of a Lie group G. The tangent space at e, T_eG, forms a Lie algebra, which we will denote by $\mathfrak{g}$. The exponential map, $\exp : \mathfrak{g} \to G$, provides a method for mapping vectors in the tangent space T_eG into G. Given a vector $\mathbf{v} \in \mathfrak{g}$, the point $\exp(\mathbf{v}) \in G$ is obtained by flowing to time 1 along the unique one-parameter subgroup emanating from e with initial velocity vector $\mathbf{v}$. When the Lie group is given a compatible Riemannian metric, this one-parameter subgroup is the unique geodesic at e with velocity $\mathbf{v}$.

84

The exponential map is a diffeomorphism on a neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in G. The inverse of the exponential map is called the log map. The geodesic distance between two points $g, h \in G$ is given by $||\log(g^{-1}h)||$.

As shown in [Fletcher et al., 2003b], the set of all medial atoms forms a Lie group $M = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbf{SO}(3) \times \mathbf{SO}(2)$, which we call the *medial group*. Likewise, the set of all m-rep models containing n medial atoms forms a Lie group M^n, i.e., the direct product of n copies of M.

9.5.2 m-rep Means

The Riemannian distance between m-rep models $\mathbf{M}_1, \mathbf{M}_2 \in M^n$ is given by

$$d(\mathbf{M}_1, \mathbf{M}_2) = ||\log(\mathbf{M}_1^{-1}\mathbf{M}_2)|| \tag{9.2}$$

Thus the intrinsic mean of a collection of m-rep models $\mathbf{M}_1, \ldots, \mathbf{M}_N$ is the minimizer of the sum-of-squared geodesic distances:

$$\mu = \arg\min_{\mathbf{M} \in M^n} \sum_{i=1}^{n} ||\log(\mathbf{M}_i^{-1}\mathbf{M})||^2$$

As shown in [Fletcher et al., 2003b], the mean model may be computed by the following iterative gradient descent algorithm (algorithm 4).

Algorithm 4 *m-rep Mean*

Input: $\mathbf{M}_1, \ldots, \mathbf{M}_n \in M^n$, m-rep models
Output: $\mu \in M^n$, the intrinsic mean
  $\mu = \mathbf{M}_1$
  Do
    $\Delta\mathbf{M}_i = \mu^{-1}\mathbf{M}_i$
    $\Delta\mu = \exp(\frac{1}{n}\sum_{i=1}^{n}\log(\Delta\mathbf{M}_i))$
    $\mu = \mu\Delta\mu$
  While $||\log(\Delta\mu)|| > \epsilon$.

9.5.3 PGA of m-reps

Principal components of Gaussian data in $\mathbb{R}^n$ are defined as the projection onto the linear subspace through the mean spanned by the eigenvectors of the covariance matrix. If we consider a general manifold, the counterpart of a line is a geodesic curve, that is, a curve with minimal length between two points. In the Lie group M^n geodesics can be computed via the exponential map. Given a tangent vector $\mathbf{v}$ in the Lie algebra $\mathfrak{m}^n$, the geodesic starting at the identity with initial velocity $\mathbf{v}$ is given by $\gamma : \mathbb{R} \to M^n$, where $\gamma(t) = \exp(t\mathbf{v})$. Similarly, the curve $x \cdot \gamma(t) = x \cdot \exp(t\mathbf{v})$ is a geodesic starting at the point $x \in M^n$.

Algorithm 5 *Algorithm: m-rep PGA*

Input: m-rep models, $\mathbf{M}_1, \ldots, \mathbf{M}_N \in M^n$

Output: Principal directions, $\mathbf{u}^{(k)} \in \mathfrak{m}^n$

       Variances, $\lambda_k \in \mathbb{R}$

  $\mu = $ mean of $\{\mathbf{M}_i\}$

  $\mathbf{x}_i = \log(\mu^{-1}\mathbf{M}_i)$

  $\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i^T$

  $\{\mathbf{u}^{(k)}, \lambda_k\} = $ eigenvectors/eigenvalues of $\mathbf{S}$.

As shown in [Fletcher et al., 2003a], the covariance structure of a Gaussian distribution on M^n may be approximated by a covariance matrix $\mathbf{\Sigma}$ in the Lie algebra $\mathfrak{m}^n$. The eigenvectors of this covariance matrix correspond via the exponential map to geodesics on M^n, called *principal geodesics*. The principal geodesic analysis (PGA) on a population of m-rep figures, $\mathbf{M}_i, \ldots, \mathbf{M}_N \in M^n$, is thus computed by algorithm 5.

Analogous to linear PCA models, we may choose a subset of the principal directions $\mathbf{u}^{(k)} \in \mathfrak{m}^n$ that is sufficient to describe the variability of the m-rep shape space. New m-rep models may be generated within this subspace of typical objects. Given a set of coefficients $\{\alpha_1, \ldots, \alpha_l\}$, we generate a new m-rep model by

$$\mathbf{M} = \mu \exp \Big(\sum_{k=1}^{l} \alpha_k \mathbf{u}^{(k)} \Big),$$

where α_k is chosen to be within $[-3\sqrt{\lambda_k}, 3\sqrt{\lambda_k}]$.

9.6   Shape Model Bootstrapping

The shape model bootstrapping method is now quite simple. The batch fitting process is used to give rough representations of each shape. The PGA is used to generate the mean model from the 46 fitted models. This mean model is then used as the starting model in the next iteration to fit the shapes using the batch fitting process, and the bootstrapping method is iterated.

The underlying assumption is that the mean of the roughly fitting models will be a better prototype than the initial Potato. As the bootstrap iterations progress the generated mean models will hopefully converge to a good prototype.

9.6.1   Bootstrapping the Potato using PGA Mean

The graph in figure 9.5 shows the development of the image matches during bootstrapping. The image matches become excellent as the bootstrap progresses — even the worst match is good. The resulting mean model can be deformed automatically into the shapes in the training collection, so we now have a prostate shape model as desired. Since the allowed deformations are exactly the same as in the initial attempt in section 9.4.1, the improved performance is exclusively due to the resulting mean shape being a better prostate prototype — as expected.
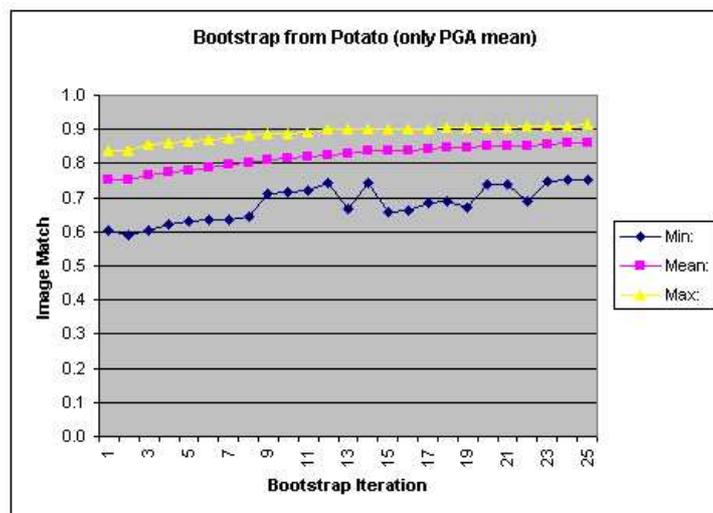
Figure 9.5: *The evolution of the worst, best, and mean image match for the 46 cases during bootstrap starting from the Potato.*

However, the method started from a reasonably good initial model being the manually constructed Potato. For the prostate, this is not a problem. However, it is not desirable if we need to divine a new suitable vegetable from which to start the bootstrapping for each new organ to be modeled.

9.6.2 Bootstrapping the Generic using PGA Mean

In order to see how dependent the bootstrap method is on the initial model, alternative m-rep models with the same 4x4 atom grid were used. The *Generic* is the default 4x4 slab model that Pablo generates as a starting model for building handcrafted models. Figure 9.6 shows how the bootstrap is virtually independent of the choice of starting model — except for the number of atoms that reflects the compactness vs accuracy issue.

9.6.3 Bootstrapping using PGA Mean and Modes

The principal geodesic analysis provides the shape mean and well as the principal modes of variation. Above, we only use the mean in the bootstrap. However, the batch fitting process can also use the PGA modes in the figural level transformation in place of the elongation.

The expectation is that these trained, global deformations steer the model to the desired image match in fewer bootstrap iterations. Figure 9.7 shows exactly that expected behavior where the 10 top PGA modes are used.
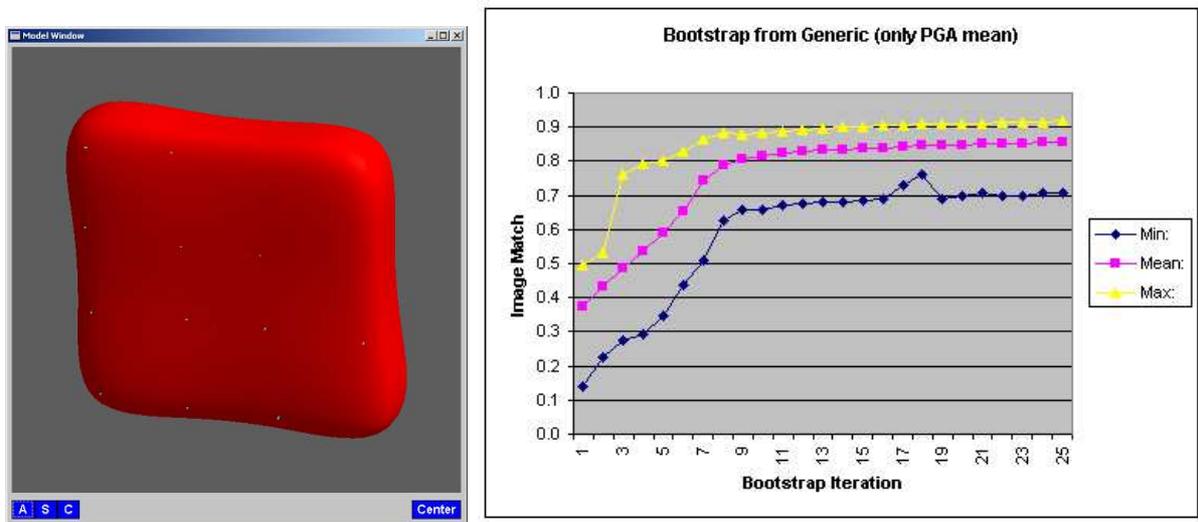
87

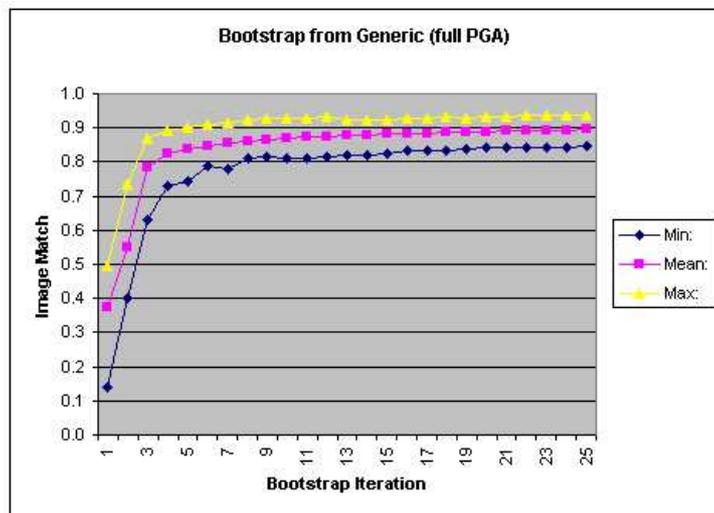Figure 9.6: *The Generic initial model and the resulting bootstrap image match evolution.*

Figure 9.7: *The bootstrap evolution starting from the Generic using global PGA modes in the figural level optimization.*

9.6.4 Convergence

In this work, we address the question of convergence superficially. The image match values above appear to be converging. Convergence of the PGA mean and modes is considered future work. Furthermore, we may want to constrain the evolution of the model such that the converging model has specific correspondence properties. Our expectation is that eventually the method will converge to the same resulting mean model independent of the choice of starting model and optimization options — using different paths through model parameter space.

Here we use the heuristic approach of running the bootstrap until the image match values cease to improve significantly. We therefore use the resulting model from 10 bootstrap iterations starting from the Generic using PGA deformations in the optimization (see figure 9.7).

9.6.5 Exclusion of Outliers

In particular during the early bootstrap iterations, the optimization process reaches poor image match values for some training cases. This is simply due to too much difference between the initial model and these training prostate shapes. As a result, the optimized model then does not correspond to a prostate shape within a reasonable precision.

Therefore, it can be argued that these non-prostate shapes should be excluded from the calculations of the PGA mean and modes. Much like outliers are often identified and excluded during statistical methods in general. This identification could simply be done by defining an image match threshold based on e.g. the variance of these values. In this work we experimented superficially with outlier exclusion and concluded that the PGA bootstrap method performs robustly without this extra step.

9.6.6 Resulting Prostate Shape Model

The resulting Prostate shape model then consists of the mean shape and the deformations illustrated in figure 9.8. The 10 modes of variation include 95% of the variation in the training collection. This ensures little need for atom optimization in the segmentation process. The automatic fitting achieves image matches in the range 0.81–0.93 with mean 0.87. That this is significantly better than the original attempt with the Potato model can be seen in figure 9.9.

The shape model is trained on binary images — therefore the resulting model is only evaluated for segmenting binary images. However, in the full m-rep segmentation framework, the shape model is combined with profile models for the local boundary instead of the just using the Gaussian derivative profile (as done in [Rao, 2003] for instance).

Apart from being directly applicable for segmentation, the shape model — and especially the condensed PGA parameterization — is also directly applicable for shape classification.

Furthermore, the hierarchical m-rep optimization framework also allow easy generalization of the model building method. The work presented here is limited to single-figure m-rep models. The shape model method could easily be extended to multi-object multi-object ensembles due to the modular optimization method where each individual step could be modeled using the PGA bootstrap approach introduced here.
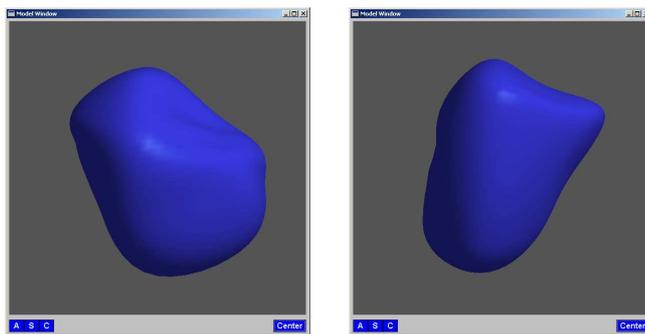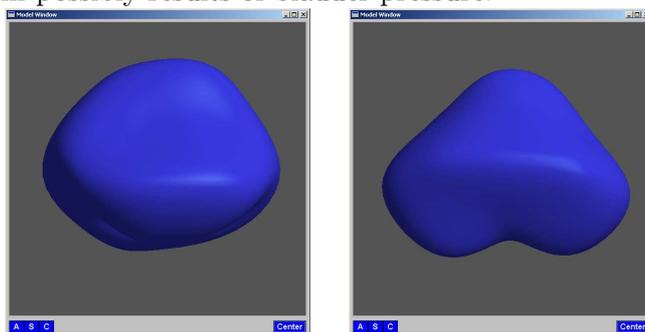
Mode 1: Sagittal view (from the side) possibly showing effects of varying pressure from bladder (is top right).
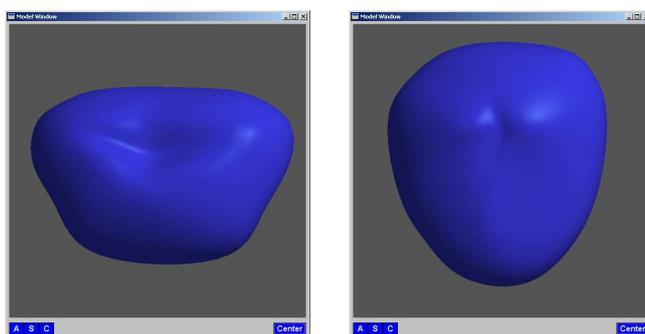
Mode 2: Coronal view: Laurel/Hardy deformation.
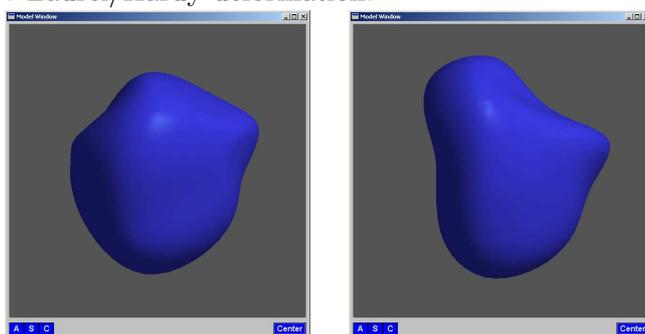
Mode 3: Sagittal: again possibly results of bladder pressure.

Mode 4: Axial view: Saddle-back where the prostate curves around the rectum — a known problematic behavior that complicates radiation treatment.

Figure 9.8: *The four primary modes of deformation in the Prostate model. Each mode is illustrated by $\pm$ 1.5 standard deviation images.*
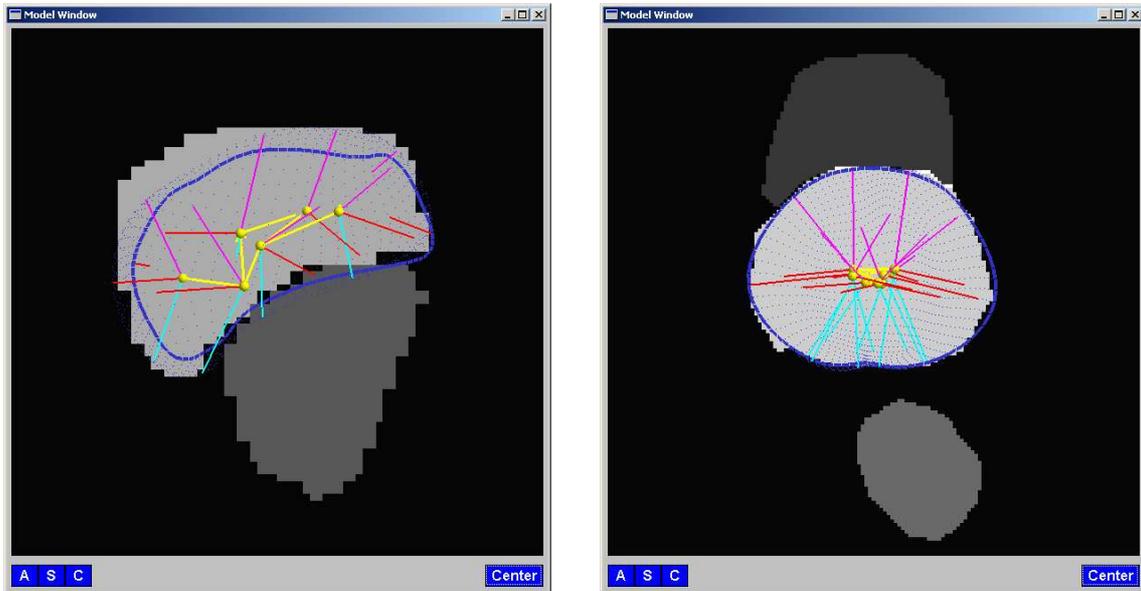
Figure 9.9: *The Prostate model on the cases from figure 9.4 with image matches 0.85 and 0.91 (was 0.60 and 0.84). Since left is a worst case this is highly satisfying.*

9.7 Conclusion

We present a novel shape model construction method using a medial shape representation. The method is essentially automatic based on an iterative bootstrap method that alternates between shape representation optimization and principal geodesic analysis of shape mean and variations. The method constructs an m-rep shape model consisting of a mean and corresponding main modes of variation.

The non-automatic step is the choice of sampling in the medial sheet. We have chosen a 4x4 atom grid that appears to be a suitable compromise between compactness and accuracy.

The method is evaluated through construction of a prostate shape model from a training collection of 46 manually segmented prostates. The resulting model has good quantitative performance. In addition, the modes of variation show deformations that corresponds intuitively well with known prostate behavior. We are especially pleased with the presence of the saddle-back variation in the fourth mode.

Future work is centered on ensuring desirable convergence of the shape model in the bootstrap iterations. Furthermore, we plan to evaluate the method against the method that uses the MDL approach to generate an ASM [Davies et al., 2002]. Central points to evaluate are compactness, correspondence, and legality (how likely are illegal models).

Finally, we look forward to applying the method on other anatomical organs (among others kidneys, hearts and various brain structures).

Acknowledgements

Recycling in this Chapter

This chapter is an edited version of [Dam et al., 2003b].

Chapter 10

Revisiting Prostate Shape Modeling

The previous chapter left the question of convergence for the iterative bootstrap PGA modeling method as an open problem. Even though the method showed great promise, convergence is obviously highly desirable for such a method.

10.1 Convergence of the Bootstrap PGA Model

In section 9.6, we conclude that the match values appear to be converging during the bootstrap iterations. This does not imply that the underlying mean model is converging (or the variations for that matter) — figure 10.1 illustrates that this is actually not the case. The observed behavior is highly undesirable not only due to the lack of model convergence. The medial sheet of the mean model does not stay inside the prostate, so the model is actually no longer medial. As a consequence, shape analysis (eg. measurements of width or bending) is not possible.

10.1.1 Relative vs Absolute Geometric Penalties

The force that drives the evolution of the mean model in the bootstrap is minimization of the objective function in the optimization step. This objective function is composed of the image match and the geometric typicality (see section 9.3 for details). The image match is maximized until the geometric penalty prevents further deformation of the model. It is important to note that this geometric restriction is relative to the current mean model in every bootstrap iteration. This means that the geometric penalties are only *relative* — they limit the change of the mean model in each bootstrap iterations, but does not limit the total deformation. In principle, the absence of *absolute* geometric penalties allows the model to deform arbitrarily given enough bootstrap iterations.

Therefore, during the bootstrap iterations, the actual driving force is the image match exclusively. The behavior displayed in figure 10.1 is then probably simply due to the fact that the image match to prostate shapes is improving when the end atoms are further away from the boundary.
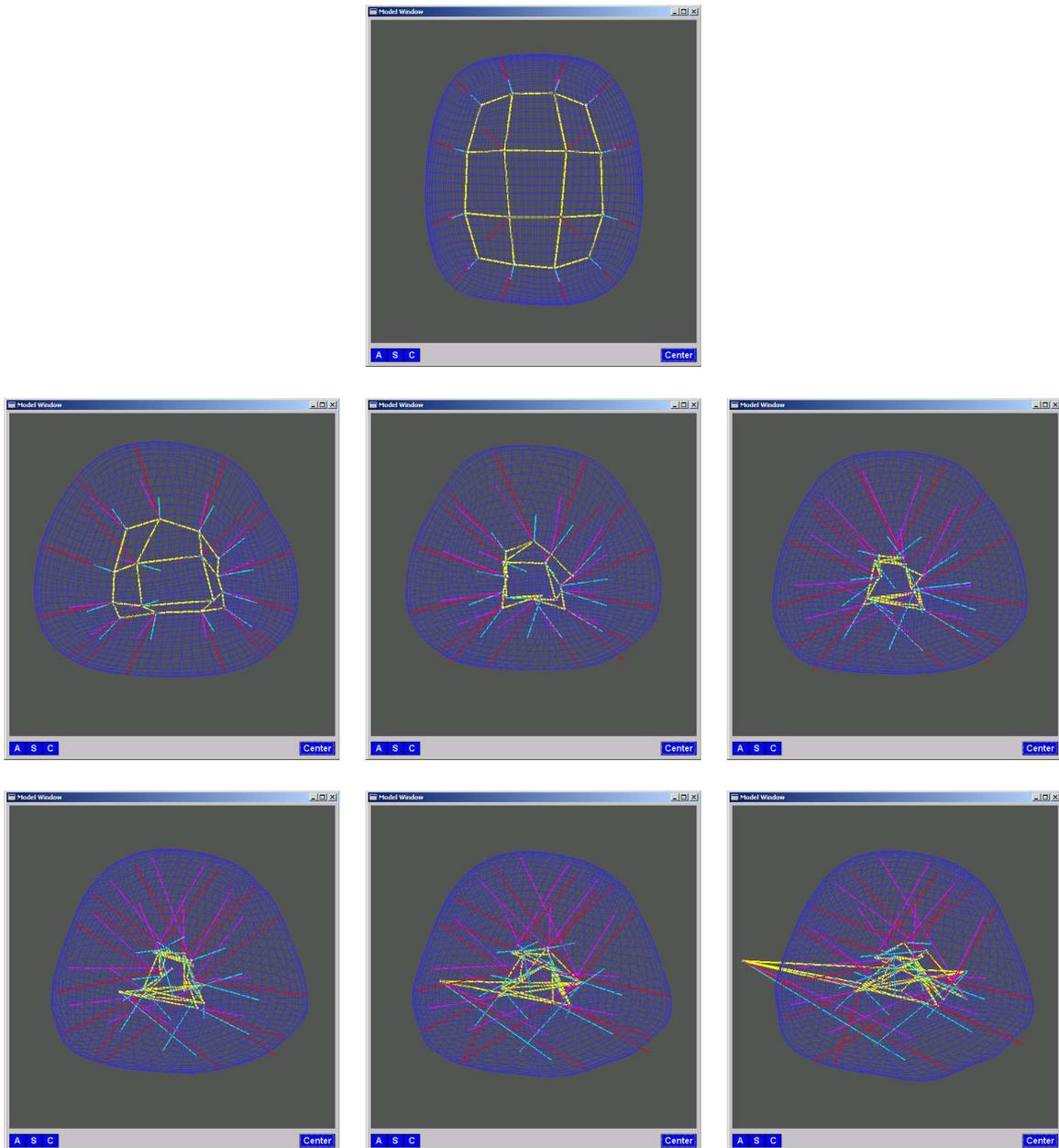
Figure 10.1: *The lack of convergence for the mean model during the bootstrap iterations. Above, the mean resulting from starting the bootstrap from the Generic model using full PGA is displayed after 1, 4, 10, 20, 30 40, and 50 iterations. The end atoms appear to be moving away from the boundary and thereby pass across the center of the model. This implosion leaves the implied boundary nicely optimized for representing the prostates but the medial sheet is not suitable for shape analysis — and not medial!*

End atoms further away from the boundary allow smaller curvature of the implied boundary, and since the prostates are not flat, the curvature at the crest (the part of the implied boundary that corresponds to the transition between top and bottom of the medial sheet) is quite low[1].

The geometry of the medial axis has been studied closely. There exists a number of theorems that describe conditions for ensuring legal geometry [Damon, 2003] for a medial axis and the implied boundary. A relevant example is a theorem that limits the local curvature of the medial axis given the local radius. However, these theorems are not enforced in the m-rep optimization.

10.1.2   Constraining the Medial Sheet

In the following the approach using absolute geometric penalties is investigated through a simple example. Instead of implementing the theorems that ensure nice properties for the medial axis and the implied geometry, a more ad hoc approach is used. The purpose is not to present a theoretically solid method, but rather to sketch a simple approach that implements absolute geometric penalties in a simple and intuitively reasonable manner.

The medial sheet in intended to be a sampled medial axis. This implies that the atoms are supposed to be approximately evenly spaced on the sheet. Furthermore, the beforementioned theorems on medial geometry implies that the curvature of the medial sheet is limited. A simple way to interpret curvature in terms of medial atoms is to look at the angles of the grid vectors across the atoms.

This results in two simple geometric penalties:

Distance:
>The atoms are to be kept approximately evenly spaced. An indirect way of doing this is to penalize the atoms from moving too close to each other. A simple measure at an atom is the minimal distance to a neighbor atom.

Angle:
>The curvature is to be limited. A qualitative way of enforcing this curvature constraint is to penalize large angles at the medial grid across the atoms. A simple measure at an atom is the maximal angle between vectors to two opposing neighbor atoms. End atoms must be treated as a special case — even though no angle can be measured across the boundary of the sheet, they indirectly cause an angle at the neighboring internal atom. This internal angle is used as the measure that penalizes large angles at the end atoms.

These simple penalties resemble the regularising penalties used for enforcing regular sampling and smoothness for point distribution boundary models.

The specific geometric penalties used are inversely proportional to the minimal distance squared and proportional to the maximal angle squared. These penalties are added to the equation for the geometric typicality (equation 9.1) in the optimization of the m-rep models.

---

[1]For those with an eye for the finer details and detailed m-rep knowledge: since the end atom elongation is kept above 1, low curvature can not be obtained by minimizing the end atom angle.

The absolute geometric penalty for large angles and for small inter-atom distances are each equipped with a proportionality factor in the equation for the geometric penalty. With *Ang* defined as the maximal angle across an atom (as described above) and *Dist* defined as the minimal distance to a neighboring atom, the full equation for the geometric penalty becomes:

$$G(\mathbf{M}) = (1 - \beta) \, P(\mathbf{M}) + \beta \, N(\mathbf{M}) + \gamma \, Ang^2 + \delta \, Dist^{-2} \qquad (10.1)$$

10.1.3 Bootstrapping using Absolute Geometric Penalties

In the following, some preliminary experiments on the use of absolute geometric penalies are presented. In order to speed up experimentation, a smaller training collection with only six prostates is used. Figure 10.2 shows how the mean evolves when the bootstrapping method is applied without absolute geometric penalties on this smaller training set. The undesirable implosion effect is evident.

The expectation is that the added geometric penalties will constrain the model and thereby limit the undesirable implosion effect. However, too high absolute geometric penalties will constrain the model too much and prevent it from deforming into the given shape. Informal experiments show that the factors $\gamma = 0.02$ and $\delta = 0.000005$ seem reasonable.

Figure 10.3 illustrates how these absolute penalties manage to prevent the implosion effect in the evolution of the model. However, the model still appears to continue wiggling slightly. Convergence is possibly not attained.

This wiggling is believed to be due to an implementation issue in Pablo [Fletcher et al., 2002]. Each medial atom implies two boundary points with corresponding surface normals. However, in order to ensure a smooth boundary, the actual normals of the surface are only required to be within a certain angle compared to the normals given by the medial atoms. This means that infinitely many atom positions can yield the same boundary. Therefore the question of convergence is not only related to the medial model and the model bootstrapping method — but also to the specific implementation of the m-reps. This is a reasonable explanation for the small drift in the model in figures 10.3. Future research will investigate this further.

Even though the absolute geometric penalties alone possibly do not ensure correspondence, the basic principle do appear to solve the main part of the problem. However, this comes with a price. The added geometric penalties constrain the m-rep model. Therefore the image match values are expected to suffer as a results of these extra constraints. Figure 10.4 illustrates this effect.

The experiments here demonstate that while absolute geometric penalties succeed in ensuring a nice medial sheet, they alone do possibly not provide convergence of the mean model in the bootstrapping method. Furthermore, the approach introduces ad hoc choices of penalty weights that seemingly keep the model nice while not lowering the image match values too much. These ad hoc penalty weights should be disposed of — and the absolute geometric penalties should be replaced by expressions based on geometry. As an example, any angle above zero is penalized above — instead only angles that implies illegal geometry (as determined by the local radius [Damon, 2003]) should be prevented.

Figure 10.2: *Evolution on training subset without absolute geometric penalties. The mean models are visualized after every 9 bootstrap iterations starting from iteration 1 and ending at 100. The implosion is present and the process is obviously not converging to an acceptable mean model.*
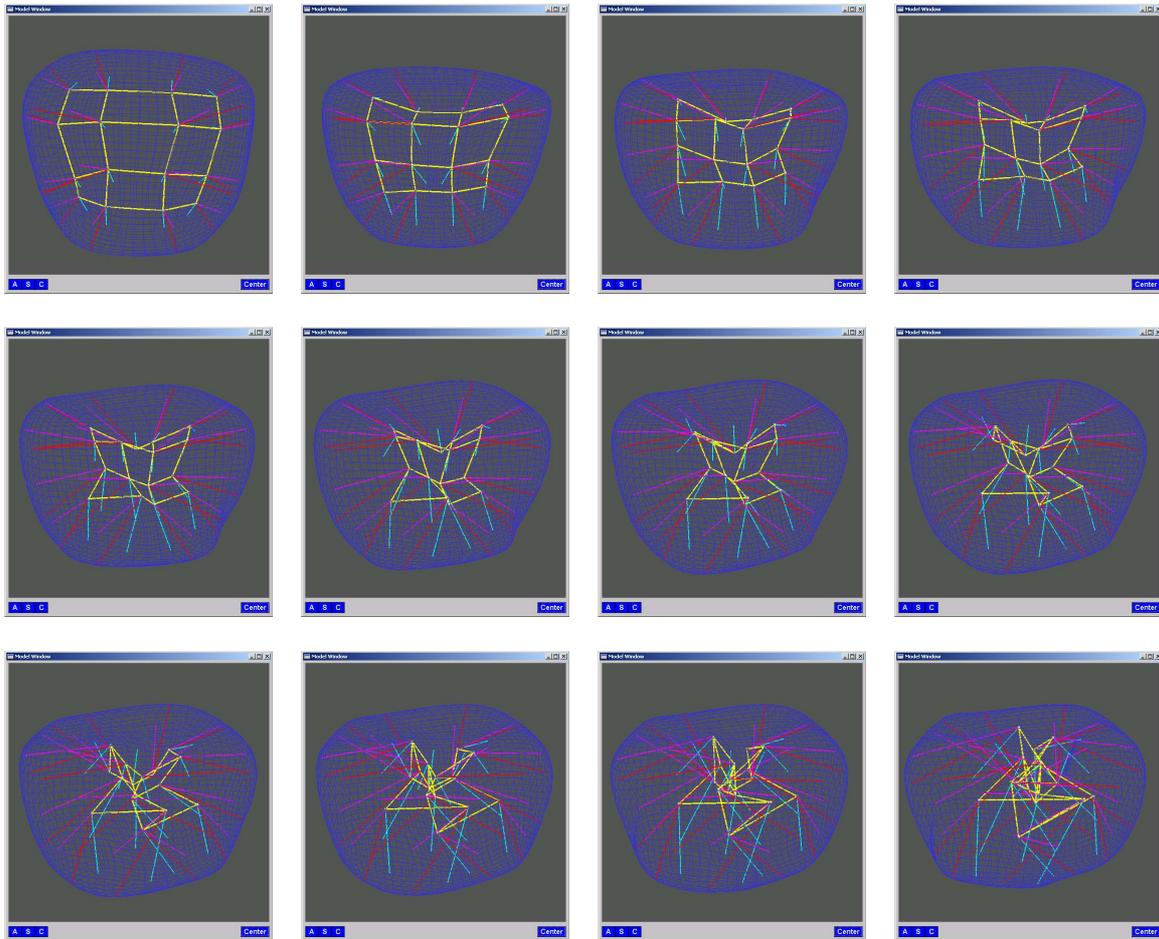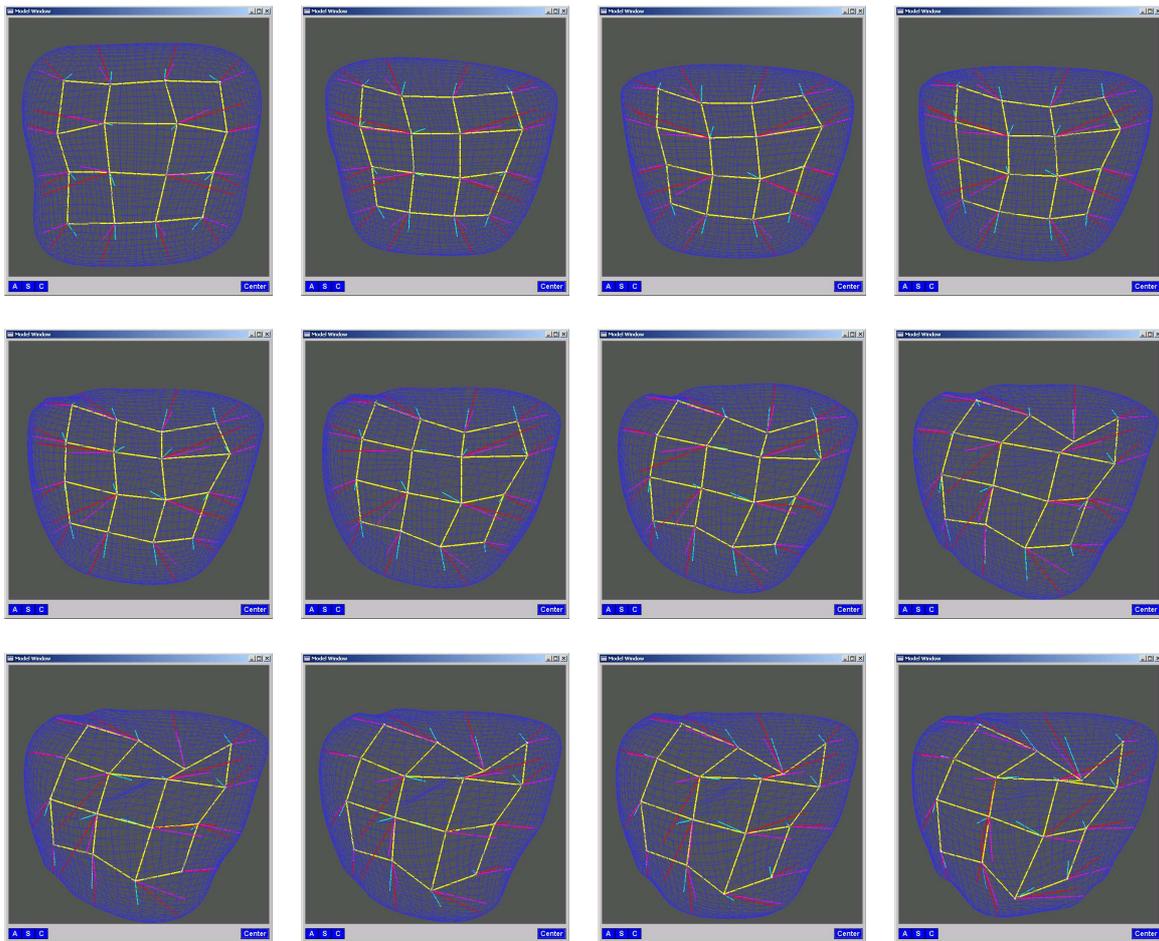
Figure 10.3: *Evolution on the training subset with absolute geometric penalties. The mean models are visualized after every 9 bootstrap iterations starting from iteration 1 and ending at 100. The process is not apparently convergent, but obviously more in control than in figure 10.2.*

Figure 10.4: *With absolute geometric penalties the model is constrained from achieving optimal image match.* **Top**: *Evolution of image match values without absolute geometric penalties.* **Bottom**: *Evolution with absolute geometric penalties constrains the model and results in worse image match values.*

10.1.4 Proving Convergence
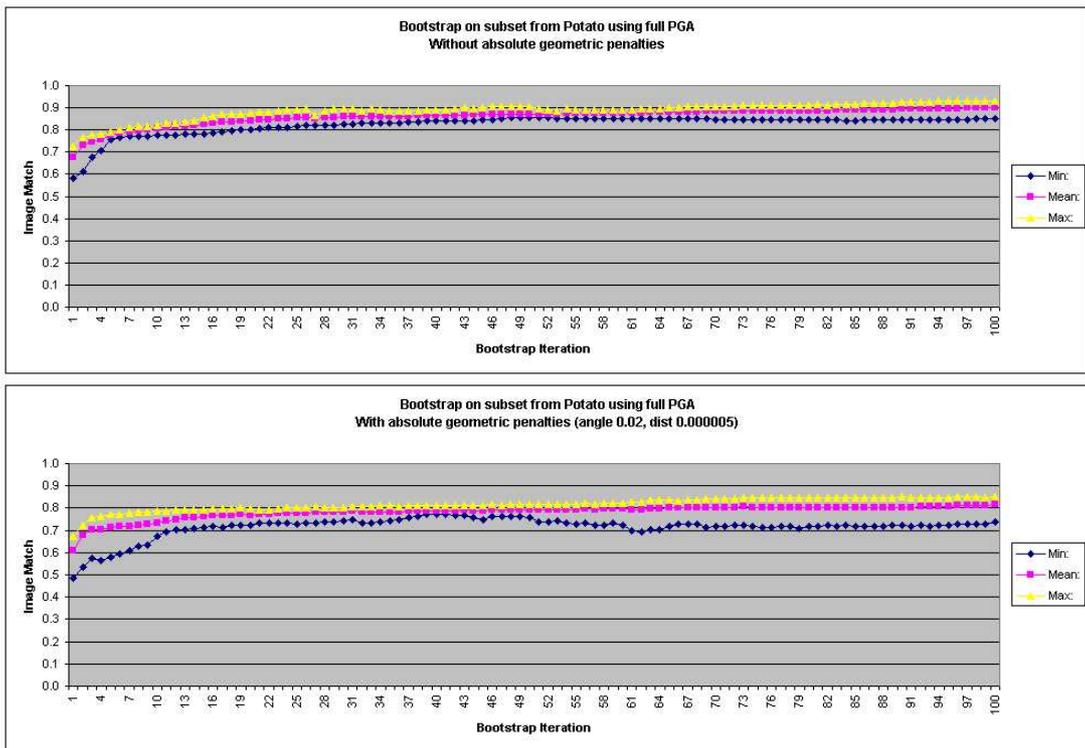
Above, convergence is treated informally with little mathematical foundation. A proper derivation of the requirements for convergence is obviously desirable.

However, proving convergence is not trivial. The fact that there is a many-to-one mapping between m-reps and implied boundaries complicates matters. This can mean that even though the implied model boundary is converging towards an optimal shape, the underlying mean model is not necessarily doing the same. As explained above, there are also implementation issues that need to be considered.

Finally, the unknown training collection also complicates proving convergence.

Therefore a formal proof will probably need to make some assumptions. For instance formulated like the following: *If the optimization of the current mean model to each training shape is converging, then the mean model will converge to the optimal mean model.*

The derivation of a formal proof is left for future work.

10.2 Conclusion

The work in this chapter offers preliminary results on ensuring convergence in the principal geodesic analysis bootstrapping method introduced in chapter 9. The approach is somewhat ad hoc, but the introduction of absolute geometric penalties appears to ensure controlled behavior of the mean model during the bootstrap evolution.

The formulation of proper absolute geometric constraints based in geometry is left for future work.

Another approach on investigating convergence indicates that if no atom deformations are allowed during the optimization process in the bootstrap iterations, the mean model is actually converging. The global deformations defined by the PGA modes during the figure stage are then the only shape deformations that allow the mean model to improve. This preliminary result is due to Tom Fletcher.

Also, research shows that although convergence is not yet attained, the method works for kidney modeling. An ad hoc approach for stopping the bootstrap when the image match seems to be converging. Thereby automatic segmentation is achieved [Rao, 2003]— this is not possible without the mean shape obtained through the PGA bootstrap modeling method.

Recycling in this Chapter

This contents of this chapter are previously unpublished.

Chapter 11

Summary

The work in this dissertation explores a path from a non-committed multi-scale segmentation method towards more committed methods for segmentation and shape modeling.

The first body of work is focused on specializing the multi-scale segmentation method through the use of non-linear diffusion. We present a *Generalized Anisotropic Non-linear* diffusion scheme with methods for optimizing the parameters and evaluating the performance. The results apply for segmentation of both 2D and 3D objects and show considerable performance improvements through the use of GAN. Upper limits for this performance gain are empirically established through experiments on artificial ideal objects. The methodology is evaluated on segmentation of brain structures in both 2D and 3D and shows significant improvements in segmentation efficiency.

Non-linear diffusion is then analyzed in more detail via the introduction of the *Diffusion Echo* that allows explicit analysis of the local filters in non-linear diffusion schemes. We use this approach to investigate the connection between linear and non-linear diffusion given by scale selection.

While non-linear diffusion allows a large degree of incorporation of prior knowledge and thereby facilitates specialization towards a specific task, the work shows that for the multi-segmentation program, the possible performance improvement due to non-linear diffusion is not unlimited. Specifically, near-automatic segmentation seems unfeasible without further commitment of the method. The most promising way of incorporating additional prior knowledge is through the use of a shape model.

The final body of work is focused on shape modeling via the medial shape representation known as the m-rep. We present an essentially automatic method for generating a statistical shape model from a training collection. We demonstrate the method for constructing a prostate shape model.

11.1 Future Work

As is often the case in research, this dissertation leaves more open ends than closed chapters.

Each of the three main bodies of work provides promising future work as outlined in the chapter conclusions throughout the dissertation. Here, only future directions that combine the bodies of work are mentioned.

The analysis of non-linear diffusion made possible by the diffusion echo has obvious potential in the multi-scale watershed segmentation framework. A simple, very pragmatic, direction is to investigate the use of downsampling in non-linear diffusion. Downsampling is essential in order to reduce the computation time needed for preprocessing of the building blocks. The diffusion echo directly offers a way to measures the change in the diffusion due to downsampling.

Another more fundamental path that combines the diffusion echo with multi-scale watershed segmentation is to pursue the design of the scale-selection mechanism outlined in chapter 7. Apart from the theoretical implications, this would also allow a combination of the best of both worlds in the multi-scale watershed segmentation method. The design of a scale-selection mechanism that allows multi-scale approximation of non-linear diffusion schemes offers computational efficiency in the preprocessing step while providing the interaction efficiency of building blocks constructed through the use of non-linear diffusion.

The best of both worlds is also available through augmentation of the multi-scale watershed segmentation method with a shape model. Thereby near-automatic segmentation combined with the intuitive direct interaction with the data through building blocks is possible. A typical problem with near-automatic methods (such as segmentation methods based on the active shape model or the m-rep) is how to handle the, say, 5% problematic cases where the automatic scheme fails — or alternatively how to interactively refine an almost correct segmentation. Selection and deselection of low scale building blocks would allow the user to handle this with little effort.

The incorporation of a shape model could either be independent of the multi-scale watershed framework — or an integrated component. A simple integration would be a shape model that uses the number of building block actions necessary to reach the current shape instead of the image match term in the optimization. This approach is possibly somewhat naive, but it is nicely in alignment with the multi-scale watershed framework.

A more ambitious and fully integrated shape model would be based on the watershed linking tree. A major problem with scale-space tree models is the heavy influence of the background on the tree representation. Non-linear diffusion could here be used to provide a better separation of object and background in the tree and thereby make the approach more robust.

Bibliography

[Abbasi et al., 1999] S. Abbasi, F. Mokhtarian, & J. Kittler. *Curvature Scale Space Image in Shape Similarity Retrieval.* Multimedia Systems, 7:467–476, 1999.

[Blum & Nagel, 1978] H. Blum & R. Nagel. *Shape Description Using Weighted Symmetric Axis Features.* Pattern Recognition, 10(3):167–180, 1978.

[Brox et al., 2003] Brox, Welk, Steidl, & Weickert. *Equivalence Results for TV Diffusion and TV Regularization.* In *Scale Space Methods in Computer Vision*, number 2695 in LNCS, 2003.

[Catté et al., 1992] F. Catté, P.-L. Lions, J.-M. Morel, & T. Coll. *Image Selective Smoothing and Edge Detection by Nonlinear Diffusion.* SIAM Journal of Numerical Analysis, 29:182–193, 1992.

[Cocosco et al., 1997] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, & A.C. Evans. *BrainWeb: Online Interface to a 3D MRI Simulated Brain Database.* In *Proceedings of 3rd International Conference on Functional Mapping of the Human Brain*, volume 5 of NeuroImage, page 425, May 1997. http://www.bic.mni.mcgill.ca/brainweb/.

[Collins et al., 1998] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, & A.C. Evans. *Design and Construction of a Realistic Digital Brain Phantom.* IEEE TMI, 17, June 1998. http://www.bic.mni.mcgill.ca/brainweb/.

[Cootes et al., 1995] T.F. Cootes, C.J. Taylor, D.H. Cooper, & J. Graham. *Active Shape Models: Their Training and Application.* CVIU, (1):38–59, 1995.

[Cremers et al., 2002a] D. Cremers, T. Kohlberger, & C. Schnörr. *Nonlinear Shape Statistics in Mumford-Shah Based Segmentation.* In *7th European Conference on Computer Vision*, volume 2351 of Springer LNCS, 2002.

[Cremers et al., 2002b] D. Cremers, F. Tischhuser, J. Weickert, & C. Schnörr. *Diffusion snakes: introducing statistical shape knowledge into the Mumford–Shah functional.* International Journal of Computer Vision, 50, 2002.

[Crouch et al., 2003] J. Crouch, S.M. Pizer, E.L. Chaney, & M. Zaider. *Medial Techniques to Automate Finite Element Analysis of Prostate Deformation.* Submitted to IEEE Transactions on Medical Imaging, 2003. http://midag.cs.unc.edu/pubs/papers/TMI03-Crouch-FEM.pdf.

[Dam & Letteboer, 2003] Erik Dam & Marloes Letteboer. *Non-linear Diffusion in 3D for Interactive Segmentation of Brain Tissue*. In *Submitted to MICCAI 2003*, 2003.

[Dam & Nielsen, 2000] Erik Dam & Mads Nielsen. *Non-Linear Diffusion for Interactive Multi-scale Watershed Segmentation*. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2000*, volume 1935 of Lecture Notes in Computer Science, pages 216—225. Springer, October 2000.

[Dam & Nielsen, 2001] Erik Dam & Mads Nielsen. *Exploring Non-Linear Diffusion: The Diffusion Echo*. In Michael Kerckhove, editor, *Scale-Space Theories in Computer Vision*, Lecture Notes in Computer Science. Springer, 2001.

[Dam & ter Haar Romeny, 2003] Erik Dam & Bart ter Haar Romeny. *Front End Vision and Multi-Scale Image Analysis*, chapter Chapters 13–15: Deep Structure I, II & III. Kluwer, 2003.

[Dam, 2000] Erik Dam. *Evaluation of Diffusion Schemes for Watershed Segmentation*. Master's thesis, University of Copenhagen, 2000. Technical report 2000/1 on http://www.diku.dk/research/techreports/2000.htm.

[Dam, 2003] Erik B. Dam. *Billedanalyse og strålebehandling*. At the science portal of the Danish Radio, 2003. http://www.dr.dk/videnskab/artikler/it/billedanalyse.asp.

[Dam et al., 2000] Erik Dam, Peter Johansen, Ole Fogh Olsen, Andreas Thomsen, Tron Darvann, Andy B. Dobrzeniecki, Nuno V. Hermann, Noriyuki Kitai, Sven Kreiborg, Per Larsen, & Mads Nielsen. *Interactive Multi-Scale Segmentation in Clinical Use*. In *European Congress of Radiology 2000*, March 2000.

[Dam et al., 2002] Erik Dam, Dr. Stephen M. Pizer, Dr. Julian Rosenman, & Gregg Tracton. *M-Rep Based Classification of Organ Conformations for Radiation Treatment Planning of Prostate Cancer*. The 17th Annual UNC Radiology Research Symposium, 2002. Reviewed abstract.

[Dam et al., 2003a] Dam, Fogh Olsen, Johansen, Lillholm, Nielsen, & Thomsen. *Nabla Vision*, 2000–2003. http://www.itu.dk/image/nablavision.

[Dam et al., 2003b] Erik Dam, P. Thomas Fletcher, Stephen M. Pizer, Gregg Tracton, & Julian Rosenman. *Prostate Shape Modeling based on Principal Geodesic Analysis Bootstrapping*. In *Submitted to ICCV workshop VLSM*, 2003.

[Dam et al., 2003c] Erik Dam, Ole Fogh Olsen, & Mads Nielsen. *Approximating Non-linear Diffusion*. In *Scale-space Methods in Computer Vision, Proc. from the 4th int. conference*, volume 2695 of LNCS. Springer, 2003.

[Damon, 1997] James Damon. *Generic Properties of Solutions to Partial Differential Equations*. Arch. Rational Mech. Anal., (140):353–403, 1997.

[Damon, 2003] James Damon. *Smoothness and Geometry of Boundaries Associated to Skeletal Structures*. Technical report, University of North Carolina, Chapel Hill, 2003. http://midag.cs.unc.edu/pubs/papers/Damon_SkelStr_I.pdf, http://midag.cs.unc.edu/pubs/papers/Damon_SkelStr_II.pdf, http://midag.cs.unc.edu/pubs/papers/Damon_SkelStr_III.pdf.

[Davies et al., 2002] R.H. Davies, C.J. Twining, T.F. Cootes, J.C. Waterton, & C.J. Taylor. *A Minimum Description Length Approach to Statistical Shape Modeling*. IEEE Transactions on Medical Imaging, 21(5), 2002.

[Demirci et al., 2003] M. Fatih Demirci, Ali Shokoufandeh, Yakov Keselman, Sven Dickinson, & Lars Bretzner. *Many-to-Many Matching of Scale-Space Hierarchies using Metric Embedding*. In *Scale Space Methods in Computer Vision, Proceedings from 4th Int. Conference*, volume 2695 of LNCS, 2003.

[Duistermaat & Kolk, 2000] J. J. Duistermaat & J. A. C. Kolk. *Lie Groups*. Springer, 2000.

[Duits, 2003] Florack Platel Duits, Felsberg. α *Scale Spaces on a Bounded Domain*. In *Scale Space Methods in Computer Vision*, number 2695 in LNCS, 2003.

[Fischl & Schwartz, 1997] B. Fischl & E.L. Schwartz. *Learning an Integral-Equation Approximation to Nonlinear Anisotropic Diffusion in Image-Processing*. PAMI, 19(4), 1997.

[Fletcher et al., 2002] Fletcher, Gash, Joshi, Liu, Lu, Pizer, Styner, Thall, Tracton, & Yushkevich. *Pablo*. http://midag.cs.unc.edu, 2002. Pablo is a product of MIDAG, University of North Carolina, Chapel Hill.

[Fletcher et al., 2003a] P. T. Fletcher, S. Joshi, C. Lu, & S. M. Pizer. *Gaussian Distributions on Lie Groups and Their Application to Statistical Shape Analysis*. To appear *Information Processing in Medical Imaging*, 2003.

[Fletcher et al., 2003b] P. T. Fletcher, C. Lu, & S. Joshi. *Statistics of Shape via Principal Geodesic Analysis on Lie Groups*. To appear *Computer Vision and Pattern Recognition*, 2003.

[Gage, 1983] M. Gage. *An isoperimetric inequality with applications to curve shortening*. Invent. Math., (76):357–364, 1983.

[Gauch & Pizer, 1993] John M. Gauch & Stephen M. Pizer. *Multiresolution Analysis of Ridges and Valleys in Grey-Scale Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(6):635–646, June 1993.

[Gauch, 1999] John M. Gauch. *Image Segmentation and Analysis via Multiscale Gradient Watershed Hierachies*. IEEE Transactions on Image Processing, 8(1):69 – 79, January 1999.

[Gomez et al., 2000] G. Gomez, J.L. Marroquin, & L.E. Sucar. *Probabilistic Estimation of Local Scale*. In *Proc. of the ICPR*, 2000.

[Griffin & Colchester, 1995] Lewis D Griffin & Alan C F Colchester. *Superficial and deep structure in linear diffusion scale space: isophotes, critical points and separatrices*. Image and Vision Computing, 13(7):543–557, September 1995.

[Griffin et al., 2003] Griffin, Lillholm, & Nielsen. *Natural Image Profiles are most likely to be Step Edges*. Vision Research, 2003. Submitted.

[ibs, 1999] *The Internet Brain Segmentation Repository*, 1999. MR brain data set 788_6_m and its manual segmentation was provided by the Center for Morphometric Analysis at Massachusetts General Hospital and is available at http://neuro-www.mgh.harvard.edu/cma/ibsr.

[Ijima, 1962] T. Ijima. *Basic theory on normalization of a pattern (in case of typical one-dimensional pattern)*. In *Bulletin of Electrotechnical Laboratory*. 1962. In japanese.

[Jackway, 1996] P.T. Jackway. *Gradient watersheds in morphological scale-space*. IEEE Trans. Image Proc., 5:913–921, 1996.

[Jaynes, 2003] E.T. Jaynes. *Probability Theory: The Logic of Science*. http://omega.albany.edu:8008/JaynesBook, 2003.

[Johansen et al., 1999] Peter Johansen, Mads Nielsen, & Sven Kreiborg. *The Computation of Natural Shape*, 1999.
`http://www.diku.dk/research-groups/image/research/NaturalShape/`.

[Joshi et al., 2002] S. Joshi, S. Pizer, P. T. Fletcher, P. Yushkevich, A. Thall, & J. S. Marron. *Multiscale deformable model segmentation and statistical shape analysis using medial descriptions*. Transactions on Medical Imaging, 21(5), 2002.

[Kanters et al., 2003] Frans Kanters, Bram Platel, Luc Florack, & Bart M. ter Haar Romeny. *Content Based Image Retrieval using Multi-scale Top Points*. In *Scale Space Methods in Computer Vision, Proceedings from 4th Int. Conference*, volume 2695 of LNCS, 2003.

[Kass & Terzopoulos, 1988] Michael Kass & Andrew Witkin Demitri Terzopoulos. *Snakes: Active Contour Models*. Int. Journal of Computer Vision, pages 321–331, 1988.

[Koenderink, 1984] Jan J. Koenderink. *The Structure of Images*. Biological Cybernetics, 50:363–370, 1984.

[Koster, 1995] André Koster. *Linking Models for Multi-scale Image Sequences*. PhD thesis, University of Utrecht, 1995.

[Kullback & Leibler, 1951] S. Kullback & R.A. Leibler. *On Information Theory and Sufficiency*. Annals of Mathematical Statistics, 22, 1951.

[Kwan et al., 1996] R.K.-S. Kwan, A.C. Evans, & G.B. Pike. *An Extensible MRI Simulator for Post-Processing Evaluation*. Lecture Notes in Computer Science, 1131:135–140, 1996. http://www.bic.mni.mcgill.ca/brainweb/.

[Letteboer et al., 2001] M. Letteboer, W. Niessen, P. Willems, E. B. Dam, & M. Viergever. *Interactive multi-scale watershed segmentation of tumors in MR brain images*. In *Proc. of the IMIVA workshop of MICCAI*, 2001.

[Letteboer et al., 2003] M. Letteboer, O. F. Olsen, E. B. Dam, P. Willems, M. Viergever, & W. Niessen. *Segmentation of Tumors in MR Brain Images using an Interactive Multi-scale Watershed Algorithm*. IEEE Transactions on Medical Imaging, 2003. Submitted.

[Lifshitz & Pizer, 1990] Lawrence Lifshitz & Stephen Pizer. *A Multiresolution Hierarchical Approach to Image Segmentation Based on Intensity Extrema*. IEEE PAMI, 12(6):529 – 540, June 1990.

[Lindeberg, 1994] Tony Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.

[Lorensen & Cline, 1987] William E. Lorensen & Harvey E. Cline. *Marching Cubes: A High Resolution 3D Surface Construction Algorithm.* Computer Graphics, 21(4):163–169, 1987.

[Lorenzen et al., 2001] P. Lorenzen, S. Joshi, G. Gerig, & E. Bullitt. *Tumor-Induced Structural Radiometric Asymmetry in Brain Images.* In *Proc. Workshop on Mathematical Methods in Biomedical Image Analysis MMBIA 2001.* IEEE Computer Society, 2001.

[Maxwell, 1870] J.C. Maxwell. *On Hills and Dales.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 4th Series, 40(269):421–425, December 1870.

[Mumford & Shah, 1985] D. Mumford & J. Shah. *Boundary Detection by Minimizing Functionals.* In *CVPR85*, pages 22–26, 1985.

[Murakami, 1998] S. Murakami. *Three and four dimensional diagnosis for TMJ.* In *Proc. of CAR 98*, pages 88–91, 1998.

[Niessen et al., 1997] Wiro J. Niessen, Koen L. Vincken, Joachim A. Weickert, & Max A. Viergever. *Nonlinear Multiscale Representations for Image Segmentation.* Computer Vision and Image Understanding, 66(2):233–245, May 1997.

[Nitzberg & Shiota, 1992] Mark Nitzberg & Takahiro Shiota. *Nonlinear Image Filtering with Edge and Corner Enhancement.* IEEE Tr. Pattern Analysis and Machine Intelligence, 14(8):826 – 833, 1992.

[Nordstrom, 1990] K.N. Nordstrom. *Biased Anisotropic Diffusion: A Unified Regularization and Diffusion Approach to Edge Detection.* IVC, 8(4):318–327, 1990.

[Olsen & Nielsen, 1997] Ole Fogh Olsen & Mads Nielsen. *Generic events for the gradient squared with application to multi-scale segmentation.* In *Scale-Space Theory in Computer Vision, Proc. 1st International Conference*, volume 1252 of Lecture Notes in Computer Science, pages 101–112, Utrecht, The Netherlands, July 1997.

[Olsen, 1996] Ole Fogh Olsen. *Multi-Scale Segmentation of Grey-Scale Images.* Technical Report 96/30, Department of Computer Science, University of Copenhagen, 1996.

[Olsen, 1997] Ole Fogh Olsen. *Multi-Scale Watershed Segmentation.* In Jon Sporring, Mads Nielsen, Luc Florack, & Peter Johansen, editors, *Gaussian Scale-Space Theory*, pages 191–200. Kluwer, 1997.

[Olsen et al., 2000] O. F. Olsen, E. B. Dam, M. Lillholm, A. Thomsen, P. Johansen, , & M. Nielsen. *Method and System for Multi-dimensional Segmentation.* Danish patent application filed October 2000, 2000.

[Ott, 1993] Edward Ott. *Chaos in Dynamical Systems.* Cambridge University Press, 1993.

[Perona & Malik, 1990] Pietro Perona & Jitendra Malik. *Scale-Space and Edge Detection Using Anisotropic Diffusion.* IEEE PAMI, 12(7):629 – 639, July 1990.

[Pizer et al., 1994] S.M. Pizer, C.A. Burbeck, J.M. Coggins, D.S. Fritsch, & B.S. Morse. *Object Shape Before Boundary Shape: Scale-Space Medial Axes.* JMIV, 4:303–313, 1994.

[Pizer et al., 1996] S.M. Pizer, D.S. Fritsch, P. Yushkevich, V. Johnson, & E. Chaney. *Segmentation, Registration, and Measurement of Shape Variation via Image Object Shape.* IEEE Transactions on Medical Imaging, 18(10), 1996.

[Pizer et al., 2003a] Pizer, Chen, Fletcher, Fridman, Fritsch, Gash, Glotzer, Jiroutek, Joshi, Muller, Thall, Tracton, Yushkevich, & Chaney. *Deformable M-Reps for 3D Medical Image Segmentation.* IJCV, 2003. To appear, see http://midag.cs.unc.edu/pubs/papers/IJCV01-Pizer-mreps.pdf.

[Pizer et al., 2003b] S.M. Pizer, K. Siddiqi, G. Székely, J. Damon, & S.W. Zucker. *Multiscale Medial Loci and Their Properties.* IJCV, 2003. To appear, see http://midag.cs.unc.edu/pubs/papers/IJCV01-Pizer-medloci.pdf.

[Press et al., 1999] Press, Teukolsky, Vetterling, & Flannery. *Numerical Recipies in C.* Cambridge University Press, 1999.

[Rao, 2003] Manjari Rao. *Analysis of a Locally Varying Intensity Template for Segmentation of Kidneys in CT images.* Master's thesis, University of North Carolina, Chapel Hill, 2003.

[Rudin et al., 1992] L. I. Rudin, S. Osher, & E Fatemi. *Nonlinear total variation based noise removal algorithms.* Physica D, pages 259–268, 1992.

[Samet, 1984] H. Samet. *The Quadtree and Related Hierarchical Data Structures.* Surveys, 16(2):187–260, June 1984.

[Sapiro, 2001] Guillermo Sapiro. *Geometric Partial Differential Equations and Image Analysis.* Cambridge, 2001.

[Scharr & Weickert, 2000] Hanno Scharr & Joachim Weickert. *An Anisotropic Diffusion Algorithm with Optimized Rotation Invariance.* In *Mustererkennung 2000, Proceedings from 22. DAGM-symposium*, pages 460–467. Springer, 2000.

[Shokoufandeh et al., 2002] Ali Shokoufandeh, Sven Dickinson, Clas Jonsson, Lars Bretzner, & Tony Lindeberg. *On the Representation and Matching of Qualitative Shape at Multiple Scales.* In *Proceedings, European Conference on Computer Vision*, LNCS, 2002.

[Siddiqi et al., 1999] K. Siddiqi, A. Shokoufandeh, S. Dickinson, & S. Zucker. *Shock Graphs and Shape Matching.* International Journal of Computer Vision, 30, 1999.

[Sporring et al., 1997] Jon Sporring, Mads Nielsen, Luc Florack, & Peter Johansen. *Gaussian Scale-Space Theory.* Kluwer Academic Publishers, Dordrecht, 1997.

[Styner & Gerig, 2001] Martin Styner & Guido Gerig. *Medial Models Incorporating Object Variability for 3D Shape Analysis.* In *Proc. of Information Processing in Medical Imaging*, 2001.

[ter Haar Romeny, 1994] Bart M. ter Haar Romeny, editor. *Geometry-Driven Diffusion in Computer Vision.* Kluwer Academic Publishers, 1994.

[Thall, 2002] A. Thall. *Fast C^2 interpolating subdivision surfaces using iterative inversion of stationary subdivision rules.* Technical report, University of North Carolina Department of Computer Science, 2002. http://midag.cs.unc.edu/pub/papers/Thall_TR02-001.pdf.

[Tracton et al., 1994] G. Tracton, E. Chaney, J. Rosenman, & S. Pizer. *MASK: combining 2D and 3D segmentation methods to enhance functionality*. In *Proceedings of Mathematical Methods in Medical Imaging III*, volume SPIE Vol 2299, 1994.

[van Hateren & van der Schaaf, 1998] van Hateren & van der Schaaf. *Independent component filters of natural images compared with simple cells in primary visual cortex*. Proceedings of the Royal Society of London Series B - Biological Sciences, 265 (1394):359–366, 1998.

[Vincken, 1995] Koen Vincken. *Probabilistic Multi-scale Image Segmentation by the Hyperstack*. PhD thesis, University of Utrecht, 1995.

[Wang, 2003] Haonan Wang. *Functional Data Analysis of Populations of Tree-structured Objects*. PhD thesis, University of North Carolina, Chapel Hill, 2003.

[Webb, 1988] Steve Webb. *The Physics of Medical Imaging*. Institute of Physics Publishing, London, 1988.

[Weickert, 1998a] Joachim Weickert. *Anisotropic Diffusion in Image Processing*. B. G. Teubner, 1998.

[Weickert, 1998b] Joachim Weickert. *Efficient Image Segmentation Using Partial Differential Equations and Morphology*. Technical Report 98-10, Department of Computer Science, University of Copenhagen, 1998.

[Witkin, 1983] Andrew P. Witkin. *Scale-Space Filtering*. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 1019–1022, Karlsruhe, Germany, 1983.