Multi-Scale Active Shape Description in Medical Imaging

Julia Anne Schnabel

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

Department of Computer Science
University College London

October 1997

ProQuest Number: U643552

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest U643552

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Shape description in medical imaging has become an increasingly important research field in recent years. Fast and high-resolution image acquisition methods like Magnetic Resonance (MR) imaging produce very detailed cross-sectional images of the human body - shape description is then a post-processing operation which abstracts quantitative descriptions of anatomically relevant object shapes. This task is usually performed by clinicians and other experts by first segmenting the shapes of interest, and then making volumetric and other quantitative measurements. High demand on expert time and inter- and intra-observer variability impose a clinical need of automating this process. Furthermore, recent studies in clinical neurology on the correspondence between disease status and degree of shape deformations necessitate the use of more sophisticated, higher-level shape description techniques.

In this work a new hierarchical tool for shape description has been developed, combining two recently developed and powerful techniques in image processing: differential invariants in scale-space, and active contour models. This tool enables quantitative and qualitative shape studies at multiple levels of image detail, exploring the extra image scale degree of freedom. Using scale-space continuity, the global object shape can be detected at a coarse level of image detail, and finer shape characteristics can be found at higher levels of detail or lower *scales*. New methods for *active shape evolution* and *focusing* have been developed for the extraction of shapes at a large set of scales using an active contour model whose energy function is regularized with respect to scale and geometric differential image invariants. The resulting set of shapes is formulated as a *multi-scale shape stack* which is analysed and described for each scale level with a large set of shape descriptors to obtain and analyse shape changes across scales. This shape stack leads naturally to several questions in regard to variable sampling and appropriate levels of detail to investigate an image. The relationship between active contour sampling precision and scale-space is addressed.

After a thorough review of modern shape description, multi-scale image processing and active contour model techniques, the novel framework for multi-scale active shape description is presented and tested on synthetic images and medical images. An interesting result is the recovery of the fractal dimension of a known fractal boundary using this framework. Medical applications addressed are grey-matter deformations occurring for patients with epilepsy, spinal cord atrophy for patients with Multiple Sclerosis, and cortical impairment for neonates. Extensions to non-linear scale-spaces, comparisons to binary curve and curvature evolution schemes as well as other hierarchical shape descriptors are discussed.

Acknowledgements

I would like to thank Dr. Simon Arridge from the Department of Computer Science at University College London for being such a unique supervisor and great friend. Without his constant moral encouragement and scientific support, kind understanding, huge enthusiasm, sharp analytic mind, multiple proof reading, fund raising for conferences and equipment, tiring writing of grant applications, and that little arm-twisting I would have never managed writing this dissertation.

Scientific acknowledgements go to the NMR Research Unit of the Institute of Neurology London and the MRI Unit of the Epilepsy Society for Epilepsy, in particular Drs. Paul Tofts, Liqun Wang, Sam Free, Louis Lemieux, and Nick Losseff, and the Medical Physics Department of UCL, in particular Dr. John Thornton and Prof. Roger Ordidge. All of them have been very helpful in supplying me with medical image data, taking the time to provide me with their scientific expertise, and even showing enthusiasm over my results.

Sincere thanks go to Dr. David Hawkes from the UMDS Radiological Sciences Image Processing Group at Guy's Hospital London, and Dr. Jim Graham from the Wolfson Image Analysis Unit at the Department of Medical Biophysics, University of Manchester, for agreeing to be my PhD examiners. Dave Hawkes was already kind enough to be my examiner in my PhD transfer viva, along with my second supervisor, Professor Andrew Todd-Pokropek, from the Medical Physics Department of UCL. I would like to thank both of them for pointing me in the right directions at an early stage in my PhD, and for their constructive criticism of my work.

Bedankt to Professor Max Viergever from the Image Sciences Center at Utrecht University, for inviting me for a visit to *AZU*, for his friendly encouragement, and for pointing me out to many relevant important directions in the medical imaging field while simultaneously warning me of scientific pitfalls. *Dank u* to the group at Image Sciences Center for making me feel so welcome, and for keeping me posted about ongoing scale-space work; especially to Wiro Niessen for his kind online help, and for showing me the best cafés and bars in Utrecht this summer. Thanks to Will Schroeder from GEC Research and Development in Schenectady, New York, for providing me with a pre-release of the *Visualization Toolkit*, his new book, and for his kind hosting in Schenectady. I would also like to thank Bryan Morse and David Eberly (both formerly from the

Department of Computer Science, University of North Carolina at Chapel Hill) for their helpful suggestions.

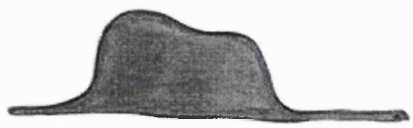
Vielen lieben Dank meinen Eltern, Ruth und Edmund Schnabel, die mich nicht nur in der Zeit meines Doktorwerdens immer lieb gehabt und unterstützt haben. *Danke, ay*, meinen Geschwistern, kleinen Neffen und vielen Berliner Freunden, die mich ebenfalls in meinem "Exil" so oft besucht und mit Roggenbrot und Gummibärchen versorgt haben.

Muchas gracias to Gerardo Ivar Sanchez-Ortiz, and *Salam* to Ahmed Shihab, along with their other pals of the Department of Computing at Imperial College, for being such good friends for me in London. *Danke* to my Blackheath flat mates, Nicole Fritz and Silke Koch, who had to bear with my blacker moods. *Cheers* to all the friends I made at UCL Computer Science, in particular to Bill Langdon for his kind advice, my surrounding friends Tina Yu, Kanta Vekaria, Chris Harris, Jorge Ortega-Arjona, Antonio Liotta, Adil Qureshi and Lee Braine, who all constantly cheered me up, my medical imaging fellow sufferers Hava Lester and Jon Davidson, my long-time *Pepper's* lunch partner Jens Dzikowski, my Greek friends Isidor Kouvelas, Thanassis Tiropanis, Natascha Vassila and Giannis Koufakis, the G11 bunch, Gavin Baily (my partner in the "environment from hell"), and Angela Sasse (who kindly invited me over to have *Linsensuppe* and *Bratwurst*). I have also very much enjoyed my little tea parties with Dave Walsh from the Medical Physics Department at UCL, philosophing about the *Gestalttheorie*, snakes, and why the hatter was mad.

Miscellaneous credits go to Tesco supermarket, for keeping the store open until 10pm, Boots, for selling cheap aspirin, Oasis, for their *champagne supernova*, the Simpsons, and London, for being "currently the hippest place on earth".

Qu'à la fin de l'envoi je touche... – merci vielmals à Daniel Rückert – *mein Spielbruder!*

If I chose to dedicate this dissertation to someone, it would be to Dirk, who was my best friend, along with all those little princes around who do not believe that the drawing below shows a hat.



DESSIN N° 1. – DRAWING NO. 1.

Le Petit Prince, Antoine de Saint-Exupéry.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

– VOICI MON SECRET. IL EST TRÈS SIMPLE: ON NE VOIT BIEN QU'AVEC LE COEUR. L'ESSENTIEL EST INVISIBLE POUR LES YEUX.

"HERE IS MY SECRET. IT IS ONLY WITH THE HEART THAT ONE CAN SEE RIGHTLY; WHAT IS ESSENTIAL IS INVISIBLE TO THE EYE."

Le Petit Prince, Antoine de Saint-Exupéry.

In recent years, a large diversity of very detailed in-vivo medical imaging modalities have been developed, allowing for the acquisition of high-resolution cross-sectional images of the human body. The obtained images can be viewed by clinicians in order to analyse them for certain complex relationships such as structural abnormalities and deformations, and distinguishing between normals and abnormals.

For example, Computer Tomography (CT) is based on ionizing X-rays, projecting the X-ray absorption coefficients of $3D$ bodies or structures in a slice-by-slice fashion onto $2D$ planes. Spiral CT as a true $3D$ acquisition is performed by collecting spirally obtained raw data, followed by interpolating planar projection data sets from the resulting volume. The resulting image intensities, which are normalized with respect to the water absorption coefficient and are measured in *Hounsfield units*, are the result of so-called hard beams, and are therefore especially useful for the investigation of bone structure and fat tissue. For an adequate acquisition of soft tissues, however, invasive contrast agents are required which causes allergic reactions in some patients. The quality of the resulting images can be degraded by aliasing due to overlapping of frequencies in the Fourier reconstruction, ringing artefacts due to wrong calibration of the scanner coordinate system, and metal artefacts caused by the reflexion of metallic structures in the body, e.g. tooth fillings or protheses.

Magnetic Resonance Imaging (MRI) is a non-invasive, non-ionizing technique which allows for high-resolution, slice-by-slice or true $3D$ image acquisition based on the characteristics of tissues when placed into a strong magnetic field, and inducing an additional magnetic impulse. In particular, Magnetic Resonance Tomography (MRT) considers hydrogen proton densities, as well as longitudinal and transversal relaxation parameters depending on the protons' neighbourhood. When exposed to a magnetic field, the protons, which have a spinning movement and an own magnetic moment characterized by a gyro-magnetic constant γ, absorb electro-magnetic waves at a so-called *Lamor frequency*, and are also affected by chemical and physical conditions of their neighbourhood, e.g. by the number of protons or neutrons in the atomic nuclei. MRI is highly suitable for the high contrast visualization of different soft tissue types, but fails in structures with low water content such as bones. Other MRI modalities include MR Spectroscopy, where nuclei other than hydrogen such as carbon, sodium or phosphorus are investigated, MR Angiography for blood flow measurements, functional MRI for the quantification of the activation in blood oxygenation, Spatial Amplitude Magnetization Modulation (SPAMM) tagged MRI, MRI Fluid Attenuated Inversion Recovery (FLAIR), and, most recently, MR Elastography.

Finally, nuclear medicine acquisition methods like Single Photon Emission Computed Tomography (SPECT), and Positron Emission Tomography (PET) are invasive functional imaging techniques using radioactive isotopes in order to localize the distribution of physiological and pathological processes rather than anatomic information. SPECT is based on the projection of emitted gamma rays, while PET is based on the *annihilation reaction* caused by the collision of emitted positrons with electrons, and resulting in the simultaneous emission of two gamma rays travelling in opposite directions. These gamma rays are then detected as pairs in coincidence by specialized detectors which are arranged in a ring around the patient. For positron emission, atoms like fluorine (for perfusion studies), oxygen, nitrogen, and carbon are used. Both techniques allow for the temporal analysis of flow and other processes. They are often used in conjunction with MRI or CT in order to register the location of physiological or pathological processes to the corresponding anatomical location.

In clinical neurology, MRI has become an increasingly popular image acquisition method due to its superiority in visualizing soft tissues in the brain such as grey or white matter, providing a good anatomical reference frame for qualitative nuclear medicine imaging techniques, and complementing CT in providing high contrast soft tissue information. In various neurological disorders and diseases such as epilepsy and Multiple Sclerosis (MS), respectively, structural deformations of certain parts of the brain tissues occur and even contribute to the disease. Finding a correspondence between the degree of deformation and the disease status, as well as monitoring shape changes over time, is therefore a mandatory task for patient diagnosis, as well as operative and drug treatment planning. A traditional way of detecting structural abnormalities is to segment the brain into its constituent parts, followed by characterizing the extracted shape in terms of volumetric measurements and other shape descriptors. Segmentation is usually performed manually or interactively by analysing the images at a medical workstation. High demand on expert time, however, as well the need to eliminate variability of the results when obtained by two different clinicians, or by one clinician at different times (so-called inter- and intra-observer variability), imposes the need to automatically post-process the extremely large volumetric data sets and image-derived shape information. Post-processing must be carefully designed and validated in order to provide the means for enhanced shape interpretation and decision-making processes based on both scientific and clinical information contained in the images.

This dissertation presents a novel and automatic multi-resolution tool for *multi-scale active shape description* in medical imaging. This tool enables quantitative and qualitative shape studies at multiple levels of image detail, exploring the scale or image resolution as an extra degree of freedom. This approach is motivated by the observation that using scale-space continuity, global object shape characteristics become best visible at a coarse level of image detail or at higher *scales*, and finer shape properties can only be found at higher levels of detail or at lower scales.

1.1 Background

The proposed technique for multi-scale active shape description is based on three major topics in image processing: shape description, scale-space theory, and active contour models. These topics are unified in this dissertation in order to access and examine shape properties of varying specificity directly from the surrounding image context.

1.1.1 Shape Description

Traditionally, a shape is represented by its external characteristics in terms of its outline, an approach which is followed in this work. Shape description is then a post-processing technique which is applied to presegmented binary shapes of high detail.

Shape representation, underlying the description process, can be divided into *local*, *global*, and *medial* (combined local and global) methods. The main difference between local and global shape representation lies in the nature of the access of shape properties: Local representation allows to extract specific detail, but does not readily capture the overall shape outline. In contrast, global techniques integrate over more general shape properties, but do not allow to access specific, localized characteristics. Medial shape representation, though difficult to obtain, combines the positive aspects of local and global representation techniques by hierarchically structuring a shape into its subparts, giving rise to both global and local shape properties simultaneously.

Shape description can analogously be divided into local, global and medial analysis, with the additional concept of shape comparison for *relative* shape investigation. Local descriptors mainly include differential properties of the shape outline, and are most commonly used when the focus is on analysing specific details such as local bending behaviour, smoothness, or degree of convolution. Global descriptors are applied when the shape properties that carry over the entirety of the shape are to be examined. Examples are boundary length and area in two dimensions, and surface area and volume in three dimensions, which form the basic shape measurements in medical imaging. Additionally, global and local relative distance measures indicate an overall or specific deviation or *mismatch* from another shape. Medial descriptors are useful if the focus is on detecting symmetric shape properties, but can only be extracted from a medial shape representation, in which case also medial relative distance measurements can be derived.

Shape description in the classic sense suffers from two major disadvantages: It can only be applied on presegmented shapes and is therefore dependent on a prior segmentation technique. This also leads to the loss of image context. Second, when analysing a binary shape of high specificity, spurious local detail compromises the analysis, while more general shape characteristics cannot be easily extracted. This leads to the need for multi-scale rather than single-scale analysis which will be briefly introduced in the following.

1.1.2 Multi-Scale Image Processing

Multi-scale image processing and analysis, first introduced by Marr [Marr, 1982] as a concept in human vision, allows to observe and describe objects at varying levels of detail. The lower limit or finest detail is determined by the image resolution or *inner scale* which corresponds to the size of the image pixels, while the field of view poses the upper limit or *outer scale*. The inner and outer scale together determine the *total number of degrees of freedom* of the measurement, which corresponds to the total number of image pixels [Koenderink, 1984]. Increasing the inner scale has a blurring effect, as finer scale details disappear. Continuously increasing the inner scale builds up a *scale-space* [Witkin, 1983]. After fixing the inner scale, all smaller image geometry is destroyed and cannot be restored, as there is no structure within a pixel. One can distinguish between two main types of scale-spaces: Linear scale-spaces are based on blurring with a Gaussian kernel as a local neighbourhood operator, but often have the undesirable effect of "blurring across edges". Nonlinear scale-spaces in contrast take higher order invariant differential properties of the image into account, leading to the preservation and even enhancement of edges while smoothing out spurious detail, but are often only well defined in two dimensions and are of considerably higher computational cost. Either approach, however, provides a useful platform for hierarchical geometric image interpretation as well as robust segmentation.

Applications in multi-scale processing can be divided into contour-based methods and imaging techniques. The former construct a contour scale-space in order to obtain a qualitative *sketch* based on multi-scale contour properties, like the concepts of scale-space *fingerprints* [Witkin, 1983] and *curvature scale-space* [Mokhtarian and Mackworth, 1986], and can also be used for contour recognition and indexing. The latter have their main use in image feature detection and multi-scale image segmentation, e.g. in terms of *edge focusing* [Bergholm, 1987], which is a robust technique to track a coarse shape estimate obtained at a high image scale down using decreasing levels of scale.

The main advantage of multi-scale over single scale techniques is their robustness towards noise and other artefacts, as well as their hierarchical nature of capturing shape characteristics of varying locality and detail. Classic multi-scale contour techniques are mathematically well defined, but require a *ground truth* shape which can be *evolved* for increasing levels of scale. This shape, if not analytically known, has to be obtained from a prior segmentation, which in turn is dependent on the chosen image scale in the segmentation process. The reverse process of *focusing* a contour is not possible, as lower scale information is lost at higher scales, and different shapes eventually all evolve to a circular shape. In contrast, multi-scale image techniques allow to segment a shape in a *coarse-to-fine* fashion by pre-computing an image scale-space. However, current image based scale-space approaches discard all but the lowest scale shape which is taken as the segmentation result.

1.1.3 Active Contour Models

Active contour models (also called *snakes*), were first presented by Kass, Witkin and Terzopoulos [Kass *et al.*, 1987b] for shape segmentation based on contour characteristics as well as the surrounding image context. The forces acting on the model are divided into internal smoothness constraints like elasticity and bending behaviour, and external image forces like salient image features which actually define the shape outline. User constraint forces can be added in order to guide the segmentation process towards a desired solution. All forces are formulated within an energy functional which is minimized by deforming the contour in an optimization process. The energy function represents the best compromise of all terms, and can be adjusted to different types of imagery and shapes by using different weighting factors for the individual forces. The optimization of active contour models is traditionally either based on a global or local technique. Global techniques guarantee reproducible and general solutions, but are of high computational cost, and might not lead to the desired solution if it does not correspond to the global minimum. Local techniques, in contrast, are much faster, but are often affected by local energy minima caused by noise and inadequate initialization. Using scale-space continuity in active contour models in order to perform a global to local optimization with respect to decreasing levels of scale has been suggested in [Kass *et al.*, 1987b], but has not been widely followed in literature.

The classic active contour model suffers from several problems: Though defined as a continuous contour, it is optimized discretely, leading to low robustness, numerical instability, as well as sensitivity towards noise and spurious image features. The formulation of the internal forces as smoothness constraints does not allow for the extraction of highly convoluted shapes (as encountered in clinical neurology, for example). Finally, active contour models are a pure segmentation tool which is usually applied at a single scale level only.

1.2 Contributions of this Dissertation

This dissertation makes several original contributions to the medical and general image processing community:

Formulation of a multi-scale active contour model. The classic active contour model is analysed in order to find solutions for its deficiencies, namely its discrete representation, its inability to extract complex shapes, and its single-scale nature. A continuous spline representation is formulated which has an inherent contour scale in terms of the spacing of the spline control points. Choosing a contour scale in correspondence to the underlying image scale is motivated by the observation that the image scale-space is less dense at higher scales. Fewer control points are therefore necessary to capture the global shape outline at high image scales, while at lower scales finer shape detail can only be adequately located if the contour scale is relatively low. Adjustment of the contour scale is performed in an adaptive sampling process based on discrete, interpolated or heuristic spline knot insertion, which are compared to alternative uniform and adaptive curvature based sampling techniques. In order to allow for the extraction of shapes with high curvature parts, a curvature matching process is developed which adjusts the model's shape to the underlying image characteristics not only in terms of its boundary location, but also with respect to its local bending behaviour with respect to the isophote image curvature. Finally, in combination with scale continuity, a local optimization is found to be sufficient for adequate shape extraction, as the image scale defines the globality or locality of the solution. Three local techniques found in the literature are modified in order to accommodate for the multi-scale spline-based representation and the curvature matching process.

Formulation of a multi-scale description process. Classic multi-scale shape representation and analysis is generalized to a *multi-scale shape stack* which is formed by a shape hierarchy in image scale-space. The shape stack is obtained using the multi-scale active contour model for *implicit* segmentation or shape regularization with respect to image scale, rather than for explicit segmentation. Shape regularization can be performed in two directions of the image scale-space, giving rise to two novel techniques:

Active shape evolution. This technique is similar to the construction of a classic contour scale-space, yet it takes the full image context defining the shape into account. Starting from the ground truth as an initial active contour model, an image scale-space is computed, and the model is tracked through it for increasing levels of scale in a *fine-to-coarse* manner.

Active shape focusing. This technique is performed in analogy to edge focusing by starting from a very coarse initial estimate (e.g. a circular or ellipse shaped model), and tracking this contour model through its surrounding image scale-space in a *coarse-to-fine* manner. In contrast to classic edge focusing, all intermediate *implicit* segmentation results are retained and investigated. In contrast to classic curve evolution, the blurring process is "reversed", yielding an approximation of the ground truth shape (which need not be known *a priori*) as a byproduct.

Dimensionalities of the multi-scale shape stack in terms of the underlying scale-space dimensionality and the nature of the implicit extraction process are formulated. *Multi-scale active shape description* then comprises the construction and analysis of a multi-scale shape stack obtained by either method. Analysis is performed by investigating classic local, global, and relative shape metrics across scale. Each layer or scale level of the multi-scale shape stack is individually quantified, and shape changes between the layers are investigated by computing mean and slope measurements for increasing or decreasing scale levels. Qualitative inspection is performed by adequate visualization of the shape stack, and the mapping of local shape information.

Application to fractal and other synthetic images. The developed techniques for active shape evolution and focusing are tested on a set of synthetic silhouette images in order to demonstrate their duality as well as their applicability to extract a variety of different shapes adequately. A new fractal measurement is derived from shape changes across scale, and is demonstrated to be correct on a true fractal structure.

Application to medical images. Three different applications in clinical neurology are addressed, namely cortical dysgenesis for patients with epilepsy, spinal cord atrophy for patients with Multiple Sclerosis, and cortical impairment for neonates. A small but representative selection of MRI data sets is tested using the newly developed technique of *active shape focusing* in absence of a ground truth. Multi-scale shape characteristics and structural differences, which are otherwise not easily perceived, extracted or quantified, are recovered using multi-scale active shape description. No large-scale clinical evaluation is performed, but instead the applicability and functionality of the developed methodologies is demonstrated.

1.3 Overview of Dissertation

This dissertation is structured in the following way:

Chapters 2-4 give a comprehensive overview and survey of classic and state-of-the-art techniques in shape description, multi-scale image processing, and active contour models, respectively, in order to formulate the work in the context of existing work. Positive and negative aspects are highlighted, and their potential for adaptation and modification in this thesis is addressed in the individual summary sections. Parts of this survey have been published in [Schnabel, 1995].

Chapter 5 forms an inter-mezzo to the main part of this dissertation, by selecting the most important techniques from the survey part as the basis for multi-scale active shape description, and presenting an overview about the interaction of the chosen techniques. Chapter 6 introduces the theoretical framework for the multi-scale active contour model, including representation, sampling, and optimization aspects in terms of illustrative examples and algorithms. Chapter 7 presents the concept of the multi-scale shape stack and formulates the techniques of active shape evolution and focusing. Multi-scale active shape description is introduced on the basis of multi-scale shape stacks of different dimensionalities, as well as on an adequate selection of classic shape metrics. An example for multi-scale active shape description is given in order to clarify the use of the presented techniques. Parts of this framework have been published in [Schnabel and Arridge, 1995] (for active shape evolution in $2D$ images), [Schnabel and Arridge, 1996b] (for active shape focusing in $2D$ images), [Schnabel and Arridge, 1996c; Schnabel *et al.*, 1996] (for active shape focusing of $3D$ MRI), and [Schnabel and Arridge, 1996a] (for the construction and visualization of multi-scale shape stacks).

Chapter 8 demonstrates the applicability of the proposed techniques on a set of fractal and other synthetic examples. The duality of active shape evolution and focusing is shown, and a new fractal measurement in terms of shape changes across scale is derived. Chapter 9 applies multi-scale active shape description on MRI data in clinical neurology. Parts of these chapters have been published in [Schnabel and Arridge, 1997a; Schnabel and Arridge, 1997b] (for the application to synthetic and fractal images), and [Schnabel and Arridge, 1997c; Schnabel and Arridge, 1997d] (for the application to MRI data).

Chapter 10 formulates future directions on the basis of limitations of the presented techniques, and concludes this dissertation. In particular, extensions to a $3D$ explicit or implicit models, the application to non-linear scale-spaces, and a true multi-scale rather than *fine-to-coarse* or *coarse-to-fine* approach are presented, comparisons to classic contour sketches are addressed and formulated, and the main achievements and results are summarized. Appendix A contains colour plates for the qualitative visualization of the multi-scale shape stacks of chapters 7-9.

1.4 Definitions and Notations

This section briefly introduces some general definitions and notations used in the remainder of this dissertation. In particular, table 1.1 lists general symbols and operators. In general, subscripts are used to denote partial derivatives with respect to the subscript, or for enumeration, e.g from

$\mathbf{x}$	Vector (bold)
$\mathbf{X}$	Matrix (bold)
$\mathbf{I}$	Identity matrix
$\mathbf{x}^T$, $\mathbf{X}^T$	Transposed vector and matrix
$\det \mathbf{X}$	Determinant of a matrix
f	Scalar function
$\mathbf{f}$	Vector-valued function
$\mathbf{t}$, θ	Tangent vector and angle
$\mathbf{n}$	Normal vector
κ	Curvature
L	Image luminance function
$L(\mathbf{x})$	N-dimensional image
$\mathcal{F}$	Fourier transform
$\mathcal{L}$	Fourier transformed image
j	$\sqrt{-1}$
u	Frequency
σ	Standard deviation or scale
$\otimes$	Convolution operator
∂	Partial derivative operator
d	Differential operator
div	Divergence operator
∇	Gradient operator
Δ	Laplacian operator
$\mathbf{H}$	Hessian matrix of second derivatives
e	Euler constant
$\mathbf{e}$, λ	Eigenvector and Eigenvalue
$\mathbf{p} = (x, y, z)$	Point (bold)
$\mathbf{P}$	Point set, e.g. $\mathbf{P} = \{\mathbf{p}_1 \cdots, \mathbf{p}_N\}$
D	Euclidean dimension, e.g. $2D$ for two-dimensional
$\|\cdot\|$	Euclidean norm
$\cdot$	Scalar product
$\wedge$	Vector product
$\|c\|$	Absolute value of a scalar c
ϕ	Angle
$\max, \min$	Maximum and minimum operators
$[\,]$, $(\,)$, $[\,)$, $(\,]$	Closed, open, and semi-open intervals

Table 1.1: General symbols, operators and notations.

$1 \cdots N$. Superscripts in brackets are used as indices with respect to an iteration, as opposed to ordinary superscripts used for raising to some power. Special terms and definitions are listed in the glossary section at the end of this dissertation, along with the page number where they are introduced and defined. An index register is included to provide a quick reference for the most important terms and techniques.

Chapter 2

Survey of Shape Description Methods

> ELLES M'ONT RÉPONDU: "POURQUOI UN CHAPEAU FERAIT-IL PEUR?" MON DESSIN NE REPRÉSENTAIT PAS UN CHAPEAU. IL REPRÉSENTAIT UN SERPENT BOA QUI DIGÉRAIT UN ÉLÉPHANT. J'AI ALORS DESSINÉ L'INTÉRIEUR DU SERPENT BOA, AFIN QUE LES GRANDES PERSONNES PUISSENT COMPRENDRE. ELLES ONT TOUJOURS BESOIN D'EXPLICATIONS.
>
> BUT THEY ANSWERED: "FRIGHTEN? WHY SHOULD ANY ONE BE FRIGHTENED BY A HAT?" MY DRAWING WAS NOT A PICTURE OF A HAT. IT WAS A PICTURE OF A BOA CONSTRICTOR DIGESTING AN ELEPHANT. BUT SINCE THE GROWN-UPS WERE NOT ABLE TO UNDERSTAND IT, I MADE ANOTHER DRAWING: I DREW THE INSIDE OF THE BOA CONSTRICTOR, SO THAT THE GROWN-UPS COULD SEE IT CLEARLY. THEY ALWAYS NEED TO HAVE THINGS EXPLAINED.
>
> *Le Petit Prince, Antoine de Saint-Exupéry.*

In order to describe an object's shape more accurately than simply using terms like *elongated, rounded, curved, with sharp edges* or *straight*, various shape description techniques have been developed for improved quantitative and qualitative measurements. Shape description is either based on segmentation, followed by analysis of external characteristics of the binary shape, or it works directly on the grey-level image when the focus is on internal shape characteristics like texture or other intensity-related features. In either case, the chosen shape descriptor should be translation and rotation invariant and insensitive to changes of scale. The aim of this dissertation is to analyse a shape boundary rather than its interior. Defining a shape boundary as a planar curve gives rise to three different curve forms: The explicit form is given by $y = y(x)$, the implicit form by $f(x, y) = 0$, and the parametric form by $\mathbf{v}(s) = (x(s), y(s))$, with $s \in [0; 1]$ as the arc length parameter.

This chapter first presents several local, global, and medial (as a combination of local and global) *shape representation* techniques based on a planar curve as the shape boundary. Then *shape description* techniques based on local, global, medial, and relative measurements are presented in order to provide a consistent and comprehensive overview of research done in this area. An overview of the methods surveyed in this chapter is illustrated in figure 2.1.

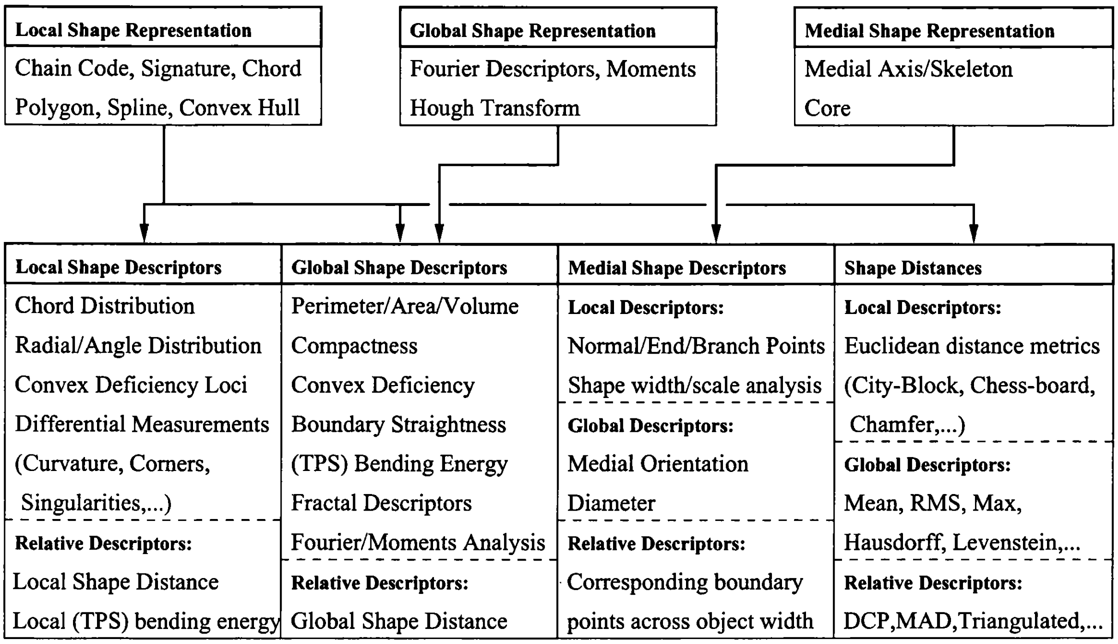


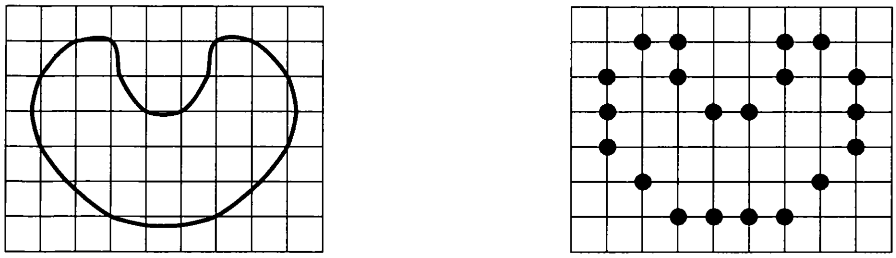
Figure 2.1: Overview of shape representation and description techniques

2.1 Shape Representation

After an image has been segmented into relevant regions, the resulting aggregate of segmented pixels needs to be represented in a form suitable for further analysis of the shape outline. Figure 2.2(a) shows a shape outline superimposed onto a sparse grid, and figure 2.2(b) shows the shape representation as a result of resampling with respect to the grid. As can be seen, the inner resolution or level of detail of a shape is dependent on the resolution of the chosen underlying coordinate system (in this case the grid). Various shape representation techniques exists, allowing to investigate a shape's external characteristics. One can distinguish between local representation, enabling direct access of the shape boundary points, global representation, representing the entirety of the shape, and medial techniques, which combine local with global representation. All representation techniques lead naturally to local, global and relative shape measurements, which will be discussed in sections 2.2–2.4. However, some of the representation techniques presented are *inherent* shape descriptors, which will be indicated when appropriate.

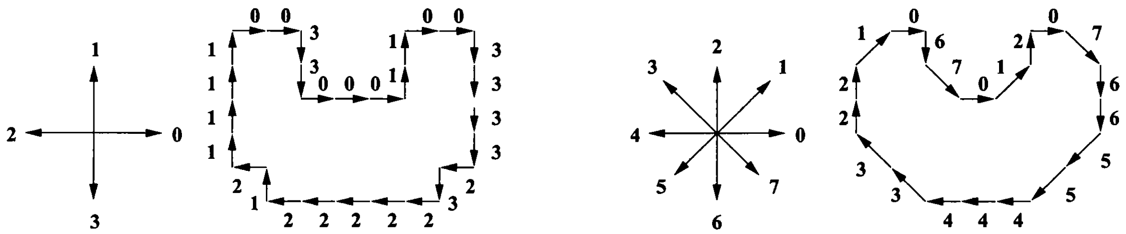
2.1.1 Local Shape Representation

Local shape representation has the advantage of providing a concise, direct access to the shape under investigation. One has to distinguish between relative and absolute techniques, where the former provide an efficient, compact, and invariant way of coding an object independently from its context, and the latter are more complex, but allow to access the object location directly. Relative representation techniques can be transformed into absolute representations by keeping a starting point address, along with orientation, offset and scaling information.



(a) (b)

Figure 2.2: Resampling of digital boundary: (a) Superimposed resampling grid. (b) Result of resampling.



(a) (b)

Figure 2.3: Chain code representation of the boundary in figure 2.2: (a) 4-directional chain code representation. (b) 8-directional chain code representation.

2.1.1.1 Chain Codes, Signatures and Chords

The *chain code* representation [Freeman, 1974] is one of the earliest representation techniques in image processing. It describes a digital boundary as a connected sequence of direction vectors based on 4- or 8-connectivity (see figure 2.3). It allows for a complete reconstruction of an object, given a reference to the absolute object location. Its disadvantages are that chain codes can get rather long for complex objects (chain codes of simple objects can be easily compressed), and that small disturbances of the boundary due to noise or a different segmentation technique can cause unwanted changes in the code which may be difficult to incorporate. A form of *generalized chain coding* encodes the boundary curvature as a function of boundary path length (curvature will be defined in section 2.2).

The *signature* of a shape [O'Rourke, 1986] can be obtained as a sequence of normal boundary distances and is computed for each boundary element as a function of the boundary path length. For each boundary point $\mathbf{a}$, the distance of an opposite border point $\mathbf{b}$ is found in direction perpendicular to the border tangent at $\mathbf{a}$, yielding a distance function $d(s)$ of boundary parameter s. Note that this is not necessarily a symmetric relation (e.g. given $d_{\mathbf{ab}}(s)$, $d_{\mathbf{a'b'}}(s)$ with $\mathbf{a'} = \mathbf{b}$ does not necessarily imply that $\mathbf{b'} = \mathbf{a}$) and can lead to difficulties for objects with concavities. Moreover, signatures are very sensitive to noise. Figure 2.4(a) shows an example of a signature construction and the symmetry problem.

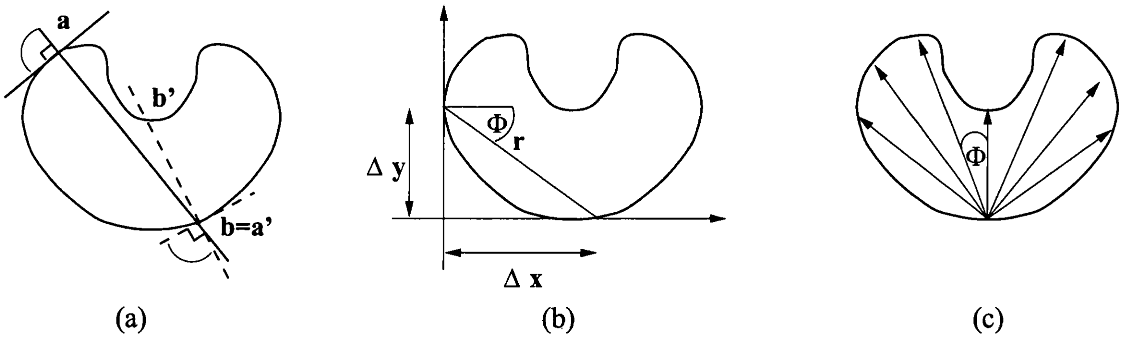


Figure 2.4: (a) Signature construction. (b) Chord distribution. (c) Rotation-independent radial distribution.

The distribution of lengths and angles of all *chords* on a shape boundary may be used for shape representation as well, where a chord is a line joining any two points of the shape boundary (figure 2.4 (b)). The chord distribution d_{chord} is computed as

$$d_{chord}(\Delta x, \Delta y) = \int_x \int_y f(x,y) f(x+\Delta x, y+\Delta y)\, \mathrm{d}x\, \mathrm{d}y, \tag{2.1}$$

where $f(x,y) = 1$ if (x,y) is a boundary point, and $f(x,y) = 0$ for all other points. To obtain a rotation-independent radial distribution $d_{radial}(r)$, the integral over all angles is computed:

$$d_{radial}(r) = \int_{-\pi/2}^{\pi/2} d_{chord}(\Delta x, \Delta y)\, r\, \mathrm{d}\theta, \tag{2.2}$$

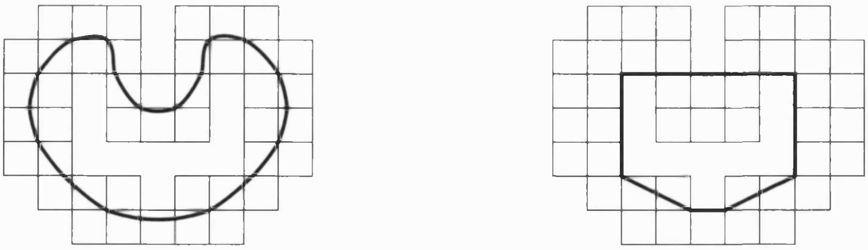
where $r = \sqrt{\Delta x^2 + \Delta y^2}$ and $\theta = \tan^{-1}\left(\frac{\Delta y}{\Delta x}\right)$. This radial distribution varies linearly with scale, while the angular distribution $d_{angular}(\theta)$ is scale-independent, with the rotation causing a proportional offset:

$$d_{angular}(\theta) = \int_0^{\max(r)} d_{chord}(\Delta x, \Delta y)\, \mathrm{d}r \tag{2.3}$$

Figure 2.4(c) shows an example for radial distribution. Arc chord distances have been used for shape partitioning by [Phillips and Rosenfeld, 1987], and chord distribution has been applied for shape matching by [Smith and Jain, 1982; You and Jain, 1984]. Combinations of chord and radial distribution have been successfully used for shape description [Cootes *et al.*, 1992a].

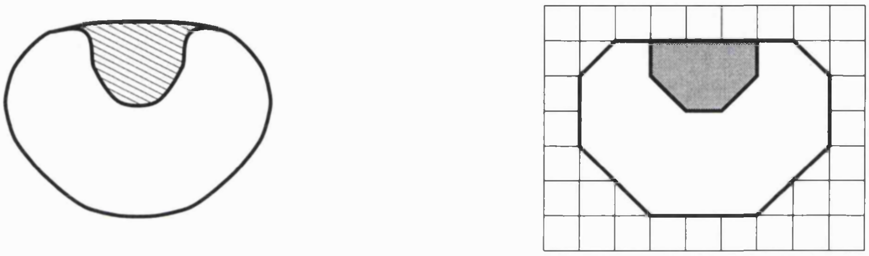
2.1.1.2 Polygonal Representation and Convex Hull

A digital boundary can be approximated with arbitrary accuracy by a polygon, where the most exact approximation is defined as the number of segments of the polygon being equal to the number of points of the boundary. A *polygonal representation* offers the possibility to describe the essential boundary shape with the fewest possible polygonal segments [Pavlidis, 1977]. An example for *curvature-based polygonal approximation* is given in [Wu and Wang, 1993]. They propose to apply a simple corner detection method to locate potential corners, followed by partitioning each curve segment between any two consecutive potential corners via standard polygonal approximation. *Minimum perimeter polygons* approximate a boundary by enclosing the boundary by a



(a) (b)

Figure 2.5: (a) Object boundary enclosed by cells. (b) Minimum perimeter polygon.



(a) (b)

Figure 2.6: (a) Boundary enclosed by convex hull with convex deficiency marked in grey. (b) Polygonal convex hull and deficiency representation.

set of concatenated cells and regarding the boundary as a rubber band contained within the inside and outside boundaries of the strip of cells, shown in figure 2.5(a). Shrinking that rubber band produces a polygon of minimum perimeter, as shown in figure 2.5(b).

The *convex hull* of a finite point set was first presented by [Graham, 1972] and later expanded among others by [Akl and Toussaint, 1978; Graham and Yao, 1983; McQueen and Toussaint, 1985; Toussaint, 1985]. It defines the minimum area convex polygon containing all points of the set. The set difference between the convex hull and the point set is called the *convex deficiency* of the set. However, as most digital objects tend to be irregular, convex deficiency has rather small scattered components. For this reason, prior to finding the convex deficiency a polygonal approximation can be used. An example of a convex hull and its convex deficiency is shown in figure 2.6. Shape description using these representation schemes can be based on the area of the convex hull, the area of its convex deficiency, the number of components of the convex deficiency (which is the number of concavities of the shape) and the relative locations of these components. Fast algorithms for the computation of the convex hull of simple polygons can be found in [Chen, 1989; Hussain, 1988].

2.1.1.3 Spline Representation

Splines refer to flexible metal strips which are used to lay out the surfaces of ships, cars and airplanes. Weights are attached in order to pull the spline to desired directions. The mathematical equivalent, *natural cubic splines* with a polynomial basis function of order 3, are used as a second-order, C^2 continuous model to smoothly interpolate a set of given points. A natural spline representation is *global*, as the coefficients of a natural cubic spline are dependent on all n control points, involving an $n+1$ by $n+1$ matrix inversion for the computation [Bartels *et al.*, 1987]. In order to obtain a computationally more efficient, *local* spline representation, a smoothed curve can be created by using a set of curve segments whose coefficients depend only on a few points. The boundary is then described by the set of curve segments $\mathbf{v} = \{\mathbf{v}_1(s), \ldots, \mathbf{v}_n(s)\}$, $s \in [0;1]$. Each curve segment $\mathbf{v}_i(s)$ describes a curve starting at point $\mathbf{p}_i = \mathbf{v}_i(0)$ and ending at point $\mathbf{p}_{i+1} = \mathbf{v}_{i+1}(0)$. For open curves, special treatment of the endpoints needs to be taken. For closed curves, the set of curve segments becomes periodic. Splines have been applied for tracking [Bartels *et al.*, 1994], smoothing [Chen and Chin, 1993; Unser *et al.*, 1993], object modelling and shape estimation [Cohen and Wang, 1994; Wang and Cohen, 1994], shape preserving approximation [Howell *et al.*, 1993], and surface matching [Szeliski and Lavallée, 1994]. In the following, two spline models, one which is commonly used for curve approximation and interpolation, and the other for surface interpolation and deformation analysis, will be briefly presented.

2.1.1.3.1 B-Splines. *B-splines* are local interpolants of the same continuity as natural cubic splines, but they do not necessarily interpolate (pass through) the control points. Each curve segment $\mathbf{v}_i$ is defined by the points $\mathbf{p}_{i-3}$, $\mathbf{p}_{i-2}$, $\mathbf{p}_{i-1}$, and $\mathbf{p}_i$ and can be described by the B-spline geometry vector, $\mathbf{G}_{Bs_i}$:

$$\mathbf{G}_{Bs_i} = \begin{pmatrix} \mathbf{p}_{i-3} \\ \mathbf{p}_{i-2} \\ \mathbf{p}_{i-1} \\ \mathbf{p}_i \end{pmatrix} \tag{2.4}$$

The B-spline basis matrix, $\mathbf{M}_{Bs}$ is derived by [Bartels *et al.*, 1987] and relates the geometrical constraints $\mathbf{G}_{Bs}$ to the blending functions and polynomial coefficients:

$$\mathbf{M}_{Bs} = \frac{1}{6} \begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{pmatrix} \tag{2.5}$$

The entire curve is generated by applying for every curve segment i and row vector $\mathbf{T} = [s^3 \quad s^2 \quad s \quad 1]$, $0 \leq s < 1$, the following transformation:

$$\mathbf{v}_i(s) = \mathbf{T} \cdot \mathbf{M}_{Bs} \cdot \mathbf{G}_{Bs_i} = \begin{pmatrix} x_i(s) \\ y_i(s) \end{pmatrix}, \tag{2.6}$$

yielding a curve in parametric form. Given the analytic representation of B-splines, the polynomials defined by this transformation can be differentiated:

$$\dot{\mathbf{v}}_i(s) = \begin{pmatrix} \dot{x}_i(s) \\ \dot{y}_i(s) \end{pmatrix} \quad \text{and} \quad \ddot{\mathbf{v}}_i(s) = \begin{pmatrix} \ddot{x}_i(s) \\ \ddot{y}_i(s) \end{pmatrix} \tag{2.7}$$

This allows for computation of the spline patch tangent

$$\mathbf{t}_i(s) = \left(\frac{\dot{x}_i(s)}{\sqrt{\dot{x}_i(s)^2 + \dot{y}_i(s)^2}}, \frac{\dot{y}_i(s)}{\sqrt{\dot{x}_i(s)^2 + \dot{y}_i(s)^2}} \right) \tag{2.8}$$

and normal

$$\mathbf{n}_i(s) = \left(\frac{-\dot{y}_i(s)}{\sqrt{\dot{x}_i(s)^2 + \dot{y}_i(s)^2}}, \frac{\dot{x}_i(s)}{\sqrt{\dot{x}_i(s)^2 + \dot{y}_i(s)^2}} \right) \tag{2.9}$$

Advanced B-spline models include the works by [Goldman and Warren, 1993], and other locally controlled spline models include the *Overhauser splines* developed by [Brewer and Anderson, 1977] of the *Catmull-Rom* family [Catmull and Rom, 1974] which smoothly interpolate a set of points. They can be formulated as higher order polynomials, but this may cause unwanted oscillations and zero-crossings of higher-order derivatives. Extensions to 3D spline surfaces are straight forward in principle, but are quite complex in terms of implementation issues.

2.1.1.3.2 Thin-Plate Splines. A different form of splines which are mainly used for shape description rather than representation are the so-called surface-interpolating *thin-plate splines* which can also be used to determine the physical bending energy of a thin metal plate on point constraints [Bookstein, 1989; Bookstein, 1991a; Bookstein, 1991b]. Interpolating over a fixed set of (possibly) irregularly spaced planar points, the bending energy is defined by a quadratic form of the heights assigned to the surface. The spline itself is expressed as the superposition of eigenvectors of the bending energy matrix over a tilted plane having no bending energy at all. Pairing the splines f_x, f_y for representing the x-y-coordinates of the plane yields an interpolation map relating to sets of *landmark points*. Both the spline maps and surfaces are composed by a linear part (which represent *affine transformations*) and a geometrically non-affine part, called *principal warps*. Defining a surface by

$$f(x,y) = -U(r) = -r^2 \log r^2 \quad \text{with} \quad r = \sqrt{x^2 + y^2}\,, \tag{2.10}$$

yields a special function U which satisfies

$$\Delta^2 U = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)^2 U \propto \delta_{(0,0)} \tag{2.11}$$

The right hand side of this equation is proportional to the "generalized function" or delta-function $\delta_{(0,0)}$ (which is only not equal to zero at the origin, and has an integral of 1). U is therefore a fundamental solution of the biharmonic equation $\Delta^2 U = 0$, which is the equation of a shape made of thin plate $f(x, y)$ lofted above the x-y-plane. Note that this basis function is the generalization of the basis function of natural cubic splines to two dimensions [Wahba, 1990]. Applying only slight bendings to the surface $f(x, y)$ yields a bending energy proportional to $\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2$ at any point $\mathbf{p}$ of that surface, minimizing

$$\int\int_{\Re^2} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 \mathrm{d}x\, \mathrm{d}y \tag{2.12}$$

in order to take a shape in which it is least bent. For modelling and analysing the deformation between two sets of *landmarks* (which can be *biological landmarks* having meaningful and reproducible biological counterparts in both sets, or *pseudo-landmarks* such as registration points or other fiducials), two separate thin-plate spline functions f_x and f_y are used to model the displacement of the landmarks in the x and y directions. This leads to the definition of a vector-valued thin-plate spline mapping function $\mathbf{f} : (x, y) \rightarrow (f_x(x, y), f_y(x, y))$ which maps all points $\mathbf{p}_i$ located in the Euclidean image plane into their homologues $\mathbf{p}_i'$. Each thin-plate spline mapping function can be written as

$$f(x, y) = a_0 + a_x x + a_y y + \sum_{i=1}^{N} w_i U(\|(x, y) - \mathbf{p}_i\|) \tag{2.13}$$

$\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n)$ can be defined as a vector with two-component weighting vectors $\mathbf{w}_i$ (one component for each spatial dimension) which sum to zero and whose crossproducts with the x- and y-coordinates of the points $\mathbf{p}_i$ are equally zero. The first part of equation (2.13) presents the affine transformation part, while the summation part forms the principal warps. Thus given a deformation of a set of points in a plane, a deformation of *all* points of the plane in terms of an interpolated value as well as the bending energy necessary for the deformation can be computed using this concept.

Applications of thin-plate splines include surface interpolation and smoothing [Dyn *et al.*, 1983], statistical interpolation of scattered data [Wahba, 1990], landmark feature spaces [Bookstein and Green, 1992], visualization and morphometric analysis of group differences [Bookstein, 1996a; Bookstein, 1996b; Dean *et al.*, 1996], medical image registration [Edwards *et al.*, 1995; Little *et al.*, 1996; Edwards *et al.*, 1997; Lester and Arridge, 1997], cardiac deformation analysis [Sanchez-Ortiz *et al.*, 1996b], and deformable templates [Rueckert and Burger, 1997].

2.1.2 Global Shape Representation

The presented local shape representation techniques, though very suitable for local shape access, suffer from the disadvantage of not providing an overall object representation. In contrast to that, global representation techniques allow to capture global characteristics inherent to the entirety of the object. The most frequently global shape representation techniques are *Fourier* techniques, statistical moments, and the *Hough* transform.

2.1.2.1 Fourier Representation

A *Fourier* representation decomposes a shape contour into its frequency components (or *Fourier descriptors*) obtained via its Fourier transform, where each Fourier descriptor describes a global property of the shape. Lower frequency components correspond to gradual changes of the contour, while higher frequency components account for example for greater curvature of the contour, leading to a *coarse-to-fine* hierarchical shape representation.

In order to compute the Fourier representation of a parameterized contour $\mathbf{v}(s)$, a complex periodic function $f(s)$ is defined which expresses the shape as a sequence of coordinates in the complex plane, namely $f(s) = x(s) + jy(s)$, with $j = \sqrt{-1}$. After having reduced a $2D$ problem into a $1D$ problem, the discrete Fourier transform of $f(s)$ is

$$\mathcal{F}(u) = \frac{1}{N} \sum_{s=0}^{N-1} f(k) e^{-2j\pi us/N} \tag{2.14}$$

for $u = 0 \cdots N-1$. The Fourier descriptors are given by the complex coefficients $\mathcal{F}(u)$. The inverse transform of $\mathcal{F}(u)$ restores the boundary $f(s)$:

$$f(s) = \sum_{u=0}^{N-1} \mathcal{F}(u) e^{2j\pi us/N} \tag{2.15}$$

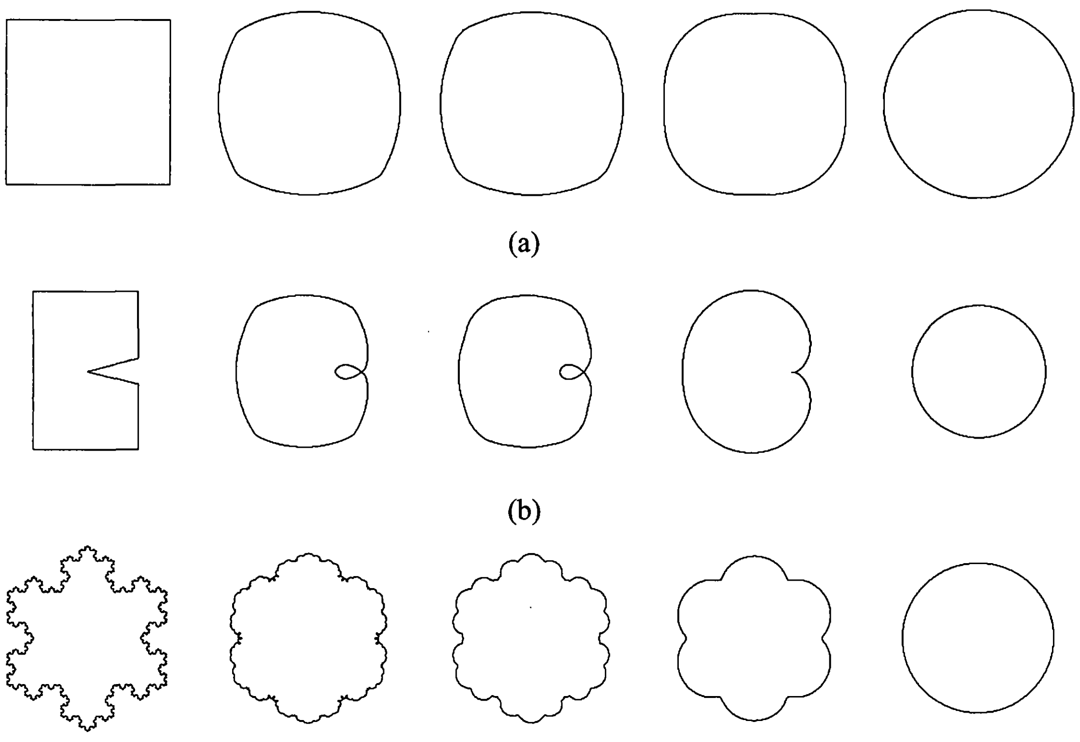
for $s = 0 \cdots N-1$.

However, when only using the first M coefficients to restore the boundary instead of the complete set of N coefficients, the boundary is approximated by

$$\hat{f}(s) = \sum_{u=0}^{M-1} \mathcal{F}(u) e^{2j\pi us/N} \tag{2.16}$$

for $s = 0 \cdots N-1$, although only M terms are used to obtain each component of $\hat{f}(s)$. The approximated boundary still consists of the same number N of points, but less Fourier coefficients are used for the reconstruction. Recall that Fourier descriptors for high u describe high-frequency components (and finer shape details) of the shape boundary, and Fourier descriptors for lower u describe low-frequency shape properties (or broader shape details). Reconstructing a contour using Fourier descriptors for increasing levels of M is a method for hierarchical, frequency-based shape representation. Figure 2.7 shows examples for Fourier shape reconstruction.

However, each Fourier descriptor represents a global rather than local property of the boundary, as the individual Fourier descriptors are computed by integrating over the entire curve. Thus local spatial information about the shape is not readily available, and the level of shape detail can only be controlled on a global basis. Moreover, globally truncating detail can sometimes result in self-intersecting reconstructions which changes the topology of a shape (see figure 2.7, middle row). A method of shape discrimination using Fourier descriptors is presented by [Persoon and Fu, 1986].



(a) (b) (c)

Figure 2.7: Examples of Fourier reconstruction. Reconstructions of: (a) Square, for $M = N = 2048, M = 1024, 32, 8, 2$. (b) Notched rectangle, for $M = N = 2048, M = 1024, 16, 4, 2$. (c) Von Koch curve presented in section 2.3.3, for $M = N = 8192, M = 2048, 32, 8, 2$. Note that the reconstructions shown in (b) and (c) lead to intersections and change of shape topology.

2.1.2.2 Moments

Moment descriptors focus on statistical properties of the object's entire contour outline by computing its statistical moments, e.g. the mean, variance, etc. The infinite set of moments gives a complete description of the shape contour in terms of all statistical aspects like centre of mass, elongation and overall orientation. The $2D$ Cartesian moment μ_{pq} of order $(p+q)$ of a contour $\mathbf{v}(s)$ is defined by the relation

$$\mu_{pq}(\mathbf{v}(s)) = \int_{s=0}^{1}\int_{s=0}^{1}(x(s)-m_x)^p \cdot (y(s)-m_y)^q \, f(\mathbf{v}(s)) \, \mathrm{d}s \, \mathrm{d}s \tag{2.17}$$

for $p, q = 0, 1, 2 \cdots$, and a density distribution function $f(x, y) = (x_f, y_f)$. m_x and m_y are defined as

$$m_x = \int_{s=0}^{1} x(s)x_f(s) \, \mathrm{d}s \qquad \text{and} \qquad m_y = \int_{s=0}^{1} y(s)y_f(s) \, \mathrm{d}s \tag{2.18}$$

f can be defined as binary function when regarding a shape as a binary *silhouette* image:

$$f(\mathbf{p}) = \begin{cases} 1 & \text{if } \mathbf{p} \text{ is inside or on the contour } \mathbf{v}(s) \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

According to the uniqueness theorem [Papoulis, 1965], if f is piecewise continuous and has nonzero values only in a finite part of the x-y-plane, then moments of all order exist and the moment set $\{\mu_{pq}(\mathbf{v}(s))\}$ is uniquely determined by $f(\mathbf{v}(s))$, and vice versa. The zeroth order moment, μ_{00}, computes the *total mass* of $\mathbf{v}(s)$ and is given by

$$\mu_{00}(\mathbf{v}) = \int_{s=0}^{1} \int_{s=0}^{1} f(\mathbf{v}(s))\, \mathrm{d}s\, \mathrm{d}s \tag{2.20}$$

and all higher order moments are computed by setting p and q in equation (2.17) accordingly. For digital curves, the mapping function $f(\mathbf{v}(s))$ becomes discrete, and the integration is replaced by discrete summation. The two first order moments, μ_{10} and μ_{01}, are the *central moments*, which locate the *centre of mass*:

$$\bar{x} = \frac{\mu_{10}}{\mu_{00}} \qquad \text{and} \qquad \bar{y} = \frac{\mu_{01}}{\mu_{00}} \tag{2.21}$$

Normalized central moments are computed via

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \qquad \text{with} \qquad \gamma = \frac{p+q}{2} + 1 \tag{2.22}$$

Second order moments measure the spread of the curve around its mean (*variance*), and third order moments measure the curve's deviation from symmetry about the mean (*skew*), and fourth order moments describe the *kurtosis* of the curve. Similar to Fourier descriptors, details can be truncated by using only the lower order, more general moment descriptors. However, moments only allow to extract global rather than local shape properties, and the complete set of moments is needed in order to reconstruct the original object. Thus moments suffer from the same problems as Fourier descriptors, as their integration of the entire boundary does not allow for local control of detail. A set of seven *invariant moments* with respect to translation, rotation, and zooming can be obtained from second and third moments according to [Hu, 1962]. Applications of moments for shape description are presented in [Sluzek, 1996], and an extensive survey of moment-based techniques for object representation and recognition can be found in [Prokop and Reeves, 1992].

2.1.2.3 Hough Transform

The *Hough transform* [Hough, 1962] was originally developed to detect lines in binary images, but it also offers a way to represent and hence describe shapes of more complex analytic or generalized structures contained in a grey-level images, as it examines the global relationship between potential contour points via their parameter or *Hough* space. Extensions of the classic Hough transform to other analytic curves have been presented in [Leavers, 1992], and a generalization to arbitrary curves has been developed by [Ballard, 1981]. For the detection of analytic curves, their general equation may be rewritten in terms of their parameters. Table 2.1 lists the general equations for lines, circles and ellipses. The parameter space is subdivided into *accumulator cells*, and the parameter equations for each point (x_i, y_i) are solved. The corresponding parameters are

Analytic form	Parameters	Parameter equation
Line (slope intercept)	a, b	$b = -xa + y$
Line (normal equation)	s, ϕ	$s = x\cos\phi + y\sin\phi$
Circle	x_r, y_r, s	$s^2 = (x - x_r)^2 + (y - y_r)^2$
Ellipse	x_r, y_r, s_x, s_y, ϕ	$1 = \frac{(x-x_r)^2}{a^2} + \frac{(y-y_r)^2}{b^2}$ (and rotation by ϕ)

Table 2.1: Analytic curves described with generalized shape parameters x_r, y_r, s_x, s_y, ϕ (adapted from [Ballard, 1981]).

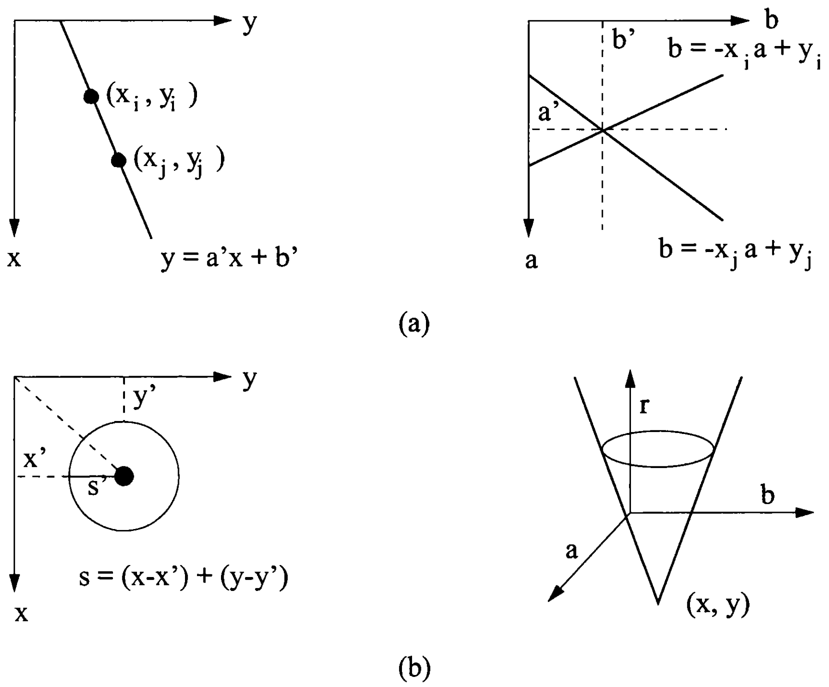


Figure 2.8: Hough transform for detecting (a) lines using slope-intercept form and (b) circles. Both are illustrated in the x-y plane and a-b-space.

used to increment the associated accumulator cell. Local maxima of the accumulator indicate shape instances for the specific parameters. The parameter space is restricted by investigating only the loci of edge pixels (*edgels*) in an image.

For example, lines can be detected by investigating the parameter space spanned by a and b, where exactly one line passing through the fixed point (x_i, y_i) is defined, and all points on this line have such a line in parameter space, intersecting the line associated with (x_i, y_i) at (a', b') (figure 2.8 (a)). As the general slope-intercept form can not be used for vertical lines, the normal ϕ-s-space should be used instead, s defining the distance the line normal to the origin, and ϕ the angle from the x-axis to the normal. Circles and ellipses are detected using higher-dimensional parameter spaces and corresponding accumulators. For instance, circles have a three-dimensional parameter space (figure 2.8(b)), which can be reduced by either fixing the radius or searching for the loci of the parameters on the circular cones of all edge pixels. Additionally, incorporating directional information associated with the edge reduces the parameter locus to a line. For the detection of ellipses, [Nair and Saunders (Jr.), 1996] have developed a fast *straight line Hough transform* (SLHT) based on the signature of the ellipse, along with other geometric ellipse properties. [Aguadao *et al.*, 1996] have used a parametric polar representation of an ellipse, decomposing the parameter space into two independent subspaces and one final histogram accumulator, and [Friedland and Adam, 1989] have applied a radial search obtain the centre of gravity of an ellipse-shaped structure and a following four parameter Hough transform.

[Ballard, 1981] has developed a generalized Hough transform to detect any arbitrary curves parameterized by $\mathbf{a} = \{\mathbf{y}, \mathbf{s}, \phi\}$, with the reference origin $\mathbf{y} = (x_r, y_r)$, orthogonal scaling factors $\mathbf{s} = (s_x, s_y)$, and orientation ϕ. These parameters can also be applied on the previously presented analytic curves (table 2.1). $\mathbf{y}$ is described using an R-table, containing all *edge-orientation reference point correspondences*, which is constructed by computing the gradient direction $\phi(\mathbf{x})$ and $\mathbf{r} = \mathbf{y} - \mathbf{x}$ for each edge pixel $\mathbf{x}$. The accumulator array is then incremented for each $\mathbf{x}$ at the corresponding loci of $\mathbf{x} + \mathbf{r}$ in the accumulator array A where $\mathbf{r}$ is a table entry indexed by ϕ. To allow for scale changes, rotation and reference point translation $R(\phi)$ can be defined as a multiple vector-valued function.

In summary, the Hough transform is able to detect (and thus represent and describe) shapes of analytical and arbitrary form by investigating the global relationship between image pixels or potential edge (and thus shape contour) pixels. This may result in a very sparse accumulator space with only few votes per cell, which can be overcome by specifying a set of nearby points rather than one point only in the accumulator array. Computational efficiency may be improved by using *a priori* information about the shape's size, location and orientation, by decomposing the parameter space, and by using edge-directional information. Other valuable features for shape description include the detection of composite shapes and even incorporation of local information by adding the edge strength or the local curvature to the accumulator cells.

2.1.3 Medial Shape Representation

An alternative approach to boundary-based shape representation is given by using a shape's *medial* or *symmetry* axis, which combines local and global shape information. One medial approach has been developed by [Blum, 1967; Blum, 1973; Blum and Nagel, 1978]. The so-called *Blum's medial axis* classifies points on that axis by the properties of their maximal discs. Thus the symmetry axis is defined by the centre points of all maximal discs fitting exactly into the object and touching its boundary at least twice. So-called *normal points* are those whose underlying disc is tangential to the object boundary at exactly two distinct places, while the underlying discs of

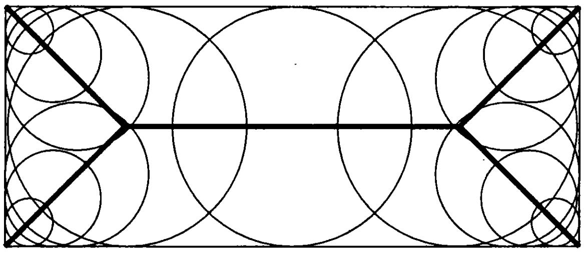


Figure 2.9: Blum's medial axis, defined as the centre points of all maximum discs within an object.

branch points touch the object boundary in three or more separate sets of points. An *end point* is one whose underlying disc only once touches the object border in one contiguous set. An example is shown in figure 2.9. The maximal discs allow to establish a relationship between boundary points across the width of an object. Taking the radius of the discs as a distance measure to the corresponding boundary points, local changes in the spatial position of the axis or centre points of the associated radius values can be used to analyse the local behaviour of the boundaries according to their orientation with respect to the axis, their curvature and other local shape descriptions.

Medial methods allow to capture local and global shape properties simultaneously, as objects are split into parts and subparts according to the relationship between several shape sections separated by the width of the object. While the information across the object provides global measurements, local specificity is still retained. Medial shape description based on the medical axis can be either based on its orientation ϕ, or on its length or *diameter*, respectively. The diameter is a measure determining the distance based on a metric $||\cdot||$ between the extreme points on the *major axis* or *medial axis* of the set B of all boundary points. It is defined as

$$\text{diameter}(B) = \max ||\mathbf{p_i} - \mathbf{p_j}|| \quad \forall(\mathbf{p_i},\mathbf{p_j}) \in B \tag{2.23}$$

The concept of medial shape representation has been extended by [Sheehy *et al.*, 1996] to medial surface construction, and [Sherbrooke *et al.*, 1996] have developed a method for a medial axis transform of three-dimensional polyhedral solids. Additionally, the concept of medialness has been used for shape *thinning* or *skeletonization* [Brandt and Algazi, 1992; Wright *et al.*, 1994], with extensions to morphological shape analysis [Reinhardt and Higgins, 1996], and scale-space [Ogniewicz, 1994]. A new form of a multi-scale, *Hough-like medial axis transform*, resulting in the so-called *core* of a shape, has been developed by [Burbeck and Pizer, 1994], with extensions by [Eberly, 1994b; Morse, 1994] and others. The concept of cores will be presented in chapter 3.

2.2 Local Shape Description

Local shape descriptors operate on a boundary shape representation. They are used when the primary focus is on locating external shape characteristics. In the following, some frequently used local shape quantifiers will be presented which are based on differential geometry of the shape.

The *curvature* κ of a planar curve at point P is defined as the instantaneous rate of change of the slope angle θ of the tangent at this point. The slope angle of the tangent of an explicit curve is given by

$$\theta(s) = tan^{-1}\frac{\mathrm{d}y}{\mathrm{d}x} \tag{2.24}$$

The curvature of any point on the curve is then given by the rate of change of $\theta(s)$:

$$\kappa(s) = \frac{\mathrm{d}\theta(s)}{\mathrm{d}s} \tag{2.25}$$

Note that this is a rotation invariant measure, as the same value for $\theta(s)$ is obtained when the x-axis is substituted by any line. However, it is not scale invariant. For example, the curvature of a circle of radius r is defined by $1/r$ at each point. This measure is halfed if the radius is doubled. κ can defined with respect to a parametric curve representation, by defining first

$$\kappa = \frac{y''}{(1+y'^2)^{3/2}} \quad \text{with} \quad y' = \frac{\mathrm{d}y}{\mathrm{d}x} \quad \text{and} \quad y'' = \frac{\mathrm{d}^2y}{\mathrm{d}x^2} \tag{2.26}$$

By representing the first and second derivatives of $x(s)$ and $y(s)$ as

$$\dot{x}(s) = \frac{\mathrm{d}x}{\mathrm{d}s} \qquad \ddot{x}(s) = \frac{\mathrm{d}^2x}{\mathrm{d}s^2} \qquad \dot{y}(s) = \frac{\mathrm{d}y}{\mathrm{d}s} \qquad \ddot{y}(s) = \frac{\mathrm{d}^2y}{\mathrm{d}s^2} \quad , \tag{2.27}$$

y' and y'' can be expressed as

$$y' = \frac{\dot{y}(s)}{\dot{x}(s)} \qquad \text{and} \qquad y'' = \frac{\dot{x}(s)\ddot{y}(s) - \dot{y}(s)\ddot{x}(s)}{\dot{x}(s)^3} \quad . \tag{2.28}$$

Thus, the curvature of the parametric curve is given by:

$$\kappa(s) = \frac{\dot{x}(s)\ddot{y}(s) - \dot{y}(s)\ddot{x}(s)}{(\dot{x}(s)^2 + \dot{y}(s)^2)^{3/2}} \tag{2.29}$$

The B-spline representation presented in section 2.1.1.3.1 allows for this analytic curvature computation. The curvature is frequently used for characterizing a shape by its points of inflection (defined as the zero-crossings of the curvature along the shape contour). When representing the curvature at all shape boundary pixels in a histogram, information about the bending behaviour can be won. In chapter 3, section 3.4.1, scale-space extensions of this scheme by [Asada and Brady, 1986; Mokhtarian and Mackworth, 1986] and others will be presented.

Corners are another important local shape characteristic, as they contain significant shape information and interpretation. They can be detected in various ways - a common approach is to detect local maxima of the absolute curvature $\|\kappa(s)\|$, which are negative and positive local maxima of the curvature. Corners can therefore be computed analytically: For $\kappa(s) > 0$, find points at s_i such that

$$\left.\frac{\partial\kappa(s)}{\partial s}\right|_{s=s_i} = 0 \quad \text{subject to} \quad \left.\frac{\partial^2\kappa(s)}{\partial s^2}\right|_{s=s_i} < 0\,, \tag{2.30}$$

and for $\kappa(s) < 0$, solve for the solutions t_i

$$\left.\frac{\partial\kappa(s)}{\partial s}\right|_{s=s_i} = 0 \quad \text{subject to} \quad \left.\frac{\partial^2\kappa(s)}{\partial s^2}\right|_{s=s_i} > 0\,, \tag{2.31}$$

which yields $\{s_i\}$ as the set of maxima of absolute curvature. One can now distinguish between concave upward corners (for positive curvature) and concave downward corners (for negative curvature). A scale-space extension of this scheme has been developed by [Rattarangsi and Chin, 1992]. Due to the higher order derivatives needed for the corner computation, however, noise along the shape boundary may be enhanced and deteriorate the quality and accuracy of corner detection. Another approach is to define corners as first order derivative discontinuities of either its $x(s)$ or $y(s)$ coordinate functions, reducing the $2D$ corner detection problem into two $1D$ detection problems which can be solved using statistical properties or numerical approximations [Chen and Chin, 1993].

Other local shape features can be directly derived from local boundary representations using various combinations of differential measurements. For example, normal and tangent deviations along the shape, the change of curvature, or curve discontinuities other than for corner detection are useful local shape descriptors.

2.3 Global Shape Description

Generally, global shape descriptors are selected when the primary focus is on the overall shape (or shape contour) rather than local shape properties. Among these, simple shape factors and quantitative measurements such as volumetric and bending measurements as well as Fourier descriptors, moments, and the Hough transform are probably among the most frequently used global methods.

2.3.1 Perimeter, Area and Compactness

The most simple descriptor is the *length* of a boundary, which can be approximated in a discrete representation by adding the pixels along the contour, or by adding the pixels on the vertical and horizontal parts of the contour and adding $\sqrt{2}$ times the number of the diagonal components, which gives the exact length in 8-connectivity (see section 2.1.1.1). For coarsely sampled boundaries, the Euclidean distances between the given boundary points may be added up instead, or, preferably, a spline interpolation may be computed to achieve dense sampling in order to be able to add up the boundary pixels in 8-connectivity. The length of a closed shape boundary is called *perimeter*.

The *area* A of a region can be determined by counting all pixels contained within its boundary. Alternatively, the polygonal area can be approximated by

$$A = \sum_{i=1}^{N} (x_i - x_{(i+1)\,\%\,N})(y_i + y_{(i+1)\,\%\,N}) \,, \tag{2.32}$$

where the boundary consists of N points ranging from $(x_1, y_1) \cdots (x_N, y_N)$, and the modulo operator $\%$ ensures cyclic summation of the closed boundary. Similar sampling treatment as for the perimeter applies here regarding sparse contour points.

While the area and the perimeter are both invariant with respect to translation and rotation, they are not invariant to scaling. A combination of both descriptors, however, can be used to obtain a dimensionless measurement which is invariant to translation, rotation and scaling: *compactness*, which is defined as $\frac{perimeter^2}{area}$, is a measure for the *roundedness* of a shape and is minimal for a circular shape. Frequently compactness is defined by $\frac{perimeter^2}{4\pi\ area}$, which yields a compactness of 1 for circles.

2.3.2 Boundary Straightness and Bending Energy

For discrete contours (e.g. a polygonal representation like in section 2.1.1.2), a *boundary scalar descriptor*, also called *boundary straightness*, is defined as the ratio between the boundary length and the number of boundary pixels where the boundary direction changes significantly. For few changes in boundary direction, the boundary scalar descriptor will be very high. In order to evaluate the change of direction of the boundary, angles between line segments positioned b boundary pixels in both directions from the evaluated boundary pixels are measured. The choice of parameter b indicates the degree of sensitivity to local changes of the boundary direction.

The *bending energy* of a shape contour is defined as the energy necessary to bend a rod to a desired shape, and is computed by integrating the squared contour curvature $\kappa(s)$ over the shape contour having border length L:

$$\mathcal{E}_{bending} = \frac{1}{L} \int_{s=0}^{1} \kappa^2(s)\, \mathrm{d}s \,, \tag{2.33}$$

κ being defined as in equation (2.29). [Duncan *et al.*, 1991] have presented a model based on the global bending energy in order to describe cardiac shape deformities by measuring bending energy as a difference in curvature. The thin-plate spline model presented in section 2.1.1.3.2 also allows to compute the deformation or bending energy necessary to transform or *warp* one shape into another, thus providing a measurement for differences in shape. In contrast to the bending energy defined by equation (2.33), the thin-plate spline bending energy takes the $2D$ geometry of the shape into account, and is invariant to affine transformations such as scaling.

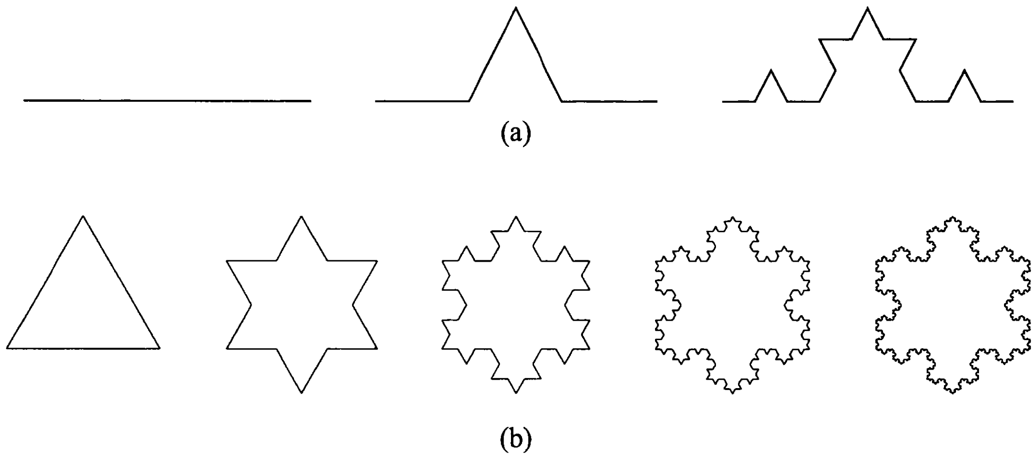


Figure 2.10: Von Koch curve (*snowflake*): (a) Generation scheme. (b) Increasing fractal resolution levels (generations).

2.3.3 Fractal Descriptors

The fractal theory introduced by [Mandelbrot, 1982] provides a mathematical description of scale-invariant structures and the relationship between measured quantity (e.g. perimeter, area or volume) and the scale at which that quantity is measured. This relationship can be described by a scalar called *fractal dimension.* A fractal structure exhibits self-similarity over all scales, thus copies of itself can be found at any scale. Fractals can be characterized by their fractal dimension which can be defined by composing the fractal structure into N distinct subsets, each being scaled down by ratio ϵ and identical to the overall structure in all statistical aspects. The fractal dimension $D_{fractal}$ is then given by the relation

$$N(\epsilon) = \frac{1}{\epsilon^{D_{fractal}}} \quad . \tag{2.34}$$

The fractal dimension of a true fractal structure is always higher than its Euclidean dimension D, and lower than its embedding dimension $(D+1)$. For example, a fractal curve, embedded in a $2D$ Euclidean space, has a fractal dimension between 1 and 2, with a circle having a fractal dimension of $D_{fractal} = 1$. The more a true fractal structure fills the space in which it is embedded, the more its fractal dimension approaches the Euclidean dimension of the embedding space.

An example of a true fractal structure is the von Koch curve or *snowflake.* It is generated by connecting a triangular *island* as an *initiator* of side length $l_{i+1} = l_i/3$, l_i being the side length of the previous step, to each side. This fractal has therefore a fractal dimension of $\log 4/\log 3 \approx 1.26$, as the length of the structure increases by factor $\frac{4}{3}$ at each generation step. The generation scheme is illustrated in figure 2.10(a), and figure 2.10(b) shows different fractal resolution levels or generations of a von Koch curve. However, natural sets behave only in a *statistically* self-similar fashion within a certain range of scales. Hence several estimation techniques have been developed in literature to allow for recovering the fractal dimension of natural sets:

Box counting method: The box counting method [Keller *et al.*, 1989] maps a given fractal structure onto a rectangular grid where the edges of each box in the grid are of length ϵ. Now the number of grid boxes $N(\epsilon)$ containing any parts of the fractal structure are counted. Performing this measurement for a whole set of box sizes ϵ, the fractal dimension can be estimated from a linear regression defined by $\log(N(\epsilon)) = -D_{fractal} \cdot \log(\epsilon) + c$, c being a constant, derived from equation (2.34).

Closely related to the box counting method is another fractal shape descriptor called *fractal signature* $S(\epsilon)$ [Peleg *et al.*, 1994], which is derived by plotting the enclosed shape area A against the scale ϵ at which it is measured, yielding a curve consisting of points $\mathbf{p}_i = (\epsilon_i, A(\epsilon_i))$. Fitting a straight line through each triple of points $\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1}$ yields a slope equal to $S(\epsilon_i) = H$ for true fractals, where H is the *Hurst* coefficient which is related to the fractal dimension by

$$D_{fractal} = (D+1) - H \;, \tag{2.35}$$

Hence the magnitude of the fractal signature indicates the amount of detail which is lost when ϵ is increased.

Power-spectral dimension: A fractal structure f can be characterized by its spectral density or power spectrum, $\mathcal{P}(f)$. The following relationship between the power spectrum and the fractal dimension holds [Pentland, 1984]:

$$\mathcal{P}(f) \propto f^{2H+1} \;, \tag{2.36}$$

Estimating the slope of the linear regression curve over the logarithm of the power spectrum as a function of f can therefore be used to estimate the fractal dimension using equation (2.35).

Fractional Brownian motion (fBm) model: The FBM model regards naturally occurring structures as the result of random walks. Assuming a D-dimensional structure f is generated by a D-dimensional fBm. The increments of a Brownian motion structure f must satisfy the following proportionality [Mandelbrot, 1982; Voss, 1985]:

$$E|\Delta f(\Delta \mathbf{x})| \propto ||\Delta \mathbf{x}||^H \quad , \tag{2.37}$$

The expected value of E of $|\Delta f(\Delta \mathbf{x})|$ can be computed by averaging and normalizing all differences of all vector pairs of corresponding distances or *fractal scales*, $\Delta \mathbf{x}$. Performing a linear regression over

$$\log\left(E|\Delta f(\Delta \mathbf{x})|\right) = H \cdot \log\left(||\Delta \mathbf{x}||\right) + c \tag{2.38}$$

yields an estimate of the fractal dimension using equation (2.35).

FIF, IFS, and FIM: Fractal interpolation functions (FIFs) [Barnsley, 1988] are self-similar functions defined on a compact interval which have more generalized spectra than the simple exponential decay found in fBm, and have also a more local scaling behaviour. Their graphs are fractal curves which allow to model natural structures and to distinguish between similar ones. [Berkner, 1997a; Berkner, 1997b] has performed wavelet-based reconstructions of FIFs for solving the inverse problem of finding maps of the corresponding iterated function systems (IFS) or fractal curves. Fractal interpolation with midpoints (FIM) interpolates between every other point of a structure, setting free parameters to ensure correct values for the FIF.

An evaluation of the different fractal dimension estimation techniques for medical images is given in [Penn and Loew, 1996]. The fractal dimension has been successfully applied in medical imaging for edge detection [Chen *et al.*, 1989; Schnabel, 1993b], segmentation [Schnabel, 1993a; Schnabel, 1994; Toennies and Schnabel, 1994], texture classification [Pentland, 1984; Chen *et al.*, 1989; Schnabel *et al.*, 1995], and shape analysis [Fortin *et al.*, 1992; Goldberger, 1992; Sakar and Chaudhuri, 1992; Samarabandu *et al.*, 1993; Brammer and Bullmore, 1994; Bullmore *et al.*, 1994; Free *et al.*, 1997].

2.4 Relative Distance Measurements

When a shape is to be evaluated not in isolation, but in comparison with another shape, both shapes can be described with the techniques presented earlier in this chapter, and the quantitative measurements obtained for each shape can be directly compared with each other. However, if local or global deviation measurements between different shapes are required (e.g. to determine the segmentation accuracy between a *gold standard* shape and a shape obtained from a segmentation technique), a *distance* measurement needs to be performed on either locally on a point-to-point basis, or globally on the whole set of points belonging to the shapes to be compared. A frequent method to compute the distance between two shapes is to use a *distance transformation* which converts a binary digital image, consisting of feature and non-feature pixels, into an image where all non-feature pixels have a value corresponding to the distance to the nearest feature pixel. Features can be points, lines, edges, or, in this case, pixels belonging to a shape under investigation. Other distance measurements operate directly on the sets of feature points. In any case, a distance measurement dist should satisfy the following conditions, given three points $\mathbf{p}_1, \mathbf{p}_2$, and $\mathbf{p}_3$:

1. Positivity: $\text{dist}(\mathbf{p}_1, \mathbf{p}_2) \geq 0$
2. Identity: $\text{dist}(\mathbf{p}_1, \mathbf{p}_2) = 0$ if and only if $\mathbf{p}_1 = \mathbf{p}_2$
3. Commutativity: $\text{dist}(\mathbf{p}_1, \mathbf{p}_2) = \text{dist}(\mathbf{p}_2, \mathbf{p}_1)$
4. Triangle inequality: $\text{dist}(\mathbf{p}_1, \mathbf{p}_3) \leq \text{dist}(\mathbf{p}_1, \mathbf{p}_2) + \text{dist}(\mathbf{p}_2, \mathbf{p}_3)$

+c5	+c4	+c3	+c4	+c5
+c4	+c2	+c1	+c2	+c4
+c3	+c1	0	+c1	+c3
+c4	+c2	+c1	+c2	+c4
+c5	+c4	+c3	+c4	+c5

(a)

+c5	+c4	+c3	+c4	+c5
+c4	+c2	+c1	+c2	+c4
+c3	+c1	0		

		0	+c1	+c3
+c3	+c2	+c1	+c2	+c4
+c5	+c4	+c3	+c4	+c5

(b)

Figure 2.11: General 5×5 neighbourhood mask for distance transformations. (a) Parallel mask. (b) Sequential forward and backward masks.

If conditions 1-3 are met, the measurement dist is a *semi-metric.* If condition 4 is also fulfilled, dist is a *metric.* In the following, the most commonly used distance measurements will be presented.

2.4.1 Local Distances

The most common local distance measurement is based on the *Euclidean* metric, which defines the distance between two points $\mathbf{p}_1 = (x_1, y_2), \mathbf{p}_2 = (x_2, y_2)$ as

$$\text{dist}(\mathbf{p}_1, \mathbf{p}_2) = \|\mathbf{p}_1 - \mathbf{p}_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{2.39}$$

Computing a *Euclidean* distance transform, however, implies a global and therefore computationally expensive search [Danielsson, 1980]. Approximations of the Euclidean distance transform take only local neighbourhoods into account, and can be efficiently implemented by propagating local distances (obtained in a local neighbourhood) either in a sequential or, preferably, in a parallel manner using a *mask* (see figure 2.11). The neighbourhood size can vary, and the mask coefficients correspond to the local distances which are propagated over the image.

In order to compute such a transform, a suitable mask is placed over each image pixel, and the mask value is added to the underlying image pixel value. The minimum value within that neighbourhood is assigned as the new image value for the centre pixel. The parallel algorithm is thus:

$$v_{i,j}^{m} = \min(v_{i+k,j+l}^{m-1} + c(k,l)) \quad \forall (k,l) \in mask \tag{2.40}$$

where $v_{i,j}^{m}$ is the value of the pixel in position (i, j) in the image at iteration m, (k, l) is the position in the mask (the centre being $(0, 0)$), and $c(k, l)$ is the local distance from the mask. The algorithm continues until the image values stop changing. In the sequential case, the mask is split into two masks, which are each passed once over the image: the forward mask from left to right, and from top to bottom, and the backward mask from right to left, and from bottom to top, as illustrated in algorithm 2.1.

// Forward:

for $(i = (size + 1)/2, \cdots, lines)$ **do**

  for $(j = (size + 1)/2, \cdots, columns)$ **do**

    $v_{i,j} = \min(v_{i+k,j+l} + c(k,l)) \quad \forall (k,l) \in mask_{forward}$

  end for

end for

// Backward:

for $(i = lines - (size - 1)/2, \cdots, 1)$ **do**

  for $(j = columns - (size - 1)/2, \cdots, 1)$ **do**

    $v_{i,j} = \min(v_{i+k,j+l} + c(k,l)) \quad \forall (k,l) \in mask_{backward}$

  end for

end for

Algorithm 2.1: Algorithm for sequential distance transformation for a $size \times size$ mask and notation as in equation (2.40).

Distance	Description	Mask values
City-block	Smallest distance in 4-neighbourhood.	$c_1 = 1, c_2 = \infty$
Chess-board	Smallest distance in 8-neighbourhood.	$c_1 = 1, c_2 = 1$
Octagonal	Alternating city-block and chess-board.	$c_1 = 1, c_2 = 1$ and $c_2 = \infty$
Chamfer	Optimal Euclidean distance transform.	See figure 2.12.

Table 2.2: Overview of Euclidean approximating distance transforms.

The most common Euclidean approximating distance transforms and their mask values are listed in table 2.2. Probably the most important distance transform is the *Chamfer* distance transform, which is the best approximation of the Euclidean distance, having a maximal difference from the Euclidean distance of less than 2%. [Borgefors, 1986] has developed a 5–7–11 transform, which is shown in figure 2.12. Examples for the Chamfer distance transform for a square contour and more complex contours are illustrated in figure 2.13. Note that large distances (corresponding to bright intensity values) can be used to detect the medial axis and the skeleton of a shape [Ge and Fitzpatrick, 1996]. An extensive survey of approximations of the Euclidean distance transform can be found in [Borgefors, 1986], with extensions to 3D in [Borgefors, 1996]

2.4.2 Global Distances

In order to compute the global distance between two shapes, any of the above local distance transforms of the first shape can be computed after setting all image pixels corresponding to the shape contour (feature pixels) to zero and the remaining pixels to infinity (or any suitably high value).

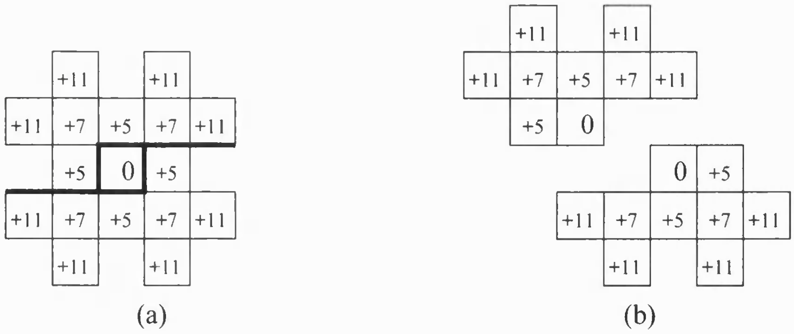


Figure 2.12: 5-7-11 masks for the Chamfer distance transform. (a) Parallel mask. (b) Sequential forward and backward masks.

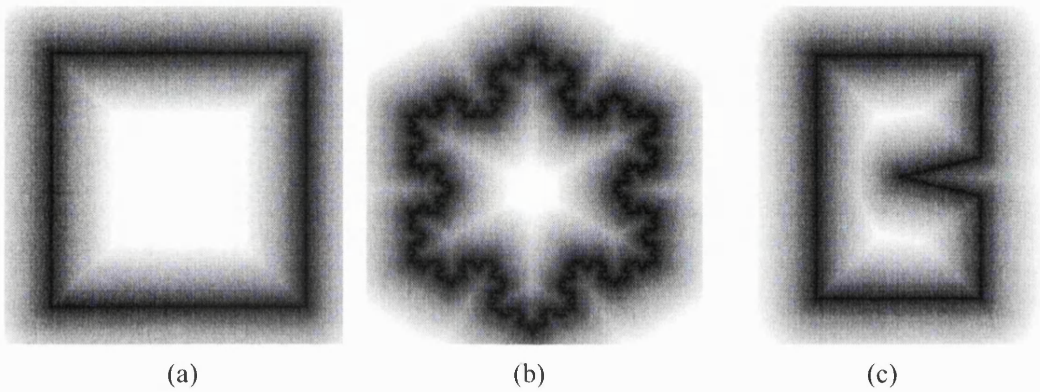


Figure 2.13: Examples for the Chamfer distance transform: (a) Square shape. (b) Von Koch curve. (c) Notched rectangle. Darker intensity values correspond to closer distances to the shape, while brighter values indicate larger distances.

Overlaying the shape to be compared onto this distance transformed image allows to obtain the local distance values underneath it. By computing the absolute error (L_1 norm), the root mean squared error (L_2 norm) or the maximum error (L_∞ norm), a global shape metric is obtained to estimate the overall deviation between the two shapes.

Alternatively, the *Hausdorff* distance can be computed as presented by [Huttenlocher *et al.*, 1993]. It measures the extent to which each point of one shape lies near some point of another shape and vice versa, which is an indication of the *degree of mismatch* between two shapes. Given two finite point sets describing the shape contours, $\mathbf{A} = \{\mathbf{a_1}, \cdots, \mathbf{a_m}\}$ and $\mathbf{B} = \{\mathbf{b_1}, \cdots, \mathbf{b_n}\}$, the Hausdorff distance is defined as

$$\text{dist}_H(\mathbf{A}, \mathbf{B}) = \max(\text{dist}_h(\mathbf{A}, \mathbf{B}), \text{dist}_h(\mathbf{B}, \mathbf{A})) \tag{2.41}$$

where the directed Hausdorff distances $\text{dist}_h(\mathbf{A}, \mathbf{B})$ and $\text{dist}_h(\mathbf{B}, \mathbf{A})$ are computed via

$$\begin{aligned} \text{dist}_h(\mathbf{A}, \mathbf{B}) &= \max\min ||\mathbf{a}_i - \mathbf{b}_j|| \quad \forall \mathbf{a}_i \in \mathbf{A}, \forall \mathbf{b}_j \in \mathbf{B} \\ \text{dist}_h(\mathbf{B}, \mathbf{A}) &= \max\min ||\mathbf{b}_i - \mathbf{a}_j|| \quad \forall \mathbf{b}_i \in \mathbf{B}, \forall \mathbf{a}_j \in \mathbf{A} \end{aligned} \tag{2.42}$$

and $||\cdot||$ is some underlying norm, e.g. the L_2 or Euclidean norm. If $\text{dist}_h(\mathbf{A}, \mathbf{B}) = d$, then each point of A must be within distance d of some point of $\mathbf{B}$, and there is some point of $\mathbf{A}$ that is exactly of distance d from the nearest point of $\mathbf{B}$. Unlike most methods comparing shapes, there is no explicit pairing of points of $\mathbf{A}$ with points of $\mathbf{B}$, as many points of $\mathbf{A}$ may be close to the *same* point of $\mathbf{B}$. The Hausdorff distance is therefore not a local measurement. [Huttenlocher *et al.*, 1993] have developed a metric which measures the mismatch between *all possible* relative positions of two shapes, where the Hausdorff distance is defined as a function of relative position with respect to translation and rigid motion.

Another global distance measurement is the *Levensthein* distance which is motivated by the fact that a shape $\mathbf{A}$ can be mapped into a shape $\mathbf{B}$ by means of three possible operations, namely substitution, insertion, and deletion. If the costs of each operation were p, q, and r, respectively, then the *weighted Levensthein distance* (WLD) between two shapes is defined as

$$\text{dist}_{WLD}(\mathbf{A}, \mathbf{B}) = \min_i(m_i p + n_i q + k_i r) \quad , \tag{2.43}$$

where the mapping of $\mathbf{A}$ into $\mathbf{B}$ requires m_i substitutions, n_i insertions, and k_i deletions. As the WLD depends on the size of the compared shapes, [Cortelazzo *et al.*, 1996] have shown that when normalizing the WLD, the *normalized weighted Levensthein distance* (NWLD) defined by

$$\text{dist}_{NWLD}(\mathbf{A}, \mathbf{B}) = \frac{\text{dist}_{WLD}(A, B)}{\max(L_{\mathbf{A}}, L_{\mathbf{B}})} \tag{2.44}$$

is independent of the shape sizes $L_{\mathbf{A}}$ and $L_{\mathbf{B}}$, respectively. The NWLD does not generally satisfy the triangle inequality and is therefore not a metric, but is still a good shape diversity measure which has been applied to the comparison of strings in pattern recognition. Both the WLD and the NWLD operate directly on a *chain-code* representation, making them computationally less complex than the Euclidean distance measurements.

2.4.3 Corresponding Distances

A common problem when comparing two shapes is the pairing of the points which is necessary to obtain the *corresponding* distances rather than the nearest distances. This is also a very important topic in *registration* or matching of structures, as well as segmentation evaluation.

[Chalana and Kim, 1996; Chalana *et al.*, 1996] have developed several distance measurements for segmentation evaluation based on the *mean* distance between two curves. The *distance to the closest point* (DCP) for point $\mathbf{a}_i$ on curve $\mathbf{A}$ to curve $\mathbf{B}$ is defined by

$$\text{dist}_{DCP}(\mathbf{a}_i, \mathbf{B}) = \min_j ||\mathbf{b}_j - \mathbf{a}_i|| \ , \tag{2.45}$$

where $\mathbf{A}$ and $\mathbf{B}$ consist of N equidistant points and $||\cdot||$ is a suitable metric. Computing this distance for all points on both curves yields the *mean absolute distance* (MAD) between the two curves:

$$\text{dist}_{MAD}(\mathbf{A},\mathbf{B}) = \frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}\text{dist}_{DCP}(\mathbf{a}_i,\mathbf{B}) + \frac{1}{N}\sum_{i=1}^{N}\text{dist}_{DCP}(\mathbf{b}_i,\mathbf{A})\right) \tag{2.46}$$

Although the DCP and hence the MAD do not fulfill the metric condition for triangle inequality, they were found to be more useful in terms of measuring the deviation between two shapes than the Hausdorff distance.

[Besl and McKay, 1992] have developed the *iterated closest points* (ICP) algorithm, which establishes an initial average curve $\mathbf{C}$ by randomly selecting a point of $\mathbf{A}$, denoted by $\mathbf{a}_1$, finding the closest point on $\mathbf{B}$, denoted by $\mathbf{b}_1$, and sequentially setting up correspondences for the remaining points. The points $\mathbf{c}_i$ on the average curve are then given by the centroid of the corresponding points $(\mathbf{a}_i + \mathbf{b}_i)/2$, and their normal intersections with both curves $\mathbf{A}$ and $\mathbf{B}$ establish a new set of point correspondences for $\mathbf{A}$ and $\mathbf{B}$, leading to a new average curve. This whole process is iterated until the average curve does not change any more. The final set of correspondences can be used to find the local and overall shape deviation between $\mathbf{A}$ and $\mathbf{B}$ using any suitable metric.

A major drawback of the corresponding distances presented so far is that the shapes to be compared must be of the same number of equidistant points. This condition may prove to be too hard for many applications in shape description, where shapes may vary in size and local point distance. A very related technique to the ICP algorithm which is independent of the lengths and point distances of the shapes is the concept of *triangulation*, used in computer graphics for rendering purposes by connecting points into shortest vertex triangular patches [Christiansen and Sederberg, 1978]. After establishing an initial single-point correspondence similar to the ICP algorithm, the remaining correspondences are found by finding the shortest connecting triangles:

$$\text{dist}_T(\mathbf{a}_i,\mathbf{b}_j) = \min(\|\mathbf{a}_i - \mathbf{b}_{j+1}\|, \|\mathbf{a}_{i+1} - \mathbf{b}_j\|) \tag{2.47}$$

This distance measurement is not only more general than the iterated conditional points algorithm, but also faster as it needs only one bidirectional pass over the contours. Similar to ICP, it allows to measure local distances via the corresponding points, and to find a global, mean measurement for shape deviation when averaging the found local distances.

The ICP and triangulation distances solve the problem of finding point correspondences *empirically*. A more generic way would be to take the structure of the shape into account in terms of either biological or pseudo-landmarks, or in terms of geometrically prominent features such as corners, points of high curvature etc. In the first case, the concept of thin-plate splines (section 2.1.1.3.2) allows to measure the local deformation between two shapes by mapping each point of one shape into its homologous counterpart in the second shape, thus yielding an interpolated local distance measurement. However, *a priori* knowledge of the shapes under investigation or appropriate expert interaction is necessary in order to define a set of suitable landmarks for each shape (which is only a small subset of the shape). In the second case, shape features can be derived by extracting all points of sufficiently high curvature, for example, and finding the correspondences between these features only, followed by a suitable pairing of the remaining non-feature points.

2.5 Summary

This chapter has presented a survey on techniques for shape representation, description and distance-based comparison. Shape representation has been divided into three categories: local, global, and medial representation.

Local shape representation techniques, though not being able to provide access to overall shape properties like end-to-end length, width, or orientation, allow to compute a large number of local, global, and relative shape characteristics. The polygonal approximation and the spline representation are probably the most common ones, as they provide a compact object representation, allow for complete shape reconstruction, and are able to approximate shapes with arbitrary accuracy via their explicit polynomial representation. Additionally, splines are smooth and continuous contour interpolants and can be analytically differentiated.

Global shape representation techniques like Fourier descriptors and moments allow for the truncation of details and allow to completely reconstruct a shape (but only if all descriptors or moments, respectively, are used). The truncation leads to a hierarchical representation through the extra degree of freedom, but leads to a global rather than local truncation of detail.

Medial representation techniques combine the properties of local and global techniques. They offer a hierarchical, *structural* way of representing a shape in terms of its subparts determined via the object width, and provide global (or rather *multi-local*) shape information while retaining spatial specificity. However, the extraction of a medial axis is a non-trivial task which outweighs its advantages.

Shape descriptors can also be categorized into local, global, and medial techniques. Additionally, relative measurements may be used.

Local shape descriptors are able to extract local spatial information such as the boundary curvature or corners, and provide therefore important information for local shape interpretation. For example, the curvature can be used to detect and analyse local structural shape characteristics.

Global shape descriptors capture shape properties that carry over the entirety of the shape. Important global descriptors in medical image analysis are 2D measurements like area and perimeter, with volume and surface area as their respective counterparts in 3D. Compactness and the fractal dimension allow to obtain a quantitative measurement for the roundness and ruggedness of a shape, and are therefore highly suitable as measures of shape complexity. Bending energy and boundary straightness characterize the global structural deformation and elongation of an object.

Medial shape descriptors, which can be only extracted from a medial representation, include global shape properties such as shape orientation and elongation, as well as local measurements like curvature or corners. Additionally, they give spatially-specific non-local measurements like shape width, changes of width, and symmetric behaviour.

Relative (distance) measurements allow for inter- rather than intra shape description by finding the dissimilarity between different shapes. For example, the spatial deviation of two shapes in terms of their distance can be obtained *locally* on a point-to-point basis (using a Euclidean metric), or *globally*, e.g. by measuring the mean or maximum deviation, or the worst mismatch via the Hausdorff distance. Additionally, *medial* distance measurements can be obtained by computing the distance of corresponding opposite boundary points. However, in order to perform appropriate distance or other dissimilarity measurements, a point correspondence needs to be set up which is a non-trivial problem (except for medial shape representation). An appropriate *corresponding* distance measurement is the distance obtained by triangulating the points of two shapes.

The following chapter will review the concept of multi-scale image processing, introducing an extra scale degree of freedom, including geometric image descriptors, and applications to multi-scale contour and image processing and analysis techniques.

Chapter 3

Survey of Multi-Scale Image Processing

– TU REGARDERAS, LA NUIT, LES ÉTOILES. C'EST TROP PETIT CHEZ MOI POUR QUE JE TE MONTRE OÙ SE TROUVE LA MIENNE. C'EST MIEUX COMME ÇA. MON ÉTOILE, ÇA SERA POUR TOI UNE DES ÉTOILES. ALORS, TOUTES LES ÉTOILES, TU AIMERAS LES REGARDER... ELLES SERONT TOUTES TES AMIES.

"AND AT NIGHT YOU WILL LOOK UP AT THE STARS. WHERE I LIVE EVERYTHING IS SO SMALL THAT I CANNOT SHOW YOU WHERE MY STAR IS TO BE FOUND. IT IS BETTER, LIKE THAT. MY STAR WILL JUST BE ONE OF THE STARS, FOR YOU. AND SO YOU WILL LOVE TO WATCH ALL THE STARS IN THE HEAVENS... THEY WILL ALL BE YOUR FRIENDS."

Le Petit Prince, Antoine de Saint-Exupéry.

Multi-scale image processing and analysis was first introduced by Marr [Marr, 1982] as a concept in human vision. Scale-space, first represented by Witkin [Witkin, 1983], represents a signal or an image by a family of signals or images on various levels of inner spatial scale by sampling with a neighbourhood operator or *kernel* which also allows to make differential measurements. Hence local averaging allows to obtain derivatives on the discrete image grid where no infinitesimal limits exist. [Koenderink, 1984; Koenderink and van Doorn, 1990] have shown that the sampling function to obtain the measurements should be the Gaussian kernel as the lowest order, rescaling operator, and its linear partial derivatives. [Koenderink, 1984] also stated the *scale-space causality principle*, implying that coarser (or higher) scales can only be caused by what what happened at finer (or lower) scales. Moreover, Gaussian convolution does not enhance minima and maxima for increasing scales (*maximum principle*), and in $1D$ no new extrema are created for increasing scales [Babaud *et al.*, 1986]. In [ter Haar Romeny, 1996; ter Haar Romeny, 1997], a tutorial-like introduction to scale-space theory is given, and the main developments and applications will be reviewed in this chapter.

3.1 Linear Diffusion

In the continuous case, the normalized Gaussian kernel in n-dimensional space is

$$G(\mathbf{x};\sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{-\|\mathbf{x}\|^2}{2\sigma^2}} \tag{3.1}$$

where σ is the scale or width of the scaling operator. Convolving an image luminance function $L(\mathbf{x}) = L(\mathbf{x}; 0)$ with a Gaussian kernel yields a smoothed version of the image,

$$L(\mathbf{x};\sigma) = G(\mathbf{x};\sigma) \otimes L(\mathbf{x}) \quad , \tag{3.2}$$

where $\otimes$ denotes a convolution over image space. As implementations of scale-space are complex and very intensive in both memory and computing time, normally a fixed discretization of scale-space is made and the blurring is carried out in the Fourier domain followed by an inverse Fourier transform. In the Fourier domain, the kernel becomes diagonal:

$$\mathcal{L}(\omega;\sigma) = \mathcal{L}(\omega) \cdot \mathcal{G}(\omega;\sigma) \tag{3.3}$$

The Gaussian kernel is the Green's function (or *propagator*) of the homogeneous *diffusion* or *heat* equation; thus convolution of the luminance function L with a Gaussian (or *blurring*) can be expressed as evolution under diffusion ([Koenderink and van Doorn, 1990]), where the scale is expressed by a time parameter:

$$\frac{\partial L}{\partial t} = \nabla \cdot c\nabla L \quad \text{with} \quad t = \frac{\sigma^2}{2c} \tag{3.4}$$

The constant conductance term c controls the rate of blurring with respect to time t. Again, the initial condition of the full solution $L(\mathbf{x}; t)$ for the continuous scale-space is given by $L(\mathbf{x}; 0) = L(\mathbf{x})$. Each time slice is a version of the original image after some amount of *linear* blurring. For discrete images, [Lindeberg, 1994] has shown that the necessary spatial blur kernels are modified Bessel functions (being fundamental solutions of a discrete version of equation (3.4)). For infinitesimal grid spacing, the solutions approach the continuous case and the modified Bessel functions approach the Gaussian kernel. Figure 3.1 shows samples of a linear scale-space for increasing diffusion times.

3.1.1 Multi-Scale Differential Invariants

Using a scale-space representation of an image, local image structure can be revealed with the help of fundamental mathematical operations like differentiation. The Gaussian kernel and its linear partial derivatives permit to construct multi-scale, orthogonal differential invariants with respect to changes of coordinate systems which allow for multi-scale shape measurements and descriptions of local image structure. [ter Haar Romeny *et al.*, 1991; ter Haar Romeny *et al.*, 1993; Florack, 1993] have presented a hierarchical set of natural scaled differential operators which are true geometric image descriptors that resemble the receptive field profiles in the human front-end visual system. They propose a local jet of order N (or N-jet) defined as a class of functions of the image intensity L sharing the same N-truncated Taylor expansion at a given point $\mathbf{x}$:

$$J^N(L(\mathbf{x})) \equiv \{L_{i_1 \cdots i_n}(\mathbf{x})\}_{n=0}^{N} \tag{3.5}$$

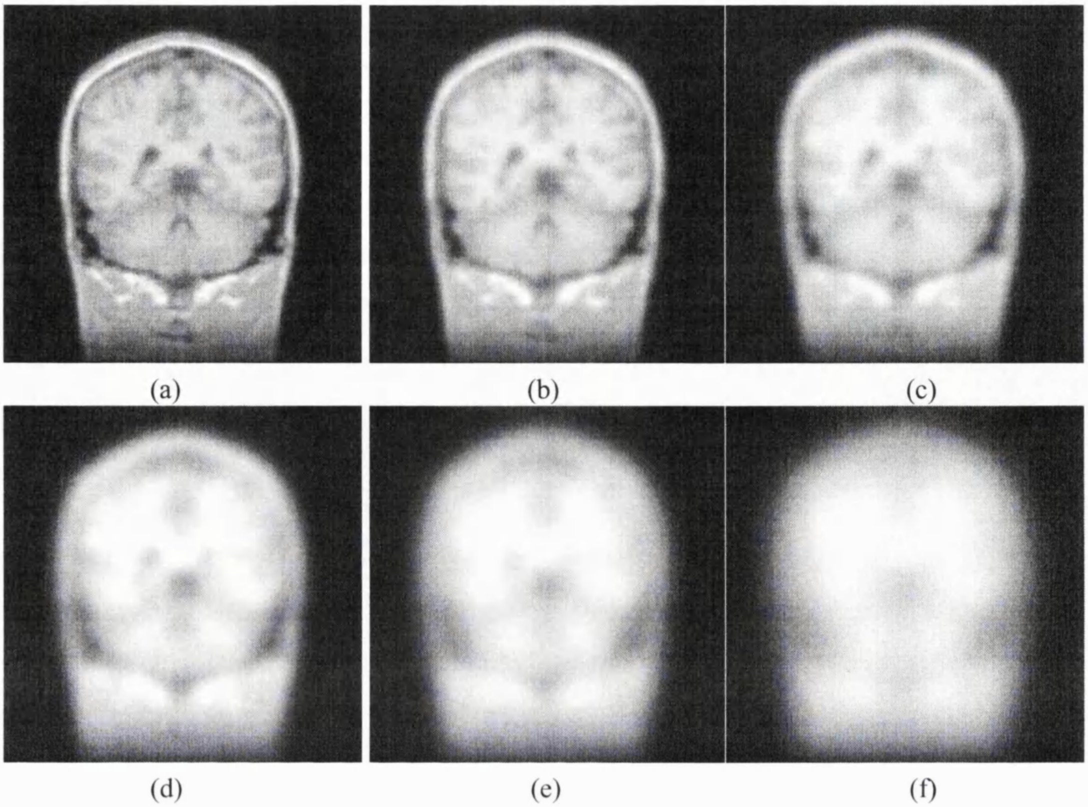


Figure 3.1: Linear scale-space samples. Upper row: (a) $t = 2$ (b) $t = 4$ (c) $t = 8$. Lower row: (d) $t = 16$ (e) $t = 32$ (f) $t = 64$. The diffusion was carried out with the explicit scheme (see section 3.3.1), with $\sigma = 0.8$ as the regularizing scale and $\Delta\tau = 0.25$ as the numerical time step. The evolution times t correspond to $\Delta\tau \cdot \tau$, with $\tau = 8, 16, 32, 64, 128, 256$ as the respective iteration steps.

where the lower subscripts of L range from $1 \cdots D$ and denote differentiation with respect to the associated spatial variables, D is the dimensionality of L. Restricting the N-jet to an orthonormal basis yields a set of independent or irreducible invariants which can express every image property. On the other hand, every image property being invariant to coordinate transformations can be given a geometric meaning. The complete hierarchy of higher order smoothed operators is thus given by

$$\{G_{i_1 \cdots i_n}(\mathbf{x}; \sigma) = \partial_{i_1 \cdots i_n} G(\mathbf{x}; \sigma)\}_{n=0}^{N} \tag{3.6}$$

In other words, any Cartesian partial derivative of order n of a rescaled image $L(\mathbf{x}; \sigma)$ is obtained by convolving the original image $L_0(\mathbf{x})$ with the corresponding partial derivative of the zero-th order Gaussian $G(\mathbf{x}; \sigma)$. Table 3.1 lists the complete set of irreducible invariants in $2D$, which are illustrated at three different scales in figure 3.2. Single derivatives like L_x, however, rely by definition on the choice of the coordinate system. Tensor calculus can be used in order to describe the transformation behaviour of tensor components like L_x. For example, rotating L_x by an angle α yields $L'_x = \cos\alpha L_x + \sin\alpha L_y$. Thus L_y must be added in order to obtain a two-component

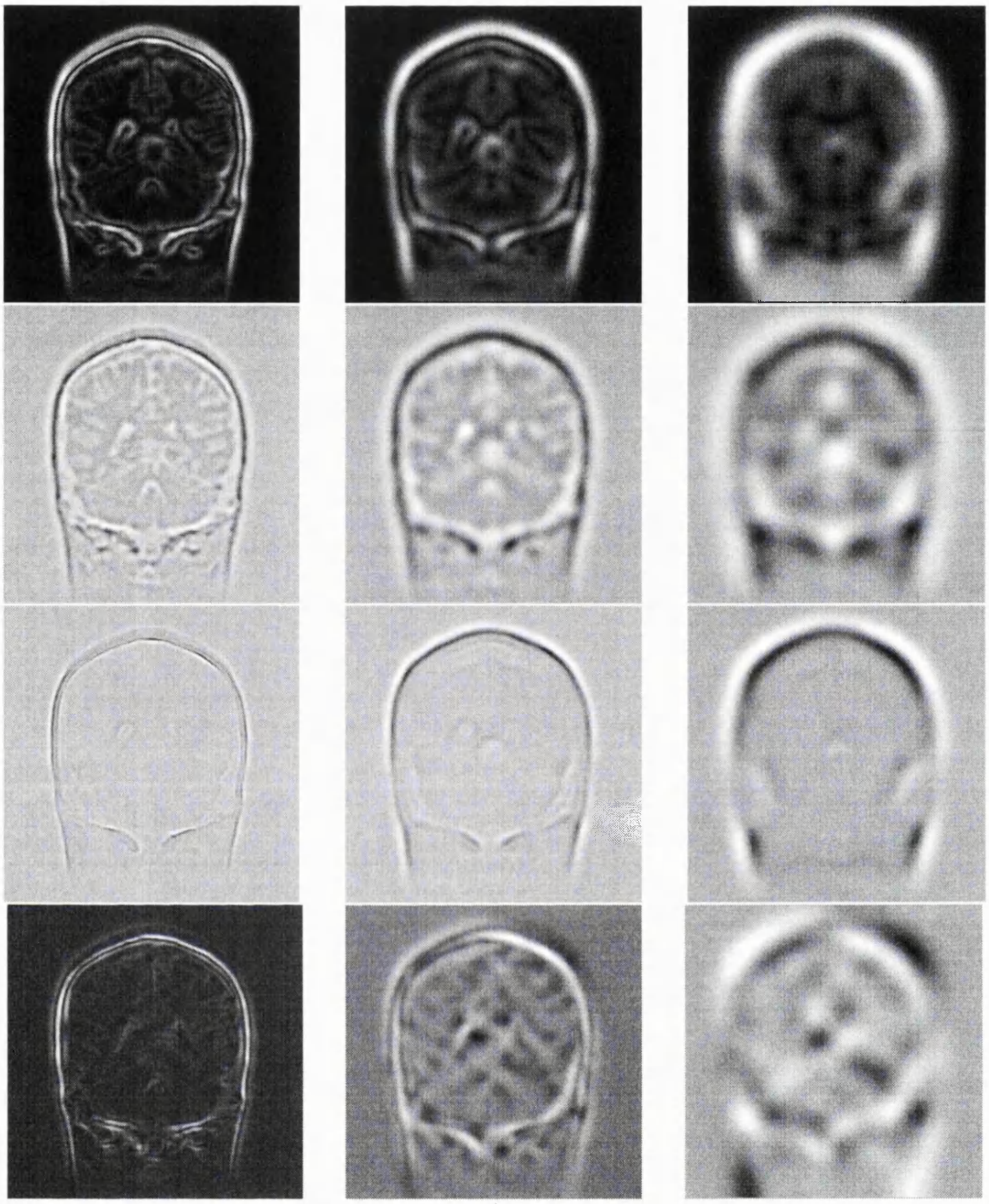


Figure 3.2: Scale-space samples for up to second order Euclidean irreducible set of differential invariants. From top to bottom: $L_i L_i$ (squared gradient), L_{ii} (Laplacian), $L_i L_{ij} L_j$ (no name), $L_{ij} L_{ji}$ (deviation from flatness), for (from left to right) evolution times $t = 4, 8, 64$, time step $\Delta\tau = 0.25$, $\sigma = 0.8$.

vector: (L_x, L_y). Important tensors are the *Hessian*, which is in $2D$ the set of all second order partial derivatives, and two constant tensors, one of which is the symmetric Kronecker tensor δ_{ij}, and the other is the Lévi-Civita tensor ϵ_{ij}:

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \epsilon_{i_1 \cdots i_D} = \begin{cases} 1 & \text{if} \quad (i_1 \cdots i_D) \quad \text{is even permutation of } 1 \cdots D \\ -1 & \text{if} \quad (i_1 \cdots i_D) \quad \text{is odd permutation of } 1 \cdots D \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

Name	Cartesian	Manifest	Gauge
Intensity	L	L	L
Gradient^2	$L_x^2 + L_y^2$	$L_i L_i$	L_w^2
Laplacian	$L_{xx} + L_{yy}$	L_{ii}	$L_{vv} + L_{ww}$
(no name)	$L_{xx}L_x^2 + 2L_{xy}L_xL_y + L_{yy}L_y^2$	$L_iL_{ij}L_j$	$L_{ww}L_w^2$
Deviation from flatness	$L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2$	$L_{ij}L_{ji}$	$L_{vv}^2 + L_{ww}^2$

Table 3.1: Set of irreducible, up to second order Euclidean image invariants in $2D$. The invariants are expressed in Cartesian, manifest invariant and gauge coordinates (Adapted from [ter Haar Romeny *et al.*, 1993]). The manifest invariant notation follows the Einstein summation convention, e.g. in $2D$: $L_iL_i \equiv \sum_{i=x,y} L_iL_i$ and $L_{ij} \equiv \sum_{i=x,y}\sum_{j=x,y} L_{ij}$.

Given a set of tensors, an invariant is obtained by fully contracting and alternating the indices in a tensor product using the constant tensors δ and ϵ.

Beside the Cartesian and manifest tensor representation, invariants can be represented by gauge coordinates singling out a particular, geometrically meaningful frame and using derivatives along the axes. For example, choosing separately for each point of the image one axis (called w-axis) along the gradient direction at that particular point, and the other axis (called v-axis) tangentially along the isophote (contour of constant image values) which is orthogonal to the gradient direction, and applying directional derivative operators to the image will yield manifest invariants. For example, the *isophote curvature* κ can be derived at point $P(v = 0, w = 0)$ of the curve (at the same time the origin of the (v, w)-system) by $\kappa = w''(0)$. The isophote passing through P is implicitly given by $L = L_P$ whose first- and second-order derivatives with respect to v are given by

$$L_v + L_w w' = 0 \quad \text{and} \quad L_{vv} + +2L_{vw}w' + L_{ww}w'^2 + L_w w'' = 0 \tag{3.8}$$

Since $L_v(0) = 0$, $w'(0) = 0$ as well, thus $\kappa_{isophote} = -\frac{L_{vv}}{L_w}$. Analogously, the flowline curvature $\kappa_{flowline}$ (the orthogonal trajectories of the isophotes) is given by $\kappa_{flowline} = -\frac{L_{vw}}{L_w}$. The isophote curvature in $3D$ is expressed by two independent invariants, the principal curvatures. In $3D$, where the gauge coordinates are a triple (u, v, w), the principal curvatures are given by $\kappa_1 = -\frac{L_{uu}}{L_w}$ and $\kappa_2 = -\frac{L_{vv}}{L_w}$, and the mean curvature κ_{mean} and Gaussian curvature $\kappa_{Gaussian}$ are given by their average $\kappa_{mean} = \frac{\kappa_1+\kappa_2}{2}$, and product $\kappa_{Gaussian} = \kappa_1 \cdot \kappa_2$, respectively. In manifest invariants, κ_{mean} and $\kappa_{Gaussian}$ are expressed as [Florack, 1993]:

$$\kappa_{mean} = \frac{1}{2}\frac{L_iL_{ij}L_j - L_iL_iL_{jj}}{(L_kL_k)^{\frac{3}{2}}} \quad \text{and} \quad \kappa_{Gaussian} = \frac{1}{2}\frac{\epsilon_{ijk}\epsilon_{lmn}L_iL_lL_{jm}L_{kn}}{(L_pL_p)^2}\,. \tag{3.9}$$

It is important to note that isophote properties are invariant under more general transformations than coordinate transformations, e.g intensity transformations such as gamma-corrections, brightness and contrast adjustments).

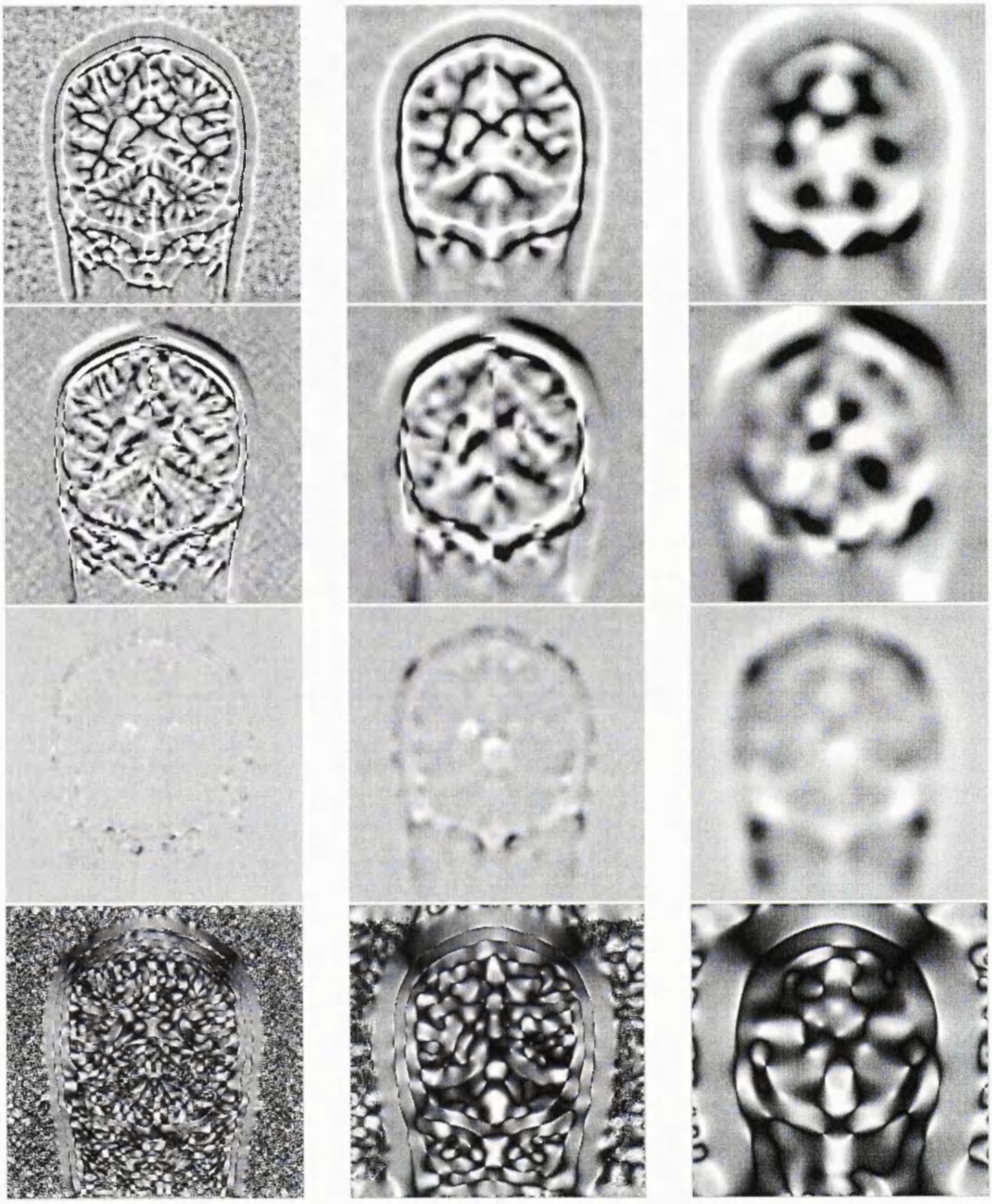


Figure 3.3: Scale-space samples for second order differential invariants. From top to bottom: $-\frac{L_{vv}}{L_w}$ (isophote curvature), $-\frac{L_{vw}}{L_w}$ (flowline curvature), $L_{vv}L_w^2$ (cornerness), $\frac{2L_{vv}L_{ww}}{L_{vv}^2+L_{ww}^2}$ (umbilicity), for (from left to right) evolution times $t = 4, 8, 64$, time step $\Delta\tau = 0.25$, $\sigma = 0.8$.

A list of some other $2D$ differential invariants (which can be obtained by combination of the irreducible invariants) of the up to second order set in $2D$ is given in table 3.2 and examples are illustrated in figure 3.3. The most frequently used invariants perhaps are those related to the notion of edges. The squared gradient magnitude, $\|\nabla L\|^2 = L_i L_i = L_w^2$, indicates the likeliness of a point presenting an edge (so-called "edgeness"), so edges can be found by looking for maxima of this invariant (which forms the basis of Canny edge-detector [Canny, 1987]). The zero-crossings of the Laplacian, ΔL, are another edge-detector (and form the basis of the Marr-Hildreth edge de-

Name	Cartesian	Manifest	Gauge
Isophote curvature	$\frac{2L_xL_yL_{xy}-L_x^2L_{yy}-L_y^2L_{xx}}{(L_x^2+L_y^2)^{\frac{3}{2}}}$	$\frac{L_iL_jL_{ij}-L_iL_iL_{jj}}{(L_kL_k)^{\frac{3}{2}}}$	$-\frac{L_{vv}}{L_w}$
Flow-line curvature	$\frac{L_xL_y(L_{yy}-L_{xx})+L_{xy}(L_x^2-L_y^2)}{(L_x^2+L_y^2)^{\frac{3}{2}}}$	$\frac{1}{2}\frac{L_iL_{ij}L_j-L_iL_iL_{jj}}{(L_kL_k)^{\frac{3}{2}}}$	$-\frac{L_{vw}}{L_w}$
Cornerness	$L_x^2L_{yy}-2L_xL_yL_{xy}+L_y^2L_{xx}$	$L_{ii}L_jL_j-L_{ij}L_iL_j$	$L_{vv}L_w^2$
Umbilicity	$\frac{-2(L_{xx}L_{yy}+L_{xy}^2)}{L_{xx}^2+2L_{xy}^2+L_{yy}^2}$	$\frac{\epsilon_{ij}\epsilon_{kl}L_{ik}L_{jl}}{L_{mn}L_{nm}}$	$\frac{2L_{vv}L_{ww}}{L_{vv}^2+L_{ww}^2}$

Table 3.2: Examples of $2D$ invariants obtained as combinations of irreducible invariants.

tector [Marr and Hildreth, 1980], though they describe edges only if the isophotes are sufficiently straight (thus corners cannot be reliably detected using this invariant). This can be clearly seen from the Laplacian's manifest and gauge representation:

$$\Delta L = L_{ii} = L_{vv} + L_{ww} = L_{ww} - \kappa L_w \tag{3.10}$$

The last equation follows by inserting the definition of the isophote curvature given above. Other important invariants are the second-order invariant called *cornerness* or corner strength $L_{vv}L_w^2$, and the third-order *bendedness* $L_{vvv}L_w^5 - 3L_{vv}L_{vw}L_w^4$. Combinations of invariants are invariants themselves. For example, [Lindeberg, 1993b] has proposed to multiply the curvature by the gradient magnitude raised to some power, m, a natural choice being $m = 3$, in order to give a stronger response near edges.

3.2 Non-linear Diffusion

Though linear smoothing greatly reduces the effect of random noise, it also "smoothes across edges", an effect which may be quite unsatisfactory for the detection of boundary locations. Hence, several *non-linear* scale-space approaches have been developed in order to preserve edges or other interesting image features.

3.2.1 Edge-Affected and Multi-Scale Diffusion

Non-linear, edge-preserving diffusion has first been developed by [Perona and Malik, 1990]. They have proposed to use a variation on the heat equation which allows the conductance parameter to vary over space and time:

$$\frac{\partial L}{\partial t} = \nabla \cdot c(\mathbf{x}, t)\nabla L \tag{3.11}$$

It is also referred to as variable heat conductance or edge-affected diffusion. Writing the conductance as a function of the image gradient yields

$$\frac{\partial L}{\partial t} = \nabla \cdot g(\|\nabla L\|)\nabla L \,. \tag{3.12}$$

The conductance g is here a bounded, positive, decreasing, non-linear function which affects equation (3.12) in limiting blurring near edges and increasing the gradient of sufficiently steep edges of the original intensity function L. [Perona and Malik, 1990] have proposed two variations for the conductance:

$$g(\|\nabla L\|) = e^{\frac{-\|\nabla L\|^2}{2k^2}} \quad \text{and} \quad g(\|\nabla L\|) = \frac{1}{1 + \frac{\|\nabla L\|^2}{k^2}} \quad , \tag{3.13}$$

The first function privileges high contrast edges over low contrast edges, and the second function privileges wider regions over smaller ones. The introduced free conductance parameter k reflects the range of gradients in the image or a chosen neighbourhood and thus controls the effect of a given gradient value. For very large k, equations (3.13) approach 1, which will result in linear diffusion. Thus k acts as a threshold for preserving or blurring edges, and can be either manually adjusted, or derived from the *noise estimator* [Canny, 1987] by setting k to the 90% value of the histogram of the absolute values of the gradient at each iteration. Figure 3.4 shows samples using this scheme. The value of 90% is somewhat arbitrary, but usually provides good results. Smaller values lead to only little blurring, and higher values lead to smoothing over weaker borders. [Saint-Marc *et al.*, 1991] have fixed the number of iterations and have used k as a scale parameter, and [Simmons, 1992] has examined the rate of change of k which decreases in a pseudo-exponential manner over time, until it reaches a constant value and no further blurring of the image occurs. This naturally leads to formulate k as a function of evolution time, e.g. by continuously increasing k over time or varying k logarithmically. Alternatively, [Simmons, 1992] has suggested to increment k when it falls below a fraction $\frac{1}{e}$ of its initial value, reflecting its pseudo-exponential decay over time. [Yoo, 1996] has proposed to use normalized local covariances as a basis for choosing local control parameters like the conductance parameter k in variable conductance diffusion processes, providing a statistically adaptive way to normalize the gradient with the expected local noise distribution.

[Alvarez *et al.*, 1992] and [Catté *et al.*, 1992] have pointed out that the method proposed by [Perona and Malik, 1990] has two major drawbacks: For noisy images, very large oscillations of the gradient are introduced. These edges are kept throughout the non-linear diffusion process. Perona and Malik proposed to smooth the image with some low pass filter prior to the diffusion process, which nevertheless results in a non-adaptive filtering causing an unwanted loss of the edges' accuracy. The other drawback lies in equations (3.13) which do not guarantee the expression in equation (3.12) is nondecreasing, which might result in non-deterministic and unstable diffusion

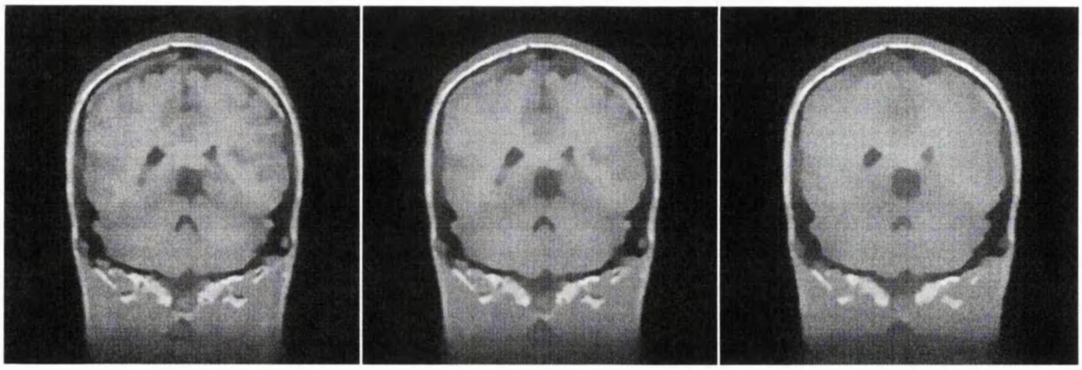


Figure 3.4: Edge-affected scale samples: (a) $t = 4$ (b) $t = 8$ (c) $t = 32$, for time step $\Delta\tau = 0.25$, and k derived from the absolute gradient histogram.

processes. [Catté *et al.*, 1992] therefore have proposed to replace the gradient of equation (3.12) by its estimate at time t and the associated scale σ:

$$\frac{\partial L}{\partial t} = \nabla \cdot g(|\nabla G_\sigma \otimes L|)\nabla L \tag{3.14}$$

where G_σ denotes linear Gaussian blurring at each step of the non-uniform diffusion process in order to obtain a reliable gradient estimate at each time step t. Thus the necessity of smoothing the image before the diffusion process vanishes, and the Gaussian kernel guarantees that the edge-affected diffusion is a monotonically decreasing process. Furthermore, to reflect increasing confidence in the non-linear blurring process, [Whitaker, 1994b] has suggested a multi-scale diffusion technique which makes the scale used for the gradient measurement dependent on the time parameter. In particular, as gradient measurements should become more and more reliable during the diffusion process, the scale should be a decreasing function of the evolution parameter t:

$$\frac{\partial L}{\partial t} = \nabla \cdot g(|\nabla G_{\sigma(t)} \otimes L|)\nabla L \quad , \tag{3.15}$$

where $\sigma(t)$ is called the *scale recipe*.

3.2.2 Geometry-Limited and Multi-Valued Diffusion

[Whitaker and Pizer, 1993; Whitaker, 1993; Whitaker, 1994b; Whitaker, 1994a; Whitaker and Gerig, 1994] have extended the multi-scale diffusion process described above in order to account for higher-order properties. The concept of non-linear diffusion is combined with geometric local shape in order to diffuse according to higher-order information describing local image geometry. This is achieved by performing a diffusion with multiple image features, called *multi-valued diffusion*:

$$\frac{\partial \mathbf{F}}{\partial t} = \nabla \cdot g(D_\sigma \mathbf{F})\nabla \mathbf{F} \tag{3.16}$$

where **F** is the *feature space* of the image, and D is the *dissimilarity* operator, a generalized form of the gradient magnitude. σ, chosen according to the evolution parameter t, is the scale for the linear diffusion process for making the dissimilarity measurements. The derivative of the feature space is given by a generalized Jacobian matrix of $\mathbf{F}$. Thus, equation (3.16) is a system of separate single-valued diffusion processes for different image features, evolving simultaneously and sharing a common conductance term. In order to choose an appropriate feature space, the truncated Taylor expansion given in equation (3.5) yields an appropriate set of partial image derivatives or differential invariants, respectively, which can be used for the *geometry-limited diffusion* process. This allows to diffuse an image with respect to its geometric characteristics, given by edges, corners, high curvature points and other desirable image features.

[Arridge and Simmons, 1997] have extended this approach to perform *probabilistic* diffusion on multi-spectral images, incorporating spatial derivatives and Bayesian classification in image feature-space. The diffusion is either performed on the *maximum a priori* (MAP) from a Bayesian feature-space classification, or on its spatial gradient. Thus boundaries are enhanced which reflect *objectness* rather than intensity differences alone. Other feature-valued techniques include applications to colour filtering, where colour images are stored as a sequence of images [Sapiro and Ringach, 1996], e.g. in red-green-blue (rgb) space or in $L^*a^*b^*$ space which is an approximately uniform colour space.

3.2.3 Geometry-Driven Diffusion

A generalized framework for geometry-driven diffusion equations where a signal can be diffused according to a specific geometric feature has been developed in [Niessen *et al.*, 1997]. Again, the concept of non-linear diffusion is combined with geometric local shape in order to diffuse according to higher-order information describing local image geometry. In order to choose an appropriate feature space, natural scaled differential operators of the image are used to diffuse an image with respect to its geometric characteristics, given by edges, corners, high curvature points and other desirable image features. Geometry-driven diffusion techniques include the *Euclidean shortening flow* which diffuses along the local image isophotes [Alvarez *et al.*, 1992; Catté *et al.*, 1992], the *affine shortening flow* [Sapiro and Tannenbaum, 1993] which additionally incorporates a gradient flow, *modified affine shortening flow* [Niessen *et al.*, 1997] which is a combination of affine shortening flow and gradient flow, and *entropy* or *reaction-diffusion* [Kimia *et al.*, 1995] which is a combination of gradient flow and Euclidean shortening flow. Table 3.3 summarizes the most important schemes, and examples are shown in figure 3.5.

Diffusion scheme	Diffusion feature
Euclidean shortening flow [Alvarez *et al.*, 1992; Catté *et al.*, 1992]	$\frac{\partial L}{\partial t} = L_{vv}$
Affine shortening flow [Sapiro and Tannenbaum, 1993]	$\frac{\partial L}{\partial t} = L_{vv}^{\frac{1}{3}} L_w^{\frac{2}{3}}$
Modified affine shortening flow [Niessen *et al.*, 1997]	$\frac{\partial L}{\partial t} = L_{vv}^{\frac{1}{3}} L_w^{\frac{2}{3}} \cdot \left(\frac{\Vert L_w \Vert}{c}\right)^{-\frac{2}{3}}$
Entropy or reaction-diffusion [Kimia *et al.*, 1995]	$\frac{\partial L}{\partial t} = \alpha \Vert L_w \Vert + \beta L_{vv}$

Table 3.3: Geometry-driven diffusion schemes.

3.2.4 Other Diffusion Schemes

The non-linear diffusion schemes presented so far rely on the differential structure of images. However, for some applications it may be desirable to incorporate higher-level knowledge into the actual diffusion process, which may be achieved by using the feature-based approach of section 3.2.2. Such a *knowledge-based diffusion* can be based on various features, such as temporal information, velocity data, known approximate position and shape of objects in an image. [Sanchez-Ortiz *et al.*, 1996a; Sanchez-Ortiz *et al.*, 1997] have developed a multi-feature and multi-dimension non-linear diffusion technique for density and velocity encoded cine MR sequences of the left heart ventricle, incorporating shape and dynamics of the heart. In [Parker *et al.*, 1998], a non-linear approach to image noise reduction has been developed which performs median and Gaussian blurring weighted by the local image entropy.

A different non-linear diffusion approach is based on *morphological diffusion*. Using the concept of the structuring functions performing *dilation* and *erosion*, a multi-scale morphological scale-space can be computed. [Jackway and Deriche, 1996] have developed such a diffusion technique based on classic structuring elements like circles of increasing diameters. *Normal* or *constant motion* [Niessen *et al.*, 1997] is based on the gradient flow, resulting in morphological erosion or dilation depending on the sign of c,

$$\frac{\partial L}{\partial t} = c \Vert L_w \Vert \quad , \tag{3.17}$$

an expression which can be obtained by setting $\beta = 0$ for the entropy diffusion equation (see table 3.3). Moreover, [Jackway and Deriche, 1996] and [Park and Lee, 1996] have shown independently that $1D$ morphological scale-spaces satisfy the causality principle that no new features are generated for increasing scale levels. [van den Boumgard and Smeulders, 1994; Maragos, 1996] have formulated differential equations of morphological scale-spaces, and [van den Boumgard, 1997] has formulated the concept of *morphological deformation curves* which are obtained by continuously deforming a shape and in the meantime measuring some geometric parameter, yielding *a measurement as a function of deformation*. [Bangham *et al.*, 1996a; Bangham *et al.*, 1996b; Harvey *et al.*, 1997] have developed a morphological scale-space called *sieves* which is equiva-

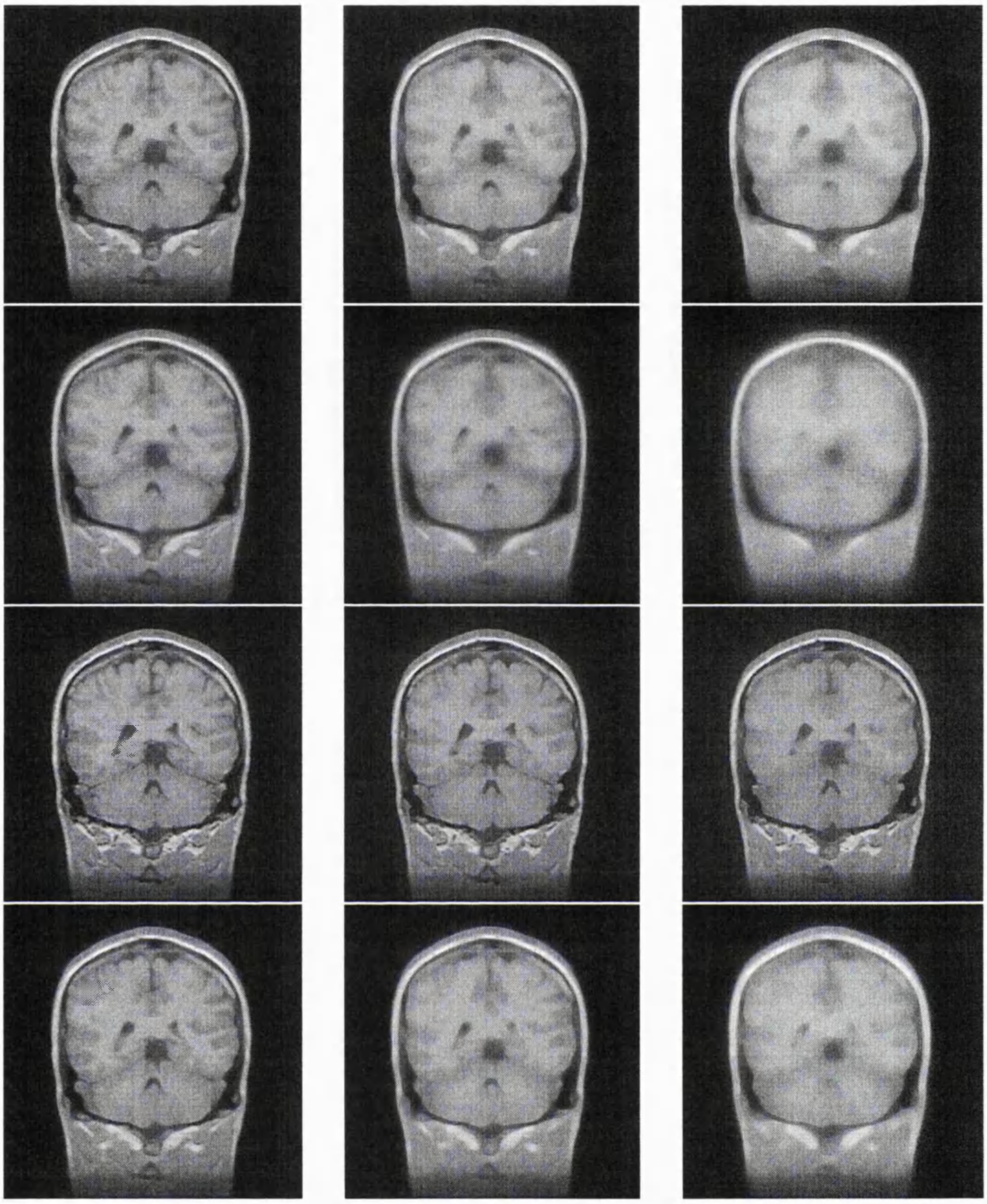


Figure 3.5: Geometry-driven scale samples. From top to bottom: Euclidean shortening flow, affine shortening flow, modified affine shortening, for entropy (reaction-diffusion), for (from left to right) evolution times $t = 3.2, 6.4, 12.8$, time step $\Delta\tau = 0.1$, $\sigma = 0.8$, $c = 0.5$, $\alpha = 0.1$, $\beta = 1$).

lent to placing a sieve of increasing "hole" sizes over the image surface and cutting all peaks of the surface which fit through the holes. Technically speaking, *sieves* present an image as a graph Γ with a set of edges describing the adjacency of the image pixels. The graph therefore defines the neighbourhood of a particular pixel $\mathbf{x}$. Defining a region $C_n(\Gamma; \mathbf{x}) = \{\eta \in C_n(\Gamma) | \mathbf{x} \in \eta\}$ over the graph that encloses $\mathbf{x}$ yields the set of connected subsets of the graph with n elements that contain $\mathbf{x}$. Diffusion based on the concept of sieves is performed by increasing n (the "holes" of the sieves) and alternating opening (ω_n) and closing (γ_n) operations which are defined by

$$\omega_n L(\mathbf{x}) = \max_{\eta \in C_n(\Gamma;\mathbf{x})} \min_{\mathbf{u} \in \eta} L(\mathbf{u}) \quad \text{and} \quad \omega_n L(\mathbf{x}) = \min_{\eta \in C_n(\Gamma;\mathbf{x})} \max_{\mathbf{u} \in \eta} L(\mathbf{u}) \tag{3.18}$$

$\gamma_n \omega_n$ denotes then grey-scale opening followed by closing, and $\omega_n \gamma_n$ is defined vice versa. The advantage of *sieves* is their fast graph-based implementation and its ability to follow extremal image regions. Moreover, *sieves* do not introduce new extrema, something which can occur under linear diffusion for dimensions greater than 1 [Lifshitz and Pizer, 1990].

3.3 Scale-Space Issues

In order to investigate an image scale-space, several implementation-related issues need to be considered. First of all, the scale-space implementation can be performed via an implicit or explicit scheme, which will be presented in the following. Another important issue that will be addressed is the definition of scale-space differentiation and metrics.

3.3.1 Explicit and Implicit Implementation

In general, the evolution of the luminance function L (the original image) is a mapping of the form $(\mathbf{x}, t) \to L(\mathbf{x}, t)$, constrained by the evolution equation:

$$\frac{\partial L}{\partial t} = \nabla F(L_i, L_{ij}, ...) \tag{3.19}$$

in which F is a function in terms of the local N-jet. There are two main ways to perform a geometrically-driven or other diffusion processes. The explicit scale-space implementation scheme is based on Euler forward steps:

$$\frac{L(\mathbf{x}, t_0 + \Delta\tau) - L(\mathbf{x}, t_0)}{\Delta\tau} = F(L_i, L_{ij}, \cdots; t_0) \tag{3.20}$$

Thus the image at time $t_0 + \Delta\tau$ can be derived by taking the image at time t_0 and adding the diffusion feature F weighted by a time step $\Delta\tau$ for numerical stability. All of the linear and non-linear schemes listed above can be implemented in this scheme.

The implicit scale-space implementation scheme needs a little more consideration. It has been noted by [Niessen *et al.*, 1997] that it is often difficult to find a rotationally invariant implicit scheme (for example, for the affine shortening flow no such scheme is feasible), but the linear and the edge-affected diffusion schemes can be implemented this way. Implicit scale-space implementation is numerically more stable than the external scheme, and therefore larger time steps can be taken. In $2D$, a $3D$ sparse (tridiagonal by blocks) matrix needs to be solved, for example by using the alternating direction implicit method. In general, the implicit scheme has the following form:

$$\frac{L(\mathbf{x}, t_0 + \Delta\tau) - L(\mathbf{x}, t_0)}{\Delta\tau} = F(L_i, L_{ij}, \cdots; t_0 + \Delta\tau) \tag{3.21}$$

This implies that $F(L_i, L_{ij}, \cdots; t_0 + \Delta\tau)$ needs already to be known for the diffusion step, as the right hand side of the equation as above is evaluated at time $t_0 + \Delta\tau$ rather than t_0. In consequence, at each time step a set of linear equations needs to be solved.

3.3.2 Differentiation and Metrics

One property of a scale-space representation is that the amplitude of spatial derivatives generally decreases for increasing scales. This property can be derived from the maximum principle [Babaud *et al.*, 1986]. [ter Haar Romeny *et al.*, 1991; Florack *et al.*, 1992] therefore introduced the notion of *normalized derivatives* which are invariant with respect to rotation, translation and scaling (or rather zooming) in terms of normalized (dimensionless) coordinates $\tilde{\mathbf{x}}$ and a natural scale parameter $\tilde{\sigma}$:

$$\tilde{\mathbf{x}} = \frac{\mathbf{x}}{\sigma} \quad \text{and} \quad \tilde{\sigma} = ln\left(\frac{\sigma}{\epsilon}\right) \quad \text{with} \quad \sigma = \sqrt{2t} \tag{3.22}$$

The natural scale parameter introduces a hidden scale ϵ which carries the dimension of a length and which allows to parameterize σ by a continuous quantity $\tilde{\sigma} \in (-\infty; \infty)$. Typically, only positive integer values for $\tilde{\sigma}$ are used (as negative values correspond to scales below the inner scale), and an equidistant natural scale sampling schedule is chosen. The normalized differential forms are then given by

$$\mathrm{d}\tilde{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\sigma} - \frac{\mathbf{x}}{\sigma}\frac{\mathrm{d}\sigma}{\sigma} \quad \text{and} \quad \mathrm{d}\tilde{\sigma} = \frac{\mathrm{d}\sigma}{\sigma} \tag{3.23}$$

which reduces to

$$\mathrm{d}\tilde{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\sigma_0} \quad \text{and} \quad \mathrm{d}\tilde{\sigma}_0 = 0 \tag{3.24}$$

for fixed-scale differential measurements. First-order dimensionless or normalized derivatives are then given by $\sigma_0 L_{\mathbf{x}_i}$ for each spatial component $\mathbf{x}_i$, and second-order dimensionless derivatives are consequently given by $\sigma_0^2 L_{\mathbf{x}_i\mathbf{x}_j}$. The gradient of L is given by the vector $\nabla L = [L_{\mathbf{x}_i}]$, and the Hessian matrix of second derivatives of L is $\mathbf{H} = [L_{\mathbf{x}_i\mathbf{x}_j}]$. The Euclidean differential of L is

$$\mathrm{d}L = \sum_{i=1}^{n} L_{\mathbf{x}_i}\mathrm{d}\mathbf{x}_i = \sum_{i=1}^{n} \sigma_0 L_{\mathbf{x}_i}\frac{\mathrm{d}\mathbf{x}_i}{\sigma_0}\,, \tag{3.25}$$

leading to the *Euclidean scale-space gradient,*

$$\tilde{\nabla} L = (\sigma_0 L_{\mathbf{x}_1}, \cdots, \sigma_0 L_{\mathbf{x}_n})\,, \tag{3.26}$$

and the *Euclidean scale-space Hessian,*

$$\tilde{\mathbf{H}} = [\sigma_0^2 L_{\mathbf{x}_{ij}}]\,. \tag{3.27}$$

In general, the dimensionless spatial derivatives are given by the unnormalized spatial derivatives multiplied by the scale to the power of the order of differentiation. The underlying *metric* for natural coordinate and scale representation is *Euclidean*, which means that the natural distance between two points $\mathbf{x}_1$ and $\mathbf{x}_2$ at a fixed scale σ_0 is given by the scale-space distance:

$$\mathrm{dist}_{scale-space}(\mathbf{x}_1;\mathbf{x}_2;\sigma_0) = \frac{\|\mathbf{x}_1 - \mathbf{x}_2\|}{\sigma_0} = \|\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2\| \tag{3.28}$$

An extension to this approach has been proposed by [Eberly, 1994a], who stated that the natural coordinate representation is only translationally invariant if the scale is assumed to be fixed. For multi-scale measurements, however, both spatial and scale differences are meaningful only in the context of the scale at which they are measured, which introduces the dimensionless spatial and scale differential forms

$$\mathrm{d}\tilde{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\sigma} \quad \text{and} \quad \mathrm{d}\tilde{\sigma} = \frac{\mathrm{d}\sigma}{\sigma} \tag{3.29}$$

For a fixed scale σ_0, these forms correspond to the natural ones given in equations (3.23) and (3.24), respectively. For a non-constant scale, however, this scale-space geometry becomes *non-Euclidean*, but hyperbolic and *Riemannian*. Let $\xi_i = \mathbf{x}_i$ for $1 \leq i \leq n$ and $\xi_{n+1} = \sigma$, the Euclidean gradient given by $\nabla L = [L_{\xi_i}]$, and the Hessian given by $\mathbf{H} = [L_{\xi_i \xi_j}]$. The differential of L is then:

$$\mathrm{d}L = \sum_{i=1}^{n+1} L_{\xi_i} \mathrm{d}\xi_i = \sum_{i=1}^{n} \sigma L_{\mathbf{x}_i} \frac{\mathrm{d}\mathbf{x}_i}{\sigma} + \sigma L_\sigma \frac{\mathrm{d}\sigma}{\sigma}\,, \tag{3.30}$$

leading to the *Riemannian scale-space gradient* of L:

$$\tilde{\nabla} L = (\sigma L_{\mathbf{x}_1}, \cdots, \sigma L_{\mathbf{x}_n}, \sigma L_\sigma) = [\sigma L_{\xi_i}]\,, \tag{3.31}$$

which reduces to equation (3.26) for fixed scales (since $\sigma_0 L_{\sigma_0} = 0$). The *Riemannian scale-space Hessian* is consequently given by

$$\tilde{\mathbf{H}} = [\sigma^2 L_{\xi_i \xi_j}] \tag{3.32}$$

The hyperbolic geometry or Riemannian metric of the scale-space suggests that the scale-space distance between two points$(\mathbf{x_1};\sigma_1)$ and $(\mathbf{x_2};\sigma_2)$ which are located at possibly different scale levels with $\sigma_1 \leq \sigma_2$, is to be computed along the geodesic curve connecting these points:

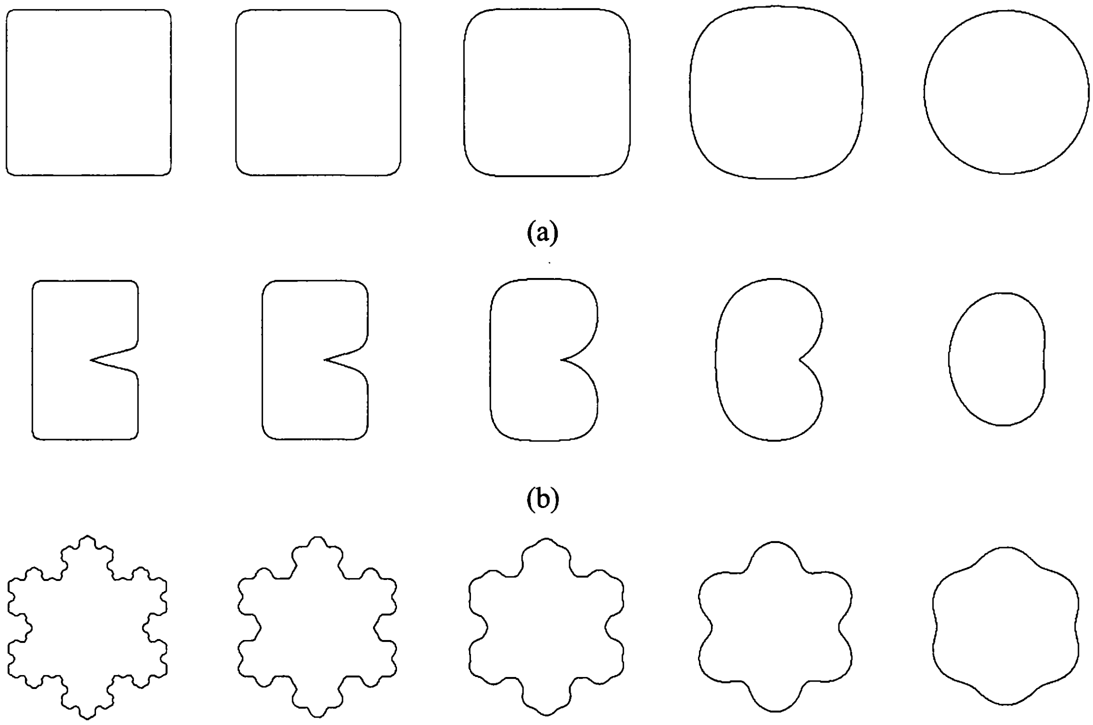
$$\mathrm{dist}_{scale-space}((\mathbf{x}_1;\sigma_1);(\mathbf{x}_2,\sigma_2)) = \log\left(\frac{\sigma_2}{\sigma_1}\frac{\left(1+\sqrt{1-(\rho\sigma_1)^2}\right)}{\left(1+\sqrt{1-(\rho\sigma_1)^2}-\rho\|\mathbf{x}_1-\mathbf{x}_2\|\right)}\right) \tag{3.33}$$

where

$$\rho = \frac{2\|\mathbf{x}_1 - \mathbf{x}_2\|}{\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + \|\mathbf{x}_1 - \mathbf{x}_2\|^2(\|\mathbf{x}_1 - \mathbf{x}_2\|^2 + 2(\sigma_1^2 - \sigma_2^2))}} \tag{3.34}$$

3.4 Applications in Contour and Image Analysis

The concept of scale-space has found many applications in shape description and image processing, the main ones being multi-scale signal analysis, e.g. scale-spaces of binary contours, multi-scale edge detection and segmentation, and multi-scale image feature detection and analysis.



(a)

(b)

(c)

Figure 3.6: Samples of a linear contour scale-space. (a) square, (b) notched rectangle, (c) Koch curve (section 2.3.3), at scales $\sigma = 16, 32, 64, 128, 256$ (from left to right). Also compare to the Fourier truncation of the same shapes in figure 2.7 in the previous chapter, section 2.1.2.1.

3.4.1 Multi-Scale Contour Analysis

In order to examine binary contours in scale-space, several scale-based techniques have been developed in literature. Convolving a contour $\mathbf{v}(s)$ with a one-dimensional Gaussian function G at increasing levels of scale yields a set of *evolved* versions of the contour:

$$\mathbf{v}(s;\sigma) = (x(s;\sigma), y(s;\sigma)) \tag{3.35}$$

with

$$x(s;\sigma) = G(s;\sigma) \otimes x(s) \quad \text{and} \quad y(s;\sigma) = G(s;\sigma) \otimes y(s) \tag{3.36}$$

Figure 3.6 illustrates a contour scale-space in terms of smoothed contour sequences. As can be seen, small or fine details of the contour disappear already at very low scales, while coarser shape characteristics are kept until larger scales. For very high scales, planar shapes eventually become circular and shrink to a point [Gage and Hamilton, 1986]. Comparing figure 3.6 to figure 2.7 in the previous chapter, section 2.1.2.1, makes also clear that similarly to the incomplete Fourier reconstruction finer details (represented by higher frequency components) can be removed by smoothing the shape with increasing scale levels. In contrast to the Fourier reconstruction, however, the higher frequency components are suppressed *locally* rather than *globally*, and the topology of the curve is maintained throughout the smoothing process. (Note however, that the topology is not always preserved for higher dimensions, as shapes embedded in an image may collapse, split or merge.) From this contour scale-space, [Witkin, 1983] has investigated the first few derivatives for general purpose qualitative shape description. In particular, the zero-crossings of a signal and its derivatives provide meaningful shape information, especially those of the second derivative (or the *Laplacian* of the signal) which denote extrema of the slope, i.e. points of inflection. Consequently, the points of inflection of a $1D$ signal at all scale levels σ are given by

$$x_{ss}(s;\sigma) = 0 \quad \text{with} \quad x_{sss}(s;\sigma) \neq 0 \tag{3.37}$$

Tracking the zero-crossings over scales yields the so-called *scale-space fingerprints* [Witkin, 1983]. Moreover, reducing the scale-space to a simple *interval tree*, the qualitative structure of the signal over all observed scales can be concisely described. Further properties of the *fingerprint* representation have been investigated by [Yuille and Poggio, 1986] who have shown that the only filter that does not introduce new zero-crossings for $1D$ signals at increasing scales is the Gaussian function.

In the previous chapter, section 2.2 it has been noted that local contour features can be derived from combinations of differential measurements in terms of partial derivatives. The same holds for multi-scale contour features, which extend this approach to a higher scale dimension. The partial nth order derivatives of a contour can be computed by convolving the signal with the nth order derivatives of the Gaussian, e.g.:

$$x_s(s;\sigma) = \frac{\partial}{\partial s}(x(s) \otimes G(s;\sigma)) = x(s) \otimes G_s(s;\sigma) \tag{3.38}$$

$$x_{ss}(s;\sigma) = \frac{\partial^2}{\partial t^2}(x(s) \otimes G(s;\sigma)) = x(s) \otimes G_{ss}(s;\sigma) \tag{3.39}$$

and $y_s(s;\sigma)$, $y_{ss}(s;\sigma)$ are computed analogously, leading to the first and second order contour derivatives $\mathbf{V}_s(s;\sigma) = (x_s(s;\sigma), y_s(s;\sigma))$ and $\mathbf{V}_{ss}(s;\sigma) = (x_{ss}(s;\sigma), y_{ss}(s;\sigma))$, respectively. A multi-scale curvature representation can be obtained analogously to equation (2.29) defined in section 2.2:

$$\kappa(s;\sigma) = \frac{x_s(s;\sigma) \cdot y_{ss}(s;\sigma) - y_s(s;\sigma) \cdot x_{ss}(s;\sigma)}{(x_s(s;\sigma)^2 + y_s(s;\sigma)^2)^{3/2}} \tag{3.40}$$

Various researchers have developed multi-scale contour description techniques based on the differential multi-scale contour properties. For example, the *curvature primal sketch* [Asada and Brady, 1986] defines a set of primitive curvature discontinuities, and then matches the multi-scale convolutions of the shape. Thus significant changes of curvature at various scales can be located and further investigated. Curvature changes that are only found at fine scales are less significant than those found or *tracked* across multiple scales, and are considered to be less geometrically significant. Several types of contour features like corners, smooth joins, ends, cranks, bumps or dents can be analysed by using curvature changes. This approach provides a method for *syntactical* shape representation and can also be used for shape reconstruction. Moreover, the invariance of scale-space curvature changes under affine transformations constitutes a *scale-space signature* of the contour, which is similar to the *scale-space fingerprint.*

[Mokhtarian and Mackworth, 1986; Mokhtarian and Mackworth, 1992] have presented a multi-scale, curvature-based shape representation technique for planar curves similar to the *scale-space fingerprint.* A so-called *curvature scale-space* (CSS) image is computed by extracting the curvature zero-crossings of the resulting curves until no further zero-crossings occur (until the curves become convex). The function implicitly defined by

$$\kappa(s;\sigma) = 0 \tag{3.41}$$

is the CSS of $\mathbf{v}(s,\sigma)$. To compute a *renormalized* curvature scale-space image, each evolved curve can be reparameterized by its normalized arc length parameter. Using increasing values for σ causes the convolved curve to shrink in direct proportion to the standard deviation. However, if the amount of movement is estimated at each point on the smoothed edges, and a vector is added to the location vector to compensate for that movement, the resulting smoothed curve is physically closer to the original curve. The curvature zero-crossings for all scales are marked in the CSS, which can then be used for scale-space tracking to determine the rate of change of curvature. For example, CSS-based applications include a silhouette-based shape recognition [Mokhtarian, 1995b; Mokhtarian, 1996], multi-scale occluded contour segmentation [Mokhtarian, 1997a], and shape indexing from large databases [Mokhtarian *et al.*, 1996; Abbasi *et al.*, 1997]. The convergence properties of the CSS have been studied in [Mokhtarian, 1995a].

3.4.2 Multi-Scale Edge Detection and Segmentation

[Canny, 1987] has suggested as an optimal edge detector the located maxima of the gradient magnitude of a Gaussian-smoothed image using different widths. This results in a *fine-to-coarse* integration of edge information at different scales. The following performance criteria need to be fulfilled by such an edge detector:

Good detection: All edge points or *edgels* should be located, and *only* those. This corresponds to maximizing the *signal-to-noise ratio* of the operator's output.

Localization: The *edgels* should be as close as possible to the centre of the true edge.

Single response: Each edge should only yield a single response by the edge detector. This is implicitly covered by the good detection criterion, since if an edge is detected twice, one of the responses must be considered false.

Different degrees of smoothing are therefore incorporated by smoothing the data adaptively on the signal-to-noise ratio.

[Bergholm, 1987] has extended Canny's approach to edge detection by performing a *coarse-to-fine tracking* of edges leading to an *edge focusing* approach with high positional accuracy and good noise reduction. This naturally yields a hierarchical segmentation scheme. Care must be taken in the choice of the step length of the scale parameter, as edge elements should not move more than one spatial pixel per focusing step to allow for a stable edge following algorithm in scale-space. The key point of this technique is that the global shape of an object can be located at a high level of blurring without any disturbances due to noise and irrelevant image detail. Subsequently focusing down the edges surrounding that object allows to track down finer scale details of the object outline also at finer scales. [Sjöberg and Bergholm, 1988] have further investigated the displacement of the extracted edges due to blurring and labelled them as diffuse and non-diffuse.

There are only a few multi-scale segmentation techniques, which is partially due to the high computational cost and amount of memory when constructing a scale-space. Moreover, they never perform a truly *multi-scale*, but rather a *coarse-to-fine* approach, as results of one scale are linked to the next lower scale in a hierarchical fashion. The *hyperstack* developed by [Koster, 1995; Vincken, 1995; Koster *et al.*, 1996; Vincken *et al.*, 1997] is a true multi-scale technique for image segmentation. It identifies the root of a segment at a high scale level, using a bottom-up linking process in which the scale-space levels are linked to another pixel-wise. This results in a tree of linkages in scale-space. For each segment in the image, such a tree is constructed and linked in a root labelling phase. This technique has been applied in medical imaging with linear and non-linear underlying scale-spaces [Vincken *et al.*, 1996]. An interesting characteristic of the hyperstack is that the so-called *partial volume effect*, caused by the limited resolution due to the acquisition method and leading to multiple object voxels (volumetric image pixels), is taken into account using a probabilistic approach.

3.4.3 Multi-Scale Image Feature Detection and Analysis

Differential geometry in image processing in combination with a multi-scale image representation offers the possibility of detecting and analysing syntactical image structures. There are three main developments in this area, notably the multi-scale detection of *blobs* or circular structures, the multi-scale description of image structure, and multi-scale medial image analysis which will be briefly reviewed.

3.4.3.1 *Blob* Detection

The detection of circular, blob-like structures has recently attracted much interest in medical imaging, e.g. as a way to locate lesions due to Multiple Sclerosis in MR brain scans [Gerig *et al.*, 1995]. [Lindeberg, 1993a; Lindeberg, 1994] has detected blobs by localizing *normalized scale-space extrema* of the normalized Laplacian (computed using scale-normalized derivatives as described above). Displaying the scale variation of the absolute value of the normalized Laplacian at each image point yields the *scale-space signatures* of

$$\|\nabla^2_{norm} L(\mathbf{x}, \sigma)\| = \|\sigma^2 \nabla^2 L(\mathbf{x}, \sigma)\| \tag{3.42}$$

with a maximum response at the scale σ_0 corresponding to the local characteristic object width. Using local directional statistics, ellipse-like structures in all orientations can be recovered from the image. The multi-scale *medialness* measurement which will be presented in more detail in section 3.4.3.3 below is also able to detect circular shapes, where again the size of the circle corresponds to the scale at which the scale-space extremum is detected. [Gerig *et al.*, 1995] have taken a slightly different approach for blob detection. They have suggested to use a non-linear scale-space computed as a Euclidean shortening flow, in which local extrema in the spatial domain are detected, corresponding to isophote level curve singularities. These singularities are subsequently tracked down in the non-linear scale-space, taking *a priori* knowledge like image contrast, expected blob size, and the standard deviation of the singularity location (radial symmetry) into account. The main problem with this otherwise very elegant approach is that the tracking of the singularities sometimes yields unconnected traces.

3.4.3.2 Multi-Scale Description of Image Structure

Multi-scale differential image invariants as formulated by [Florack, 1993] provide a powerful approach for investigating the *differential* structure of an image, leading to the notion of a continuum of structures on an interval of scales (*deep image structure*), and at a single scale only (*superficial image structure*). A causal hierarchy can be observed as finer structures determine coarser structures (but not vice versa). [Florack, 1993] distinguishes between the *full* differential image structure (given by the local N-jet), and subclasses of image structure, namely image isophotes, which do not change under arbitrary, invertible intensity transformations. Many classical image operators, such as the Marr-Hildreth edge detector or Canny's edge detector, are embedded into the framework of image invariants, which provide meaningful input for higher-level image processing tasks.

Ridges (and their dual, creases) provide another powerful image analysis tool, combining edge-based and region-based methods. [Eberly *et al.*, 1993; Eberly, 1994a; Eberly, 1996] have presented various geometric techniques for ridge detection and analysis, including *height* ridges

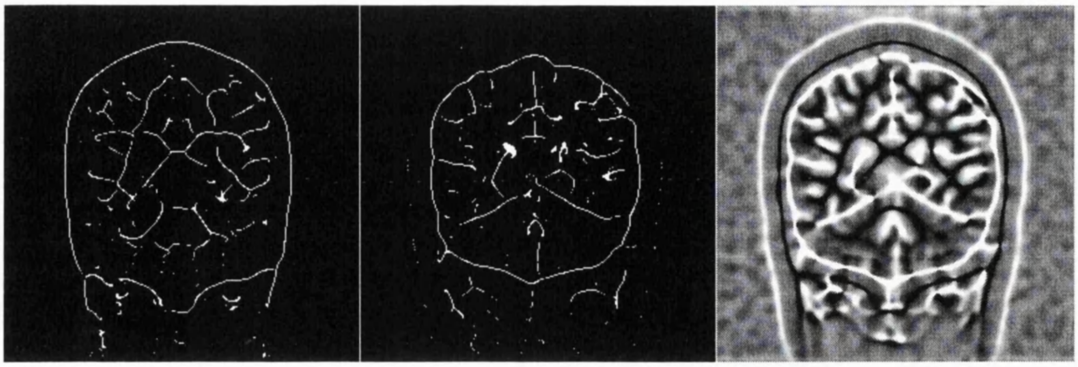


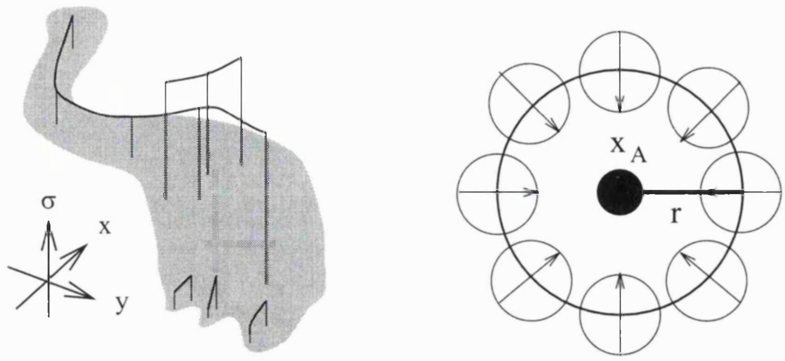
Figure 3.7: Ridge measures of the image intensity. (a) Binary ridges of the intensity. (b) Binary creases of the intensity. (c) Fuzzy ridgeness of the intensity (where ridges have high values and creases have low values). The image used throughout this chapter was blurred at scale $\sigma = 4$.

which are computed via local extrema of e.g. the image intensity function. More specifically, ridge points of a function f are defined as local maxima of f along the direction of the greatest principal curvature of f. Computing for an ND-dimensional function the eigenvalues $||\lambda_1|| \geq \cdots \geq ||\lambda_n||$ of the Hessian matrix $\mathbf{H}$ of second derivatives (see section 3.3.2) and their corresponding eigenvectors $\mathbf{e}_1 \cdots \mathbf{e}_n$, a point is defined to be on an m-dimensional ridge, $m < n$, if for all $i < n - m$

$$\lambda_i < 0 \quad \text{and} \quad \mathbf{e}_i \cdot \nabla f = 0 \quad . \tag{3.43}$$

[Eberly and Pizer, 1994] have extended this ridge definition for hierarchical image segmentation by computing ridges of image scale-spaces (using the Riemannian scale-space Hessian), and segmenting each level of the scale-space by decomposing the ridges into curvilinear segments, followed by labelling and constructing a region for each ridge segment based on a ridge flow model. Figures 3.7 (a) and (b) show the extracted ridges and creases of the image intensity, respectively, for an intermediate scale level $\sigma = 4$. A fuzzy (non-binary) *ridgeness* measure for medical image registration has been suggested by [van den Elsen *et al.*, 1995; Maintz *et al.*, 1996a; Maintz, 1996]. Their measurement is based on the L_{vv} operator which represents the second order derivative in the direction perpendicular to the local gradient direction. They have extended this measurement to $3D$ which is non-trivial as in $3D$ the v direction, being perpendicular to the gradient, needs another constraint to be properly defined [Maintz *et al.*, 1996b]. Other fuzzy ridgeness measurements include the negated isophote curvature (L_{vv}/L_w), and the more general operator $L_{vv}L_w^{-\alpha}$. The negated L_{vv}/L_w image is illustrated in figure 3.3, and the L_{vv} image for scale $\sigma = 4$ is shown in figure 3.7(c).

Related to ridges are watersheds (point loci dividing image areas that drain to different minima and watercourses (the converse to watersheds), which together divide the image into dales (areas



(a) (b)

Figure 3.8: (a) Multi-scale medial axis shown in scale-space. Height indicates the object width as a function of axis position. (b) Individual boundary-sensitive directional operators combine to produce a medial response.

that drain to the same minimum) and hills (analogous areas for maxima). Watershed transforms of the image gradient have become very popular as a multi-scale, morphological segmentation tool [Jackway, 1996; Choy and Jin, 1996; Najman and Schmitt, 1996; Maes *et al.*, 1995], providing a generic and accurate image decomposition method. However, it has also been observed that they tend to over-segment the image, leading to the necessity of grouping regions in a hierarchical segmentation process or a *minimum description length* (MDL) criterion which selectively merges neighbouring regions such that the total boundary length is reduced while the relevant object regions are maintained.

[Griffin, 1995; Griffin and Colchester, 1994; Griffin and Colchester, 1995] have extensively investigated the superficial and deep structure of the image surface in terms of inter-relationships of image isophotes and their dual, the flowlines, as well as critical points and *separatrices* in linear scale-spaces. Regarding the image intensity as height above a plane, separatrices of two categories are defined: those running *uphill* from isophote saddle points to intensity maxima, and those running downhill from saddle points to intensity minima. Mapping all separatrices onto the image planes divides the image into *districts*. [Griffin and Colchester, 1995] have identified separatrices as a superset to watersheds and watercourses: from each saddle, there are two separatrices running uphill, which as a pair form a watershed if they separate two drainage basins. Otherwise they are *virtual separatrices*. An analogous virtual definition applies to downhill separatrices. This frameworks provides a generic platform for multi-scale image description and inter-scale linkage segmentation, e.g. for the tracking of hierarchical partitions across scales and linking them into a *multi-scale n-ary hierarchy* based on constraint-guided correspondences [Griffin, 1995].

3.4.3.3 Multi-Scale Medial Axis (*Cores*)

An important approach to representing the structural shape of a region is to reduce it to its *medial axis* as presented in section 2.1.3. A recent development is a new type of medial axis called *multi-scale medial axis* (MMA) or *core*, originally presented by [Burbeck and Pizer, 1994], with extensions by [Eberly, 1994b; Morse, 1994; Fritsch, 1994]. The MMA is computed by measurements taken at multiple scales, particularly scales proportional to the object width. The idea of this approach is to describe object boundaries at multiple scales by pairing corresponding sides of the boundary. This concept is illustrated in figure 3.8(a). Fuzzy boundary measures are used to compute fuzzy medial measures, and medial axis points are identified as *height ridges* in this fuzzy medial space. Therefore the fuzzy nature of the image is retained until the highest level possible, which avoids the loss of information at an early stage. The extraction of *cores* is performed in two steps:

Boundariness and Medialness Measurement:

After computing a linear image scale-space, the *boundariness* or boundary-like behaviour of each point with respect to the scale is determined by applying a simple boundary function B, e.g. $B(\mathbf{x};\sigma) = \sigma\nabla L(\mathbf{x};\sigma)$. Then the *medialness* of each pixel with respect to specific object widths is derived by letting each boundariness measurement vote for possible medialness in a manner similar to the Hough transform(see section 2.1.2.3). Alternatively, the normalized Laplacian can be used. Points are medial with respect to a certain half-width r if there are at least two boundary points at distance r from that point. Consequently, the medial response $M(\mathbf{x}_A, r)$ is related to the amount of boundary response for the set of points $\{\mathbf{x}_B\}$ on a circle of radius r centred at $\mathbf{x}_A$ (see figure 3.8(b)). Additionally, $M(\mathbf{x}_A, r)$ depends on the directional response for all points on the circle in direction $\mathbf{x}_B - \mathbf{x}_A$. The integrated directional response along the points of the circle yields the medial response, which is the basis for the multi-scale *Hough-like medial axis transform* (HMAT):

$$M(\mathbf{x}_A, r) = \int_{\mathbf{u}} B(\mathbf{x}_A - r\mathbf{u}; \mathbf{u}; \sigma) d\mathbf{u} \tag{3.44}$$

where C is the unit circle, and $\mathbf{u}$ are the points on that circle. The half-width r can be related to scale by

$$r = k \cdot \sigma \quad \text{or, in practice,} \quad r = k \cdot \sigma - c\,, \tag{3.45}$$

where k is a proportionality factor and c is a small constant.

Analysis of Medial Response Space

The second step is to examine the medial response space to identify the *ridge points* which define the object *core*. The ridge points are found by computing the Riemannian scale-space Hessian matrix of second derivatives and its eigenvectors and sorted eigenvalues as described above. If the $n-1$ largest eigenvalues for any pixel are negative, a ray is projected through the point in direction of that eigenvector until it intersects a face of the cube formed by the neighbours of the point. The face's value is then interpolated by using the neighbours' values and compared to the value of the interpolating neighbouring point.

Cores were originally developed for symmetry detection and medical image registration. Recent applications of *cores* in medical imaging include stimulated cores [Fritsch *et al.*, 1995], shape-based segmentation and description [Lepard and Robb, 1996], and scale-space boundary evolution [McAuliffe *et al.*, 1996]. These applications are all based on the *boundary at the scale of the core* (BASOC) which is obtained by back-projecting the core to the shape boundary in scale-space using equation (3.45), thus providing a multi-scale boundary representation which can be refined using standard segmentation methodologies.

3.5 Summary

This chapter has presented a survey on multi-scale image processing techniques. It has been divided into the concepts of linear and non-linear image scale-spaces, geometrically meaningful multi-scale differential invariants, scale-related issues concerning implementation, differentiation and metrics, and applications in image processing.

Linear scale-space theory and its associated multi-scale differential invariants are powerful tools for geometric image interpretation. While suppressing noise and spurious image features, the global spatial relationship of image pixels at any desired detail can be described. Non-linear diffusion eliminates some of the problems caused by linear smoothing, as it allows to preserve edges and high curvature parts of an image even at higher scales, while simultaneously smoothing homogeneous regions. This concept provides a very useful platform for image interpretation and segmentation, extending the image dimensionality to an extra scale degree of freedom.

Besides spatial convolution and multiplication in the Fourier domain, scale-spaces can be computed via finite differencing, and implicit and explicit diffusion. The latter is often the preferred technique, as it allows to diffuse an image with respect to a large variety of geometric image features based on differential invariants. When performing computations *on* an image scale-space, either a Euclidean setting (for a constant scale) or a Riemannian setting (for varying scale) has to be taken into account before performing differential or spatial measurements.

Applications of multi-scale techniques can be divided into contour and image methods:

Multi-scale contour representation and analysis has the advantage over higher dimensional multi-scale techniques that no new extrema are created for increasing scales. Valuable multi-scale contour-based techniques are scale-space fingerprints of the Laplacian zero-crossings of a contour, and their curvature counterpart, the curvature scale-space (CSS) image. Both representations establish a hierarchical shape representation, which can be further explored at each scale level: Differential measurements of a multi-scale contour representation allow for local scale-based contour shape description (e.g. in terms of points of inflection and corners), global multi-scale shape measurements (e.g. the change of perimeter over scale), and relative distance measurements in scale-space. To construct a contour scale-space, however, the contour must be available at its zero scale.

Multi-scale image representation and analysis have their main application in feature detection (in terms of scale-space extrema) and segmentation. *Coarse-to-fine* segmentation strategies allow to focus objects in an image robustly down by tracking features across scales. These approaches use a hierarchical image representation, but discard all but the final results obtained at the lowest scale level. In contrast, true *multi-scale* segmentation techniques allow to obtain an object representation at varying scale levels by reducing the image scale-space to its most meaningful spatial and scale locations, like the *hyperstack*, the *multi-scale n-ary hierarchy*, and techniques based on the *BASOC*. The latter is also very interesting in terms of shape description, as it is derived from the *core* representation of an object which is a multi-scale medial shape representation. *Cores*, however, are computationally expensive to compute, and suffer from ridge linkage problems.

Integrating multi-scale image processing techniques into the shape extraction or segmentation process allows to construct a shape hierarchy in scale-space. This dissertation follows this idea by integrating multi-scale differential invariants into a recently developed powerful segmentation tool: active contour models. The following chapter introduces the theory of active contours, followed by chapter 6 which presents a multi-scale active contour model for segmentation and shape description as a novel approach of this dissertation.

Chapter 4

Survey of Active Contour Models (*Snakes)*

– QUELLE EST CETTE HISTOIRE-LÀ! TU PARLES MAINTENANT AVEC LES SERPENTS!
"WHAT DOES THIS MEAN? WHY ARE YOU TALKING WITH SNAKES?"
Le Petit Prince, Antoine de Saint-Exupéry.

Active contour models, first introduced by Kass, Witkin and Terzopoulos [Kass *et al.*, 1987b; Kass *et al.*, 1987a], represent a special form of the more general multi-dimensional deformable model developed by [Terzopoulos, 1986a; Terzopoulos, 1986b]. Active contour models are often referred to as the classic *snake* or *deformable contour model.* They are energy-minimizing splines guided by internal shape forces, external constraint forces, and external image forces like edges that pull them towards images features during an optimization process. They dynamically segment an image by locking onto nearby edges and localizing them accurately. Applications of active contour models include line and edge detection, detection of subjective contours, motion tracking, stereo matching, and interactive interpretation of image scenes with user-imposed constraints, in the areas of computer vision, computer graphics, computer-aided geometric design, and more recently in computer-assisted medical image analysis. An extensive survey of current research in this area is given in [McInerney and Terzopoulos, 1996; Terzopoulos, 1996].

The key point of active contour models is the design and optimization of a suitable energy function whose local minima comprise a set of alternative solutions which can be based on *a priori* knowledge of the approximate shape, size, location, and motion of the object under investigation, or on a user-defined initial estimate. In lack of such a mechanism, interactive approaches like the *snake pit* [Kass *et al.*, 1987b] can be used, providing an interactive mechanism for defining pushing and pulling forces in the image scene via *spring* and *volcano* forces. The classic model is based on a spline with controlled continuity [Terzopoulos, 1986b], providing piecewise smoothness constraints as internal spline forces and thus regularizing the deformation of the model in terms of its elasticity and bending. The representation of the classic active contour model, however, is not spline-based during the deformation process, but only for the final interpolation of the result. The image forces push the model towards salient image features such as lines or edges, and the external constraint forces are responsible for pulling the model near a desired local energy minimum using appropriate user interaction, automatic attentional mechanisms, or high-level interpretation. During a following optimization process which was originally formulated within a *Euler-Lagrangian* setting for the classic model, the internal, and external image and constraint forces are adjusted by higher level processes to find the desired local optimum causing a suitable deformation of the active contour model. Scale-space continuation, as presented in [Witkin, 1983; Witkin *et al.*, 1987], can be incorporated into the deformation process in order to enlarge the capture region of image features, and to perform a *coarse-to-fine* tracking of image features.

In the following, the theoretical framework for the classic active contour model as well as its recent developments in terms of the representation, design of an energy function and different optimization techniques will be briefly summarized.

4.1 Representation

An active contour model is based on a parameterized contour $\mathbf{v}(s)$, but is normally represented by a discrete set of points or *snaxels*$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, with $\mathbf{v}_i = (x_i, y_i)$. Closed contours are obtained by making the contour periodic, e.g. by setting $\mathbf{v}(0) = \mathbf{v}(1)$ in the parametric form or $\mathbf{v}_1 = \mathbf{v}_{n+1}$ in the discrete form. Each *snaxel* has two neighbours (in the closed case), and otherwise at least one neighbour. An energy function is formulated to obtain an estimate of the quality of the model in terms of its internal (autonomous) shape, and external forces, e.g. underlying image forces and user-constraint forces. The energy function integrates a weighted linear combination of the internal and external forces over the spline contour:

$$\mathcal{E} = \int_0^1 (\mathcal{E}_{internal}(\mathbf{v}(s)) + \mathcal{E}_{image}(\mathbf{v}(s)) + \mathcal{E}_{constraint}(\mathbf{v}(s)))\,\mathrm{d}s, \tag{4.1}$$

This energy function can also be regarded as the compromise between internal and external contour shape quality. Moving the snaxels leads to a change in energy, which transforms the segmentation problem into an optimization task. The continuous energy function $\mathcal{E}$ is usually discretized by replacing the integrals by summation, leading to a discrete energy function $\mathcal{E}^*$.

Extending an active contour model to a three-dimensional active surface model is straightforward, yet computationally complex. An active surface which is parameterized by (s, r), with arc length parameters $s \in [0; 1], r \in [0; 1]$, can be represented as $\mathbf{v}(s, r) = ((x(s, r), y(s, r), z(s, r))$ with coordinate functions x, y, z, and the energy function is given by

$$\mathcal{E}_{snake} = \int_0^1 \int_0^1 (\mathcal{E}_{internal}(\mathbf{v}(s, r)) + \mathcal{E}_{image}(\mathbf{v}(s, r)) + \mathcal{E}_{constraint}(\mathbf{v}(s, r)))\,\mathrm{d}s\,\mathrm{d}r\,, \tag{4.2}$$

This dissertation concentrates on $2D$ models, and briefly discusses the main developments of $3D$ models in section 4.4.4.

4.2 Energy Function

The classic formulations of internal, and external image and constraint forces as well as more recent formulations and developments will be summarized in the following.

4.2.1 Internal Energy Terms

$\mathcal{E}_{internal}$ represents the internal energy of the contour with respect to elastic deformations and the bending of the snake:

$$\begin{aligned}\mathcal{E}_{internal}(\mathbf{v}(s)) &= \alpha_{elasticity}(s)\ \mathcal{E}_{elasticity}(\mathbf{v}(s)) + \alpha_{bending}(s)\ \mathcal{E}_{bending}(\mathbf{v}(s)) \\ &= \alpha_{elasticity}(s)\ \|\mathbf{v}_s(s)\|^2 + \alpha_{bending}(s)\ \|\mathbf{v}_{ss}(s)\|^2 . \end{aligned} \tag{4.3}$$

The first order derivative term, $\mathbf{v}_s(s)$, makes the snake behave like a membrane and represents the elastic energy of the contour, and the second order derivative term, $\mathbf{v}_{ss}(s)$, makes the snake act like a thin plate and represents the contour's bending energy. Decreasing $\alpha_{elasticity}$ allows the contour to develop gaps, while increasing $\alpha_{elasticity}$ increases the *tension* of the model by reducing its length. Decreasing $\alpha_{bending}$ allows the active contour model to develop corners, and increasing $\alpha_{bending}$ increases the bending *rigidity*, making the model smoother and less flexible. Setting either of the weighting coefficients to zero permits first and second order discontinuities, respectively.

Both [Kass *et al.*, 1987b] and [Amini *et al.*, 1990] have suggested to approximate the derivatives in equation (4.3) by finite differences. The first order elasticity term then becomes

$$\begin{aligned}\mathcal{E}^*_{elasticity}(\mathbf{v}_i) &\approx \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^2 \\ &= (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \end{aligned} \tag{4.4}$$

This elasticity term minimizes the distance between the snake points, causing the active contour model to shrink during the optimization process in absence of appropriate external image or constraint forces. Analogously, the second order term minimizing the bending of the active contour model can be discretely approximated by

$$\begin{aligned}\mathcal{E}^*_{bending}(\mathbf{v}_i) &\approx \|\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}\|^2 \\ &= (x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2 \end{aligned} \tag{4.5}$$

[Williams and Shah, 1992] have pointed out that equation (4.4) is made under the assumption that the snaxels of the active contour model are evenly spaced. As this might not be always the case, they have proposed to subtract the continuity term from the average distance $\|\bar{\mathbf{d}}\|$ of the snaxels, as otherwise the energy expression will be larger for points which are farther apart. This hard constraint forces the points to be more evenly spaced, and avoids a possible contraction of the

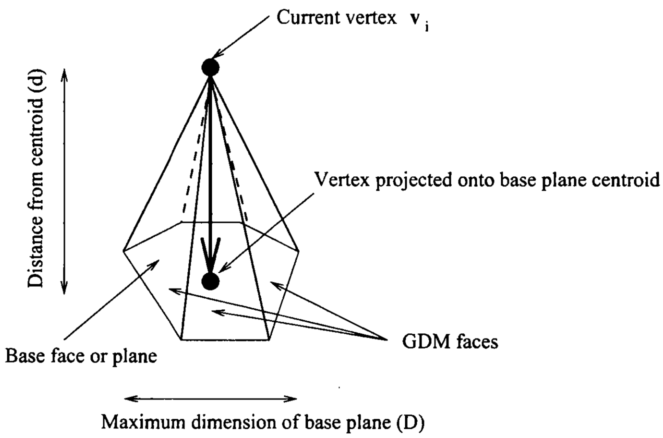


Figure 4.1: Local topology constraint of a solid model cut out of a GDM. The ratio between the distance d between current vertex and the centroid of its neighbours, and the maximum dimension D of the base plane gives an estimate of the local vertex curvature.

snake:

$$\mathcal{E}_{elasticity}(\mathbf{v}_i) \approx \bar{d} - \left((x_i - x_{i-1})^2 + (y_i - y_{i-1})^2\right) \tag{4.6}$$

Alternatively to the *backward* difference, a *forward* difference $||\mathbf{v}_i - \mathbf{v}_{i+1}||^2$ or a *centred* difference $\frac{||\mathbf{v}_{i-1} - \mathbf{v}_{i+1}||^2}{2}$ can be taken. Moreover, they have argued that the discrete approximation of the quantity $||\mathbf{v}_{ss}|| = \sqrt{\ddot{x}^2(s) + \ddot{y}^2(s)}$ only measures the curvature if the active contour model is arc length parameterized, otherwise it is given by

$$\kappa_i = \frac{||\dot{x}_i \ddot{y}_i - \ddot{x}_i \dot{y}_i||}{(\dot{x}_i^2 + \dot{y}_i^2)^{3/2}} \tag{4.7}$$

as presented in section 2.3. Besides the discrete bending approximations of equations (4.5) and (4.7) Williams and Shah have investigated three more curvature measurements: The mathematical formulation of curvature given by equation (2.25) defines curvature as the rate of change of the angle θ of the curve tangent. This measurement depends linearly on the angle $\Delta\theta$ between the vectors $\mathbf{u}_i = \mathbf{v}_i - \mathbf{v}_{i-1}$ and $\mathbf{u}_{i+1} = \mathbf{v}_{i+1} - \mathbf{v}_i$. Thus, a discrete approximation of bending energy can be obtained by

$$\left(\frac{\mathrm{d}\theta}{\mathrm{d}s}\right)^2 \approx \left(\frac{\Delta\theta}{\Delta s}\right)^2 \quad \text{with} \quad \Delta\theta = cos^{-1}\left(\frac{\mathbf{u}_i \cdot \mathbf{u}_{i+1}}{||\mathbf{u}_i||\ ||\mathbf{u}_{i+1}||}\right) \quad \text{and} \quad \Delta s = \frac{||\mathbf{u}_{i+1}|| + ||\mathbf{u}_i||}{2} \tag{4.8}$$

A computationally more efficient way is to compute $||\mathbf{u}_{i+1} - \mathbf{u}_i||^2$ which reflects the difference of direction between two contour arcs as well as their difference in length. Finally, the two vectors can be normalized which removes the length differential, making the measurement solely dependent on the direction information. The bending energy is computed as $(\mathbf{u}_{i+1}/||\mathbf{u}_{i+1}|| - \mathbf{u}_i/||\mathbf{u}_i||)^2$, a concept which has also been adopted in [Lobregt and Viergever, 1995].

Finally, [Miller *et al.*, 1991] have developed an internal topological constraint term for their *geometrically deformed model* (GDM) whose concept will be explained later in section 4.4.1. This constraint term enforces surface continuity and hence topological integrity, endangered by incomplete object boundaries and noise. *Vertices* are associated with the contour points, which are connected to their neighbours with edges (see figure 4.1). Measuring the distance between each vertex $\mathbf{v}_i$ and the centroid of the base plane formed by its neighbours, and computing the ratio between this distance and the maximum dimension of the base plane gives an estimate of the local vertex curvature:

$$\mathcal{E}_{topology}(\mathbf{v}_i) = \frac{\|\mathbf{v}_i - \frac{1}{m}\sum_{j=1}^{m}\mathbf{v}_j\|}{\max_{j,k}(\|\mathbf{v}_j - \mathbf{v}_k\|)} , \tag{4.9}$$

where m is the number of neighbours $\mathbf{v}_j, \mathbf{v}_k$ of the current point. This topology constraint is scale invariant and enforces the vertex to move onto the plane formed by its neighbours, defaulting to a spherical mesh in absence of other constraints. Note that this topology measurement is formulated for a higher dimensional model, e.g. a surface model, but it can also be applied to a contour model (which reduces the neighbours of $\mathbf{v}_i$ to $\mathbf{v}_{i-1}$ and $\mathbf{v}_{i+1}$).

4.2.2 Image Energy Terms

The external image energy term $\mathcal{E}_{image}$ represents the energy due to image forces like lines, edges and terminations of line segments and corners. In [Kass *et al.*, 1987b], the following image terms are suggested, consisting of a weighted sum of the terms

$$\mathcal{E}_{image}(\mathbf{v}(s)) = \alpha_{line}(s)\mathcal{E}_{line}(\mathbf{v}(s)) + \alpha_{edge}(s)\mathcal{E}_{edge}(\mathbf{v}(s)) + \alpha_{term}(s)\mathcal{E}_{term}(\mathbf{v}(s)) \tag{4.10}$$

The simplest useful image functional is the image intensity or luminance function L. In the discrete case,

$$\mathcal{E}_{line}(\mathbf{v}_i) = L(\mathbf{v}_i) , \tag{4.11}$$

attracting the active contour model either to dark or to light lines depending on the sign of the weighting coefficient α_{line}. The simplest, discrete edge functional can be used by setting

$$\mathcal{E}_{edge}(\mathbf{v}_i) = -\|\nabla L(\mathbf{v}_i)\|^2 , \tag{4.12}$$

where the negative sign produces low energy values for high gradient values. Squaring the gradient narrows the edge response. In order to find terminations of line segments and corners, [Kass *et al.*, 1987b] have proposed to use the curvature of level lines in a slightly smoothed image, corresponding to the isophote image curvature presented in chapter 3. Let $L(\sigma) = G_\sigma \otimes L$ be a slightly smoothed version of the image, and let $\theta_i = tan^{-1}(L_y(\mathbf{v}_i)/L_x(\mathbf{v}_i))$ be the gradient angle at snaxels $\mathbf{v}_i$, and let $\mathbf{T}_i = (\cos\theta_i, \sin\theta_i)$ and $\mathbf{N}_i = (-\sin\theta_i, \cos\theta_i)$ be the unit normal and tangent vectors to the image isophotes at the snaxel. Then the discrete curvature of the level contours in L can be written as

$$\mathcal{E}_{term}(\mathbf{v}_i) = \frac{\partial\,\theta_i}{\partial\,\mathbf{T}_i}$$

$$= \frac{\partial^2 L(\mathbf{v}_i)/\partial\, \mathbf{T}_i^2}{\partial L(\mathbf{v}_i)/\partial\, \mathbf{N}_i}$$
$$= \frac{2L_x(\mathbf{v}_i)L_y(\mathbf{v}_i)L_{xy}(\mathbf{v}_i) - L_x^2(\mathbf{v}_i)L_{yy}(\mathbf{v}_i) - L_y^2(\mathbf{v}_i)L_{xx}(\mathbf{v}_i)}{(L_x^2(\mathbf{v}_i) + L_y^2(\mathbf{v}_i))^{\frac{3}{2}}} \tag{4.13}$$

By combining E_{edge} and E_{term}, the snake is attracted to edges or terminations. However, as the isophote image curvature is a signed measurement of the bending behaviour, an absolute value should rather be chosen for E_{term}. As an alternative to the line functional, [Miller *et al.*, 1991] have used a simple event detector such as a discrete weighted threshold operator of the image intensity:

$$\mathcal{E}_{threshold}(\mathbf{v}_i) = \begin{cases} 0 & \text{if } L(\mathbf{v}_i) < \text{ threshold} \\ L(\mathbf{v}_i) - \text{ threshold} & \text{otherwise} \end{cases} \tag{4.14}$$

Finally, [Cohen and Cohen, 1993] have suggested to use the distance transform of the zero-crossings of the Laplacian ΔL as an edge potential.

4.2.3 External Constraint Terms

The classic snake model incorporates user constraints, allowing the user to attach springs between points of the contour and fixed positions in the image plane:

$$\mathcal{E}_{spring}(\mathbf{v}_i) = -\alpha_{spring}(\mathbf{v}_i - \mathbf{x})^2 \tag{4.15}$$

This term attracts contour point $\mathbf{v}_i$ to a point $\mathbf{x}$ in the plane, with α_{spring} as the *spring constant.* Depending on the sign of α_{spring}, the active contour model is attracted by the spring or repelled, in which case the reverse effect of a volcano force takes place. In [Kass *et al.*, 1987b], a similar constraint force for matching stereo contour models has been suggested:

$$\mathcal{E}_{stereo}(\mathbf{v}_i^{L,R}) = \alpha_{stereo}(\mathbf{v}_i^L - \mathbf{v}_i^R)^2 \;, \tag{4.16}$$

where $\mathbf{v}^L$ and $\mathbf{v}^R$ represent the left and right contours of a stereo pair, which couples the left and the right snake and enforces the disparity to vary slowly along the contour. Additionally, a *balloon* or inflation force as presented by [Cohen, 1990; Cohen, 1995] can be used to expand or contract the active contour model in lack of other forces until it is locked:

$$\mathcal{E}_{balloon}(\mathbf{v}_i) = \alpha_{balloon}\mathbf{n}_i \;, \tag{4.17}$$

where $\mathbf{n}_i$ is a normal unitary vector at $\mathbf{v}_i$, enforcing an expansion of the contour point in direction of its normal. The weighting factor $\alpha_{balloon}$ for the balloon force should be chosen slightly smaller than the weighting factors for the other energy terms in order to allow the active contour model to stop in the presence of these forces. For coefficients of the same order, for example, weak edges will be surpassed, while strong edges will stop the expansion of the model. If $\alpha_{balloon}$ changes its sign, the effect will be deflation instead of inflation.

4.2.4 Region Constraint Terms

The classic active contour model is based only on the boundary characteristics of a shape, disregarding the enclosed region pixels. Several approaches have been developed to constrain the model's shape to its enclosed region homogeneity, the most important being *statistical snakes* or *active region models* by [Ivins and Porrill, 1993b; Ivins and Porrill, 1994b; Ivins and Porrill, 1994a] who also have also presented an extensive overview about existing active contour model techniques [Ivins and Porrill, 1993a]. Active region models start from a user-defined homogeneous *seed region* (or template region) whose mean μ and variance σ are computed, and then grows with the help of an inflation or *pressure* force until it encounters pixels whose intensities change the variance of the region's intensity significantly. The discrete pressure force is defined by:

$$\mathcal{E}_{pressure}(\mathbf{v}_i) = \mathbf{n}_i \left(\frac{\|L(\mathbf{v}_i) - \mu\| - k\sigma}{k\sigma} \right)^2 , \tag{4.18}$$

where k is the constant defining the significance of a change in variance, and the pressure force is normalized by the scaling term $(k\sigma)^2$. This scheme is equivalent to weighting the balloon force by the mean pixel intensity at each boundary pixel. A similar approach has been developed by [Bascle and Deriche, 1995], who have proposed to use a normalized correlation criterion, measuring the differences of grey-levels in the current region and a *template* region. An alternative approach has been developed by [Poon *et al.*, 1994a; Poon *et al.*, 1994b] who have suggested to incorporate a total region energy term into the energy functional by the use of a discriminant function D_{region} based on the intra-region variance and a region-related constraint H_{region} imposed by the user. [Chakraborty and Duncan, 1995; Chakraborty *et al.*, 1996] have integrated region homogeneity information into a deformable model using a Bayesian framework, and [Zhu and Yuille, 1996] have a developed a *region competition* algorithm combining aspects of statistical region growing and balloon models.

4.2.5 Combinations of Internal and Image Energy Terms

The classic active contour model strictly distinguishes between internal, *autonomous* contour forces, and external image and other constraint forces, making the contour model intrinsically *non-adaptive* with respect to the underlying image data. This leads to several problems: for example, incorporating the contour curvature term into the energy functional enforces a minimization of the contour curvature which might not always be desirable for objects having parts of high curvature. Choosing adaptive weighting parameters, $\alpha_{elasticity}(s)$ and $\alpha_{bending}(s)$, which model adequate material densities along the contour $\mathbf{v}(s)$, seems to be the appropriate solution to this problem, yet introduces even more complex problems regarding the derivation of such parameters. Several researchers have therefore suggested to relate the elasticity and curvature properties of the contour to the underlying differential image structure.

[Cohen and Cohen, 1993] have used local implicit expressions for the material densities depending on the intrinsic metric (the inverse contour curvature) as well as on first and second order variations of the image gradient potential along the model. The optimal elasticity and bending densities are obtained by minimizing the segmentation error, defined for a given precision $p \in \Re^+$ as the residual $(\|\nabla L\|^2 - p)^2$. In order to obtain these densities, the energy of the model is minimized using constant material densities $\alpha_{elasticity_0}, \alpha_{bending_0}$, and then the model is refined with the optimal values. The material densities are therefore always of same magnitude order as the edge potential.

[Rougon and Prêteux, 1993a; Rougon and Prêteux, 1993b] have reviewed previously developed bending densities, such as (in the continuous case)

$$\alpha_{bending}(s) = \frac{\alpha_{bending_0}}{(1 + \|\mathbf{v}_{ss}\|^2)^3} \tag{4.19}$$

and

$$\alpha_{bending}(s) = \frac{\alpha_{bending_0}}{(1 + \|\nabla L(\mathbf{v}(s))\|^2)^P}\,, \tag{4.20}$$

where the bending weighting coefficient in equation (4.19) is purely geometric and data-independent, with $\alpha_{bending}(s)$ approaching 0 for high curvature points, and a constant value $\alpha_{bending_0}$ whenever the tangent of the contour varies slowly. The bending coefficient in equation (4.20) is data-dependent, approaching 0 at high image gradient magnitude values, and $\alpha_{bending_0}$ in homogeneous image areas [Samadani, 1991].

[Williams and Shah, 1992] have suggested a combination of these two schemes, by adjusting the bending weighting term $\alpha_{bending}(s)$ with respect to the local bending energy value $\mathcal{E}_{bending}(\mathbf{v}(s))$ and the local gradient magnitude $\|\nabla L(\mathbf{v}(s))\|$. If these values exceed the predefined thresholds $thres_{bending}$ and $thres_{magnitude}$, $\alpha_{bending}(s)$ is set to zero for the next iteration, providing a higher level feedback to the energy optimization process and preventing the formation of corners until the contour is sufficiently close to an edge.

Starting from these methods, [Rougon, 1993] has investigated the intrinsic geometric differential properties of deformable models, and [Rougon and Prêteux, 1993b] have extended this approach by introducing *oriented anisotropic adaptive constraints* relating the differential properties of the internal constraints of the model to those of the underlying image intensity. Using appropriate, image related weighting tensors results in weakening the internal constraints within image regions where the image surface becomes non-stationary, and inhibiting them in case of discontinuities of the image surface. The continuous formulation of the internal energy term is

$$\mathcal{E}_{intern}(\mathbf{v}(s)) = \int_0^1 \mathbf{v}_s^T(s) W_1 \mathbf{v}_s(s) \mathrm{d}s + \int_0^1 \mathbf{v}_{ss}^T(s) W_2 \mathbf{v}_{ss}(s) \mathrm{d}s \tag{4.21}$$

with weighting tensors W_1 and W_2 for the elasticity and bending constraints, respectively. W_1 is

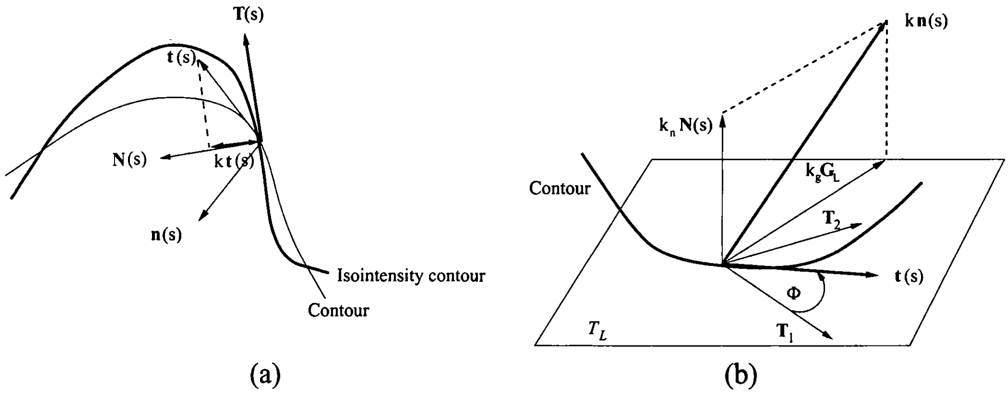


Figure 4.2: Geometric interpretation of elasticity and bending constraints, adapted from [Rougon and Prêteux, 1993b]. (a) The quantity $k\mathbf{t}(s)$ of the contour tangent vector along the normal $\mathbf{N}(s)$ of an iso-intensity contour is minimized. (b) The component $k_g \mathbf{G}_L$ of the vector $k\mathbf{n}$ within the tangent plane $\mathcal{T}_L$ is minimized along the principal direction $\mathbf{T}_1$ of the intensity surface.

given by

$$W_1(s) = \alpha_{elasticity}(s) \nabla L \cdot \nabla L^T = \alpha_{elasticity}(s) \begin{pmatrix} L_x^2 & L_x L_y \\ L_x L_y & L_y^2 \end{pmatrix}, \tag{4.22}$$

aligning the normal $\mathbf{n}(s)$ of the contour $\mathbf{v}(s)$ to the normal $\mathbf{N}(s)$ of the underlying image isophotes, depending on first-order partial derivatives of the image intensity L evaluated at $\mathbf{v}(s)$. W_2 is defined as

$$W_2(s) = \alpha_{bending}(s) \mathbf{H}(L)^2 = \alpha_{bending}(s) \begin{pmatrix} L_{yy}^2 + L_{xy}^2 & -L_{xy}(L_{xx} + L_{yy}) \\ -L_{xy}(L_{xx} + L_{yy}) & L_{xx} + L_{xy}^2 \end{pmatrix}, \tag{4.23}$$

which is based on the squared Hessian matrix of second derivatives $\mathbf{H}$. This tensor enforces alignment of the contour normal $\mathbf{n}(s)$ to the minimum and maximum principal directions of the image, which are determined by the unit tangent of the contour $\mathbf{t}(s)$, and the unit normal $\mathbf{N}(s)$ of the image isophotes and the geodesic normal vector $\mathbf{G}(s)$ to the contour, where $\mathbf{G}(s) = \mathbf{N}(s) \wedge \mathbf{t}(s)$. This relates curvature properties along the contour with curvature properties in L. Figure 4.2 illustrates this concept. This approach satisfies three important conditions: consistency (first-order metric contour properties are related to first-order partial derivatives of the image, and second-order curvature contour properties are related to curvature properties of the image, expressed by second-order partial derivatives), positive definiteness (the tensors W_1 and W_2 are constrained to be quadratic with respect to first- and second-order partial derivatives of the image, respectively), and rigid invariance (the internal energy terms are invariant with respect to coordinate transformations). Note that this adaptive constraint scheme is made under the assumption that the contour is parameterized by arc length, and that the classic active contour model is obtained from this framework by multiplying the parameters $\alpha_{elasticity}(s)$ and $\alpha_{bending}(s)$ with the two-dimensional identity matrix $\mathbf{I}$.

4.2.6 Scale-Space Incorporation

For the classic model it was suggested to incorporate scale-space continuity into the image energy functional in order to enlarge the capture region of image features. Spatially smoothing of the edge or line energy terms enables the active contour model to come to equilibrium on a very blurry energy functional, thus getting attracted to very distant and prominent edges. By slowly reducing the blurring, similar to Bergholm's edge-focusing approach presented in the previous chapter, section 3.4.2, a finer adjustment of the active contour model can be obtained. This is equivalent to minimization by scale continuation as proposed in [Witkin, 1983; Witkin *et al.*, 1987]. Following the Marr-Hildreth theory of edge detection [Marr and Hildreth, 1980], the Laplacian-of-Gaussian can be chosen as an appropriate, discrete edge term [Kass *et al.*, 1987b]:

$$\mathcal{E}_{edge}(\mathbf{v}_i) = (G(\mathbf{v}_i; \sigma) \otimes \Delta L(\mathbf{v}_i))^2 \tag{4.24}$$

where G is the Gaussian function with standard deviation σ. As minima of the smoothed image gradient $G \otimes \nabla L$ lie on zero-crossings of $G \otimes \Delta L$, using this term as an edge functional results in attracting the active contour model to zero-crossings.

4.3 Optimization

Active contour models are *active* because the minimization of its energy functional causes the model to change dynamically. The slithering movement of the contour during the minimization process is the reason why they are also called *snakes*. The deformation of the active contour is controlled by an optimization or energy minimization process. In the following, the classic optimization technique and other energy minimizing techniques applied for active contour models are presented.

4.3.1 Variational Approach

The classic model is embedded in a Euler-Lagrangian setting, using variational calculus in order to derive a differential equation solved by an iterative minimization technique using sparse matrix methods. Each iteration performs implicit Euler forward steps with respect to the internal energy, and explicit Euler steps with respect to the external image and constraint energy terms, yielding a semi-implicit method which allows to travel the entire length of the model in a single $O(n)$ iteration. The minimization process remains stable in the presence of very large internal forces, but in the presence of large external and image forces, small time steps must be taken. Let $\mathcal{E}_{extern} = \mathcal{E}_{image} + \mathcal{E}_{constraint}$. The contour which minimizes the total snake energy must satisfy the following two independent Euler-Lagrange equations:

$$\alpha_{elasticity}(s)x_{ss}(s) + \alpha_{bending}(s)x_{ssss}(s) + \frac{\partial \mathcal{E}_{extern}(\mathbf{v}(s))}{\partial x} = 0$$

$$\alpha_{elasticity}(s)y_{ss}(s) + \alpha_{bending}(s)y_{ssss}(s) + \frac{\partial \mathcal{E}_{extern}(\mathbf{v}(s))}{\partial y} = 0 \tag{4.25}$$

The internal energy terms are discretely approximated as in equations (4.4) and (4.5). Let $f_x(s) = \partial \mathcal{E}_{extern}(\mathbf{v}(s))/\partial x$ and $f_y(s) = \partial \mathcal{E}_{extern}(\mathbf{v}(s))/\partial y$ which is discretely approximated by $f_x(i) = \partial \mathcal{E}_{extern}(\mathbf{v}_i)/\partial x_i$ and $f_y(i) = \partial \mathcal{E}_{extern}(\mathbf{v}_i)/\partial y_i$ unless analytical derivatives exist. The corresponding Euler equations are

$$\begin{aligned}
\alpha_{elasticity_i}(\mathbf{v}_i - \mathbf{v}_{i-1}) - \alpha_{elasticity_{i+1}}(\mathbf{v}_{i+1} - \mathbf{v}_i) &\quad + \\
\alpha_{bending_{i-1}}(\mathbf{v}_{i-2} - 2\mathbf{v}_{i-1} + \mathbf{v}_i) &\quad - \\
2\,\alpha_{bending_i}(\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}) &\quad + \\
\alpha_{bending_{i+1}}(\mathbf{v}_i - 2\mathbf{v}_{i+1} + \mathbf{v}_{i+2}) &\quad + \\
f_x(i) + f_y(i) &\quad = \quad 0
\end{aligned} \tag{4.26}$$

These equations can be written in matrix form as

$$\begin{aligned}
\mathbf{A}\cdot\mathbf{x} + \mathbf{f_x}(\mathbf{x},\mathbf{y}) &= 0 \\
\mathbf{A}\cdot\mathbf{y} + \mathbf{f_y}(\mathbf{x},\mathbf{y}) &= 0
\end{aligned} \tag{4.27}$$

where $\mathbf{A}$ is a pentadiagonal banded matrix. To solve equations (4.27) by successive over-relaxation, the right hand side of the equations is set to the product of a step size γ and the negative time derivative of the left hand sides. Assuming that $\mathbf{f_x}$ and $\mathbf{f_y}$ are constant during one time step cuts down the computational cost of otherwise changing $\mathbf{A}$ every iteration. Thus the Euler method is explicit with respect to the external forces, but implicit for the internal forces as they are completely specified by the banded matrix. The resulting equations for evaluating the time derivative at time (t) rather than at time $(t-1)$ are

$$\begin{aligned}
\mathbf{A}\cdot\mathbf{x}^{(t)} + \mathbf{f_x}\left(\mathbf{x}^{(t-1)},\mathbf{y}^{(t-1)}\right) &= -\gamma\left(\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\right) \\
\mathbf{A}\cdot\mathbf{y}^{(t)} + \mathbf{f_y}\left(\mathbf{x}^{(t-1)},\mathbf{y}^{(t-1)}\right) &= -\gamma\left(\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\right)
\end{aligned} \tag{4.28}$$

At equilibrium, the time derivative vanishes and the Euler equations (4.27) are solved. Equations (4.28) can then be solved by matrix inversion:

$$\begin{aligned}
\mathbf{x}^{(t)} &= (\mathbf{A}+\gamma\mathbf{I})^{-1}\cdot\left(\mathbf{x}^{(t-1)} - \mathbf{f_x}\left(\mathbf{x}^{(t-1)},\mathbf{y}^{(t-1)}\right)\right) \\
\mathbf{y}^{(t)} &= (\mathbf{A}+\gamma\mathbf{I})^{-1}\cdot\left(\mathbf{y}^{(t-1)} - \mathbf{f_y}\left(\mathbf{x}^{(t-1)},\mathbf{y}^{(t-1)}\right)\right)
\end{aligned} \tag{4.29}$$

As the matrix $\mathbf{A}+\gamma\mathbf{I}$ is also a pentadiagonal banded matrix, its inverse can be computed in $O(n)$ time. However, this variational approach does not guarantee global optimality of the solution, and requires estimates of higher order derivatives of the discrete data. Moreover, hard constraints (which are a restriction on the range of $\mathbf{v}(s)$ or its derivatives) cannot be directly enforced, unless the constraints are differentiable, in which case higher dimensional spaces are required for more unknowns. Given a desired constraint term like a mean or minimum *snaxel* spacing, it can only be enforced by increasing the associated weighting term, which will force more effect on this constraint, but on the cost of other terms. Another disadvantage of the variational approach in the context of active contour models is the numerical instability with respect to the explicit part of the Euler forward step, and the tendency for points to bunch up on a strong portion of an edge.

4.3.2 Dynamic Programming

[Amini *et al.*, 1990] have proposed dynamic programming (DP) as an approach to solving variational problems in vision with a special application to energy-minimizing active contour models. In contrast to the Euler-Lagrangian setting presented in the previous section, their approach allows to enforce hard constraints on the behaviour of the solution directly and naturally, ensuring a globally optimal solution with respect to the search space, and numerical stability by moving the contour points on a discrete grid without any approximization requirements. However, their method is rather slow and they have large memory requirements of $O(nm^2)$ and time complexity of $O(nm^3)$, where n is the number of points on the contour and m is the number of potential locations to which every point can move during one optimization step (often referred to as *neighbourhood size*). The optimization problem is viewed as a discrete multi-stage decision process and is solved by a *time-delayed* discrete dynamic programming algorithm. Dynamic programming bypasses local minima as it is embedding the minimization problem in a family of related problems. This is achieved by replacing the minimization of the total energy measure by the problem of minimizing a function of the form

$$\mathcal{E}(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n) = \mathcal{E}_1(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) + \mathcal{E}_2(\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) + \cdots \mathcal{E}_{n-2}(\mathbf{v}_{n-2}, \mathbf{v}_{n-1}, \mathbf{v}_n) \tag{4.30}$$

where each variable is allowed to take only m possible values and

$$\mathcal{E}_{i-1}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) = \mathcal{E}_{intern}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) + \mathcal{E}_{extern}(\mathbf{v}_i) \tag{4.31}$$

with

$$\mathcal{E}_{intern}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) = \mathcal{E}_{elasticity}(\mathbf{v}_{i-1}, \mathbf{v}_i) + \mathcal{E}_{bending}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) \tag{4.32}$$

where $\mathcal{E}_{elasticity}$ and $\mathcal{E}_{bending}$ are computed as in equations (4.4) and (4.5). In order to apply dynamic programming to equation (4.31), a two element vector of state variables, $(\mathbf{v}_{i+1}, \mathbf{v}_i)$, is fixed and a recurrent optimal value function based on two adjacent points is formulated:

$$S_i(\mathbf{v}_{i+1}, \mathbf{v}_i) = \min_{\mathbf{v}_{i-1}} S_{i-1}(\mathbf{v}_i, \mathbf{v}_{i-1}) + \mathcal{E}_{intern}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) + \mathcal{E}_{extern}(\mathbf{v}_i) \tag{4.33}$$

Apart from the energy matrix corresponding to the optimal value function S_i, a position matrix is also needed so the value of $\mathbf{v}_i$ minimizing equation (4.33) can be stored. The optimal contour

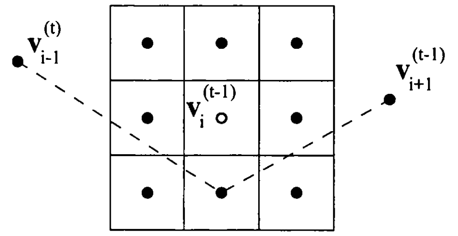


Figure 4.3: *Greedy* optimization and local neighbourhood search. The energy function is evaluated at $\mathbf{v}_i^{(t-1)}$ and each of its eight neighbours, using the points $\mathbf{v}_{i-1}^{(t)}$ and $\mathbf{v}_{i+1}^{(t-1)}$ for computing the internal terms. The location with the lowest energy is chosen to be the new position $\mathbf{v}_i^{(t)}$.

of minimum energy $\mathcal{E}_{min}(s)$ can be found by *backtracking* in the position matrix, with

$$\mathcal{E}_{min}(s) = \min_{\mathbf{v}_n} S_{n-1}(\mathbf{v}_n) \, . \tag{4.34}$$

This process is iterated until the total energy does not change significantly any longer, where one iteration consists of a forward pass, deriving the minimal energy values of each $\mathbf{v}_i$, and a backward path, finding the minimum energy path in the position matrix. With dynamic programming, hard constraints limit the set of potential solutions and reduce the computational complexity and cost. The concept of dynamic programming for solving variational problems has been extended to two dimensions in [Amini *et al.*, 1995b; Amini *et al.*, 1995a].

4.3.3 *Greedy* Algorithm

The *greedy* algorithm developed by [Williams and Shah, 1992] is a very stable, fast and flexible optimization technique for active contour models. It allows to incorporate hard constraints as described in [Amini *et al.*, 1990], while having a much lower computational complexity of $O(nm)$ for a contour of n points which are allowed to move in a neighbourhood of size m. Unlike dynamic programming, this algorithm does not guarantee to find the global minimum within its search space, but has proven to be a good and efficient optimization technique. The key point of this algorithm is the approximation of the first-order energy term as given in equation (4.6) enforcing even point spacing, and the different curvature measurements as discussed above (section 4.2.1). Additionally, several normalization and reparameterization strategies are employed, including the local normalization of the energy values with respect to the local search space, and adjustment of the weighting parameter $\alpha_{bending}(s)$ with respect to a bending and gradient magnitude threshold (section 4.2.5).

Figure 4.3 illustrates the local neighbourhood search of the *greedy* algorithm. The energy function for the current location of $\mathbf{v}_i^{(t-1)}$ and each of its neighbours is computed under consideration of the adjacent contour points $\mathbf{v}_{i-1}^{(t)}$ and $\mathbf{v}_{i+1}^{(t-1)}$. The location with the smallest energy value is chosen as the new position $\mathbf{v}_i^{(t)}$. Note that $\mathbf{v}_{i-1}^{(t)}$ has already been updated to its new position in the current iteration over the contour, while $\mathbf{v}_{i+1}^{(t-1)}$ will be updated next. In the first stage of the algorithm, all contour points are sequentially updated within one iteration. In the second stage the forming of corners is determined by recomputing the bending energy terms with the updated points, and adjusting the weighting parameter $\alpha_{bending_i}$ for each contour point accordingly.

This otherwise very efficient and straight forward technique introduces several problematic issues regarding the sequentiality of the optimization, and the choice of appropriate thresholds, weighting adjustments for corner formations, and normalization strategies, none of which can be easily resolved. However, a change of the energy functional does not directly affect the optimization strategy, which makes this algorithm a very flexible and efficient local optimization technique for problem-oriented tasks. Modifications of this algorithm will be presented in chapter 6 of this dissertation.

4.3.4 Stochastic and Probabilistic Relaxation

Simulated annealing (SA) is a stochastic relaxation technique which is based on the physical process of annealing a metal: At high temperatures the atoms are randomly distributed. With decreasing temperatures they arrange themselves in a crystalline state minimizing their energy. This model has been successfully used for global optimization purposes [Geman and Geman, 1984]. For active contours, the associated energy functional can be optimized similarly: Assuming that the position of a contour control point depends only on itself and its direct neighbouring control points, the active contour model can be regarded as a $1D$ Markov random field (MRF). The SA algorithm generates new configurations of the contour points in a defined neighbourhood sampling from a Gibbs probability distribution of the MRF given by

$$P(X = \omega) = \frac{1}{Z(T)} e^{-\frac{\mathcal{E}(\omega)}{T}}, \tag{4.35}$$

where $Z(T)$ is a normalization factor and T is a control parameter, called temperature, which influences the form of the probability distribution. New configurations are accepted with a certain acceptance probability $H(T)$ depending on the temperature:

$$H(T) = e^{-\frac{\Delta\mathcal{E}}{T}} \tag{4.36}$$

Since increases of energy can be accepted, the algorithm is able to escape local energy minima. [Geman and Geman, 1984] have shown that the algorithm converges to a global energy minimum, if the temperature at iteration k is

$$T(k) \geq \frac{c}{\log(1+k)} \tag{4.37}$$

where c is a constant. Active contour optimization using Simulated Annealing as a Bayesian approach has been first presented by [Rueckert, 1993; Toennies and Rueckert, 1994]. Additionally, SA has been applied to active contour models in [Storvik, 1994; Grzeszczuk and Levin, 1994]. It has been pointed out in [Rueckert and Burger, 1995a] that stochastic relaxation optimization techniques like SA make high computational demands. Iterated conditional modes (ICM) [Besag, 1986] is a probabilistic relaxation technique which makes deterministic instead of random changes by maximizing the conditional probability based on a provisional estimate $\hat{\mathbf{v}}$. For active contours, each control point $\mathbf{v}_i$ can be replaced by a point in a defined neighbourhood by maximizing the conditional probability of ω_i. Modelling the active contour as a $1D$ MRF, the probability $P(\omega_i|\hat{\mathbf{v}})$ is again given by the Gibbs distribution (equation (4.35)). ICM is equivalent to SA with *instantaneous freezing* and therefore converges much faster. In contrast to SA, the contour optimized by ICM depends on the initial estimate and is therefore a local optimization technique rather than a global one. In [Rueckert and Burger, 1995b; Rueckert *et al.*, 1997], SA and ICM have been combined for the optimization of active contours by segmenting only the first time frame of a set of cine MR images using SA, and propagating the result to the next slice. As the initial estimate for the next slice is sufficiently close to the global minimum, optimization using ICM is used, the result of which is copied to the next slice. This process is repeated until all slices are segmented.

4.4 Other Models

A vast diversity of active contour models has been developed in recent years. Apart from the classic snake model presented in [Kass *et al.*, 1987b], and its main derivations in [Amini *et al.*, 1990; Williams and Shah, 1992; Cohen, 1990; Cohen, 1995] which have been described above, some other models will be briefly reviewed in the following. They can be categorized as geometrically deformable (or deformed) models, shape-based models, spline-based models, level set models, and extensions to $3D$. Other approaches, including deformable templates, exist, but are beyond the scope of this survey.

4.4.1 Geometrically Deformed Models

Geometrically deformed (or deformable) models, called GDMs, have been originally developed by [Miller *et al.*, 1991]. GDMs can be viewed as a semi-permeable balloon placed within an object, reaching its boundary by a local, geometry-driven relaxation mechanism. The balloon is actually a collection of discrete polygons or a polygonal mesh which ensures that the data is sampled only at the vertices of the polygon. Permeability is needed to make irrelevant image features and noise pass through the model without blocking it. Similar to the classic active contour model, the vertices are evaluated using an associated cost function which is minimized during the expansion of the polygonal mesh. The cost function is a linear combination of a balloon mechanism, an image threshold and a topological constraint:

$$\mathcal{E}(\mathbf{v}_i) = \alpha_{balloon}\mathcal{E}_{balloon}(\mathbf{v}_i) + \alpha_{threshold}\mathcal{E}_{threshold}(\mathbf{v}_i) + \alpha_{topology}\mathcal{E}_{topology}(\mathbf{v}_i) \,, \qquad (4.38)$$

where the single energy terms have already been presented earlier in this chapter in equations (4.17), (4.14), and (4.9), respectively. The optimization of this energy functional is performed locally, following a steepest descent algorithm by moving each vertex $\mathbf{v}_i$ of the model in opposite direction of the direction of the gradient of the energy function. Each relaxation step is followed by a global sampling step which adds new vertices between the existing ones. Alternating relaxation and sampling yields a high-resolution, accurate model of the object under investigation.

The concept of GDMs has been extended in [Rueckert, 1993] to global energy optimization using Simulated Annealing. In [Bulpitt and Efford, 1994; Bulpitt and Efford, 1995], a local mesh refinement is performed in order to obtain a model with self-optimizing topology. Both approaches combine the deformation term and the sub-sampling process of the GDM with energy terms of the classic active contour model. Finally, in [Lobregt and Viergever, 1995], an extension of the cost function is suggested, incorporating a velocity and acceleration force of the vertices in order to describe the dynamic state of each vertex along with its position. As the model may oscillate instead of coming to an effective standstill, a weighted damping force is introduced which is proportional to the velocity. Additionally, a constant mass is associated with each vertex, allowing to model a physical solid. This model therefore belongs to the class of *dynamic contour models*, of which the classic active contour model is a special case.

4.4.2 Shape-Based Models

All of the models described above are not able to incorporate any specific prior knowledge of the object shape. Shape-based models are a statistical tool based on a linear *point distribution model* (PDM) originally developed by [Cootes *et al.*, 1992b]. PDMs are defined via a training set of N shapes, which are example shapes for the object under investigation. Each shape is described by n landmark points. Aligning the set of shapes $\mathbf{v}_i = \{((x_{i1}, y_{i1}), \cdots, ((x_{in}, y_{in})\}^T$ to the mean shape $\bar{\mathbf{v}}$, which is calculated by

$$\bar{\mathbf{v}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{v}_i\,, \tag{4.39}$$

allows to analyse the differences from the mean shape using principal component analysis, where for each aligned shape vector $\mathbf{v}_i$ a vector $d\mathbf{v}_i = \mathbf{v}_i - \bar{\mathbf{v}}$ is computed. Calculating the $2n \times 2n$ covariance matrix $\mathbf{S}$ by

$$\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N} d\mathbf{v}_i d\mathbf{v}_i^T \tag{4.40}$$

yields the modes of variation of the training data described by the unit eigenvectors $\mathbf{p}_k$, $k = 1 \cdots 2n$, of $\mathbf{S}$ such that

$$\mathbf{S}\mathbf{p}_k = \lambda_k \mathbf{p}_k \tag{4.41}$$

where λ_k is the k-th eigenvalue of $\mathbf{S}$ and $\lambda_k \geq \lambda_{k+1}$, and $\mathbf{p}_k^T\mathbf{p}_k = 1$. The resulting model consists of the mean shape $\bar{\mathbf{v}}$, and the subset of t eigenvectors corresponding to the t largest eigenvalues which correspond to the most significant modes of variation in the training data. Most of the variation can be explained by a rather small number of modes $t < 2n$ which can be chosen such that the sum of their variances explain a sufficiently large proportion of the shape variability λ_T. Any shape of the training set can be approximated by the mean shape and a weighted sum of the first t modes, $\mathbf{v} = \bar{\mathbf{v}} + \mathbf{P}\mathbf{b}$, where $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_t)$ is the matrix obtained from the first t eigenvectors, and $\mathbf{b} = (b_1, b_2, \cdots, b_t)^T$ is the vector of weights for each eigenvector. Varying the linearly independent parameters b_i allows to generate new examples of the shape in a shape-constraint fashion. This linear PDM forms the basis of *active shape models* developed by [Cootes and Taylor, 1992; Cootes *et al.*, 1995]. It allows to make an initial guess in terms of the shape, position, orientation and scaling of the object. The deformation of the PDM provides a powerful technique for refinement of objects of any shape, given an appropriate training set. Optimization is performed by applying *genetic algorithms* (GA) [Goldberg, 1989], which generate new populations or generations of solutions at each optimization step, using genetic operators such as crossover between existing solutions, as well as mutation in order to escape from local minima. Extensions of this model have been developed by [Baumberg and Hogg, 1995], whose *eigenshape model* can incorporate arbitrary numbers of landmark points using a parametric spline representation, and by [Heap and Hogg, 1995] who have developed a *hybrid cartesian-polar PDM* which allows for accurate modelling of non-linear bending and pivotal deformations. Statistical models based on PDMs have been successfully used for segmentation as well as for deformation analysis in medical imaging by comparing the shape under investigation with the training set [Cootes *et al.*, 1993; Ruff *et al.*, 1996]. It should be noted that in contrast to the classic active contour model, statistical models do not operate on a local basis, as the deformation which is performed by varying the weights of the eigenvectors affects the model globally. Another problematic aspect of statistical models is that there might not be sufficient training data available for a specific problem, or that one special case varies too much from the training data (as might be the case for medical data of an abnormal).

4.4.3 Spline-Based Models

A discrete contour representation of an active contour model has, despite its computational efficiency and speed, two major disadvantages: first, the numerical internal derivatives have to be approximated discretely via finite differences, leading to numerical instability and lack of precision. Second, the discrete nature of the model implies that there is no knowledge about the shape *between* contour points, leading to a lack of robustness. Using a spline representation as presented in chapter 2, section 2.1.1.3, allows to calculate derivatives analytically and to interpolate between contour points (in this case, spline control points) with any desired precision. However, there are only few examples of spline-based active contour models, most of which only use the interpolated spline contour to obtain a smoothly interpolated resulting contour and to control the contour resolution. In the following, an overview of spline-based deformable models is given:

The classic active contour model by [Kass *et al.*, 1987b] incorporates a controlled continuity spline as a generalization of a so-called *Tikonov stabilizer*. However, the optimization technique is based on a sparse matrix, formulating a discrete, finite-difference setting for the model.

[Menet *et al.*, 1990b; Menet *et al.*, 1990a; Saint-Marc *et al.*, 1993] have developed a model called *B-snakes* based on parametric B-splines. The contour control point are formulated as a set of control vertices, where moving one of the vertices results in only local changes of the contour. The model optimization is performed in two stages: first, the energy of the vertices is discretely minimized using a gradient descent approach. Then the final model is obtained by performing a least-square fit of the data by the B-spline, using the analytic spline derivatives. This approach has been modified by [Bascle and Deriche, 1992; Bascle and Deriche, 1993] who first perform a least square fit of the data in form of extracted edgels, averaging the intensity gradient along the B-spline contour, and then refine the obtained contour using a steepest descent optimization technique.

In [Cipolla and Blake, 1992; Cham and Cipolla, 1996], *B-spline active contours* for real-time tracking are used for curve fitting for chains of extracted edgels obtained from an edge detection and linking algorithm. The internal forces are only implicitly defined by a simple differential system (the variables of which are the B-spline control points), as the spline regularization is regarded as *intrinsic*. The B-spline active contour is used as a tool for least square regression fitting of extracted edgels. In [LeGoualher *et al.*, 1996], a similar *snake-spline model*, originally developed in [Leitner *et al.*, 1990], is used whose internal energy is again only implicitly defined by the B-spline representation, and which is optimized in a discrete time evolution process, similar to the variational approach used in [Kass *et al.*, 1987b]. A three-stage B-spline model has been presented recently by [Wang *et al.*, 1996], which has also only implicitly defined smoothness constraints. After a coarse local adjustment stage, the spline control points are *redistributed* according to the local spline curvature in the second stage. In the final stage, a *global fine-tuning* is performed, which also handles insertion of additional *knots* at corners.

[Amini *et al.*, 1995b] have introduced so-called *DP B-snakes* for representing tag lines in radial and SPAMM tagged MRI, where the B-spline control points represent the junction of vertical and horizontal tags, yielding a sparse data representation, as only few control points are needed to represent the whole grid, while achieving sub-pixel accuracy for tag detection over time. Dynamic programming (DP) is used for the global, yet discrete control point optimization.

Finally, [Rueckert and Burger, 1995a] have presented an *adaptive spline model* based on Overhauser polynomial splines and the concept of GDMs. The vertices of the GDM are modelled as interpolated spline control points, and optimization is carried out by minimizing the definite integrals of the internal spline elasticity and bending, while interpolating image forces and other constraints along the spline contour.

Only the last of these spline-based approaches uses the interpolating nature of the spline for the optimization of the model (yet allowing no internal constraints to be included), while the other approaches use the spline only for a least square fitting of discrete edgels, or perform a discrete optimization on the spline control points, and computing the internal spline forces either analytically or considering them as intrinsic.

4.4.4 $3D$ Deformable Models

The first volumetric model was developed as a polygonal mesh by [Miller *et al.*, 1991]. [Cohen *et al.*, 1992; Cohen and Cohen, 1993; McInerney and Terzopoulos, 1995a] have developed analytic active surfaces based on finite elements and a thin plate under tension surface spline, controlling the stretching and bending of the surface (see chapter 2, section 2.1.1.3.2). [Sclaroff, 1995; Nastar and Ayache, 1994] have developed physics-based models on a modal basis (see section 4.4.2). Parametric surface models for geometric matching and shape description based on a Fourier representation have been developed by [Staib and Duncan, 1992; Brechbühler *et al.*, 1995] (see chapter 2, section 2.1.2.1 for details on Fourier representations and descriptors). Other work on $3D$ deformable models can be found in [Delingette *et al.*, 1992; Bardinet *et al.*, 1996b; Bardinet *et al.*, 1996a], and the following section presents implicitly defined surface models.

4.4.5 Implicit and Topological Models

All the models presented so far rely on an explicit shape representation, relying on a close initialization and not being able to handle topological changes easily. Reparameterization and user interaction are often necessary which leads to numerical instability and inaccuracy. As an alternative, a topological snake model and several implicit approaches have been developed recently, acting directly on the grey-level curves and surfaces contained in the image.

[Whitaker, 1994a; Whitaker and Chen, 1994] have introduced the concept of implicit volumetric deformable models or *active blobs* which are described by the level sets (sets of isophotes) of the image for visualizing and segmenting volumetric data. Simultaneously, [Caselles *et al.*, 1993; Malladi *et al.*, 1995; Malladi *et al.*, 1996; Caselles *et al.*, 1997] have developed a very similar *geometric level set model*. Level set models are based on an image scale-space and an implicitly defined multi-scale energy functional. They are based on their intrinsic geometry and are derived by embedding them as levels sets of a $3D$ scalar function $F : \Re^3 \rightarrow \Re$ evolving over time, hence removing the parameterization needed for the active surface model in section 4.1:

$$\frac{\partial F}{\partial t} = \nabla f \, \|\nabla F\| \text{ div} \left(\frac{\nabla F}{\|\nabla F\|} \right) + \alpha_{image} \nabla f \, \|\nabla F\| \tag{4.42}$$

with initial condition $F(0, \mathbf{x}) = F_0(\mathbf{x})$. Equation (4.42) acts on level sets of the image, treating each set as an individual surface under evolution of its own constraints. $\text{div}\left(\frac{\nabla F}{\|\nabla F\|}\right)$ is the sum of the two principle curvatures of the level sets of F (and twice the mean curvature), and f is a scalar field function defined over the range of the model representing image features of interest (in particular edges). The term $\alpha_{image}\|\nabla F\|$ corresponds to an image feature force, with ∇f acting as a stopping criterion, which must be zero for the surface to stop (e.g. in the event of edges). The sign of α_{image} determines whether the surface is to propagate inwards or outwards, causing a balloon-like deflation or inflation of the model. The main advantage of level set methods is that they can incorporate scale continuation (in a similar fashion as presented in chapter 3, section 3.4.2) and their topological flexibility, as they can split and form multiple objects. The final surfaces can be recovered by extracting them via iso-surface rendering. An extensive introduction to the theory and application of level set methods can be found in [Sethian, 1996].

[Tek and Kimia, 1994; Tek and Kimia, 1995b; Tek and Kimia, 1995a; Tek and Kimia, 1997] have taken a slightly different approach. Their concept of *bubbles* relies on a *shock-based* representation of shape, as developed by [Kimia *et al.*, 1995; Kimia and Siddiqi, 1996]. The *shape from deformation* framework enables to derive the shape representation directly from the data in an inverse process, based on detecting the shocks of first (orientation discontinuity), second (curvature discontinuity), third (collapsing of distinct boundaries), and fourth (collapsing of a boundary to a single point) order which are formed when the shape is evolved in a *reaction-diffusion* process (see chapter 3, section 3.2.3 on how a reaction-diffusion scale-space is computed). Bubbles (or small spherical deformable structures) are randomly initialized as fourth-order shocks in the homogeneous areas of a volumetric image, and allowed to grow, shrink, merge, split, and generally deform under physically motivated forces, and eventually coming to halt near specific differential image structures (e.g. at a high image gradient magnitude). This is achieved by first smoothing the image by a shock-based reaction-diffusion technique, randomly initializing bubbles as $F_0(\mathbf{x})$ in homogeneous image areas, and then solving equation

$$\frac{\partial F}{\partial t} = \nabla f(\mathbf{x})(\beta_0 - \beta_1 \kappa(\mathbf{x}))\|\nabla F\| \, , \tag{4.43}$$

where $f(x, y)$ is again a scalar potential field denoting image features, F is the evolving active level set surface as parameterized by Caselles et al. and Malladi et al., κ is its curvature (e.g. the Gaussian, mean, or one of the principal curvatures), and β_0 and β_1 are scalar weighting factors. The advantage of bubbles in comparison to the level set evolution presented above is that no prior knowledge is necessary about how many level sets are needed for a complete segmentation, that narrow regions can be captured more easily even in the presence of gaps in the data, and that bubbles are a completely automatic technique for segmentation.

[McInerney and Terzopoulos, 1995b; McInerney and Terzopoulos, 1995c] have developed a $2D$ topologically adaptable or adaptive snake model which retains all the features of the classical snake model, but overcomes many of its limitations. By superimposing a grid over the image, an implicit formulation of the snake is obtained which allows to iteratively reparameterize the deforming snake model, enabling it to flow into complex shapes and dynamically changing its topology by branching. The snake is based on a set of nodes and springs which does not remain constant during the deformation. The decomposition of the image into a grid of discrete cells allows to compute the intersections of the model with the grid, and to disconnect or reconnect the snake nodes accordingly. The topological snake model has the functionality of the implicit level-set models presented above, but does not require any mathematic formulation beyond that of classic snakes, retaining their parametric form.

4.5 Summary

This chapter has presented the original or classic active contour model by [Kass *et al.*, 1987b; Kass *et al.*, 1987a] as well as its main derivations with respect to the choice of energy functional and optimization technique. The concept of active contour models is a very efficient tool for shape extraction and segmentation in terms of combining autonomous shape forces, image-data dependent terms and other constraints, but it usually suffers from several problems which will be briefly stated and discussed in the following.

4.5.1 Representation

All major active contour models are discretely represented as a set contour points, *snaxels*, vertices or spline control points. This sparse representation allows to optimize a contour model efficiently, yet leads to problems related to number of snaxels and spacing. [Williams and Shah, 1992] have suggested an enforced uniform spacing scheme, where the spacing distance corresponds to the contour *resolution*. [Miller *et al.*, 1991] have proposed to refine the resolution iteratively with each optimization step by subsampling, inserting new contour points in the middle between each neighbouring pair of contour points until the desired, uniform contour resolution is reached. This imposes the need to perform a very dense sampling of the contour points in order to capture all finer shape structure, leading to a large number of redundant contour points and therefore to high computational costs. [Bulpitt and Efford, 1995] have adjusted the sampling density of the control points to the local curvature of the model, yielding dense sampling at shape parts of high curvature, while straight parts are modelled by a few points. However, this sampling schedule is still of a discrete nature, as the model is only evaluated at the control points, and no knowledge of the contour parts between are given. This often leads to a lack of robustness, as noise and spurious edges can affect the model quite considerably. Section 4.4.3 discussed several spline-based models, which allow to evaluate the model not only at its discrete spline control points, but also along the interpolated spline contour.

4.5.2 Energy Function

The discrete nature of active contour models not only influences the robustness and resolution of the model, but also poses problems regarding the computation of the internal contour derivatives needed for the local elasticity and curvature estimation of the model. Discrete derivatives are usually computed using finite differences, which is no problem in itself. However, non-uniform or very sparse spacing causes numerical inaccuracy as discussed by [Williams and Shah, 1992]. Solutions to this problem are either using dense uniform sampling, causing a huge computational cost, or exploiting a spline representation. Several spline-based approaches have been developed in the literature (section 4.4.3), yet they only evaluate the derivatives at the knots of the contour model.

4.5.3 Optimization

Based on the chosen discrete contour representation and the associated discrete derivative approximation, several optimization techniques have been presented in section 4.3, including the variational approach, DP, the *greedy* algorithm, and stochastic and probabilistic relaxation. Techniques like GA (section 4.4.2), and level set evolution (section 4.4.5) have been briefly discussed. They suffer from several disadvantages, one of them being that the discrete formulation of the model to be optimized implies for most techniques that derivatives of the energy functionals must exist, e.g. for the variational approach and DP. Although the energy derivatives can be estimated using finite differences, the same reservations as in section 4.5.2 apply here. In particular, the choice of a suitable time step for the variational approach is difficult, as small choices slow down the technique considerably, and larger steps cause numerical instability. Another frequently encountered problem is that the energy function cannot be modified freely - in particular the incorporation of hard constraints is limited. This implies that it is difficult to relate geometric image structure or other constraints into the internal model representation. However, the *greedy* algorithm, DP, SA and ICM as well as GA can minimize any given energy function, which make these methods very flexible.

The optimization techniques discussed so far are either local or global. Local methods include the variational approach, level set evolution, the *greedy algorithm*, and ICM, while DP, SA and GA are all global techniques. Local optimization implies that an active contour model locks to nearby edges and other image features with respect to its internal forces, given a good initial estimate. It should be noted though that for all techniques a trade-off between computational cost and the locality or globality of the solution needs to be performed. The larger the search space, e.g the local neighbourhood size of DP or the *greedy* algorithm, the more global is the final solution (but DP always ensures the global solution *within* its search space), but the higher is the computational cost and complexity. Global optimization is performed if no initial estimate is available, or when user-independence and absolute reproducibility are required. If the image object under investigation does not correspond to the global energy minimum, i.e. the global solution is not the desired solution, as it is most frequently the case, local techniques are used. These in turn rely on a good initial contour estimate or appropriate external energy forces, causing considerable user- and initialization-dependent variation. As it is still desirable to obtain reproducible and initialization-insensitive results when using the efficient local techniques, scale-space continuity as suggested by [Kass *et al.*, 1987b] (section 4.2.6) can be used to attract the model at a high image scale to distant, but very prominent image features like edges, and slowly focusing it down for decreasing image scale. This ensures a robust *coarse-to-fine tracking* as discussed in the previous chapter (section 3.4.2). The initial scale defines the *locality* of the final solution [Whitaker, 1994a], as at very large scales, the model gets attracted to large distant objects only, and at finer starting scales, it locks to nearby objects, often disturbed by noise and spurious edge responses.

4.5.4 Incorporation of Other Knowledge

Though formulations of shapes exists and have been discussed in chapter 2 of this dissertation, few approaches in active contour models take into account any prior knowledge one might have of the shape which is segmented. *A priori* shape knowledge not only speeds up the optimization process, but also allows to use deformable models to locally regularize similar shapes and investigate the internal and external shape forces with respect to the model's energy terms. It is possible to use techniques like the Hough transform to derive a close initial estimate, and the *active shape models* and their variations (section 4.4.2) deform in ways characteristic of the class of objects they represent, which also enables a measurement of differences in shape. In applications to time sequences like in [Rueckert and Burger, 1995b], results from earlier time slices obtained via global SA relaxation are used as initial estimates for the next time slices which allow to use local ICM relaxation, thus improving the optimization speed. Deformable templates, a concept which goes beyond the scope of this dissertation, allow to constrain the deformation of a model to a specific shape. The use of scale-space continuation, discussed in the previous section, is another good example for the direct incorporation of *a priori* shape knowledge, as the estimates obtained at higher image scales provide a coarse approximation of the model for the next lower scale, speeding up the optimization, while providing robust, global to local (or *coarse-to-fine*) shape extraction.

Geometric image structure, revealed by differential invariants of the image intensity as presented in the previous chapter (3.1.1), allows to analyse the underlying shapes contained in an image. It is therefore highly desirable to adjust the shape of the deformable model to the underlying shape which it is to extract. This is partially achieved by locking the model to the edges of the shape outline in the image or via one of the image energy terms in section 4.2.2. Yet there are various other image features which might be taken into account. The level set methods in section 4.4.5 formulate the deformable model to be represented by image isophotes or iso-intensity contours and surfaces which directly relate the deformable model to the underlying image intensity and shape. However, despite several other forces like an edge potential, no higher-differential image force is incorporated or taken into account. A different approach is listed in section 4.2.5 above, where the internal forces of a deformable model are constrained to the underlying image structure using weighting tensors of similar differential structure, but this technique has only been tested on images containing very smooth and simple shapes, making it difficult to judge how well the weighting tensors perform. Moreover, the choice of the introduced free parameters of this technique and a possible extension to image scale have not been addressed.

The following chapter introduces the new framework for shape description using active contour models and scale-space. A new active contour model will be presented, which is based on the items discussed in this section. Scale continuation is incorporated in order to adjust and regularize the contour model locally to the underlying geometric image shape, where the locality is related to scale. Differences in the locality of the solution, such as intermediate results within an image scale-space will be investigated as a valid method for *multi-scale active shape description*.

Chapter 5

Introduction to Multi-Scale Active Shape Description

> – JE TE DIS ÇA... C'EST À CAUSE AUSSI DU SERPENT. IL NE FAUT PAS QU'IL TE MORDE... LES SERPENTS, C'EST MÉCHANT. ÇA PEUT MORDRE POUR LE PLAISIR...
>
> "I TELL YOU– IT IS ALSO BECAUSE OF THE SNAKE. HE MUST NOT BITE YOU. SNAKES– THEY ARE MALICIOUS CREATURES. THIS ONE MIGHT BITE YOU JUST FOR FUN..."
>
> *Le Petit Prince, Antoine de Saint-Exupéry.*

In chapters 2-4 of this dissertation, three important topics in image processing have been reviewed: shape description, multi-scale image processing, and active contour models. In this part a new framework is presented which combines these techniques into a hierarchical tool for *multi-scale active shape description*. The motivation behind this framework is to investigate a shape in its image scale-space rather than its contour scale-space, using an active contour model as a scale-based shape regularization tool. Before an overview of the framework is given, the surveyed methods from the previous chapters used for the development of the new shape description tool will be briefly summarized, and their relevant aspects and potential for application in this context will be highlighted.

5.1 Methods

In order to formulate the proposed new shape description tool, it is necessary to define a suitable, active contour model based shape representation embedded in an underlying image scale-space. After discussing the choice of shape representation and description, the concept of a hierarchical, *multi-scale shape stack* will be presented which is directly derived from the concepts of contour scale-space and image scale-space. Finally, an extension of active contour models to a multi-scale shape regularization tool, and associated topics regarding multi-scale representation and energy minimization are introduced.

5.1.1 Shape Representation and Description

Before designing a new shape description tool, two choices have to be made, namely which shape representation is to be chosen, and what kind of measurements are to be taken. Let us consider the representation aspect first: As was stated in the introduction of chapter 2, this dissertation focuses on planar shape outlines. Several representation schemes have been presented in section 2.1, which are categorized into local, global and medial representation methods, which in turn allow to obtain local, global, and medial shape measurements, respectively. This dissertation favours a local representation (but in a global-to-local scale-space setting) for the following reasons:

- Global representation techniques like Fourier descriptors and statistical moments have been considered, being *intrinsically* hierarchical representation techniques as they allow for truncation of shape information. However, such a truncation affects the shape globally, making it impossible to locate characteristic shape features, and may lead to topological changes, as illustrated in figure 2.7 (chapter 2, page 38).
- Medial representation techniques like the medial axis, though being attractive in terms of providing a structural, multi-local (combined global and local) approach to hierarchical representation and description, are difficult to obtain as they are based on the symmetric behaviour of a shape rather than its outline. Additionally, small changes in the object shape can lead to branching or fragmentation of the medial axis.
- Local representation techniques are possibly the most general way of representing an object, as they allow to compute a large range of local and other shape measurements. They directly encode the shape outline, allowing for easy and direct access.

Of the local representation techniques, the concept of splines is the most versatile. It allows to change a shape locally, while at the same time enables to interpolate or approximate a shape based on a sparse representation. Moreover, splines are also a hierarchical technique as the spacing of the spline control points can be used as to adjust the accuracy of the spline contour: Larger spacing of the control points naturally leads to a coarser representation, while smaller spacing allows to represent a shape in more detail. In combination with their analytic nature this makes them an ideal underlying shape representation for this dissertation.

Let us now consider which of the local, global, and relative shape description techniques reviewed in chapter 2, sections 2.2-2.4, will be used to describe a spline-represented shape. None of these shape description techniques alone is sufficient for a complete shape description, but one can expect that a balanced combination is able to capture the most important shape characteristics. The following global, local and relative shape quantifiers have been selected in this work:

- In medical imaging, clinicians are generally interested in volumetric and surface measurements (corresponding to area and boundary length for planar shapes), and the complexity of objects. Thus, a combination of area, perimeter, compactness, and statistical self-similarity in terms of fractal quantifiers seems to be an appropriate set of descriptors for *global shape measurements.*

- Recent research in clinical neurology suggests that local deformations of the brain are related to the neurological disease status. Hence a local shape interpretation of deformation and bending behaviour is desirable, and can be performed in terms of differential geometric *local shape measurements* of the shape boundary, e.g. in terms of the local curvature behaviour.

- Finally, it is often important for clinicians to perform *relative shape measurements* for inter-shape comparison, e.g. to monitor shape changes over time series or to compare shapes obtained from different imaging modalities. In order to derive the differences between two shapes, an alignment or *registration* needs to be performed in order to find the correct shape correspondence. Without such a correspondence, no accurate shape deviation measurement can be performed. This is a very important and wide topic in medical imaging, going beyond the scope of this dissertation. Some global shape deviation and corresponding distance measurements, however, have been presented in section 2.4 and can be used when a shape correspondence is available.

All of these descriptors can be efficiently and accurately computed from a local spline representation. The last point, suggesting the need for relative *inter-shape description*, can be extended to the concept of *intra-shape description.* Obviously, intra-shape description can only be performed if a shape is represented in form of a shape hierarchy. In this hierarchy, shapes are inherently registered. This dissertation aims to investigate the multi-scale properties of a shape, thus a multi-scale hierarchy can be used to apply the shape descriptors and distance measurements across scale.

5.1.2 Multi-Scale Techniques

Chapter 3 has presented two different multi-scale representation techniques: multi-scale contour representation, and multi-scale image representation. Both representation schemes allow to perform hierarchical or structural descriptions at different levels of detail or resolution, capturing global features at high scales and finer detail at lower scales. This dissertation proposes a multi-scale shape representation by uniting contour and image scale-space representation methods. This approach is motivated by the following two facts:

- The classic multi-scale contour representation is obtained by smoothing a binary shape outline for increasing levels of scale. However, a segmentation from an image at a chosen scale is necessary to obtain the initial contour unless an analytic contour representation is available. The choice of this scale is not obvious and its role is disregarded in the following binary smoothing process.

- Constructing a contour scale-space means that the *image context* is ignored. However, a contour is not only defined by its intrinsic multi-scale representation, but also by its associated image scale-space in which it is embedded. For example, the image isophotes define the shape, the image gradient separates the shape from the image background, and other geometric image features define particular shape characteristics at all scale levels.

Both items naturally lead to the core idea in this dissertation: Extracting a shape at range of *image scales* rather than at a single scale yields a novel multi-scale shape representation which will be termed *multi-scale shape stack.* The embedding of a shape in its image scale-space is directly associated with two main techniques in multi-scale image analysis which are mutually dependent on each other:

- *Multi-scale differential image invariants* describe various image features like the *edgeness*, *cornerness* or *curvature.* Consequently, they provide a geometric and structural way to characterize a shape embedded in its surrounding image scale-space, and are therefore the first step for a *multi-scale shape extraction* scheme. This directly leads to the necessity of computing a suitable *image feature scale-space* from a given image scale-space, comprising commonly used edge detectors like the image gradient or the zero-crossings of the image's Laplacian to describe the shape outline, as well as features which are used only rarely in image processing, like the isophote image curvature.

- The computation of a feature image scale-space forms the basis of *multi-scale feature tracking*, which is a fundamental technique for following a shape across scales. For example, *edge focusing* is a *coarse-to-fine* tracking technique which can be used to detect a coarse shape outline at a high level of image scale, and to focus the shape subsequently down in order to track finer scale details of the shape. Obviously, such a process can be reversed in order to permit a shape, whose zero-scale version is known (e.g. from a manual segmentation by an expert, or from an analytic representation), to *evolve* by tracking its outline for increasing levels of scale in a *fine-to-coarse* tracking process. Other image features may be considered, both for *evolution* or *focusing* of a shape.

Both techniques combined provide the means for shape tracking through image scale-space. The novelty in this approach is that in contrast to classic techniques like edge focusing which discards all but the lowest scale shape, this approach formulates the shape tracked at each intermediate scale level as layers in a multi-scale shape stack (which is organized either in a *fine-to-coarse* or a *coarse-to-fine* fashion).

5.1.3 Active Contour Models or *Snakes*

The remaining issue for the proposed multi-scale shape description technique is the actual shape extraction, as the shape is still embedded in the image and not yet available in its binary form (which is required in order to describe it). In contrast to pure segmentation techniques, this work proposes to track a shape through its image feature scale-space by implicitly segmenting it at each scale level. This implicit segmentation can be characterized as regularization with respect to scale, or, depending which direction in the scale-space is taken, as image data-driven shape evolution or focusing. While many techniques for (implicit) segmentation may seem to be suitable, the model of active contours or *snakes* has been adopted as it provides a flexible framework for incorporating internal shape characteristics and a large variety of external image features into an energy function, transforming the shape regularization problem into an energy minimization problem. Moreover, scale-space continuity can be readily integrated into this model. The development of a suitable multi-scale active contour model is carried out in three stages:

1. The shape *representation* for the classic active contour model is discrete (and continuous only for initial least-squares fitting to a set of edgels, or a final smooth shape interpolation - see chapter 4, section 4.4.3 for more details). As discussed earlier, this leads to several problems concerning the model's numerical stability, robustness, and sampling resolution. An arc length parameterized local spline representation, as proposed in section 5.1.1, can be used to circumvent these problems, as it allows for analytic differentiation, a continuous interpolating representation and evaluation, as well as natural sampling (an important characteristic which will be explained in more detail in the next chapter).

2. The design of the *energy function* is a crucial point for the development of a new active contour model, as it describes the quality of the model in terms of various image and contour features, and to provide a weighted balance or *compromise* between them. It serves as a heuristic for the following optimization or energy minimization process for the model deformation. This dissertation will expand the classic energy function in the following ways: The relationship between internal (analytic) elasticity and contour resolution will be investigated, and an improvement of the classic minimizing curvature term will be developed, leading to a combination of internal and external forces. More specifically, the analytic spline curvature will be matched to the isophote image curvature in a *curvature matching process*, adjusting the shape of the model to the shape of the underlying image isophotes. Finally, an additional set of image features will be integrated into the energy function.

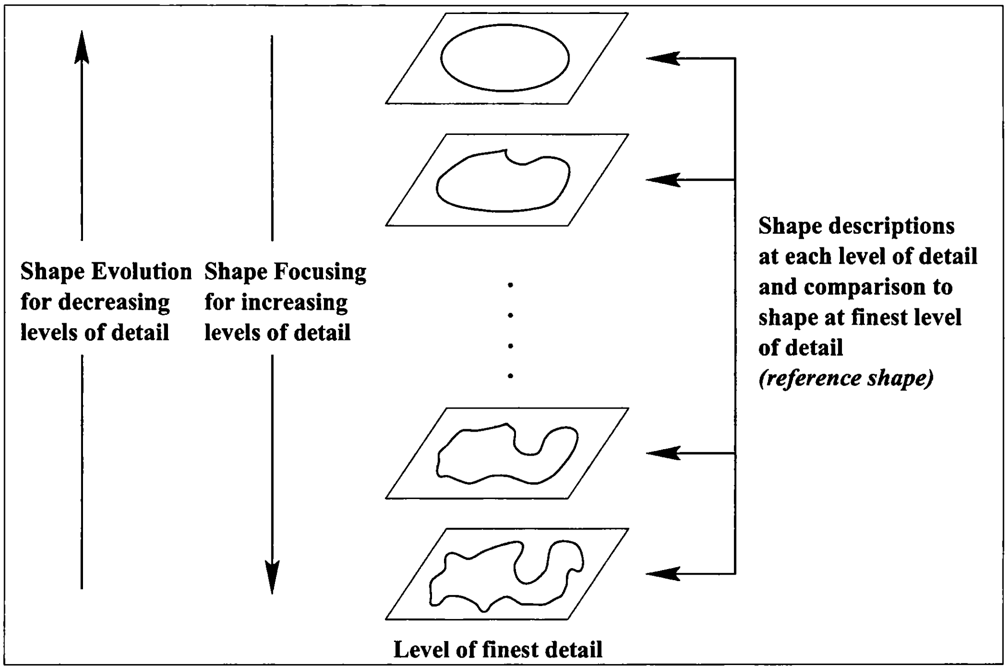


Figure 5.1: Overview of multi-scale active shape description.

3. The *optimization* techniques for active contour models presented in section 4.5.3 are either local (having initialization problems and often leading to local, undesired energy minima) or global (needing infinite time to ensure an optimal solution). This dissertation proposes to use the concept of scale-continuity which naturally regularizes the locality or globality of the solution. Hence it is sufficient to use a fast and efficient local optimization technique, which is attracted at high scales to more global image features, and locks at low scales to more local, nearby features.

5.2 Overview

The above summary lists the methods that will be employed for the proposed multi-scale active shape description technique, yet their interaction and global relationship needs some further explanation. The core of the technique lies in providing a hierarchical tool for shape analysis based on the notion of scale-space. The aim is to describe a shape not at a single scale, but rather at a large set of scales in order to detect shape characteristics and changes *across scales*. In order to do so, first an image scale-space and a corresponding image feature scale-space will be constructed. Then the shape of interest will be extracted using a novel multi-scale active contour model. This model is used for shape tracking or regularization in scale-space rather than for traditional segmentation, a process which is referred to as *implicit segmentation* in this context.

If the true shape outline is known (e.g. by an analytic function), or assumed to be a *gold standard* (e.g. using a prior manual segmentation by a clinician), a *fine-to-coarse* strategy similar to constructing a contour scale-space can be applied. This process will therefore be called *active shape evolution*, as it corresponds to the classic contour scale-space, yet at each level of scale the image context is taken into account. Alternatively, a *coarse-to-fine* technique similar to edge focusing can be employed which eliminates the need of a prior segmentation, as the initial model can be very coarse. Consequently, this process will be referred to as *active shape focusing*. Both techniques are performed by propagating the shape contour through image scale-space and regularizing the active contour model's energy function with respect to scale. The set of initial, intermediate, and final evolution or focusing results forms a multi-scale, image data-driven representation of the shape, and thus composes a *multi-scale shape stack*. This stack is then evaluated at each scale level with the shape measurements described above. Figure 5.1 illustrates this concept.

In the following chapter, the theoretical framework for the multi-scale active contour model used for the new shape description technique will be presented in more detail. The construction of the multi-scale shape stack via active shape evolution and focusing for shape description will be presented and discussed in chapter 7, followed by the application of the presented techniques in chapters 8 and 9.

Chapter 6

Multi-Scale Active Contour Model

MAIS QUELQUE CHOSE LE RASSURA:
– C'EST VRAI QU'ILS N'ONT PAS LE VENIN POUR LA SECONDE MORSURE...

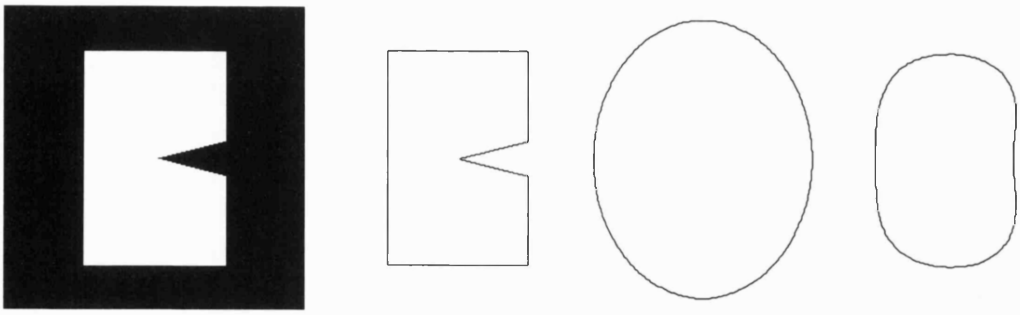
BUT A THOUGHT CAME TO REASSURE HIM:
"IT IS TRUE THAT THEY HAVE NO MORE POISON FOR A SECOND BITE."

Le Petit Prince, Antoine de Saint-Exupéry.

This chapter presents the theoretical framework for a multi-scale active contour model for shape description, and discusses details regarding its representation, energy function, and optimization. The motivation behind this model is twofold: first, the classic model and its recent developments suffer from several drawbacks, including their numerical instability, lack of robustness, and inability to extract strongly curved, complex shapes such as brain contours in Magnetic Resonance Imaging (MRI). Second, though scale-space continuity has been suggested in [Kass *et al.*, 1987a], such an approach has not been further developed in the literature, except for the level set methods presented in chapter 4 (section 4.4.5) whose implicit representation, however, conflicts with the adapted concept in this dissertation of shape as an explicit contour.

This dissertation aims to investigate and develop enhancements of the classic active contour model in terms of continuous spline representation, contour shape adjustment to the underlying image shape, and integration of other geometric image structure into the energy function. Moreover, the incorporation of the notion of scale-space into the representation, energy function and optimization process will be explored in order to regularize a shape in its image context at varying levels of image scale. This leads to the application of computationally more efficient local rather than global optimization techniques where the choice of the image scale determines the globality of the solution.

Examples and illustrations will be used to complement the theoretical formulation of the proposed model and to demonstrate the usefulness and validity of this approach. Figure 6.1 shows a test image, *notched rectangle*, its true shape outline, an ellipse-shaped model used as a rough estimate, and an initial model obtained by optimizing the ellipse-shape model at a high scale level. In the



(a) (b) (c) (d)

Figure 6.1: (a) Test image *notched rectangle*. (b) Known shape outline. (c) Ellipse-shaped model. (d) Initial model, obtained by optimizing shape (c) at scale $\sigma_0 = 32$.

following, the multi-scale contour representation as well as associated sampling issues will be investigated.

6.1 Representation

As mentioned in the previous chapter, a local spline representation is adopted as the underlying representation for the new active contour model. Cubic C^2 continuous B-splines (see chapter 2, section 2.1.1.3.1) have been applied in the literature for various active contour models (see chapter 4, section 4.4.3), and have been generally found very suitable for shape representation, approximation and analysis, because of their smooth and continuous approximation, their efficient decoupling of the coordinates $x(s)$ and $y(s)$, their analytic differentiability, and their generative nature and local control.

Extending the classic active contour model which is based on a discrete set of *snaxels* to a B-spline representation requires the redefinition of the *snaxels* as *B-spline control points* $\mathbf{v} = \{\mathbf{v}_1, \cdots, \mathbf{v}_N\}$. This allows to approximate a spline contour via a set of spline patches defined by $\mathbf{v}(s) = \{\mathbf{v}_1(s_1), \cdots, \mathbf{v}_N(s_N)\}$. Each spline patch $\mathbf{v}_i(s_i)$ with $s_i \in [0; 1)$ is based on four control points $\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}$ in modulo notation, and is related to the overall contour by

$$\mathbf{v}(s) = \mathbf{v}_i(0) \quad \text{if} \quad s = \frac{i}{n} \quad \text{with} \quad s \in [0; 1) \tag{6.1}$$

The set of B-spline control points can be more sparse than the discrete set of *snaxels*, as the spline representation allows to approximate (and therefore evaluate) contour parts in between the control points with sub-pixel precision. Note that B-splines are only interpolating if the control points are doubled, and almost interpolating if the controls points are very close, or if they lie along a linear part of the curve. The control points can therefore be used to *steer* the active contour model rather than to directly *define* it. It is therefore sufficient to optimize the positions of the set of snake control points, by evaluating the internal differential energy terms and external image and constraint terms along the spline patches. The number of control points chosen defines the degree of freedom or flexibility of the contour in the optimization process, and the separate parametric representation of the spline coordinates allows to perform a fast and efficient optimization. The local control implies that local changes in the B-spline based active contour model due to movements of the control points in the optimization process are reflected only by changes of B-spline parameters local to that change, allowing for a fast recomputation and evaluation.

A B-spline representation for active contour models naturally leads to questions concerning the choice of the number and distance of the control points of the model and associated sampling strategies in a scale-space setting. In the following, fixed-scale sampling strategies will be first developed and discussed and then further extended to multiple scales.

6.1.1 Fixed-Scale Sampling

The problem of determining the appropriate number and local distance of contour points for active contour model optimization has been addressed by many researchers (see chapter 4, section 4.5.1), and has been approached in three main ways:

1. [Kass *et al.*, 1987a] have originally developed an intrinsic uniform sampling scheme of a fixed number of points which is enforced by the internal elasticity term weighted by $\alpha_{elasticity}$ (equation (4.4), chapter 4), and [Williams and Shah, 1992] have additionally suggested to integrate the mean distance into the elasticity energy term (equation (4.6), chapter 4). However, for very complex shapes a larger number of control points might be necessary for accurate local adjustment than for very smooth shapes. Determining the appropriate number of points is not trivial, leading usually to a compromise between desired accuracy of the solution and computational speed.

2. [Bulpitt and Efford, 1995] have addressed the problem that a shape may vary locally in its detail, i.e. it may be smooth at some parts, and complex elsewhere. They have therefore proposed to adjust the local sampling density with respect to the local internal curvature of the model. The main problem of this approach is that a model might not gain higher local curvature and hence recover finer detailed structures *unless* it is already densely sampled, making this an *inverse problem*.

3. [Miller *et al.*, 1991] have developed a scheme which starts with only few points, allowing for an efficient computation of an initial, coarse model, whose density is subsequently refined by inserting new points. However, smooth shapes tend to be *oversampled* this way, leading to increased computational complexity.

All three approaches are based on a fixed-scale image setting, and either enforce a fixed or variable *contour scale* ς, which is defined by the spacing of the control points. The models suffer from

Properties	Uniform	Variable	Refinement
Intrinsic sampling	$+$	$\square$	$-$
Adaptive number of points	$-$	$+$	$+$
Adaptive level of detail	$-$	$+$	$+$
Local refinement	$-$	$+$	$\square$
Efficiency	$\square$	$\square$	$+$
Flexibility	$-$	$\square$	$\square$
Accuracy	$\square$	$\square$	$+$
Notation	ς	$\varsigma(s,r)$	ς_i
Sampling schedule	$\alpha_{elasticity} \cdot \mathbf{v}_s(s)$ (*) $\alpha_{elasticity} \cdot (\lvert\bar{\varsigma} - \mathbf{v}_s(s)\rvert)$ (**)	$\frac{1}{\lvert\kappa(s,r)\rvert}$	$[\varsigma_1 \cdots \varsigma_n]$

Table 6.1: Qualitative comparison, notations, and sampling schedules of fixed-scale sampling strategies for deformable models. *Uniform* sampling refers to the strategies of [Kass *et al.*, 1987b] (*) and [Williams and Shah, 1992] (**) using intrinsic enforcement of uniform snaxel spacing via $\alpha_{elasticity}$ and mean distance constraints, respectively. *Variable* sampling defines local mesh refinement as a function of local curvature [Bulpitt and Efford, 1995]. *Refinement* sampling denotes an iterative point insertion as a *coarse-to-fine* approach [Miller *et al.*, 1991], which gradually increases the resolution. The symbols $+, -, \square$ denote fulfilment, no fulfilment, and partial fulfilment of the listed properties.

the problems of determining the appropriate number of points [Williams and Shah, 1992], and the range and adjustment of ς [Bulpitt and Efford, 1995; Miller *et al.*, 1991]. The choice of these parameters is obviously related to the size and the complexity of the object to be segmented, which is not known *a priori*. Table 6.1 summarizes the properties of these three fixed-scale approaches.

The concept of scale-space presented in chapter 3 provides a framework to represent an image (and objects contained in the image) at different levels of detail. In this dissertation, it is proposed to relate the shape complexity to the choice of the image scale, and to generalize the fixed-scale sampling approaches to multiple scales, allowing for flexible contour scale limits and number of contour points related to the notion of scale. In the following, the multi-scale B-spline representation with associated suitable sampling strategies is presented.

6.1.2 Multi-Scale Sampling

In a scale-space setting, coarse shape details can be captured at higher scales, and all finer scale details contributing to the shape complexity become more apparent at lower scales. Therefore, objects which are segmented at a high level of image scale tend to have a smooth, simple shape outline, and can therefore be represented by fewer points than objects segmented at low scales, which in general have a more complex shape outline. This dissertation proposes to relate the distance of the spline control points, or the contour scale ς, and the associated number of control points to the underlying image scale σ. Moreover, distance and differential measurements should be based on the geometry of scale-space.

Let us consider the scale-space distance measurements: Given two points $\mathbf{v}_1$ and $\mathbf{v}_2$ of Euclidean distance $\|\mathbf{v}_1 - \mathbf{v}_2\|$, their distance in a Euclidean scale-space at a fixed scale σ_0 is given by the distance of their normalized (dimensionless) coordinates (equation (3.28)):

$$\mathrm{dist}_{scale-space}(\mathbf{v}_1, \mathbf{v}_2; \sigma_0) = \|\tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_2\| \tag{6.2}$$

In the Riemannian setting, equation (3.33) (chapter 3, page 70) reduces for a fixed scale measurement to

$$\mathrm{dist}_{scale-space}(\mathbf{v}_1, \mathbf{v}_2; \sigma_0) = \log\left(\frac{1+\sqrt{1-(\rho\sigma_0)^2}}{1+\sqrt{1-(\rho\sigma_0)^2}-\rho\|\mathbf{v}_1-\mathbf{v}_2\|}\right) \tag{6.3}$$

with $\rho = 2/\|\mathbf{v}_1 - \mathbf{v}_2\|$. For example, given a Euclidean distance $D = 8$, at image scales $1, 8, 16$ the Euclidean and Riemannian fixed-scale distance measurements yield scale-space distances of $8, 1, 0.5$ (Euclidean) and $4.189424, 0.962424, 0.494933$ (Riemannian). Both scale-space distances have been motivated in literature (see [Florack, 1993] and [Eberly, 1994a]) by the observation that the image is *less dense* at higher scales, leading to an decrease in the distance between points. This implies that distance measurements (e.g. when measuring differential image structure) need to take the local, decreased density into account. This can be achieved by performing measurements over larger distances at high scales, which does not cause any loss in spatial information, as the capture region of the image is increased with respect to the image scale. Observing now two points at increasing distances $1, 8, 16$ and at corresponding scales $1, 8, 16$ keeps all measurements at a comparable level in the Euclidean scale-space setting. In the Riemannian setting, the shortest path between points is along geodesic curves which takes the scale-space curvature into account, and points become even closer (or more dense) at lower scales due to the logarithmic nature of the distance measurement. In this dissertation the Euclidean setting is therefore chosen for fixed-scale measurements, and the Riemannian setting for multi-scale measurements.

Let us now investigate the issue of the desired contour scale ς in an image scale-space: The increase in density of spatial information of the image scale-space at lower scale layers leads naturally to an increase in shape complexity. One has two choices when performing distance measurements:

1. The distance of points in scale-space can be computed according to equations (6.2) and (6.3), respectively, depending on a fixed-scale or multi-scale setting (i.e. depending on whether two points are located at the same scale or at different scale levels in the image scale-space). This implies that all differential image measurements like the image gradient are effectively computed over larger distances at higher scales than at lower scales, leading to a local adjustment of the scale-space density.

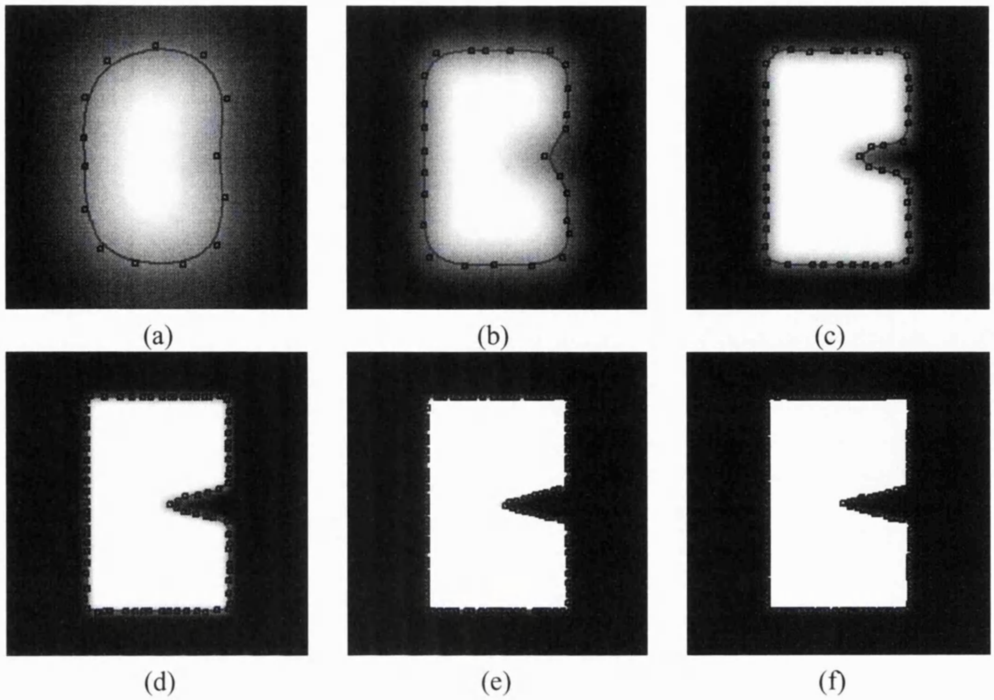


Figure 6.2: Snake optimization for different levels of image scale and related contour scale. Optimization results of the test image and initial model of figure 6.1 for scales σ (a) 32, (b) 16, (c) 8, (d) 4, (e) 2, (f) 1, with corresponding contour scales ς. The resulting contours are shown superimposed on the blurred images, and have 14, 26, 49, 83, 120, 162 knots, respectively.

2. Alternatively, the control points $\mathbf{v}_i$ can be expressed by their normalized coordinates $\tilde{\mathbf{v}}_i$, and all image related differential measurements can be *scaled*, like the scale-space gradient and Hessian presented in chapter 3, section 3.3.2 in the Euclidean or Riemannian setting. This naturally leads to an adjustment of the image scale-space density, which can be efficiently precomputed. Fixing the spacing of the normalized control points $\tilde{\mathbf{v}}_i$ for all scales naturally leads to a less dense spacing of the control points $\mathbf{v}_i$ at higher image scales, but no loss of spatial information.

In this dissertation the second approach is adopted, which relates the spacing of the contour control points $\mathbf{v}_i$ to the image scale. Figure 6.2 illustrates a direct scale-related sampling scheme for the *notched rectangle* for six samples of an image scale-space, setting $\varsigma = \sigma$ at each scale level, where ς represents the spacing of the control points $\mathbf{v}_i$. In order to illustrate the usefulness of this scheme, figure 6.3 shows the extreme case of segmenting a contour of low scale in a high scale image, and vice versa. In the former case, the low contour scale model in figure 6.3 (a) yields a very similar result to the high contour scale model once more shown in figure 6.3 (b) (both were optimized at the same high image scale), thus the much larger number of control points only contributes to an increase in *redundant information*, rather than in a more accurate segmentation result. In contrast to that, the contour model in figure 6.3 (c) has a too high contour

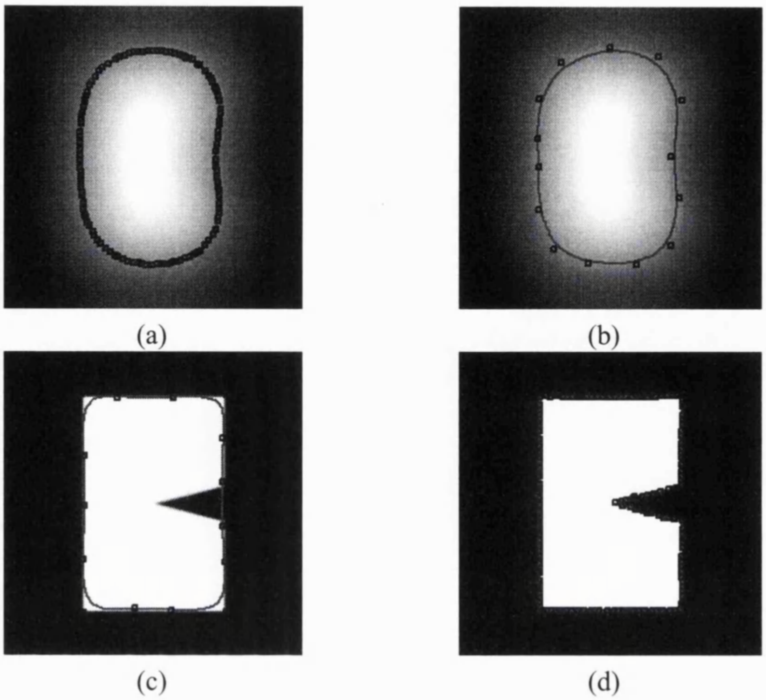


Figure 6.3: Snake optimization for low contour scale and high image scale and vice versa. Results for (a) $\sigma = 32$ and $\varsigma = 3.5$, with 121 knots (b) Same as figure 6.2 (a). (c) $\sigma = 1$ and $\varsigma = 32$, with 15 knots. (d) Same as figure 6.2 (f). All contours are superimposed on the corresponding image scale-space slice.

scale and therefore cannot extract the object properly. The corresponding low contour scale result is shown in figure 6.3 (d).

6.1.2.1 Knot Insertion Techniques

Before the equivalent multi-scale techniques to the fixed-scale sampling strategies of section 6.1.1 will be presented, the issue of adequate spline control point or *knot* insertion needs to be briefly addressed. As was mentioned above, the classic snake model by [Kass *et al.*, 1987a] as well as its extension by [Williams and Shah, 1992] enforce uniform *snaxel* spacing in an *intrinsic* manner, as it is controlled by the energy function rather than by some external adjustment. If the number of *snaxels* needs to be locally increased, e.g. as is the case for the variable and refinement sampling strategies above, sampling is performed in an *extrinsic* fashion. Moreover, using a B-spline representation, it is desirable that the approximated spline contour changes as little as possible. Three different B-spline based techniques for knot insertion have been developed in this dissertation: *discrete*, *interpolated*, and *heuristic*.

Figure 6.4 illustrates all three concepts for knot insertion for three spline patches defined by 10 points in total with quadruple end and start points (6.4 (a)). When using the discrete insertion scheme (6.4 (b)), one can observe that the new two spline patches, replacing the middle original

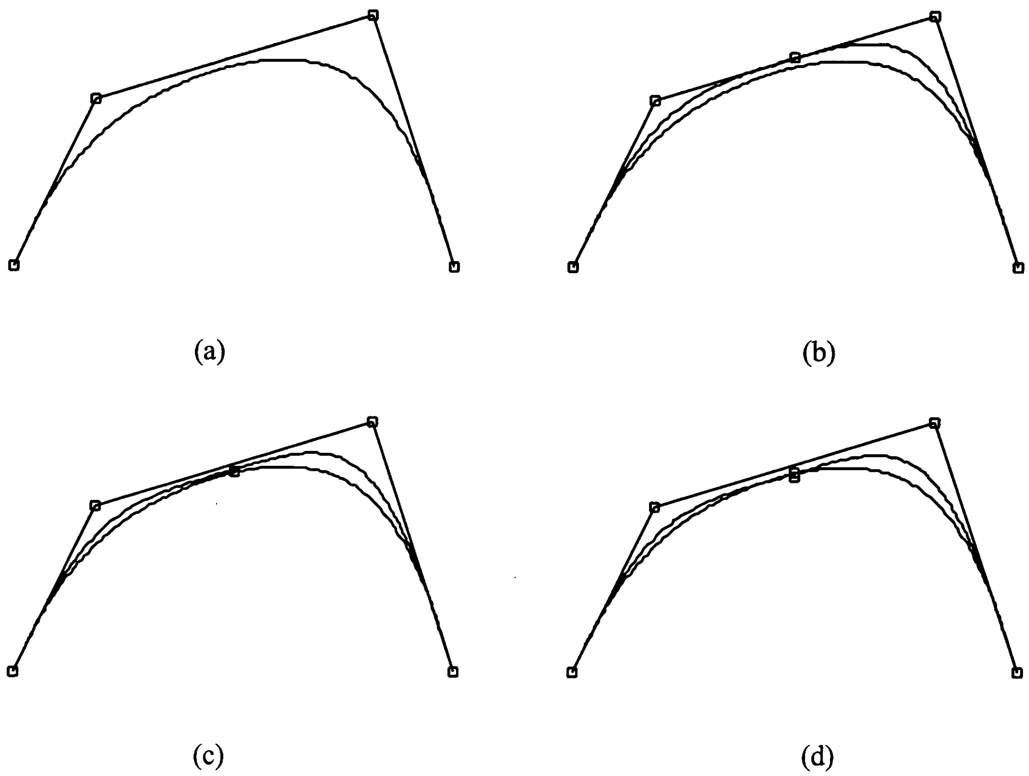


Figure 6.4: Different B-spline knot insertion strategies. (a) Three spline patches defined by 10 points, with quadruple end and start points. The four control points of the middle spline patch are $\mathbf{v}_{i-1} = (0,0), \mathbf{v}_i = (0.25, 0.4), \mathbf{v}_{i+1} = (0.75, 0.55), \mathbf{v}_{i+2} = (1,0)$. b) Discrete knot insertion of point $\mathbf{v}_{i,i+1} = (0.5, 0.475)$. (c) Interpolated knot insertion of point $\mathbf{v}_{i,i+1} = (0.5, 0.459375)$. (d) Heuristic knot insertion of point $\mathbf{v}_{i,i+1} = (0.501731, 0.451757)$.

patch, move very close to the polygon connecting the control points. Using the interpolated or heuristic scheme (6.4 (c) and (d)), the deviation between the new splines and the old one can be decreased, with the heuristic scheme yielding only a slightly better result. However, it is always possible to find a better parameterization when choosing a suitable knot for the parameterization [Farin, 1993]. With respect to the presented knot insertion schemes, one can improve the shape of the curve at the expense of computation time, by the following hierarchy of methods: discrete, interpolated, and heuristic, with probably the best compromise between computational cost and quality being achieved using the interpolated method. In the following, the three knot insertion techniques are explained in more detail.

Discrete knot insertion between two neighbouring control points $\mathbf{v}_i, \mathbf{v}_{i+1}$ is performed by computing the new control point $\mathbf{v}_{i,i+1} = \frac{\mathbf{v}_i + \mathbf{v}_{i+1}}{2}$, i.e. the point on the middle of the line between $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$. This might be in many cases a reasonable choice, especially if the control points interpolate the curve. This scheme has the advantage that it is very fast, needing no spline approximation for the insertion. However, this insertion scheme might lead to *short cuts* at highly curved parts of the overall shape.

Interpolated knot insertion is achieved by by setting $\mathbf{v}_{i,i+1} = \mathbf{v}_i(s_i)$ at $s_i = 0.5$. This results in a smoother spline than using the discrete insertion scheme, with a curve closer to the curve without the new point.

Heuristic knot insertion is directly based on the interpolated insertion scheme, and uses a curve fitting approach in order to change the shape of the overall curve only minimally. This is achieved by first interpolating a point $\mathbf{p} = \mathbf{v}_i(s_i)$ at $s_i = 0.5$, followed by computing the spline patches $\mathbf{v}_i^{(1)}(s_i^{(1)}), \mathbf{v}_i^{(2)}(s_i^{(2)}), \mathbf{v}_i^{(3)}(s_i^{(3)})$ (where superscripts are used for enumeration of the splines and their parameters). These spline patches are defined by the control points $\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{p}, \mathbf{v}_{i+1}$, and $\mathbf{v}_i, \mathbf{p}, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}$, and $\mathbf{p}, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}, \mathbf{v}_{i+3}$, respectively. Additionally, the starting and end points of the original splines are computed, with $\mathbf{p}_{start}$ at $\mathbf{v}_i(0)$, and $\mathbf{p}_{end}$ at $\mathbf{v}_{i+1}(0)$. The corresponding points to $\mathbf{p}_{start}, \mathbf{p}, \mathbf{p}_{end}$ on the new spline patches are then given by $\mathbf{p}'_{start}$ which is interpolated at $\mathbf{v}_i^{(1)}(0)$, $\mathbf{p}'$ at $\mathbf{v}_i^{(2)}(0)$, and $\mathbf{p}'_{end}$ at $\mathbf{v}_i^{(3)}(0)$. Computing the following distance,

$$D_{spline} = \|\mathbf{p}_{start} - \mathbf{p}'_{start}\|^2 + \|\mathbf{p} - \mathbf{p}'\|^2 + \|\mathbf{p}_{end} - \mathbf{p}'_{end}\|^2 \tag{6.4}$$

allows for an estimate of the closeness of the new spline patches to the original one. Obviously, more than three pairs of points can be taken into account, leading to a better fitting, but also to higher computational cost. Moving now $\mathbf{p}$ along the normal $\mathbf{n}(0.5)$ of the original spline allows to compute new spline patches, from which new points $\mathbf{p}'_{start}, \mathbf{p}', \mathbf{p}'_{end}$ can be obtained and evaluated using equation (6.4). The point along the normal which minimizes the distance between the original spline patch and the new two spline patches is taken as the new control point for insertion. This is a general scheme that allows for insertion of arbitrary points along the spline. One can then restrict the search space along the spline normal using the formula

$$\Delta = \begin{cases} s_i & \cdot\|\mathbf{v}_i - \mathbf{v}_{i+1}\| \quad \text{if} \quad s_i \in [0; 0.5) \\ (1 - s_i) & \cdot\|\mathbf{v}_i - \mathbf{v}_{i+1}\| \quad \text{if} \quad s_i \in [0.5; 1) \end{cases} \tag{6.5}$$

as points for s_i closer to $s_i = 0$ or approaching $s_i = 1$ lie closer to $\mathbf{p}_{start}$ or $\mathbf{p}_{end}$, respectively.

Having now set up a relationship between contour scale ς and image scale σ, and defined suitable knot insertion schemes, the question arises how existing fixed-scale sampling schemes like the ones in section 6.1.1 can be adapted and generalized to multi-scale sampling strategies. In particular, uniform, variable, and refinement sampling strategies are considered in the following.

6.1.2.2 Uniform Sampling

Let us first consider the approach by [Williams and Shah, 1992], which was developed in order to enforce uniform spacing of the *snaxels* to prevent snaxels from bunching up at strong portions of

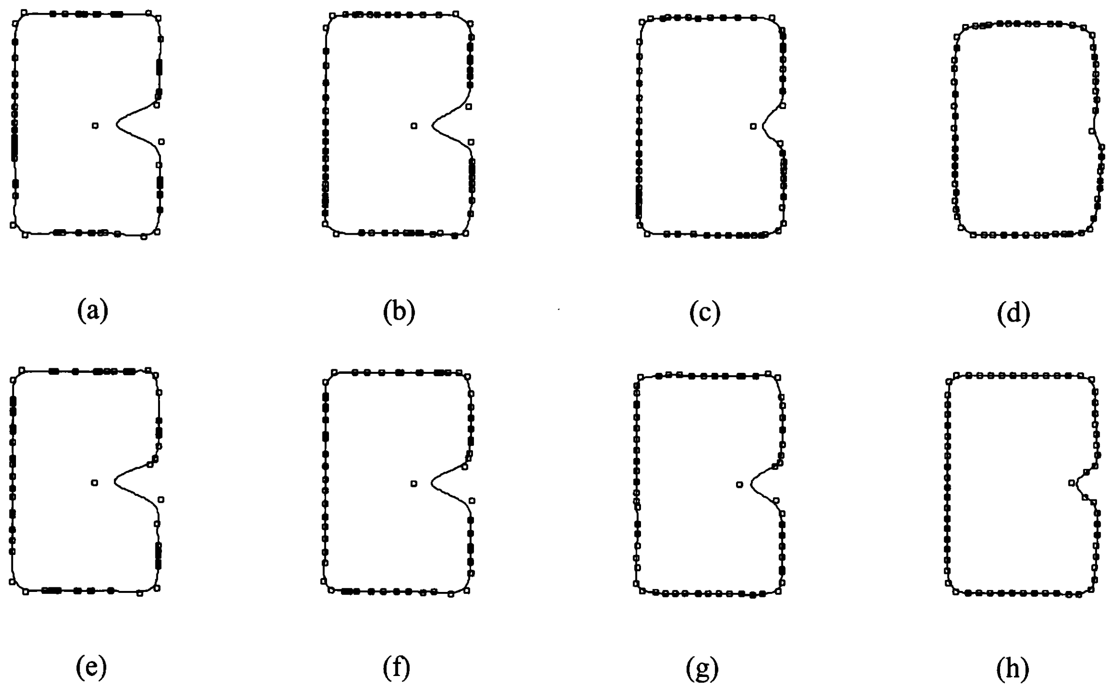


Figure 6.5: Optimization results for the models of Kass et al. (top) and Williams and Shah (bottom). The model of figure 6.1 (d) with 65 knots and $\varsigma = 8$ was optimized for (from left to right) $\alpha_{elasticity} = 0.01, 0.001, 0.0001, 0.00001$ at scale $\sigma = 8$. See text for further details.

an edge (see equation (4.6)). Using such a mean distance constraint, however, only yields evenly spaced points if the influence of the elasticity term in terms of the size of the energy weighting factor $\alpha_{elasticity}$ is chosen adequately. A similar result can be produced for the classic model by [Kass *et al.*, 1987a]. If this weighting factor is chosen too high, the shape cannot be accurately located since the control points are too constrained to deform sufficiently. In contrast, if the elasticity weighting term is chosen too low, the control points tend to move together at edges, while moving apart at other parts of the contour despite the mean distance constraint. Hence a compromise between accuracy of the optimization and uniformity of the control point distribution must be made. Figure 6.5 illustrates this problem for the models of [Kass *et al.*, 1987a] and [Williams and Shah, 1992]. It can be observed that the latter performs only slightly better than the former model in terms of enforcing a uniform spacing of control points. Another problem with this scheme may occur when the mean distance of the control points during the optimization increases or decreases, depending on whether the model expands or shrinks when deforming. Given a very dense initial model this may lead to the formation of intersections and an inability to deform. Using an initial model with only few points may lead to a too sparsely sampled model which is unable to locate the shape of interest at the level of detail needed. The appropriate control point distance therefore needs to be chosen with respect to the complexity of the shape of interest. The classic model suffers from similar problems as the number of control points also

Properties	Uniform	Adaptive Uniform	Variable	Refinement
Intrinsic sampling	$-$	$-$	$-$	$-$
Adaptive number of points	$+$	$+$	$+$	$+$
Adaptive level of detail	$\square$	$\square$	$+$	$+$
Local refinement	$-$	$\square$	$+$	$\square$
Efficiency	$\square$	$\square$	$\square$	$+$
Flexibility	$-$	$+$	$\square$	$\square$
Accuracy	$\square$	$+$	$\square$	$+$
Notation	ς	ς	$\varsigma(s)$	$\varsigma_{\tilde{\sigma}}$
Sampling schedule	σ	$\sigma \pm \Delta\sigma$	$\frac{1}{\lvert\tilde{\kappa}(s)\rvert}$ $\frac{1}{\lvert\tilde{\kappa}_{image}(s)\rvert}$ $\frac{1}{\lvert\tilde{\kappa}(s)\tilde{\kappa}_{image}(s)\rvert}$	$\sigma_{\tilde{\sigma}}$

Table 6.2: Qualitative comparison, notations, and sampling schedules of multi-scale sampling strategies for deformable models. All schemes are based on a natural coordinate representation of a B-spline which is related to the underlying image scale. *Uniform* sampling refers the redistribution of the spline control points on the approximated contour, *adaptive uniform* sampling to the discrete, interpolated or heuristic knot insertion or removal scheme, *variable* sampling to the local contour refinement as a function of multi-scale contour and image curvature, and *refinement* sampling to a *coarse-to-fine* approach based on *adaptive uniform* sampling. The symbols $+$, $-$, $\square$ denote fulfilment, no fulfilment, and partial fulfilment of the listed properties.

remains fixed during the optimization process.

In order to overcome the first problem, in this dissertation the mean distance is fixed to a desired value rather than fixing the number of control points, enforcing an even spacing of the control points in terms of this fixed distance. Obviously, an appropriate contour point spacing is related to the underlying image scale, leading to an image scale-related *natural contour scale* throughout the optimization process, while the number of control points may increase or decrease. Two main sampling strategies to adjust the spacing of the control points $\mathbf{v}_i$ have been developed in this dissertation: *uniform sampling* and *adaptive uniform sampling*. An qualitative overview of all schemes is given in table 6.2.

Uniform sampling: After each complete iteration step during the optimization process, the contour is approximated via its spline representation, and at equidistant intervals (corresponding to the image scale), points are marked as the new control points, while the previous control points are discarded. Figure 6.6 (a) shows an example of uniform sampling with very regular spacing using only a low elasticity weighting term ($\alpha_{elasticity} = 0.00001$), but which also illustrates the main problem with such a scheme: The model is not able to locate finer detail structures, like the corners and the inward notch of the notch in the test image, as details are prevented from forming due to the frequent redistribution of control points. Moreover, the approximating rather than interpolating nature of B-splines leads to

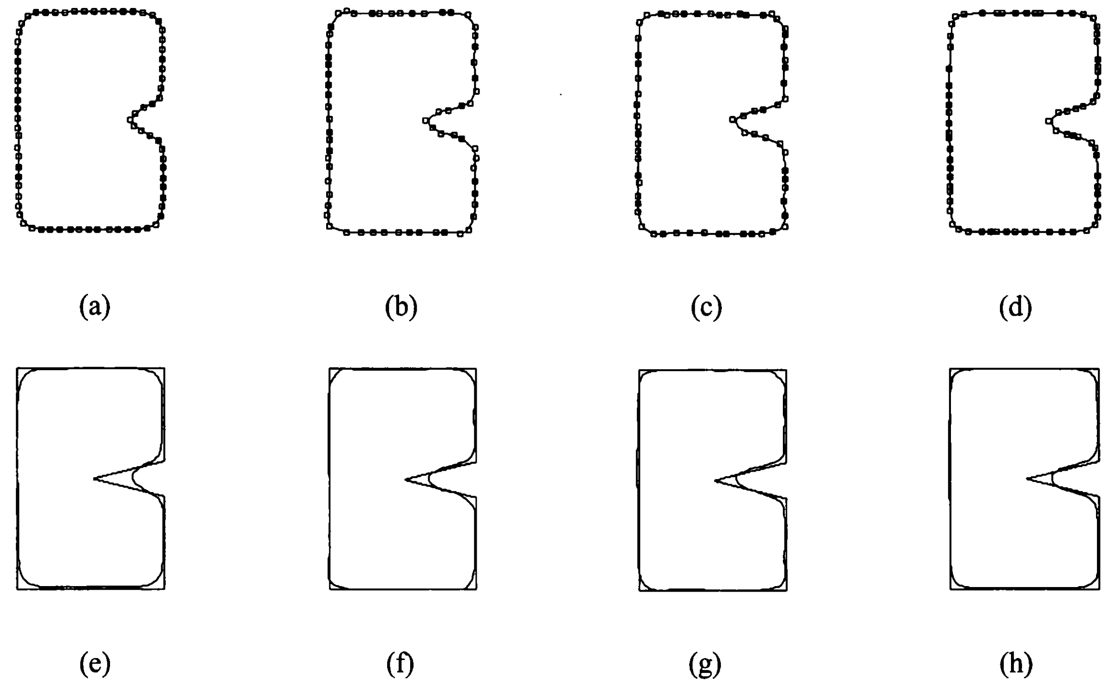


Figure 6.6: Uniform B-spline sampling strategies for active contour model optimization. (a) Uniform sampling (with 75 final control points). (b) Adaptive discrete sampling (65 control points). (c) Adaptive interpolated sampling (67 control points). (d) Adaptive heuristic sampling (73 control points). Figures (e)-(h) show once more the optimization results from figures (a)-(d) with the known shape outline from figure 6.1 (b) superimposed. The contour scale of the initial model of figure 6.1 (d)) was $\varsigma = 8$, in direct relation to the image scale, with a tolerance of $\Delta\varsigma = \pm\frac{\varsigma}{2}$ for the adaptive schemes.

a shrinkage of the overall model, as the control points do not necessarily lie on the spline curve, while for straight, linear contour parts or a very high chosen model scale the redistribution of the control points along the spline contour causes only minor changes in the overall shape.

Adaptive uniform sampling: This scheme is based on control point insertion and removal rather than redistribution, and has the advantage that control points of adequate spacing are kept intact, and control points whose spacing is too low are removed. In case they are too far apart, new control points are inserted on the basis of the spline defined by the existing control points using one of the knot insertion strategies presented in section 6.1.2.1. Algorithm 6.1 illustrates this scheme in a simplified form (as the number of control points increases or decreases during the sampling process, and several new control points between a pair of control points may be inserted). Figure 6.6 (b), (c), and (d) shows example optimization results using this algorithm based on *discrete*, *interpolated*, and *heuristic* knot insertion. All results are superior to uniform sampling (shown in figure 6.6 (a)), and the interpolated and heuristic insertion schemes allow to locate corners more accurately than the discrete

```
// Loop for adaptive sampling of N control points (modulo arithmetic) with respect to ς
for i = 1 to N do
    // Compute spatial distance
    set distance := ||v_i − v_{i+1}||
    // Insert control point if spacing is too large
    if (distance >= ς + Δς) do
        Insert (v_{i,i+1}) // Insert knot (new control point) between v_i, v_{i+1}
    // Remove control point if spacing is too small
    else if (distance <= ς − Δς) do
        Remove (v_i)
    end if
end for
```

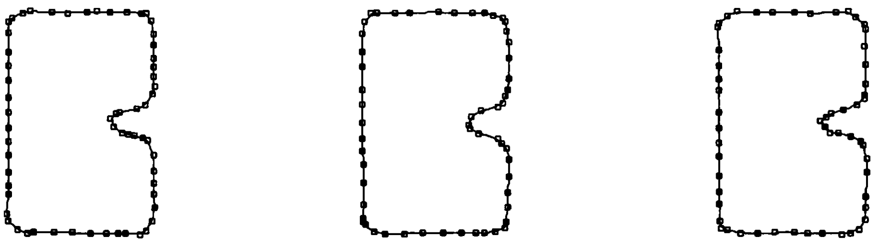
Algorithm 6.1: Algorithm for adaptive uniform sampling.

scheme.

6.1.2.3 Variable Sampling

The previous section has focused on uniform sampling techniques which, though solving the problem of determining the number of points needed, still disregard the problem of choosing the appropriate *local* scale of the model. Choosing a too low value leads to a higher vulnerability of the model towards the formation of intersections and local minima, while choosing too large a value leads to short cuts and an inability of the model to detect structure below the chosen scale. Clearly, there is no optimal global sampling distance for shapes, but one would like to have a locally sparse representation at smooth parts of the shape, and a denser representation for more complex parts, where the *sampling density* is related to the degree of shape complexity and image scale. In other words, given a fixed-scale setting, one would like to have a more flexible contour scale with respect to local shape detail. This leads to the formulation of three variable sampling strategies, which are based on the spline curvature, the underlying isophote image curvature, and both:

Adaptive spline curvature sampling: This scheme is based on the scheme by [Bulpitt and Efford, 1995] to refine the active contour with respect to its local complexity in terms of its internal curvature. This can be easily incorporated into the adaptive sampling algorithm (algorithm 6.1) by replacing the constant contour scale parameter ς with a variable term $\varsigma(s)$, being a function of the local curvature $\kappa(s)$, e.g. an intuitive way of doing this is to set $\varsigma(s) = \frac{1}{|\kappa(s)|}$. In other words, for each spline patch, the inverse absolute value for $\kappa(s)$ is computed as the interpolated curvature on the middle of the spline patch ($s = 0.5$). For small contour curvature values, this results in a high contour scale (or low contour resolution) which is sufficient for the representation of smooth shape outlines of low curvature, whereas for high contour curvature values, the local contour scale decreases, leading to a



(a) (b) (c)

Figure 6.7: Different curvature-based B-spline sampling strategies for active contour model optimization. (a) Adaptive spline curvature sampling (with 62 final control points). (b) Adaptive image curvature sampling (57 control points). (c) Adaptive spline and image curvature sampling (57 control points). The distance stabilizers were chosen to be $\epsilon_1 = \frac{1}{2\sigma_0}$ (to allow for maximum sampling of $\varsigma = 16$) and $\epsilon_2 = 2$. Optimization was carried out using the initial model of figure 6.1 (d) of $\varsigma = 8$ at $\sigma_0 = 8$.

higher local flexibility needed to adjust to finer details and complex structure. Care must be taken when the local curvature approaches 0 or very high values. In this case the range of resulting distances must be restricted to a maximum and minimum local scale (which can be expressed by distance stabilizers ϵ_1 and ϵ_2, respectively):

$$\varsigma(s) = \frac{1}{|\kappa(s)|^m + \epsilon_1} + \epsilon_2 \tag{6.6}$$

with m chosen to be $m < 1$ to stretch distance values locally, and $m > 1$ for greater local compactness. A typical value is $m = 0.5$. Equation (6.6) can be related to the image scale σ_0 by setting $\epsilon_1 = \frac{1}{\sigma_0}$. This ensures that for very low curvature values, the image scale corresponds to the upper contour scale limit. ϵ_2 is accordingly set to some lower contour scale limit for very high curvature values. Figure 6.7 (a) shows an example optimization using this variable curvature-based scheme. As can be seen, the distance of the control points at straight parts of the contour is larger than at corners, leading to a more flexible (since dense) representation at more detailed parts of the shape. However, this sampling scheme is only indirectly dependent on the image context, leading to an inability of the model to deform in lack of strong enough image forces.

Adaptive image curvature sampling: This is a more consistent approach which is based on the isophote image curvature $-\frac{L_{vv}}{L_w}$ (see chapter 3) at the point $\mathbf{v}(0.5)$. Replacing $\kappa(s)$ in equation (6.6) by

$$\kappa_{image}(s) = -\frac{L_{vv}(\mathbf{v}(s))}{L_w(\mathbf{v}(s))} \tag{6.7}$$

relates the local sampling distances directly to the image context of the shape to be extracted. At image parts of low image curvature, only few points are needed, and at parts of higher image curvature, more points are inserted to decrease the local distance and to increase the local flexibility of the spline patches. Figure 6.7 (b) shows the result of the optimization using this scheme. Again, corners are represented more densely than straight parts of the shape, but problems occur at low curvature parts of the image. In particular, the low image curvature region just before the notch of the test image causes problems, as the sampling is not dense enough.

Adaptive spline and image curvature sampling: This is an alternative approach combining internal contour curvature and external image curvature, e.g. by multiplication: $(|\kappa(s)| \cdot |\kappa_{image}(s)|)^m$. Figure 6.7 (c) illustrates an example using this combination for adaptive curvature-based sampling, yielding a better result at the notch of the test image.

However, all three presented approaches for variable, curvature-based sampling suffer from two major disadvantages: first, the active contour model might not gain higher local curvature and corresponding lower scale at finer detailed shape structure *unless* it is densely sampled and of already high curvature (the inverse problem), and second, this scheme uses two parameters ϵ_1 and ϵ_2 which are related to the maximum and minimum desired contour scale ς, and whose choice is not quite obvious. In particular, the dependency of ϵ_1 of the image scale needs further investigation. In the following, a hierarchical refinement will be discussed which attempts to overcome both problems.

6.1.2.4 Refinement Sampling

[Miller *et al.*, 1991] have developed a hierarchical scheme for subsequent refinement. Starting with an initial coarse model a refinement or subsampling step is performed after each iteration of the optimization until the desired resolution is reached for all snaxels. This refinement sampling strategy can be easily integrated into the spline based active contour model, by inserting after each iteration new control points at $s = 0.5$ between the existing ones using any of the schemes for knot insertion presented in section 6.1.2.2. The adaptive sampling algorithm (algorithm 6.1) can be easily modified by halving the contour scale value ς for each subsequent refinement step. This simultaneously enforces equidistant spacing after each subdivision step. However, as figure 6.3 illustrates, it is not desirable to decouple the contour scale from the image scale. At high scales, subsampling will only lead to redundant refinement, and at low scales, a coarse initial model may lead to inaccuracy in the localization. In fact, sections 6.1.2.2 and 6.1.2.3 have already presented scale-based sampling schemes and illustrations in a fixed-scale setting, which will now be extended to a multi-scale setting.

Integrating a hierarchical contour refinement step such as the one proposed by [Miller *et al.*, 1991] into an image scale-space setting allows to relate the contour scale directly to the image scale. A natural image scale-space can be computed via exponential scale sampling, e.g.

$$\sigma_{\tilde{\sigma}} = \epsilon \cdot e^{\tilde{\sigma}} \tag{6.8}$$

with natural scale $\tilde{\sigma}$ and hidden scale parameter ϵ (see equation (3.22) in chapter 3), or more generally by

$$\sigma_i = \sigma_0 \cdot f^i \quad \text{with} \quad f = \left(\frac{\sigma_{n-1}}{\sigma_0}\right)^{\frac{1}{n-1}} \tag{6.9}$$

where the scale change factor f defines the ratio of the outer scale σ_{n-1} and inner scale σ_0. Natural sampling can be obtained by setting $\epsilon = \sigma_0$. Equations (6.8) and (6.9) can be used to relate the contour scale ς defined by the distance between the contour control points *directly* to the image scale σ, e.g. by setting

$$\varsigma_{\tilde{\sigma}} = \sigma_{\tilde{\sigma}} \quad \text{or} \quad \varsigma_i = \sigma_i \ , \tag{6.10}$$

respectively. Note that this actually is equivalent to using normalized coordinates and scaled image features. For practical reasons, ς is defined with an offset relationship by setting

$$\varsigma_{\tilde{\sigma}} = \epsilon_1 \cdot \sigma_{\tilde{\sigma}} + \epsilon_2 \quad \text{or} \quad \varsigma_i = \epsilon_1 \cdot \sigma_i + \epsilon_2 \tag{6.11}$$

with scale stabilizers ϵ_1 and ϵ_2, whose meaning can be interpreted in the following way: ϵ_1 indicates the general *agreement* between distance and scale especially at high scales, and a logical choice is to set $\epsilon_1 = 1$ or only slightly smaller. ϵ_2 causes a fixed offset whose influence is especially high at very small scales, causing a minimum distance between the control points to prevent intersections. A reasonable choice for ϵ_2 is a value between 1.5 and 3.5, depending on the size of the *search space* (the local neighbourhood within which each control point is allowed to move during an optimization process). Hence the role of ϵ_1 has changed from its previous scale relationship in section 6.1.2.3 to a proportional constant, while the meaning of ϵ_2 remains unchanged.

Going back to figure 6.2 at the beginning of this chapter, one can see that this scheme has been applied in terms of choosing the scale change factor as $f = 0.5$ for equation (6.9), with stabilizers $\epsilon_1 = 0.8$ and $\epsilon_2 = 3.5$, $n = 8$ scale samples and final and starting scales $\sigma_{n-1} = 1$ and $\sigma_0 = 32$, respectively. Sampling was performed using algorithm 6.1 replacing the constant contour scale ς by a scale-sampled ς_i.

6.1.3 Summary

This section has presented several scale-based sampling schemes for a spline-based active contour model. Notably, a general adaptive sampling algorithm was presented which can be applied to fixed-scale as well as multi-scale sampling strategies. It enforces adaptive uniform or variable spacing of the control points with respect to a given contour scale ς which is dependent on the image scale σ. Adaptive sampling is performed as a spline knot removal and insertion process, with a tolerance span to increase local variability. Three knot insertion schemes have been developed and investigated for this purpose. From the conducted experiments the interpolated knot insertion scheme was found to be the most suitable one in terms of its superiority over the discrete scheme, and its lower computational complexity in comparison with the heuristic insertion strategy. Fixed-scale adaptive uniform sampling prevents the control points from drifting too far apart or too close together and enforces equidistant spacing with respect to the fixed image scale. It generally performs better than uniform sampling based on redistributing new control points on the approximated spline contour. Fixed-scale adaptive variable sampling distributes the control points with respect to the shape complexity which is also limited by the image scale. In a multi-scale setting, adaptive uniform or variable sampling allow to locate shape details at multiple levels of image detail, and to refine the contour representation by a scale-related subsampling.

The main advantages of scale-based adaptive sampling are the following:

- At high image scales or smooth shape outlines, fewer control points are needed for an adequate contour representation and accurate shape localization. In particular, at a large image scale, the image becomes less dense, thus choosing an appropriate larger control point spacing is sufficient. Computational efficiency is improved by removing redundant control points.
- Increasing the distance between control points has a smoothing effect on the contour, a process that is similar to the underlying image smoothing process. Inserting new control points increases the contour flexibility and allows to track finer details at lower scales or more complex parts of the shape.
- The normalized representation of the spline contour as $\tilde{\mathbf{v}}(s)$ allows to relate a natural contour scale $\tilde{\varsigma}$ to the natural underlying image scale $\tilde{\sigma}$ in an adaptive sampling process. This provides the means to automatically derive the number of control points needed for the shape extraction, an issue that so far has never been investigated, as the choice of the number and related spacing of control points as well as the choice of the image scale is usually performed empirically.

Adaptive uniform sampling using the interpolated knot insertion strategy in a fixed-scale as well as multi-scale setting will be adopted as a pragmatic solution for the sampling problem in active contour models. Variable sampling based on curvature properties of the contour or image was found unsuitable due to the inverse nature of the problem. Further investigation of sampling schemes go beyond the scope of this dissertation, but a multi-scale variable sampling scheme for active contour model optimization, selecting locally adequate image scales and related contour scales, will be developed and discussed in chapter 10.

6.2 Energy Function

The spline representation and associated multi-scale sampling schemes presented in section 6.1 form the basis for the multi-scale energy function and multi-scale optimization of the spline-based active contour model developed in this chapter. The spline representation not only allows for the continuous computation and evaluation of internal and external energy terms, but also for analytic differentiation of the contour for the internal elasticity and curvature energy computation. The spline-based formulation of internal and external energy terms will be presented in the following. In particular, the incorporation of scale continuation, and a novel curvature matching process will be presented. To ease the understanding of the multi-scale setting of the model, a normalized representation $\tilde{\mathbf{v}}(s) = (\tilde{x}(s), \tilde{y}(s))$ is chosen, with $\tilde{\mathbf{v}}(s) = \left(\frac{x(s)}{\sigma_{\tilde{\sigma}}}, \frac{y(s)}{\sigma_{\tilde{\sigma}}}\right)$.

6.2.1 Internal Energy Terms

The *elasticity* energy term is traditionally a first-order tension term of the contour, controlled by a weighting factor $\alpha_{elasticity}$ which enforces uniform sampling and smoothness of the contour outline. Section 6.1.2.2 discussed the selection of a suitable weighting factor and the tradeoff between smoothness and accuracy of the contour. The presented adaptive uniform resampling scheme addresses these issues in a multi-scale setting, giving rise to a multi-scale elasticity energy defined by

$$\tilde{\mathcal{E}}_{elasticity}(\tilde{\mathbf{v}}(s)) = (\tilde{\mathbf{v}}_s(s))^2 \tag{6.12}$$

normalized with respect to the normalized contour scale $\tilde{\varsigma}$ and the associated image scale $\sigma_{\tilde{\sigma}}$. The analytic normalized elasticity is given by the first order derivatives of the normalized coordinate functions:

$$\tilde{\mathbf{v}}_s(s) = \sqrt{\tilde{x}_s(s)^2 + \tilde{y}_s(s)^2} \tag{6.13}$$

The influence of this term can be chosen very low as the adaptive uniform resampling additionally enforces uniform spacing of the spline control points.

The *bending* energy of the classic model is defined as the squared second derivative $\mathbf{v}_{ss}^2(s)$. [Williams and Shah, 1992] pointed out that this term only holds for arc length parameterization and strictly uniform spacing of the discrete set of contour points. Otherwise the absolute curvature is given by equation (4.7) (chapter 4, section 4.2.1). The bending energy of the multi-scale active contour model is given by

$$\tilde{\mathcal{E}}_{bending}(\tilde{\mathbf{v}}(s)) = \tilde{\kappa}^2(s) \tag{6.14}$$

where the curvature is computed by equation (2.29) (chapter 2), or in a scale setting by the normalized derivatives of the contour:

$$\tilde{\kappa}(s) = \frac{\tilde{x}_s(s)\tilde{y}_{ss}(s) - \tilde{x}_{ss}(s)\tilde{y}_s(s)}{(\tilde{x}_s(s)^2 + \tilde{y}_s(s))^{\frac{3}{2}}} \tag{6.15}$$

In general, curvature minimization might not be desirable when segmenting shapes with high curvature parts. In the following, a *curvature matching process* is developed which instead adjusts the contour curvature to the underlying image shape.

6.2.1.1 Curvature Matching Process

To avoid minimizing of the contour curvature at high curvature parts of the shape to be segmented, a novel curvature process has been developed which adjusts the contour curvature $\kappa(s)$ to the underlying isophote image curvature, which is computed as $-\frac{L_{vv}}{L_w}$ along the contour $\mathbf{v}(s)$. Hence, not the contour curvature is minimized, but its *deviation* from the underlying image curvature. Note that the *scale-space isophote curvature* has to be computed via the normalized partial derivatives of the image L, yielding $-\frac{\tilde{L}_{vv}}{\tilde{L}_w}$. This term is obtained by multiplying the first order derivatives of the image with σ, and the second order derivatives with σ^2. The energy term enforcing the curvature matching can then be formulated as

$$\tilde{\mathcal{E}}_{bending}(\tilde{\mathbf{v}}(s)) = (\tilde{\kappa}(s) \pm \tilde{\kappa}_{image}(s))^2 \tag{6.16}$$

with

$$\tilde{\kappa}_{image}(s) = -\frac{\tilde{L}_{vv}(\tilde{\mathbf{v}}(s))}{\tilde{L}_w(\tilde{\mathbf{v}}(s))} \tag{6.17}$$

This is a novel approach, not only because the relationship between these two curvature measurements has to our knowledge not been investigated in the context of active contour models, but also in respect to the traditionally strict separation between internal and external terms. The combination of these terms implements a local, non-parametric scale-based constraint into the no longer autonomous shape forces. Most importantly, not only the absolute curvature values of contour and image are matched, but also their *curvature behaviour*. The use of the $-\frac{\tilde{L}_{vv}}{\tilde{L}_w}$ operator is now further investigated in terms of its sign (denoted by the $\pm$ operator in equation (6.16)), its robustness, its performance in the shape extraction process, and its dependence on contour and image scale:

Sign: The sign of the scale-space isophote image curvature depends on the chosen normal direction in the local gauge coordinate system, i.e. whether the normal to the isophote is pointing inward or outward. This choice is directly related to the image *contrast*, or whether the shape to be extracted is lighter or darker than the background. For light shapes on dark background, $-\frac{\tilde{L}_{vv}}{\tilde{L}_w}$ is chosen, otherwise $\frac{\tilde{L}_{vv}}{\tilde{L}_w}$.

Robustness: The robustness of the isophote curvature operator can decrease in areas of low gradient values, expressed by $\tilde{L}_w$ approaching zero. [Gerig *et al.*, 1995] have presented a solution to this problem which is adopted in this dissertation. They suggest to compute the isophote curvature with lower limits to the partial derivatives L_x and L_y (the scale-space

Curvature matching	$\sigma = \varsigma$	RMS error
On	32	0.038642
Off	32	0.037548
On	16	0.132589
Off	16	0.111322
On	8	0.140109
Off	8	0.502091
On	4	0.429516
Off	4	1.010320
On	2	1.740510
Off	2	8.430570

Table 6.3: Curvature deviation results in pixel units with and without curvature matching computed as root-mean-squared errors for decreasing image and contour scales.

derivatives $\tilde{L}_x$ and $\tilde{L}_y$ in this case) to avoid numerical instability. A parameter ϵ governs the stabilizing viscosity effect for weak solutions ([Evans and Spruck, 1991]). In 2D, the following image feature is computed:

$$-\frac{\tilde{L}_{vv}}{\tilde{L}_w} = -\frac{(\tilde{L}_x + \epsilon)^2 \tilde{L}_{yy} + (\tilde{L}_y + \epsilon)^2 \tilde{L}_{xx} - 2\tilde{L}_x \tilde{L}_y \tilde{L}_{xy}}{((\tilde{L}_x + \epsilon)^2 + (\tilde{L}_y + \epsilon)^2)^{3/2}} \tag{6.18}$$

For low values of $\tilde{L}_x$ and $\tilde{L}_y$, this expression approaches $\tilde{\Delta} L = \tilde{L}_{xx} + \tilde{L}_{yy}$, the scale-space Laplacian. A reasonable choice is $\epsilon = 1$. Similarly, the normalized contour curvature is computed using the same stabilizer:

$$\tilde{\kappa}(s) = \frac{(\tilde{x}_s(s) + \epsilon)\tilde{y}_{ss}(s) - \tilde{x}_{ss}(s)(\tilde{y}_s(s) + \epsilon)}{((\tilde{x}_s(s) + \epsilon)^2 + (\tilde{y}_s(s) + \epsilon))^{\frac{3}{2}}} \tag{6.19}$$

Performance: Figures 6.8 and 6.9 illustrate the deviation between contour and image curvature for the optimization of the *notched rectangle* test image at decreasing levels of image scale and associated contour scale, using the bending energy terms based on equations (6.16) (for curvature matching) and (6.14) (for curvature minimization), respectively. The high positive curvature peak which can be observed in these figures describes the notch of the notched rectangle test image, and the negative peaks describe the corners of the rectangle, as well as the corners at the entrance to the notch. Differences in curvature deviation are subtle and need to be further quantified in terms of the overall error. Table 6.3 lists the root-mean-squared (RMS) error for both cases. It can be observed from the figures and the associated errors listed in the table that at high scales the curvature deviation is slightly lower when no curvature matching is performed. This is due to the fact that the image object is less complex at higher scales, and therefore minimizing the bending is almost equivalent to matching the bending to the underlying low image curvature. At higher scales, however, the RMS error for the curvature minimization becomes much higher than for the curvature

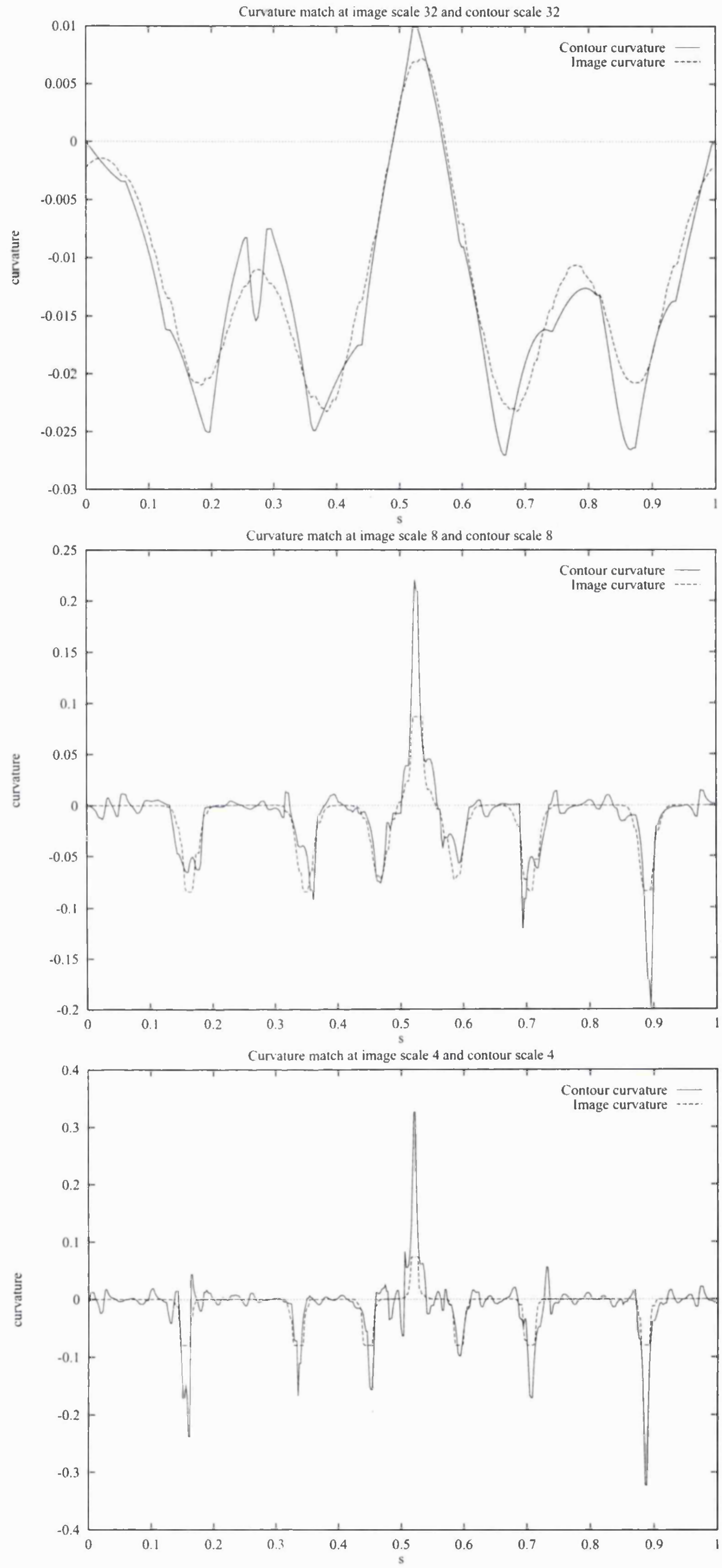


Figure 6.8: Results for curvature matching at decreasing scales. From top to bottom: $\sigma = \varsigma = 32, 8, 4$.

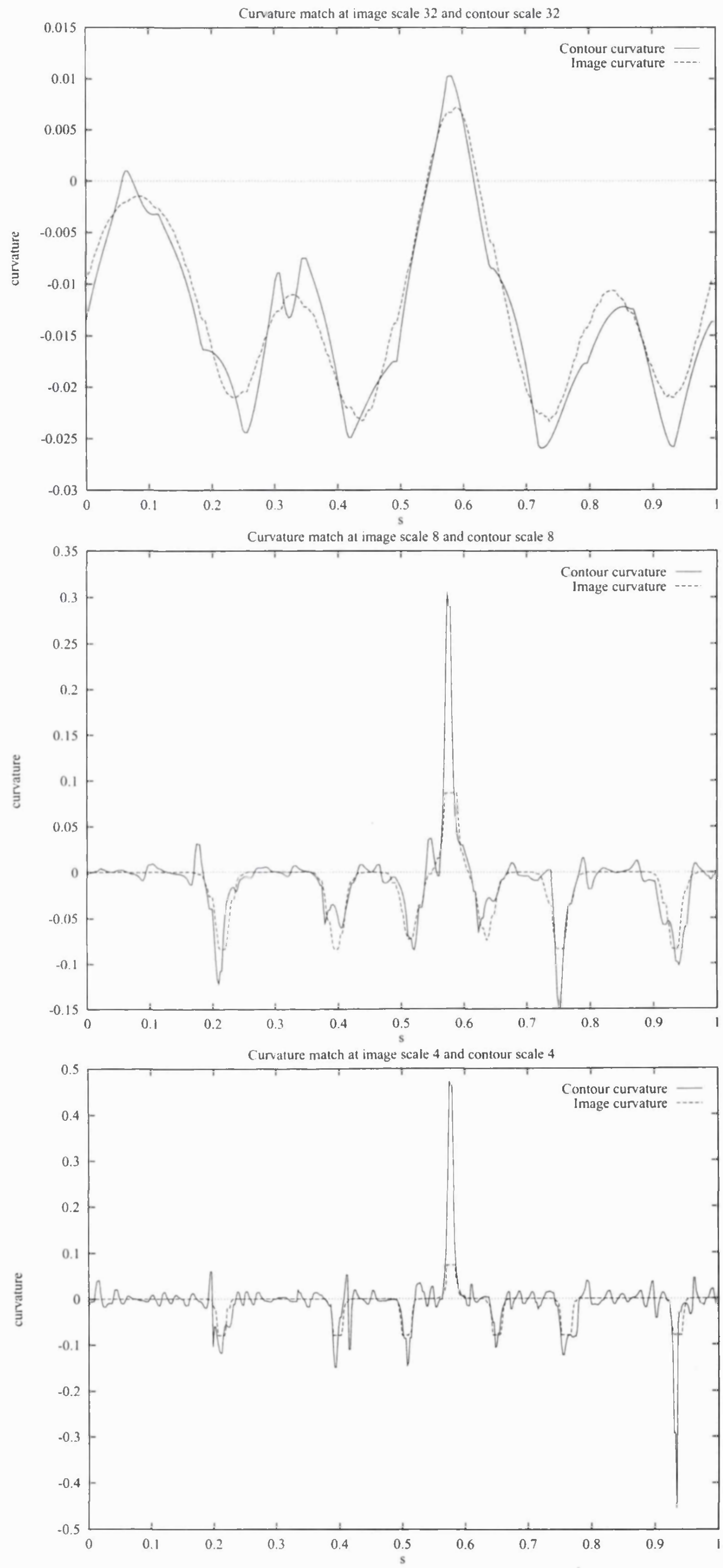


Figure 6.9: Results for curvature minimization at decreasing scales. From top to bottom: $\sigma = \varsigma = 32, 8, 4$.

RMS error	$\sigma = 32$	$\sigma = 16$	$\sigma = 8$	$\sigma = 4$
$\varsigma = 32$	0.038642	0.098777	0.604771	-
$\varsigma = 16$	0.069287	0.132589	0.376344	0.913324
$\varsigma = 8$	0.140109	0.186350	0.396742	0.672205
$\varsigma = 4$	0.429516	0.374140	0.866096	0.780066

Table 6.4: Curvature deviation results as root-mean-squared errors for different levels of image and contour scale. The best matches are achieved if contour scale ς and image scale σ correspond, or of ς is slightly higher. Figure 6.10 illustrates the matching results for $\sigma = 32$ using smaller contour scales.

matching. For both cases, the RMS error generally increases for decreasing scales due to the increasing complexity of the object. It should be noted that the contours along which the contour and image curvature values have been obtained are not identical for the curvature minimizing and curvature matching results, and that the curvature deviation in both cases is greatly influenced by the balance of the other internal and external energy forces acting on the model.

Scale dependence: Both the normalized contour curvature $\tilde{\kappa}(s)$ and the underlying scale-space isophote image curvature $\tilde{\kappa}_{image}(s)$ are dependent on scale in two respects: first, the image curvature is computed at a sampled image scale $\sigma_{\tilde{\sigma}}$, or rather from a blurred image $L(\mathbf{x}; \sigma_{\tilde{\sigma}})$. Second, using the normalized representation $\tilde{\mathbf{v}}(s)$ yields a dependence on the contour scale $\varsigma_{\tilde{\sigma}}$, related to the image scale by equation (6.11). The necessity of relating contour and image scale is further motivated by the curvature matching process itself. Figure 6.10 illustrates the effect of choosing the contour scale is chosen smaller than the image scale - the contour is *oversampled* and its curvature is oscillating with respect to the image curvature, since it is computed over smaller distances than the image curvature. In table 6.4 this effect is quantified in terms of the RMS error between contour and image curvature for varying levels of image and contour scale. Note that for the particular case of $\sigma = 4, \varsigma = 32$ no adequate curvature match could be obtained at all. The best matching results are obtained if both contour and image scale correspond, or if the contour scale is chosen slightly higher. The fact that slightly higher contour scales appear to produce better matching results is due to the tolerance span in the adaptive sampling process, which allows *snaxel* spacing in the open interval $(\varsigma - \Delta\varsigma; \varsigma + \Delta\varsigma)$ (see algorithm 6.1). Contour scales larger than the sampling contour scale have only a minor effect on the matching quality, while the oscillating behaviour occurring at contour scales below the sampling contour scale produces much higher errors. This effect may be corrected by sampling the contour scale with an offset (see equation (6.11)).

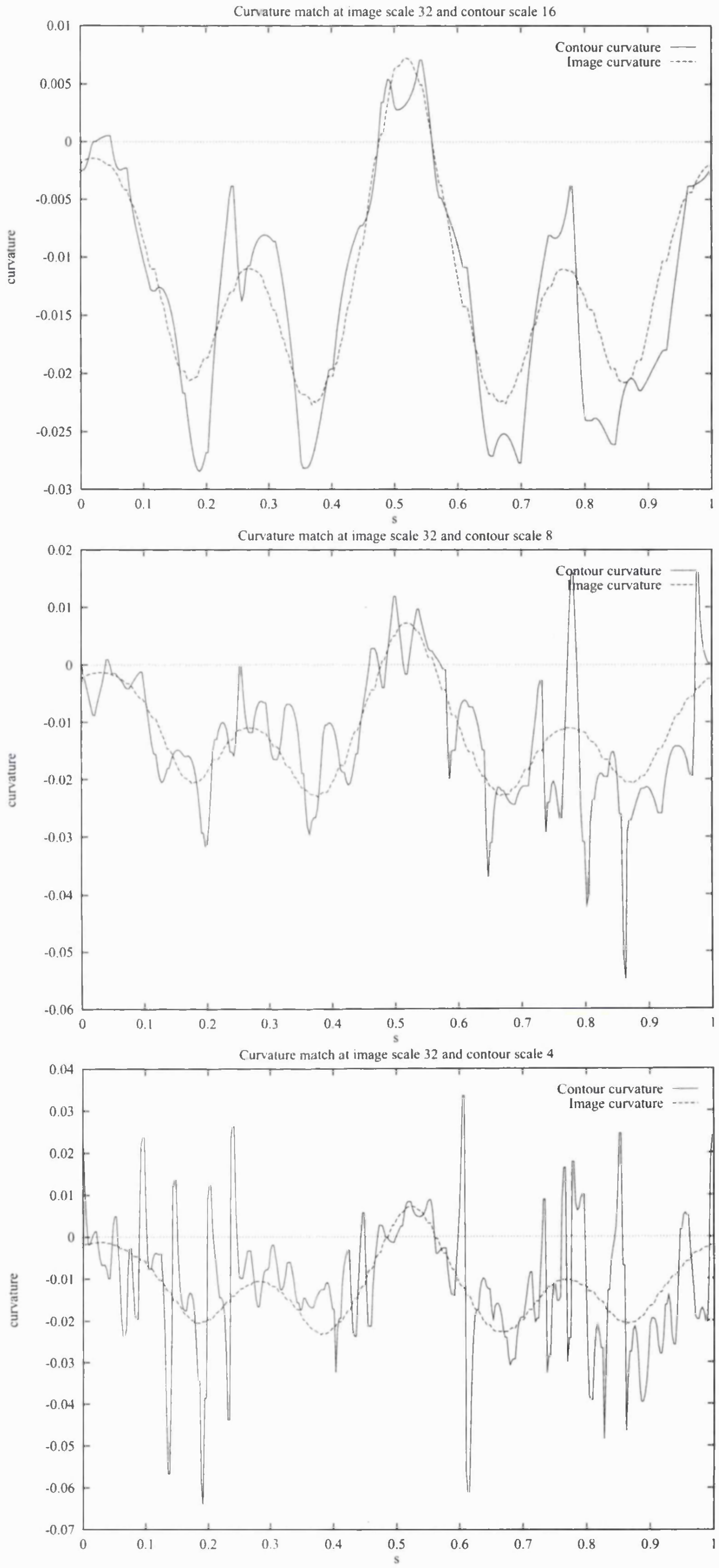


Figure 6.10: Curvature matching dependence on scale. Curvature deviation at $\sigma = 32$ and contour scales $\varsigma = 16$ (top), $\varsigma = 8$ (middle) and $\varsigma = 4$ (bottom), resulting in *oversampling* and oscillations of the contour curvature.

6.2.2 Image Energy terms

In order to attract the multi-scale active contour model to edges, the image scale-space gradient in a Euclidean setting, $\tilde{\nabla}L$, and the ridges R of the magnitude of the scale-space gradient, $||\tilde{\nabla}L||$, are chosen as the basis for suitable image potentials. Squaring and negating the scale-space gradient yields a term $-\tilde{L}_w^2$ which creates a high attraction to high negative edge values. Alternatively, the values of $\tilde{L}_w^2$ can be inverted, leading to high values in homogeneous regions, and low values at edges. Computing the ridges of $\tilde{L}_w$ allows to locate the centres of the image edges, similar to the zero-crossings of the scale-space Laplacian. The ridges are computed from $\tilde{L}_w$ as binary ridges in $2D$ for every scale $\sigma_{\tilde{\sigma}}$ using equation (3.43) (chapter 3), but fuzzy ridgeness measures like the ones presented in section 3.4.3.2 with extensions to $3D$ may also be used. The $2D$ binary ridges are then distance transformed in order to form an attraction potential. The $5-7-11$ Chamfer distance transform (chapter 2, section 2.4.1) was chosen as the most suitable method due to its good approximation quality. Using this scheme, the multi-scale active contour model is attracted to low values of the distance transformed ridge potential, $\tilde{R}^2_{dist}(\tilde{L}_w)$, indicating closeness to an edge, with zero values for the centre of edges. Using the ridge potential additionally to the gradient magnitude potential has the advantage of providing a continuous attraction potential over large distances even in homogeneous regions of an image. In order to achieve an adequate balance of these two image potentials to the internal terms presented above, they are normalized. (Note that this normalization can instead be integrated into the associated energy weighting parameters). The $\tilde{L}_w^2$ potential is normalized and inverted by setting

$$\tilde{L}_w^2(\tilde{\mathbf{v}}(s)) = \left(\frac{\tilde{L}_w(\tilde{\mathbf{v}}(s)) - \max(\tilde{L}_w)}{\max(\tilde{L}_w) - \min(\tilde{L}_w)}\right)^2 , \tag{6.20}$$

and the ridge potential is normalized between $[0;1]$ by computing

$$R^2_{dist}(\tilde{L}_w(\tilde{\mathbf{v}}(s))) \left(\frac{R_{dist}(\tilde{L}_w(\tilde{\mathbf{v}}(s))) - \min(R_{dist}(\tilde{L}_w))}{\max(R_{dist}(\tilde{L}_w)) - \min(R_{dist}(\tilde{L}_w))}\right)^2 . \tag{6.21}$$

These two normalized image potentials are weighted with the deviation of the scale-space gradient direction of $\tilde{\nabla}L$ from the contour unit normal direction of $\mathbf{n}(s)$, similar to the boundariness measure based on *directional tuning* in [Morse, 1994]. Figure 6.11 illustrates this scheme. The directional tuning is achieved by multiplying the edge terms in equations (6.20) and (6.21), respectively, with the product of $\mathbf{n}(s)$ and the unit gradient vector $\tilde{\nabla}L/\tilde{L}_w$:

$$\tilde{\mathcal{E}}_{gradient}(\tilde{\mathbf{v}}(s)) = \pm\tilde{L}_w^2(\tilde{\mathbf{v}}(s)) \left(\frac{\tilde{\nabla}L(\tilde{\mathbf{v}}(s))}{\tilde{L}_w(\tilde{\mathbf{v}}(s))} \cdot \tilde{\mathbf{n}}(s)\right)^m \tag{6.22}$$

Weighting the angular falloff by raising it to the power of m allows for narrowing or broadening the directional tuning operator (by choosing $m > 1$ or $m < 1$, respectively). Selecting $m = 0$

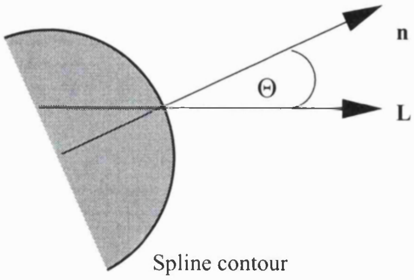


Figure 6.11: Directional tuning.

models an operator without any directional tuning, i.e. the squared normalized scale-space gradient magnitude. Similarly, the ridge potential is aligned to the contour normal direction by

$$\tilde{\mathcal{E}}_{ridge}(\tilde{\mathbf{v}}(s)) = \pm R^2_{dist}(\tilde{L}_w(\tilde{\mathbf{v}}(s))) \left(\frac{\tilde{\nabla} L(\tilde{\mathbf{v}}(s))}{\tilde{L}_w(\tilde{\mathbf{v}}(s))} \cdot \mathbf{n}(s) \right)^m \tag{6.23}$$

The image contrast needs again to be taken into account, which is achieved by changing the sign of the weighted angular falloff.

6.2.3 Integration into the Energy Function

The normalized elasticity $\tilde{\mathbf{v}}_s(s)$ and curvature $\tilde{\kappa}(s)$ of the contour, as well as the contour normal $\mathbf{n}(s)$ are computed analytically from the normalized parametric B-spline representation $\tilde{\mathbf{v}}(s)$ for all interpolated points along the contour. All image terms, including the scale-space image curvature, are computed discretely by interpolating the points of the spline contour and evaluating the image terms at these points. The energy function determining the total energy of the multi-scale active contour model is defined as a weighted linear combination of the terms given in equations (6.12), (6.16), (6.22), and (6.23):

$$\begin{aligned}\tilde{\mathcal{E}} = \frac{1}{\mathrm{L}(\tilde{\mathbf{v}}(s))} \int_0^1 \Big(\; & \alpha_{elasticity}\, \tilde{\mathcal{E}}_{elasticity}(\tilde{\mathbf{v}}(s)) \; + \alpha_{bending}\, \tilde{\mathcal{E}}_{bending}(\tilde{\mathbf{v}}(s)) \; + \\ & \alpha_{gradient}\, \tilde{\mathcal{E}}_{gradient}(\tilde{\mathbf{v}}(s)) \; + \alpha_{ridge}\, \tilde{\mathcal{E}}_{ridge}(\tilde{\mathbf{v}}(s)) \Big)\, \mathrm{d}s\end{aligned} \tag{6.24}$$

where $\mathrm{L}(\tilde{\mathbf{v}}(s))$ is the length of the overall spline contour. The integral over the spline contour can be discretized using summation with an interval increment of $\Delta s = \frac{1}{\mathrm{L}(\mathbf{v}(s))}$:

$$\begin{aligned}\tilde{\mathcal{E}}^* = \frac{1}{\mathrm{L}(\tilde{\mathbf{v}}(s))} \sum_{s=0;\Delta s}^{1} \Big(\; & \alpha_{elasticity}\, \tilde{\mathcal{E}}^*_{elasticity}(\tilde{\mathbf{v}}(s)) \; + \alpha_{bending}\, \tilde{\mathcal{E}}^*_{bending}(\tilde{\mathbf{v}}(s)) \; + \\ & \alpha_{gradient}\, \tilde{\mathcal{E}}^*_{gradient}(\tilde{\mathbf{v}}(s)) \; + \alpha_{ridge}\, \tilde{\mathcal{E}}^*_{ridge}(\tilde{\mathbf{v}}(s)) \Big)\end{aligned} \tag{6.25}$$

The normalization of the energy by the length of the contour is necessary as otherwise shorter contours are favoured over longer, more complex ones. Equations (6.24) and (6.25) compute the overall snake energy. Local optimization techniques, however, may only require the calculation of the energy of each spline patch. i.e.

$$\tilde{\mathcal{E}}(\tilde{\mathbf{v}}_i) = \frac{1}{\mathrm{L}(\tilde{\mathbf{v}}_i(s_i))} \int_0^1 \Big(\alpha_{elasticity}\, \tilde{\mathcal{E}}_{elasticity}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{bending}\, \tilde{\mathcal{E}}_{bending}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{gradient}\, \tilde{\mathcal{E}}_{gradient}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{ridge}\, \tilde{\mathcal{E}}_{ridge}(\tilde{\mathbf{v}}_i(s_i))\, \mathrm{d}s_i \Big) \tag{6.26}$$

which is discretely approximated by

$$\tilde{\mathcal{E}}^*(\tilde{\mathbf{v}}_i) = \frac{1}{\mathrm{L}(\tilde{\mathbf{v}}_i(s_i))} \sum_{s_i=0;\Delta s_i}^{1} \Big(\alpha_{elasticity}\, \tilde{\mathcal{E}}^*_{elasticity}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{bending}\, \tilde{\mathcal{E}}^*_{bending}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{gradient}\, \tilde{\mathcal{E}}^*_{gradient}(\tilde{\mathbf{v}}_i(s_i)) + \alpha_{ridge}\, \tilde{\mathcal{E}}^*_{ridge}(\tilde{\mathbf{v}}_i(s_i)) \Big) \tag{6.27}$$

with $\Delta s_i = \frac{1}{\mathrm{L}(\mathbf{v}_i(s_i))}$. In the following, several modified optimization techniques based on this energy functional will be presented.

6.3 Optimization

Since the goal of the multi-scale active contour model is the tracking of shapes through an image scale-space rather than their explicit segmentation, the choice of potential optimization routines can be restricted to local ones. This is motivated by the use of the natural contour representation which yields at all scale levels the same level of locality of the solution with respect to the image scale. The spline representation and the proposed curvature matching process rule out the classic variational approach (chapter 4, section 4.3.1) which is based on an implicit discrete Euler step with respect to the internal terms and explicit steps with respect to all external energy terms. This implies that such a system cannot easily be solved with respect to local, non-parametric image curvature constraints. In this dissertation three other local techniques have been investigated which are all based on explicit constraints: the *greedy* algorithm, Iterated Conditional Modes (ICM), and a Euler forward scheme. In the following, these modified optimization techniques will be presented.

6.3.1 *Greedy* Optimization

The *greedy* algorithm (section 4.3.3) performs a local neighbourhood search and a sequential updating of the snaxels. Additionally, the adjustment of the bending energy weighting term for corner detection as suggested by [Williams and Shah, 1992] is replaced by a fixed weighting term for the proposed curvature matching process, and the edge potentials are normalized over the whole image rather than locally. A possible sequential dependence of the greedy algorithm is avoided using a simultaneous or quasi-parallel updating of the control points. This is achieved by *freezing* the neighbouring control points for each spline patch optimization, and updating only after all spline patches have been evaluated. This might cause a slight total energy increase as the internal energy terms (elasticity, curvature and normal direction used for the directional tuning) may change due to the simultaneous updating of the neighbouring control points, but was not found to affect the final optimization result. Algorithm 6.2 illustrates the strategy for the modified, multi-scale *greedy* optimization with simultaneous updating. At the beginning of each iteration of the optimization process, the active contour model is resampled using algorithm 6.1 with respect to the underlying image scale. Then all *snaxels* or spline control points are visited once, and the local neighbourhood of size $M \times M$ is investigated. For each point in this neighbourhood, the corresponding new spline patches obtained by replacing the current control point with the point of the neighbourhood are evaluated using the energy function of equation (6.27), and the point is marked as a new potential control point if the associated local energy is lower than the energy associated with the current point or previously marked potential new points. Note that theoretically four spline patches have to be evaluated for each control point due to the local control associated with V-splines. However, investigating only the two middle spline patches was found to be quite sufficient, and additionally improved computational complexity. After all control points have been visited, they are updated according to the marked new points, and the next iteration of the optimization is performed. The optimization terminates if no new points of better (lower) energy are found, or if the total energy falls below a certain threshold, or if a maximum number of iterations is reached. As the absolute number of points may vary due to the adaptive sampling, a relative number in percent was found to be a more appropriate stopping criterion, while an absolute energy threshold needs to be chosen due to the simultaneous updating and sampling scheme.

6.3.2 Iterated Conditional Modes Optimization

Iterated Conditional Modes (ICM) (chapter 4, section 4.3.4) is a probabilistic relaxation technique and corresponds to Simulated Annealing with instantaneous freezing. It has a deterministic behaviour and therefore converges very fast. However, the search space needs to be restricted in order to improve computational efficiency, which can be achieved by only searching along the normal direction at each snaxel. Additionally, the expansion in the normal can be constrained. In this form ICM is similar to the *greedy* algorithm, and algorithm 6.2 can be modified for the ICM algorithm by changing the search space definition from $M \times M$ to the points along the contour normal. Each control point $\tilde{\mathbf{v}}_i$ is optimized according to the energy function in equation (6.25) and is replaced by a point along its normal $\mathbf{n}_i$ by maximizing the conditional probability $P(\omega_i|\hat{\mathbf{v}})$ based on the provisional estimates $\hat{\mathbf{v}}$ of the search space. Though fewer iterations than for the *greedy* algorithm are necessary, each iteration takes longer since the normal computation for the search space is computationally more expensive.

6.3.3 Euler Forward Optimization

As mentioned earlier, the classical variational approach by [Kass *et al.*, 1987b] takes a combination of implicit and explicit Euler steps. Formulating the internal energy terms also as external

while *no_points_moved* $>$ *thres_no* **and** $\mathcal{E}^* >$ *thres_energy* **and** *counter* $<$ *thres_iterations* **do**

reset *no_points_moved*

increment *counter*

// Adaptive resampling using algorithm 6.1 with respect to $\varsigma_{\tilde{\sigma}}$

sample_adaptive $(\tilde{\mathbf{v}})$

// Loop to optimize control points without updating

for $i = 1$ **to** N **do**

set $\tilde{\mathbf{v}}_{new_i} = \tilde{\mathbf{v}}_i$

// Search local neighbourhood of size $M \times M$

for $j = -\frac{M}{2}$ **to** $\frac{M}{2}$ **do**

for $k = -\frac{M}{2}$ **to** $\frac{M}{2}$ **do**

set $\tilde{\mathbf{v}}_{new} = (\tilde{x}_i + j, \tilde{y}_i + k)$

if $\tilde{\mathcal{E}}^*(\tilde{\mathbf{v}}_{new}) < \tilde{\mathcal{E}}^*(\tilde{\mathbf{v}}_{new_i})$ **then**

set $\tilde{\mathbf{v}}_{new_i} = \tilde{\mathbf{v}}_{new}$

end if

end for

end for

end for

// Simultaneous updating

for $i = 1$ **to** N **do**

if $(\tilde{\mathbf{v}}_i \neq \tilde{\mathbf{v}}_{new_i})$ **then**

set $\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_{new_i}$

increment *no_points_moved*

end if

end for

end while

Algorithm 6.2: Algorithm for modified multi-scale *greedy* optimization.

steps, an explicit Euler forward scheme is performed. Let

$$f_{\tilde{x}}(i) = \frac{\partial \tilde{\mathcal{E}}^*(\tilde{\mathbf{v}}_i(s_i))}{\partial \tilde{x}_i} \quad \text{and} \quad f_{\tilde{y}}(i) = \frac{\partial \tilde{\mathcal{E}}^*(\tilde{\mathbf{v}}_i(s_i))}{\partial \tilde{y}_i} \tag{6.28}$$

where the derivatives are computed by discrete approximizations and $\mathcal{E}^*$ is computed via equation (6.25). This leads to two independent Euler equations which can be minimized using a step size γ:

$$\begin{aligned} \mathbf{f}_{\tilde{\mathbf{x}}}\left(\tilde{\mathbf{x}}^{(n-1)}, \tilde{\mathbf{y}}^{(n-1)}\right) &= -\gamma\left(\tilde{\mathbf{x}}^{(n)} - \tilde{\mathbf{x}}^{(n-1)}\right) \\ \mathbf{f}_{\tilde{\mathbf{y}}}\left(\tilde{\mathbf{x}}^{(n-1)}, \tilde{\mathbf{y}}^{(n-1)}\right) &= -\gamma\left(\tilde{\mathbf{y}}^{(n)} - \tilde{\mathbf{y}}^{(n-1)}\right) \end{aligned} \tag{6.29}$$

At equilibrium, the time derivative vanishes and equation (6.29) is solved, which can be achieved by iteratively solving

$$\tilde{\mathbf{x}}^{(n)} = \tilde{\mathbf{x}}^{(n-1)} + \frac{\mathbf{f}_{\tilde{\mathbf{x}}}\left(\tilde{\mathbf{x}}^{(n-1)}, \tilde{\mathbf{y}}^{(n-1)}\right)}{\gamma} \quad \text{and} \quad \tilde{\mathbf{y}}^{(n)} = \tilde{\mathbf{y}}^{(n-1)} + \frac{\mathbf{f}_{\tilde{\mathbf{y}}}\left(\tilde{\mathbf{x}}^{(n-1)}, \tilde{\mathbf{y}}^{(n-1)}\right)}{\gamma} \tag{6.30}$$

Parameter name	Parameter value
$\alpha_{elasticity}$	0.00001
$\alpha_{bending}$	1
$\alpha_{gradient}$	1
α_{ridge}	0.1
Number of iterations	up to 30
Energy threshold	0.001
Number of points moved	1%
Greedy search space	up to 7×7 pixel units
ICM normal search space	20 pixel units
Euler forward step size	$\gamma = 1$

Table 6.5: Energy function parameters.

This technique is also called *simultaneous over relaxation.* Numerical details of this technique can be found in [Press *et al.*, 1992]. This purely explicit method allows for incorporation of hard constraints into the energy function, in particular the curvature matching process as a combination of internal and external energy terms. Moreover, it can be continuously evaluated along the contour, rather than at discrete *snaxels* only, whereas the classic scheme only allowed for continuous evaluation of the external energy terms. Termination criteria are chosen similar to those of the *greedy* and ICM algorithms. The main disadvantage of this techniques is that only small step sizes γ can be taken, as otherwise the technique becomes numerically unstable. This implies that although each iteration is actually faster than for the *greedy* and ICM algorithms (as for each control point only the two derivative steps need to be calculated, rather than evaluating an $M \times M$ neighbourhood, or the normal search space), more iterations are necessary.

Table 6.5 lists the fixed, general parameters of the energy terms for all three optimization techniques along with the associated optimization parameters. The energy parameters were selected empirically via a set of experiments, and then fixed for all results in this dissertation. The choice of the energy weighting parameters reflects the emphasis on the curvature matching process and the edge terms, while the weighting parameter for the elasticity terms is chosen rather low, as the adaptive sampling strategy enforces uniform *snaxel* spacing.

This section has formulated three suitable extensions for scale-space integration into existing local optimization techniques, where the locality of the respective solutions is given by the image and contour scale. Though all three techniques have been found suitable and yielding similar results, the multi-scale *greedy* algorithm is favoured in this dissertation due to its computation speed, flexible search space, numerical stability, and good convergence behaviour in practice.

6.4 Summary

This chapter has presented the theoretical framework for a novel multi-scale active contour model. In particular, a suitable continuous spline representation has been chosen to overcome problems encountered in the classic discrete active contour model. This representation has been formulated in a multi-scale setting by adapting the internal contour scale in terms of the control point spacing to the image scale using several newly developed multi-scale sampling strategies. A multi-scale energy function has been formulated to include differential invariants in scale-space, as well as normalized edge potentials in terms of the scale-space gradient and the distance transformed ridges of its magnitude. A curvature matching process of the contour curvature to the underlying isophote image curvature has been developed and evaluated in terms of its dependence on the image contrast, robustness, performance and scale-dependence, and has been found to perform better in extracting shapes of high curvature parts than using the classic energy function which minimizes the contour curvature. It was shown that the concept of a spline-based multi-scale active contour can be formulated in a local optimization framework. In the following chapter, the application of this model as an implicit segmentation tool for shape description will be developed, and the concept of the resulting multi-scale shape stack will be presented.

Chapter 7

Multi-Scale Shape Stack

– PETIT BONHOMME, N'EST-CE PAS QUE C'EST UN MAUVAIS RÊVE CETTE HISTOIRE DE SERPENT ET DE RENDEZ-VOUS ET D'ÉTOILE...

"LITTLE MAN, TELL ME THAT IT IS ONLY A BAD DREAM – THIS AFFAIR OF THE SNAKE, AND THE MEETING-PLACE, AND THE STAR..."

Le Petit Prince, Antoine de Saint-Exupéry.

The multi-scale active contour model presented in the previous chapter provides a tool for shape regularization and description in scale-space. The idea behind this is that a shape is represented and tracked by an active contour model through an image feature scale-space, consisting of the scale-space isophote image curvature, scale-space gradient magnitude and direction, and the distance transformed ridges of the scale-space gradient magnitude. At each level of this scale-space, the shape is quantified with respect to its size, curvature and other shape measurements. Furthermore, the whole set of regularized shapes is tested with respect to shape changes *across* scales. This is achieved by formulating the set of shapes in a hierarchical manner as a *multi-scale shape stack*, where each level of the stack represents the level of contour and image scale at which the shape has been regularized. The shape stack can be obtained via two different processes: *active shape evolution*, and *active shape focusing*. The analysis of the shape stack is referred to as *active shape description*.

This chapter first introduces some scale-space notations for the different possible dimensionalities of the image scale-space, multi-scale segmentation, and multi-scale shape stack. The techniques for active shape evolution, focusing, and description based on the concept of a multi-scale shape stack will be presented, and an example section will illustrate the interaction of these techniques.

7.1 Scale-Space Notations

Applying an active contour model to $2D$ or $3D$ image data and corresponding scale-spaces requires the introduction of a suitable notation for image scale-spaces, multi-scale segmentation, and the resulting multi-scale shape stack. Table 7.1 presents an overview of the different scale-

<table>
<tr><th colspan="2">Properties</th><th>$2\frac{1}{2}$D stack</th><th>3D stack</th><th>$3\frac{1}{4}$D stack</th><th>$3\frac{1}{2}$D stack</th></tr>
<tr><td rowspan="3">Original image</td><td>Dimension</td><td>2D</td><td colspan="3">3D</td></tr>
<tr><td>Notation</td><td rowspan="2">$L(x, y)$</td><td colspan="3">$L(x, y, z)$</td></tr>
<tr><td>Slice</td><td colspan="3">$L(x, y, z_k)$</td></tr>
<tr><td rowspan="4">Image scale-space</td><td>Dimension</td><td>3D</td><td>$3\frac{1}{2}$D</td><td colspan="2">4D</td></tr>
<tr><td>Notation</td><td>$L(x, y, \sigma)$</td><td>$\{L_{z_k}(x, y; \sigma)\}$</td><td colspan="2">$L(x, y, z; \sigma)$</td></tr>
<tr><td>Sample</td><td rowspan="2">$L(x, y, \sigma_i)$</td><td>$\{L_{z_k}(x, y; \sigma_i)\}$</td><td colspan="2">$L(x, y, z; \sigma_i)$</td></tr>
<tr><td>Slice</td><td>$L_{z_k}(x, y; \sigma_i)$</td><td colspan="2">$L(x, y, z_k; \sigma_i)$</td></tr>
<tr><td rowspan="3">Segmentation</td><td>Dimension</td><td>2D</td><td>$2\frac{1}{2}$D</td><td>$2\frac{3}{4}$D</td><td>3D</td></tr>
<tr><td>Notation</td><td>$\tilde{\mathbf{v}}(s; \sigma)$</td><td colspan="2">$\{\tilde{\mathbf{v}}_{z_k}(s; \sigma)\}$</td><td>$\tilde{\mathbf{v}}(s, r; \sigma)$</td></tr>
<tr><td>Sample</td><td>$\tilde{\mathbf{v}}(s; \sigma_i)$</td><td colspan="2">$\{\tilde{\mathbf{v}}_{z_k}(s; \sigma_i)\}$</td><td>$\tilde{\mathbf{v}}(s, r; \sigma_i)$</td></tr>
</table>

Table 7.1: Overview of scale-space dimensionalities of the multi-scale shape stack based on the different image scale-space dimensionalities and segmentation methods.

space dimensionalities which will be further discussed below. Let $L(x, y)$ denote a $2D$ image and $L(x, y, z)$ a $3D$ or volumetric image. Individual slices of a volumetric image are denoted by $L(x, y, z_k)$. Furthermore, one can distinguish between image scale-spaces of three different dimensionalities (note that the more general writing of σ_i rather than $\sigma_{\tilde{\sigma}}$ for the individual scale samples is used):

- **3D**. A $3D$ image scale-space $L(x, y; \sigma)$ can be computed of a $2D$ image, where the image scale σ is treated as an extra degree of freedom. Individual scale samples or slices of the scale-space are denoted by $L(x, y; \sigma_i)$.
- **4D**. A $4D$ image scale-space $L(x, y, z; \sigma)$ can be analogously computed of a volumetric image. A scale sample of this image scale-space is denoted by $L(x, y, z; \sigma_i)$, and a sample with respect to an image slice is denoted by $L(x, y, z_k; \sigma_i)$.
- **$3\frac{1}{2}$D**. A $3\frac{1}{2}D$ or *slice-by-slice* image scale-space is obtained by computing for each image slice of a volumetric image a separate $3D$ image scale-space, yielding a set $\{L_{z_k}(x, y; \sigma)\}$. A scale sample for a certain image slice is denoted by $L_{z_k}(x, y; \sigma_i)$, and for the whole set by $\{L_{z_k}(x, y; \sigma_i)\}$. The scale-space structure is either organized as a $4D$ scale-space, or, more commonly, as a list of $3D$ scale-spaces with one entry for each image slice.

The first two approaches can be generally formulated as an *(N+1)-D* scale-space $L(\mathbf{x}; \sigma)$ of an *N-D* image $L(\mathbf{x})$. Scale samples of this scale-space are consequently denoted by $L(\mathbf{x}; \sigma_i)$. In general, slice-by-slice approaches for the scale-space computation of *N-D* images are less memory exhaustive and computationally more efficient than their true *(N+1)-D* counterparts, but they loose the correlation between neighbouring image slices.

An active contour model is by its nature a $2D$ technique. If applied to a volumetric image, however, it can be optimized on the individual slices of the $3D$ data set. This is denoted by changing the notation of a $2D$ active contour model $\tilde{\mathbf{v}}(s)$ to a set $\{\tilde{\mathbf{v}}_{z_k}(s)\}$ of $2D$ active contour models, where the subscript denotes the image slice. In combination with image scale-space techniques for $2D$ and $3D$ images, one can generally distinguish between four segmentation dimensionalities [Vincken, 1995]:

- **2D**. The full $3D$ scale-space information of a $2D$ image is used for the segmentation of a planar, $2D$ shape.
- $\mathbf{2\frac{1}{2}D}$. A slice-by-slice, $3\frac{1}{2}D$ scale-space of a $3D$ image reformatted in order to obtain for each image slice z_k an associated $3D$ scale-space $L_{z_k}(x, y; \sigma)$ which is then segmented individually. The result can be formulated as a set of planar shapes (one for each image slice), or as a volumetric shape using a suitable concatenation of the individual results, e.g. using triangulation.
- $\mathbf{2\frac{3}{4}D}$. This segmentation is based on a true $4D$ scale-space of a $3D$ image, which is reformatted in order to obtain a $3D$ scale-space $L(x, y, z_k; \sigma)$ for each image slice z_k. Segmentation is then carried out in analogy to the $2\frac{1}{2}D$ approach. The higher dimensionality is motivated by the fact that the correlation between neighbouring image slices due to the higher order image scale-space influences the segmentation result.
- **3D**. A true $4D$ image scale-space of a $3D$ image is used to segment a volumetric shape. A true $3D$ segmentation requires a volumetric technique, e.g. an active surface model $\tilde{\mathbf{v}}(s, r)$. This approach is not followed in this dissertation due to the high computational complexity and memory demands involved, but necessary extensions will be discussed in chapter 10.

All segmentation dimensionalities except for the last one can be achieved using active contour models. It is important to note that the aim of this dissertation is not multi-scale segmentation, but shape regularization with respect to scale. This is in contrast to other multi-scale segmentation techniques like edge focusing or the hyperstack (chapter 3, section 3.4.2), which are either only interested in the lowest scale result only, or in a final downward projection after establishing suitable links between the different scale levels.

This dissertation proposes to investigate each scale level individually, and regards only the full set of shapes at all scales as a complete shape representation. This set is obtained using the multi-scale active model in scale-space, where it gains an extra scale dimension, and optimizing it on the individual scale-space slices in a *slice-by-slice* fashion. Additionally, if the model is applied to a volumetric image, it can be optimized on the individual slices of the $3D$ data set, as well as on the individual scale slices. Extending now the representation from $\tilde{\mathbf{v}}(s)$ to $\tilde{\mathbf{v}}(s;\sigma)$, with $\mathbf{v}(s;\sigma_{\tilde{\sigma}})$ (for a natural scale-space) or $\mathbf{v}(s;\sigma_i)$ (in the general form) at a particular scale level, allows to incorporate the extra scale dimension. Moreover, at each individual scale level the model has also an individual contour scale $\varsigma_{\tilde{\sigma}}$ or ς_i, respectively. When applying this model to volumetric images and associated $3\frac{1}{2}$ or $4D$ scale-spaces, each slice z_k can be optimized with an individual multi-scale model, which is then denoted by $\tilde{\mathbf{v}}_{z_k}(s;\sigma)$, with analogous notation for the individual scale levels. The resulting set of shapes $\mathbf{v}(s;\sigma)$ (for $2D$ images) and $\{\mathbf{v}_{z_k}(s;\sigma)\}$ (for $3D$ images) are formulated as multi-scale shape stacks which are of one of the following dimensionalities:

- $\mathbf{2\frac{1}{2}D}$. A $2\frac{1}{2}D$ multi-scale shape stack is based on instances of a planar shape in a $3D$ image scale-space, consisting of all intermediate results of a $2D$ multi-scale segmentation process. It is obtained by optimizing a model $\tilde{\mathbf{v}}(s;\sigma)$ in each slice of the scale-space of a $2D$ image separately. The organization of the stack is as a set of $2D$ planar shapes, or as a $3D$ structure obtained by concatenating the individual scale results.

- **3D**. A $3D$ shape stack is based on instances of a series of planar shapes (one for each image slice) in a $3\frac{1}{2}D$ image scale-space. A multi-scale active contour model $\tilde{\mathbf{v}}_{z_k}(s;\sigma)$ is optimized separately in a $2\frac{1}{2}D$ fashion in each $3D$ scale-space associated with each image slice. The resulting shapes for each scale can be transformed back into a set of volumetric shapes using standard concatenation techniques. The $3D$ stack can therefore be structured in two different ways:
  - as a set of $2\frac{1}{2}D$ multi-scale shape stacks, one for each image slice, or
  - as a $3\frac{1}{2}D$ multi-scale shape stack (see below), where each layer represents a single volumetric scale result.

- $\mathbf{3\frac{1}{4}D}$. A $3\frac{1}{4}D$ shape stack is similar to a $3D$ stack, but it is based on a true $4D$ image scale-space and a $2\frac{3}{4}D$ multi-scale segmentation of a model $\tilde{\mathbf{v}}_{z_k}(s;\sigma)$ which contains spatial and scale information about the shapes in the neighbouring slices. The individual results can again be concatenated to a volumetric shape. This stack is of higher dimensionality than $3D$ due to the higher order image scale-space.

- $\mathbf{3\frac{1}{2}D}$. A $3\frac{1}{2}D$ stack would be based on instances of a volumetric shape in a $4D$ image scale space. Such a stack can only be constructed using a multi-scale volumetric segmentation technique, e.g. by a multi-scale active surface model $\tilde{\mathbf{v}}(s,r;\sigma)$, which will be discussed in chapter 10.

The multi-scale shape stacks developed in this dissertation assume that the shapes under investigation are planar, or can be formulated as a set of planar shapes. They add an extra scale dimension, and extra dimensions for instances of shapes in neighbouring slices. A true multi-scale

```
// For all image slices
for z_k = 1 to N do
    // Initialize with true (ground truth) model
    set ṽ_{z_k}(s; σ_0) = ṽ_{z_k}^{(true)}(s)
    // Optimize for increasing scale levels
    for i = 1 to n do
        set ṽ_{z_k}(σ_i) = Optimize (ṽ_{z_k}(s; σ_{i-1}))
    end for
end for
```

Algorithm 7.1: Algorithm for active shape evolution.

approach where the model is optimized as a space curve in scale-space rather than as a planar curve at each individual scale level will be discussed in chapter 10.

In this dissertation two main types of shape stacks are investigated: those obtained using active shape evolution, and those obtained via active shape focusing. Both concepts will be presented in the following.

7.2 Active Shape Evolution

Active shape evolution is similar to the classic multi-scale contour evolution and analysis (chapter 3, section 3.4.1), as it starts from the "original" shape obtained via prior manual outlining by an expert or a suitable segmentation tool, or through some other form of higher level knowledge (e.g. for analytical or artificial data). This shape is also called the true model or simply *ground truth*. Instead of directly blurring the shape contour as in the classic multi-scale contour analysis, the shape contour is embedded in its image context, and is taken as an initial active contour model. An image scale-space rather than a contour scale-space is constructed, along with an associated image feature scale-space. The multi-scale active contour model is then tracked through this scale-space in a so-called *fine-to-coarse* fashion, for increasing levels of image scale and associated contour scale. In this way the model is attracted from the finest shape details to more and more global, higher scale image features. Active shape evolution in conjunction with any of the presented scale-space dimensionalities can be formulated as a multi-scale, *implicit* segmentation process of $2D$, $2\frac{1}{2}D$, or $2\frac{3}{4}D$ dimensionality. The resulting shape stack, consisting of all intermediate scale results, is consequently of dimensionality $2\frac{1}{2}D$, $3D$, or $3\frac{1}{4}D$, respectively. Algorithm 7.1 illustrates the technique for active shape evolution of $2\frac{1}{2}D$ or $2\frac{3}{4}D$ dimensionality (differing only in the dimensionality of the underlying image scale-space), which will in the following be explained in more detail.

Given an initial ground truth or true model $\mathbf{v}_{z_k}^{(true)}(s)$ at zero scale σ_0 for each image slice

$L(x, y, z_k)$, all ground truth models are optimized independently for increasing scales σ_i. As an optimization routine, any of the presented techniques of the previous chapter can be used. For each scale level, the image energy terms of the energy functional are based on slice i of the $3D$, $3\frac{1}{2}D$, or $4D$ image scale-space. Consequently, each model $\tilde{\mathbf{v}}_{z_k}(s; \sigma_i)$ at scale level σ_i is of image-scale related contour scale ς_i, enforced by an adaptive uniform sampling process (see section 6.1.2.2). The *fine-to-coarse* tracking is performed by taking at each slice i of the image scale-space the optimized contour models from the previous, next lower scale slice $i - 1$ as an initial estimate, until the highest or coarsest level of scale, n, is reached. The resulting set of shapes $\{\tilde{\mathbf{v}}_{z_k}(s; \sigma_i)\}$ represent the *fine-to-coarse* multi-scale shape stack, and can consequently be structured in three different ways:

- as a single $2\frac{1}{2}D$ shape stack, denoted by the set $\{\tilde{\mathbf{v}}(s; \sigma_i) | i = 0, \cdots, n - 1\}$, or
- as a set of $2\frac{1}{2}D$ shape stacks, organized as the set of all intermediate evolution results for each image slice z_k, where each $2\frac{1}{2}D$ stack is denoted by $\{\tilde{\mathbf{v}}_{z_k}(s; \sigma_i) | i = 0, \cdots, n - 1\}$, or
- as a single $3\frac{1}{2}D$ shape stack, which is given by the set of concatenated, volumetric shapes for each ascending scale level σ_i, denoted by $\{\tilde{\mathbf{v}}_{z_k}(s; \sigma_i) | z_k = 1, \cdots, N\}$.

The scale samples are formulated in *ascending* order. Note that the latter two structures merely refer to the re-organization of the resulting shapes, and not to a different active shape evolution method. However, in a following active shape description process, they give rise to different kinds of descriptors, which will be further discussed below.

7.3 Active Shape Focusing

Active shape focusing is the dual technique to active shape evolution, as it is performed in a *coarse-to-fine* fashion, similar to classic edge focusing (chapter 3, section 3.4.2). The true shape outline need not be known, as a very coarse initial estimate (e.g. a circle or an ellipse) is sufficient to capture the global shape outline at an adequately large scale. Taking such a coarse estimate as an initial model, this model is regularized or focused down for decreasing levels of image scale. It is again important to note that the final result of this active shape focusing process is also an *implicit* rather than *explicit* multi-scale segmentation result of $2D$, $2\frac{1}{2}D$, or $2\frac{3}{4}D$ dimensionality, as only the most prominent shape outline is followed in the *coarse-to-fine* tracking process. The resulting multi-scale shape stack is consequently of the same dimensionality as when based on active shape evolution, yet it differs in the ordering of the scales. Algorithm 7.2 illustrates the active shape focusing method of $2\frac{1}{2}$ or $2\frac{3}{4}D$ dimensionality, which will be further explained in the following.

// For all image slices

for $z_k = 1$ **to** N **do**

// Initialize with estimated model

set $\tilde{\mathbf{v}}_{z_k}(s;\sigma_n) = \tilde{\mathbf{v}}_{z_k}^{(est)}(s)$

// Optimize for decreasing scale levels

for $i = n-1$ **to** 0 **do**

set $\tilde{\mathbf{v}}_{z_k}(s;\sigma_i) =$ **Optimize** $(\tilde{\mathbf{v}}_{z_k}(s;\sigma_{i+1}))$

end for

end for

Algorithm 7.2: Algorithm for active shape focusing.

If an initial estimate $\mathbf{v}_{z_k}^{(est)}(s)$ is available for each image slice, each of these models can be optimized individually using one of the optimization techniques of section 6.3. In contrast to active shape evolution, however, a reverse order of the image scale σ_i, with $i \in [n-1;0]$ is used. Again, the image energy terms of each model are computed from the associated scale sample i of the $3D$, $3\frac{1}{2}D$, or $4D$ image scale-space. The contour scale ς_i is increasing for increasing image scales, and is again enforced by adaptive uniform sampling. The model used for active shape focusing is therefore of the same form as for evolution, $\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)$. *Coarse-to-fine* tracking is performed analogously to *fine-to-coarse* tracking by taking at each scale slice i the optimized models from the previous, *next higher* scale levels $i+1$ as initial estimates for the current scale levels, a process which is repeated until the finest or lowest level of scale, σ_0 is reached. The set of shapes at all intermediate scale levels $\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)\}$ constitute the multi-scale shape stack, which is structured like the shape stack obtained from active shape evolution, but with reverse scale ordering:

- as a single, $2\frac{1}{2}D$ shape stack, denoted by the set $\{\tilde{\mathbf{v}}_{z_k=1}(s;\sigma_i)|i = n-1,\cdots,0\}$, or
- as a set of $2\frac{1}{2}D$ shape stacks, organized as the set of all intermediate focusing results for each image slice z_k, which is individually denoted by $\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)|i = n-1,\cdots,0\}$, or
- as a single $3\frac{1}{2}D$ shape stack, which is given by the set of concatenated, volumetric shapes for each descending scale level σ_i, denoted by $\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)|z_k = 1,\cdots,N)$.

The same remarks for the actual dimensionality of the *coarse-to-fine* shape stacks as for the *fine-to-coarse* shape stacks presented in the previous section apply here. Note that here the scale ordering is *descending*, and that instead of the *ground truth*, a very coarse initial estimate for each image slice is sufficient. As was mentioned earlier in this chapter, the initial estimate can be very general, for example circular- or ellipse-shaped. Yet it may be a very tedious and time consuming task to provide initial estimates for a large volumetric dataset, as it may not be sufficient to use the same estimate at all image slices. In this case, the concept of *shape propagation*, illustrated

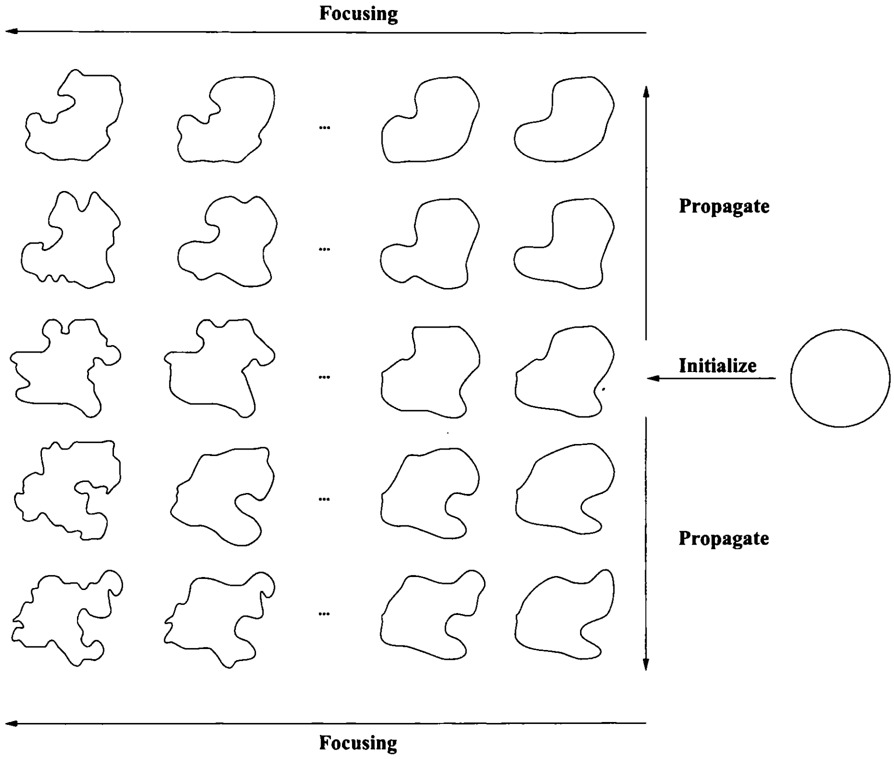


Figure 7.1: Shape propagation and focusing.

in figure 7.1 is used in combination with active shape focusing, which will be presented in the following.

7.3.1 Shape Propagation

Shape propagation is based on the assumption that the shape information contained in neighbouring image slices is highly correlated. The concepts of active shape evolution and focusing are already based on the high correlation in neighbouring scale slices, as well as on the correlation between neighbouring image slices when using a $4D$ image scale-space. This high correlation is especially true at large image scales, when finer detailed structure, defining the local difference between neighbouring image slices, has disappeared. Only the global shape outline is accessible in the adjacent image slices at high scales, which varies only little in neighbouring image slices. This observation allows to reduce the number of initial estimates needed, i.e. the number of image slices, to a single initial estimate in a suitable, intermediate image slice k. This model is denoted by $\tilde{\mathbf{v}}^{(model)}_{z_k=k}(s)$. Optimizing this estimate in its associated image slice blurred at a high level of scale yields a refined initial estimate, denoted by $\tilde{\mathbf{v}}^{(est)}_{z_k=k}(s;\sigma_n)$. This model can then be used as an initial estimate for the neighbouring slices, $k-1$ and $k+1$. Refining both estimates in their respective image slices at the same level of scale yields again initial estimates for the next neighbouring image slices, which in turn can be refined, and so forth. Figure 7.1 illustrates this concept, and the algorithm for shape propagation is given in algorithm 7.3.

// Optimize initial model to be propagated

set $\tilde{\mathbf{v}}_{z_k=k}(s;\sigma_n) =$ **Optimize** $\left(\tilde{\mathbf{v}}^{(model)}_{z_k=k}(s)\right)$

// Propagate up

for $z_k = k+1$ **to** N **do**

  set $\tilde{\mathbf{v}}_{z_k}(s;\sigma_n) =$ **Optimize** $(\tilde{\mathbf{v}}_{z_k-1}(s;\sigma_n))$

end for

// Propagate down

for $z_k = k-1$ **to** 1 **do**

  set $\tilde{\mathbf{v}}_{z_k}(s;\sigma_n) =$ **Optimize** $(\tilde{\mathbf{v}}_{z_k+1}(s;\sigma_n))$

end for

Algorithm 7.3: Algorithm for shape propagation.

As a result of the shape propagation, an initial model for each image slice is obtained in a semi-automatic manner, as the only interaction is the initialization of a single model $\tilde{\mathbf{v}}^{(model)}_{z_k=k}(s)$. Active shape focusing can then be formulated as a two stage process: shape propagation using algorithm 7.3 is performed to obtain the set of initial models for all slices, $\{\tilde{\mathbf{v}}^{(est)}_{z_k}(s)|z_k = 1,\cdots,N\}$, followed by individually focusing down each of these models using algorithm 7.2.

7.4 Active Shape Description

The previous two sections have presented the *construction* of a multi-scale shape stack in a *fine-to-coarse* and *coarse-to-fine* fashion, respectively. This section presents methodologies for the *description* of either type of stack. These methodologies can be roughly divided into techniques investigating a single or a set of $2\frac{1}{2}D$ shape stacks, and those investigating a $3\frac{1}{2}D$ shape stack. Recall that the former are based on instances of planar shapes in the scale-spaces of their associated image slices (which may be a reordered higher-order image scale-space), and the latter consist of instances of a volumetric shape (as a concatenation of planar shapes) in a true $4D$ or in a slice-by-slice $3\frac{1}{2}D$ image scale-space. Consequently, a lower order shape stacks can only be analysed using shape descriptors for planar shapes, while for a higher-order shape stack also volumetric measurements can be performed.

Shape descriptors have been categorized into global, local, and relative (chapter 2, sections 2.2–2.4). Table 7.2 lists the selected set of suitable shape descriptors. The first three descriptors (perimeter, area, and compactness) are global quantifiers based on planar shapes, but can also be computed as mean and slope measurements across scales for a set of planar shapes. The following three measurements (surface area, volume, and volumetric compactness) are based on a volumetric representation of a set of shapes extracted from a $3D$ image, e.g. using concatenation by triangulation [Christiansen and Sederberg, 1978]. Figure 7.2 illustrates the process of triangu-

Descriptor	Symbol	Definition
Perimeter	$P(\mathbf{v}(s))$	Length of the closed shape boundary
Area	$A(\mathbf{v}(s))$	Area of the enclosed region
Compactness	$C(\mathbf{v}(s))$	Dimensionless "roundness" measure
Surface area	$S(\{\mathbf{v}_{z_k}(s)\})$	Surface of a set of shapes
Volume	$V(\{\mathbf{v}_{z_k}(s)\})$	Enclosed volume of a set of shapes
Volumetric compactness	$C_V(\{\mathbf{v}_{z_k}(s)\})$	Dimensionless "sphereness" measure
Curvature	$\kappa(s)$	Local curvature
Hausdorff distance	$\text{dist}_H(\mathbf{v}(s), \mathbf{v}^{(ref)}(s))$	Worst mismatch
Chamfer distance	$\text{dist}_C(\mathbf{v}(s), \mathbf{v}^{(ref)}(s))$	Nearest distance
Triangulation distance	$\text{dist}_T(\mathbf{v}(s), \mathbf{v}^{(ref)}(s))$	Corresponding distance

Table 7.2: Selected set of global, local, and relative (distance) shape descriptors.

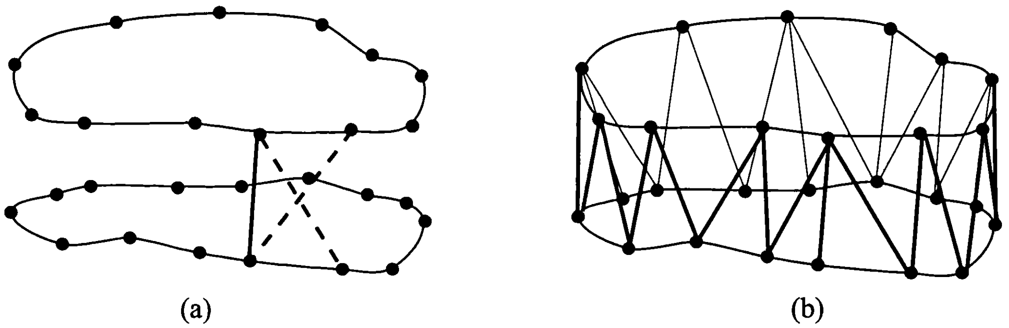


Figure 7.2: Triangulation of two shape contours. (a) Initial correspondence. (b) Triangulation result.

lating two shape contours by establishing an initial shortest-distance correspondence (figure 7.2 (a)), followed by iteratively connecting the neighbouring contour points with triangular patches (figure 7.2 (b)). A corresponding distance measurement arising from this technique, based on the shortest distance vertices, was briefly presented in chapter 2, section 2.4.3. The surface area of such a structure can be computed by summing the area of all triangular patches. Consequently, the volume of a closed structure obtained via triangulation can be computed by summing the volumes of all tetrahedrons which are formed by one of the triangle patches, and the same point of origin which must be located inside the volume. Let point $\mathbf{o}$ denote the origin, and points $\mathbf{a}, \mathbf{b}, \mathbf{c}$ denote the corners of one of the triangle patches. The squared volume of such a tetrahedron is then given by [Bronstein and Semendjajew, 1989]:

$$V^2 = \frac{1}{288} \cdot \det \begin{vmatrix} 0 & ||\mathbf{a}-\mathbf{b}||^2 & ||\mathbf{c}-\mathbf{a}||^2 & ||\mathbf{o}-\mathbf{a}||^2 & 1 \\ ||\mathbf{a}-\mathbf{b}||^2 & 0 & ||\mathbf{b}-\mathbf{c}||^2 & ||\mathbf{o}-\mathbf{b}||^2 & 1 \\ ||\mathbf{c}-\mathbf{a}||^2 & ||\mathbf{b}-\mathbf{c}||^2 & 0 & ||\mathbf{o}-\mathbf{c}||^2 & 1 \\ ||\mathbf{o}-\mathbf{a}||^2 & ||\mathbf{o}-\mathbf{b}||^2 & ||\mathbf{o}-\mathbf{c}||^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} \tag{7.1}$$

The determinant can be efficiently computed with standard numerical methods [Press *et al.*, 1992]. The "sphereness" measurement C_V is defined in analogy to the planar compactness measurement as a dimensionless quantity of the surface area S and the volume V of a volumetric structure:

$$C_V = \frac{S^{\frac{3}{2}}}{V} \tag{7.2}$$

The curvature $\kappa(s)$ is the only truly local shape measurement used. The distance measurements (Hausdorff, Chamfer, and triangulation distance, see section 2.4) allow for shape comparisons with respect to a reference shape $\mathbf{v}^{(ref)}(s)$. For active shape evolution, this shape is equivalent to the ground truth, $\mathbf{v}^{(true)}(s)$, while for active shape focusing, the lowest scale model obtained through the focusing process, $\tilde{\mathbf{v}}(s;\sigma_0)$, is used as a reference model in the absence of any ground truth. The distance measures are computed in order to derive the global deviation of each shape from this reference shape: the *Hausdorff* distance indicates the worst mismatch between each shape of the shape stack and the reference shape, the *Chamfer* distance transform of the reference shape yields the mean and RMS distances between each shape and the reference shape, and the *triangulation* distance measure obtained via concatenating each shape with the reference shape measures the mean and RMS distance of the connected minimum-edge vertices. The last two distance measurements also compute a pointwise local shape deviation from the reference shape. The iterated closest point (ICP) algorithm, also presented in section 2.4, has been considered as a suitable corresponding distance measurement. However, it requires that the shapes under comparison have the same number of points which cannot always be guaranteed. On the contrary, due the multi-scale contour representation in the shape stack, the snaxel spacing of shapes extracted at higher image scales is considerably larger than for the known or low-scale reference shape, leading to a lower number of points at high scales.

A general approach to shape description is to analyse the lowest-scale or reference shape only. The approach pursued in this dissertation, however, is to investigate a complete shape stack, taking the extra scale dimension into account. For each contour $\tilde{\mathbf{v}}_{z_k}(s)$, the construction of an associated $2\frac{1}{2}D$ shape stack yields a multi-scale planar shape representation, $\tilde{\mathbf{v}}_{z_k}(s;\sigma)$, which for active shape evolution may be written as a set $\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)|i = 0, \cdots, n-1\}$ (with reverse scale ordering for active shape focusing). Choosing one of the planar descriptors listed in table 7.2, e.g. the perimeter measurement P, does not only yield the perimeter of the reference shape, but a stack of measurements, e.g. $\{P(\tilde{\mathbf{v}}_{z_k}(s;\sigma_i))|i = 0,\cdots,n-1\}$. Consequently, for each set of contours $\{\tilde{\mathbf{v}}_{z_k}(s)|z_k = 1,\cdots,N\}$, the construction of a shape stack organized as a $3\frac{1}{2}D$ stack yields a multi-scale volumetric shape representation. Choosing the surface area measurement S from the volumetric descriptors of table 7.2 does not only yield the reference surface area, but a stack of surface area measures, e.g. for active shape evolution: $\{S(\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)|z_k = 1,\cdots,N\})|i = 0,\cdots,n-1\}$ (with reverse scale ordering for active shape focusing). It is of course also possible to compute mean shape measurements across scale for the multi-scale planar or volumetric shape sets, e.g. for a planar shape as $\frac{1}{n}\sum_{i=0}^{n-1} P(\tilde{\mathbf{v}}_{z_k}(s;\sigma_i))$, as well as to compute planar shape metrics from a set of shape contours. Additionally, the rate of the shape changes across scale can be characterized by their slope. For example, the change of perimeter with respect to scale is expressed as $\frac{\Delta P(\tilde{\mathbf{v}}_{z_k}(s;\sigma))}{\Delta\sigma}$ which can be estimated using standard linear regression techniques.

At this point it may not immediately be apparent why this approach is better than the traditional approach of describing one planar shape or one shape volume only at a single, preferably low scale to capture all finer detailed structure. It may seem as if active shape description increases the amount of data at least by a factor related to the number of scale samples, and the number of (possibly combined) shape descriptors chosen. Reformatting of the shape stack from a set of $2\frac{1}{2}D$ stacks to a single $3\frac{1}{2}D$ stack leads even to a further set of volumetric shape measurements. The motivation for performing active shape description nonetheless is based on several important observations:

- Given a ground truth shape or a lowest scale result obtained by another process, e.g. by outlining by an expert or via a suitable segmentation technique, traditional shape description is not able to capture all meaningful shape features, and to distinguish between shape properties arising from distinct physical processes embedded in the image [Marr, 1982]. Though all global shape features are inherent in the *raw* shape (the shape at lowest scale), they cannot be directly recovered. Moreover, fine-scale noise influences the quality of local derivatives computed to locate local extrema, a process which is necessary to obtain a qualitative description or *sketch* of a shape [Witkin, 1983].

- The introduction of scale as an extra parameter yields a continuum of descriptions, where no one scale of description can be defined as being correct, or being more important than the others. In fact, every scale setting yields a different description, which leads to the need of organizing and possibly simplifying the increased amount of shape information. One example solution for this problem are the concepts of *scale-space fingerprints* [Witkin, 1983] and *curvature scale-space* (CSS) [Mokhtarian and Mackworth, 1986] (chapter 3, section 3.4.1). They offer a structural way of accessing, tracking and interpreting meaningful qualitative shape properties in a curve evolution process. However, restricting a description to qualitative properties only might not be sufficient in a clinical environment where other important shape properties like global volumetric measurements, as well as relative distance measurements need to be analysed as well.

- Active shape evolution is an alternative technique to classic curve evolution, which allows to construct a multi-scale shape stack which may contain all types of local, global, and relative shape information. Its most important difference to curve evolution lies in its image data-driven approach, locking a shape to the image context by which it is defined, and therefore leading to different, but possibly more meaningful curve evolution results. Additionally, sparser scale steps are taken at higher scales due to the decrease in density of the scale-space than for the construction of classic fingerprints or the CSS, which are both based on dense linear scale sampling. Active shape evolution therefore allows to decrease the number of samples needed to obtain the highest scale result, and introduces the concept of an image scale-related contour scale.

- Both traditional, single-scale shape description and classic curve evolution processes as discussed above depend on the existence of a ground truth, or a suitable low scale shape, as it is not possible to analytically reverse the Gaussian blurring process. For real world data like medical images, however, no ground truth is available *a priori*, and therefore a reference shape needs to be acquired through some other process. Active shape focusing, having the same data-driven functionality and sampling process as active shape evolution, is based on a quasi reversed blurring process (which is implemented by computing increasingly blurred scale samples of an image, and tracking down in this image scale-space in a *coarse-to-fine* manner). This technique can be used to find this reference shape, while additionally providing the intermediate scale results at no extra computational cost. In other words, active shape focusing yields the lowest scale shape as a *byproduct* of the process.

- In order to perform a more concise yet more complete active shape description process, the collected shape information can be investigated *across* scale rather than only at each scale individually. This can be achieved by monitoring shape changes between adjacent scale levels, or with respect to the reference shape. Following this idea, the next chapter will present a new multi-scale shape metric, yielding a single shape measurement for a planar shape or set of planar shapes in scale-space.

- Finally, an important topic in shape interpretation is the visualization of qualitative shape information. The concept of the shape stack allows to visualize external shape changes in terms of the boundary location with respect to scale, and to map corresponding local shape measurements obtained by the active shape description process. Section 7.5 will illustrate this visualization method.

In the following, an example active shape evolution, focusing and description process based on the *notched rectangle* test image will be presented. In particular, a suitable representation and visualization of a multi-scale shape stack will be discussed.

7.5 Example

Figure 7.3 illustrates the intermediate active shape evolution and focusing results for the *notched rectangle* test image (chapter 6, figure 6.1 (a)), as well as the corresponding image and normalized feature scale-spaces used in the multi-scale energy function (chapter 6, section 6.2). In this example, the initial models are given by the *ground truth* (figure 6.1 (b)) for active shape evolution, and by a coarse estimate in form of an ellipse (figure 6.1 (c)) for active shape focusing. The evolution process starts at the lowest image scale, while the focusing process is directed in the opposite direction and starts at the highest image scale. The intermediate results shown in columns (d) and (e) are remarkably similar, giving rise to the observation that active shape evolution and focusing are dual techniques. This is only true if the shapes under investigation are isolated, and preferably available as silhouettes, since adjacent objects, as well as noise and artefacts can influence both processes. Results may still be quite similar, but may locally differ at some parts. What is also notable about the results in figure 7.3 is the comparison to the also hierarchical, since incomplete, Fourier reconstruction of the *notched rectangle*, as shown in figure 2.7 (b) (chapter 2); unlike the Fourier reconstruction, active shape evolution and focusing maintain the shape's topology throughout all hierarchical levels. A comparison to classic contour evolution of the same shape as presented in figure 3.6 (chapter 3) shows that the shrinkage of the shape, caused by the blurring via the diffusion or heat equation [Gage and Hamilton, 1986], is much lower for the active shape evolution and diffusion techniques at higher scales. Both processes can be described as image-data driven curve diffusion techniques. They tend to preserve the overall shape size longer by locking the shape to its surrounding image context, which is also affected by the diffusion.

Columns (d) and (e) of figure 7.3 illustrate samples of the $2\frac{1}{2}D$ multi-scale shape stacks obtained via active shape evolution and focusing, respectively. In total, each stack contains $n = 32$ slices or scale levels (26 more samples than illustrated), where each level consists of a regularized shape or *intermediate* active shape evolution or focusing result at that particular scale level. Scale sampling was chosen to be very dense, with a maximum scale difference of $\Delta\sigma_{max} = \sigma_{31} - \sigma_{30} \approx 3.3848$, and constant scale-change factors (equation 6.9, chapter 6) $f_{evolution} = (\sigma_{31}/\sigma_0)^{1/31} \approx 1.11829$ and $f_{focusing} = f_{evolution}^{-1} \approx 0.89423$. This is already a rather large difference in the

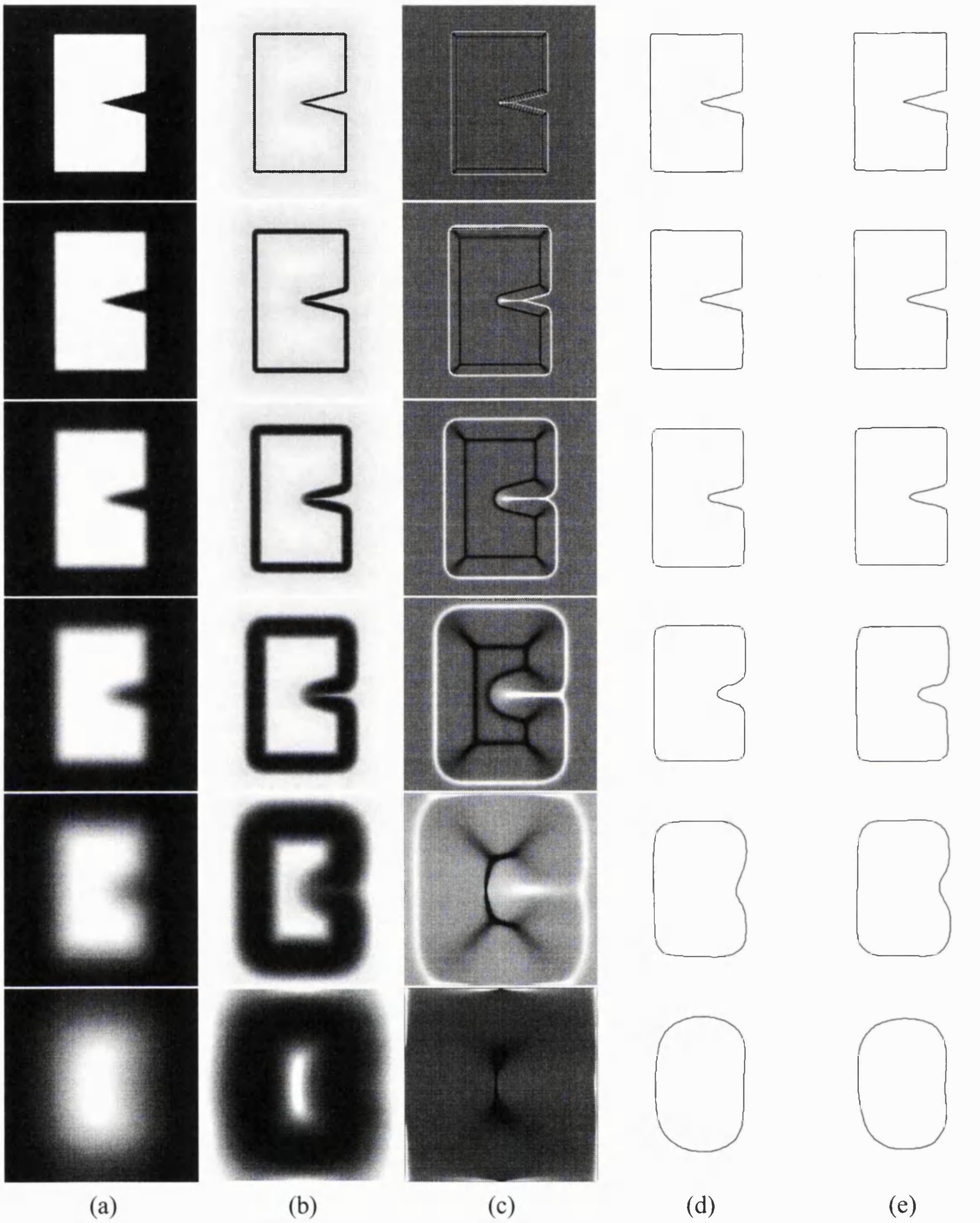


Figure 7.3: Active shape evolution and focusing of the *notched rectangle* image. Columns: (a) Image scale-space. (b) Normalized edge potential scale-space. (c) Normalized isophote curvature scale-space. (d) Active shape evolution results. (e) Active shape focusing results. The examples illustrated are scale samples for (from top to bottom) $\sigma_i = 1, 2.19, 4.28, 8.37, 16.36, 32$ with $\sigma_0 = 1, \sigma_{n-1} = 32$, $n = 32$, and scale sampling based on equation (6.9). Active shape evolution is performed using figure 6.1 (b) as an initial model starting at σ_0, and active shape focusing using figure 6.1 (c) as an initial model starting at σ_{n-1}. Also compare to hierarchical Fourier reconstruction and classic contour evolution representations in figures 2.7 (b) and 3.6 (b) (chapters 2 and 3, respectively.)

Method	$\overline{A}$	$\overline{P}$	$\overline{C}$	$\overline{\text{dist}_H}$	$\overline{\text{dist}_{C_{RMS}}}$	$\overline{\text{dist}_{T_{RMS}}}$
Evolution	20579	631.462	19.4717	72.0837	1.26972	20.7318
Focusing	20497	630.344	19.4841	137.5120	1.40287	21.8417

Table 7.3: Mean shape descriptors for active shape evolution and focusing of the *notched rectangle* image with respect to scale. The size measurements for the mean area $\overline{A}$, mean perimeter $\overline{P}$, as well as for the mean Hausdorff, mean RMS Chamfer, and mean RMS triangulation distance measurements are in pixel units, while the mean compactness $\overline{C}$ is dimensionless.

Method	$\frac{\Delta A}{\Delta\sigma}$	$\frac{\Delta P}{\Delta\sigma}$	$\frac{\Delta C}{\Delta\sigma}$	$\frac{\Delta \text{dist}_{C_{RMS}}}{\Delta\sigma}$	$\frac{\Delta \text{dist}_{T_{RMS}}}{\Delta\sigma}$
Evolution	-57.7880	-6.47411	-0.336987	0.168989	3.05724
Focusing	-61.6986	-6.60550	-0.342116	0.170671	3.04643

Table 7.4: Slope measurements of shape descriptors for active shape evolution and focusing of the *notched rectangle* image with respect to scale. The slopes are estimated using linear regression on the graphs plotted in figures 7.4 and 7.5.

sense of Bergholm [Bergholm, 1987], who analysed the displacement of edges between adjacent scale levels as being close to $|\Delta\sigma|$. In order to ensure that the edges are still tracked correctly, the maximum displacement must still be well within the search space of the tracking technique. Therefore a search space of 7×7 pixel units was chosen for the modified *greedy* algorithm (section 6.3.1), enabling to track displacements of up to 3.5 pixel units which is in the range of $\Delta\sigma_{max}$. Alternatively, a larger number of scale samples can be chosen, leading to a smaller maximum scale change. In this case, a smaller search space is sufficient which improves computational efficiency, but which may be outweighed by the increased number of optimizations and scale-space computations. This observation leads to the conclusion that scale sampling and the size of the search space of a *fine-to-coarse* or *coarse-to-fine* tracking technique are intrinsically related to each other. The search space should at least be as large as the maximum possible edge displacement, defined by the maximum scale change between adjacent scale levels. It would also be possible to formulate the size of the search space as a function of scale, allowing for very coarse sampling at higher scales with very large search spaces, and leading to smaller and hence computationally more efficient search spaces for decreasing scales. However, search space sizes larger than 7×7 are computationally very expensive and time demanding, and may lead to overlapping of neighbouring search spaces as the contour scale is related to the image scale. It is therefore better to regularize the necessary search space size by the highest possible edge displacement in the complete tracking process, rather than at each scale level individually.

Figures 7.4 and 7.5 show the plots of the global planar shape descriptors listed in table 7.2 of the *notched rectangle* across scale, where each descriptor has been obtained via active shape evolu-

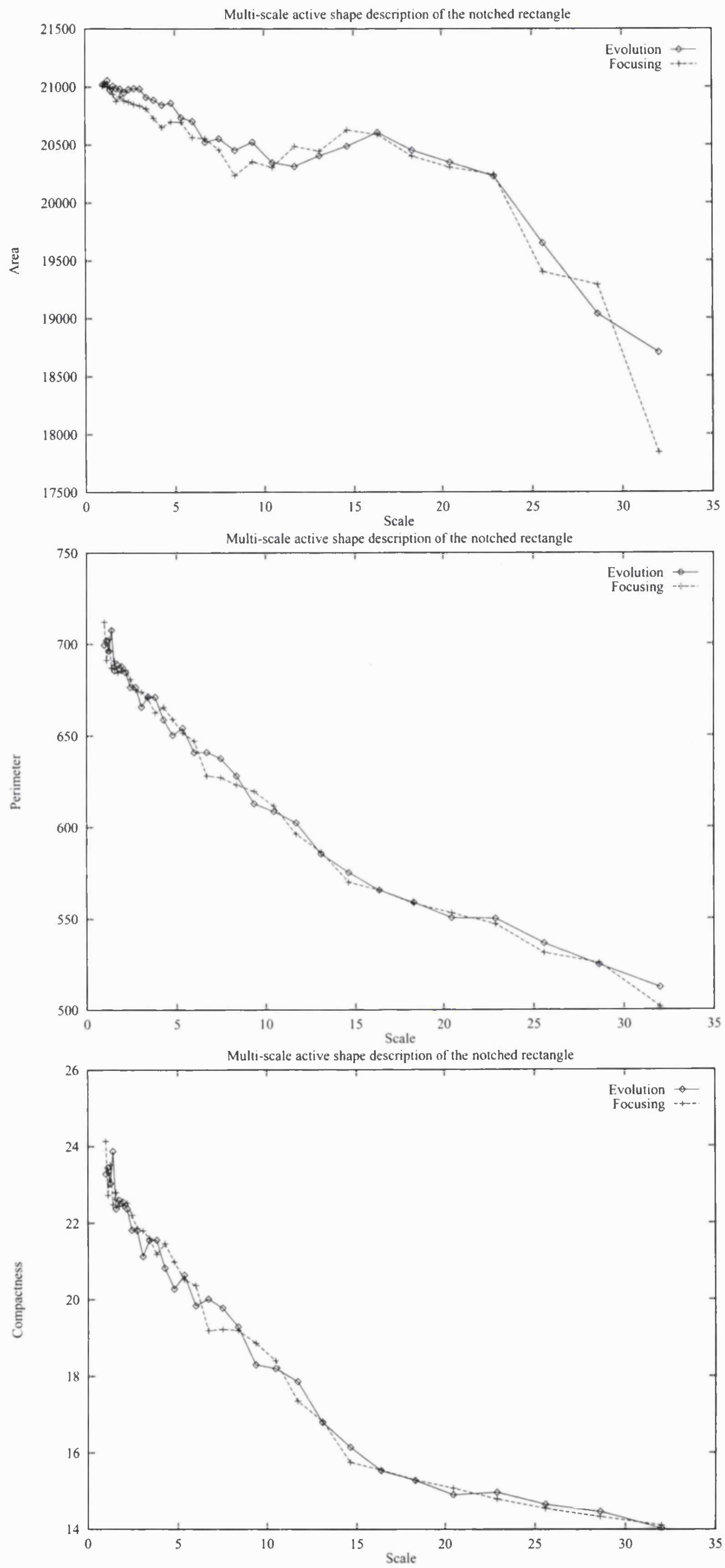


Figure 7.4: Area, perimeter and compactness of the notched rectangle across scales for active shape evolution and focusing.

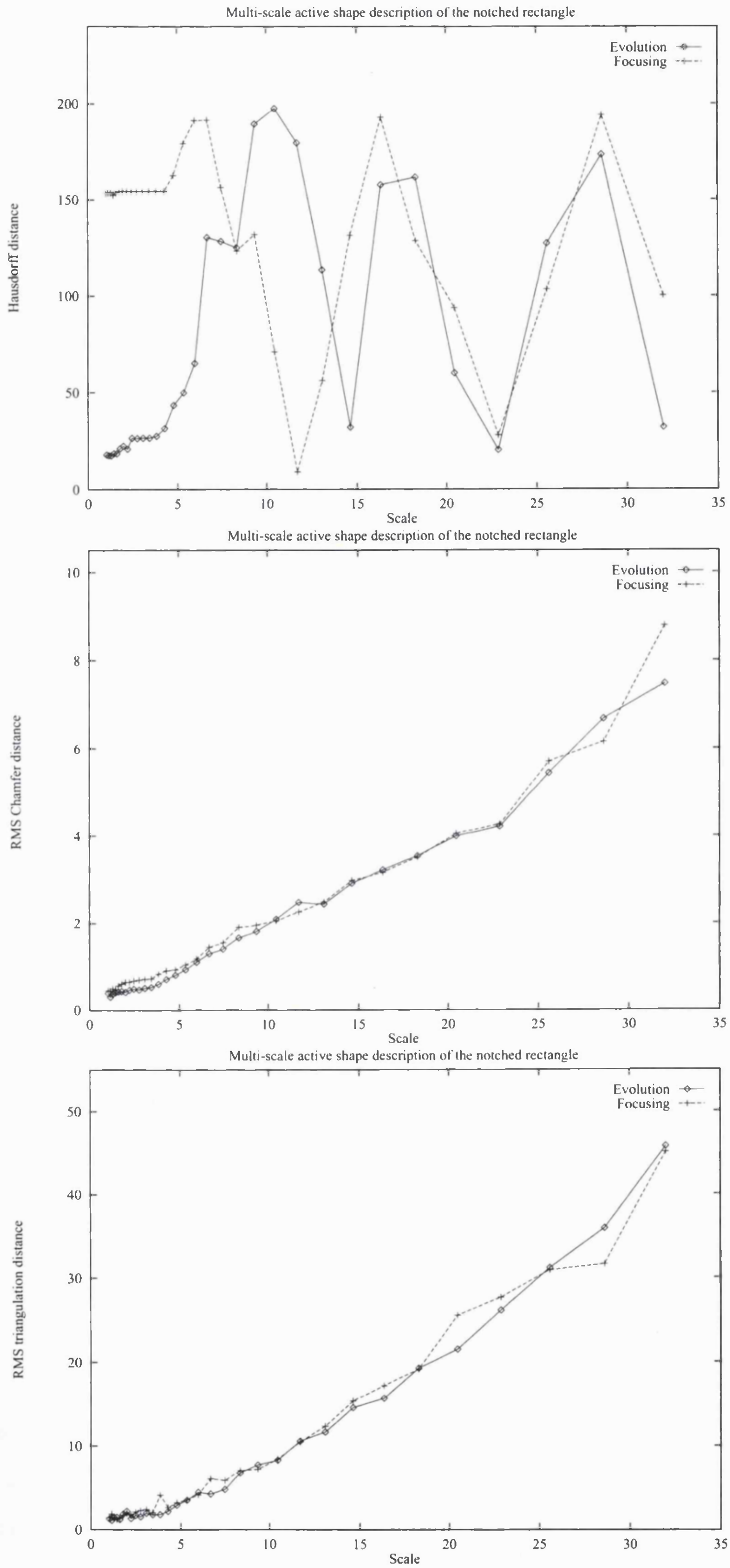


Figure 7.5: Hausdorff, RMS Chamfer, and RMS triangulation distances of the *notched rectangle* across scales for active shape evolution and focusing with respect to the known shape.

tion and focusing, respectively. Table 7.3 shows the respective mean values of these shape descriptors with respect to scale, and table 7.4 lists the respective shape changes across scale in terms of their slopes. It can be observed that the shape descriptors are of very similar values for active shape evolution and description (with the exception of the Hausdorff distance due to its nature of computing the *worst* mismatch between two shapes, rather than a *corresponding* or *closest* mismatch), as could be expected when comparing the stack of shapes for either process as shown in columns (d) and (e) of figure 7.3. It is also apparent that all global quantifiers increase for decreasing scales (characterized by a negative slope) when adjusting to more detailed or complex structures of the shape, and that the relative RMS Chamfer and triangulation distances continuously decrease for decreasing scales (which is characterized by a positive slope) when approaching the (same) known shape for decreasing scales (or increase when evolving from the known shape). Other global shape characteristics become only apparent when looking at shape changes across scales, rather than looking at the mean, slope, or lowest scale values only: at higher scales, there is an almost continuous decrease of the global quantifiers due to the shrinking effect of the diffusion or heat equation. At smaller scales, finer adjustment of the shapes leads to a local increase or decrease of the measurements. For example, it can be observed that at intermediate scales, around $\sigma = 10$, the area measurement has a local minimum. This happens at the scale at which the active contour model just leaves the notch (for active shape evolution) or locks to the notch (for active shape focusing), leading to a temporary decrease in area. The RMS Chamfer and triangulation distances, though showing a similar behaviour across scales, are in a completely different range with respect to each other; this is due to the nature of the Chamfer distance of measuring the closest rather than the corresponding distance between two shapes, leading to an underestimation of the actual shape deviation from the known shape. In this case, the main occurrence of this underestimation is at the deep protrusion of the *notched rectangle*. At the notch, the Chamfer distance measures the distance of each shape of the stack to points of the reference shape which lie nearby the notch, yielding significantly lower distance values. The triangulation distance, which is based on iteratively connecting all points of the shapes of the stack with the known shape, yields therefore much higher values at such protrusions.

In order to perform local (and therefore more qualitative) shape description across scales, it becomes obvious that a better form of visualization of a multi-scale shape stack is needed than illustrating and locally quantifying each shape separately as in figure 7.3. Recall that this particular shape stack, based on either active shape evolution or focusing, has been defined as a $2.5D$ structure. Ordering the regularized shapes contained in each stack with respect to the scale at which they have been obtained allows to concatenate this sequence of $2D$ shapes to a quasi $3D$ structure. Concatenation can be carried out using triangulation [Christiansen and Sederberg, 1978], and the result can be visualized using standard surface rendering techniques. In this dissertation, a freely available visualization package, *The Visualization Toolkit* [Schroeder *et al.*, 1997], is used for all rendering processes. Two types of surface rendering are used: wire frame rendering, which allows to perceive the triangular patches and the increasing contour sampling density for decreasing scales, and smoothly interpolated surface rendering, which continuously interpolates between the snaxels as well as between the scale levels, and which uses standard shading techniques. Figure 7.6 shows wire-frames and interpolated shaded surface renderings of the shape stacks obtained via active shape evolution and focusing of the *notched rectangle* test image, visualizing the shape stacks illustrated as individual contours in figure 7.3. Again it can be observed in this particular example that the shape stacks for active shape evolution and focusing are very similar, although they are computed by tracking from the opposite sides of the image scale-space. The *coarse-to-fine* view shown in column (a) of figure 7.6 shows the sparser sampling at higher scales, in contrast to the dense sampling which can be better perceived in the *fine-to-coarse* view illustrated in column (b). Higher-order scale-spaces can be visualized as a sequence of $2\frac{1}{2}D$ stacks, or by rendering the concatenated volumes at each scale individually. Chapter 9 will show renderings of $3\frac{1}{2}D$ shape stacks of volumetric medical data, as well as sequences of $2\frac{1}{2}D$ stacks of singular image slices.

Based on such a visualization of the multi-scale shape stack, local shape metrics can be mapped onto the rendered surface. Figure A.1 in appendix A shows the surface rendered shape stacks of column (c) in figure 7.6 with mapped colour coding of local shape metrics. Colour mapping can be performed using the *Hue-Saturation-Value* (HSV) model, illustrated in figure 7.7 [Foley *et al.*, 1990]. The hue (corresponding to the wavelength of a colour) can be adjusted to the size of a local shape metric according to the range $[0^o; 360^o]$ of all possible colours or within $[0; 1]$, respectively, the saturation (corresponding to the purity of a colour) can be set to 1 (100%), and the value (corresponding to the intensity of a colour) is given by the grey-level shading of the surface rendering and lies also in the range $[0; 1]$. Scaling a given list of shape metrics within a suitable subset of possible hue values allows to add local shape information, which is performed by creating a so-called *colour lookup table.* For example, column (a) of figure A.1 shows the mapping results for the local curvature $\kappa(\tilde{\mathrm{v}}(s); \sigma)$, which is coded between the colours green (for positive curvature values or outward bending of the shape) and red (for negative curvature values corresponding to inward bending behaviour), and yellow represents low positive or negative curvature values. Zero curvature parts (within a small tolerance span) are not colour coded at all. Note that the high inward (negative) curvature inside the notch, and the high outward (positive) curvature at the corners of the *notched rectangle* can be easily visually perceived, and a similar visual impact of the curvature mapping for active shape evolution and focusing is achieved. Column (b) of figure A.1 illustrates the mapping of the local Chamfer distance to the known shape of the *notched rectangle* between yellow and red (as only positive distances can occur), with no colour coding for very

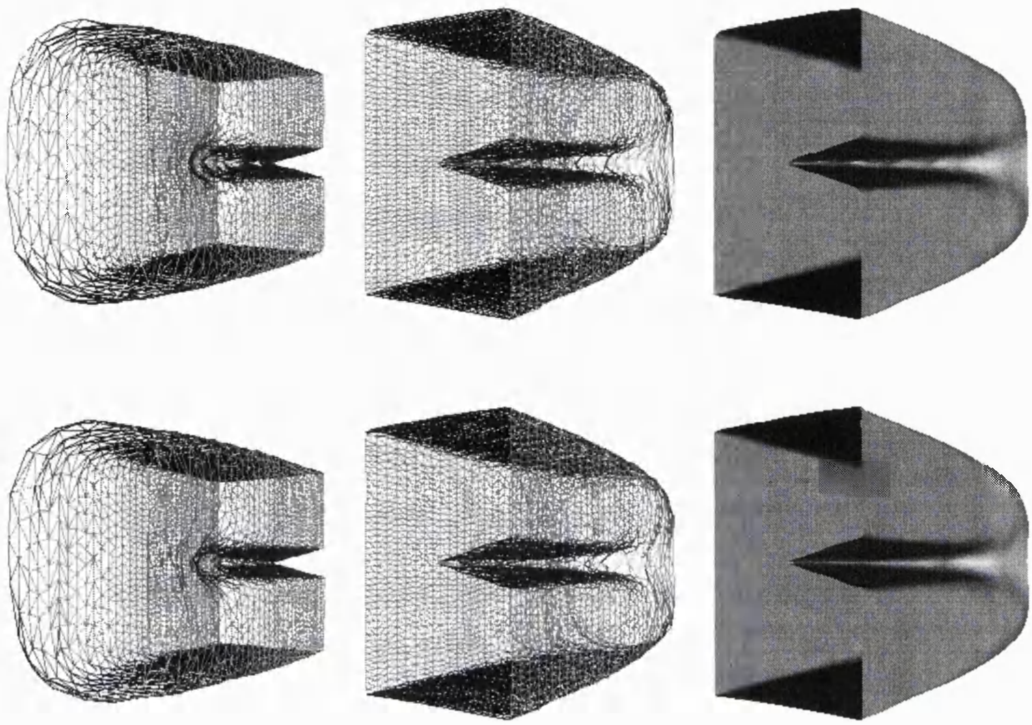


Figure 7.6: Wire-frame and interpolated shaded surface renderings of the multi-scale shape stack of the *notched rectangle* image obtained via active shape evolution (upper row) and active shape focusing (lower row). Column (a) Wire frame rendering in *coarse-to-fine* view. Column (b) Wire-frame rendering in *fine-to-coarse view.* Column (c) Surface rendering in *fine-to-coarse* view. See text for further details.

small distances. Column (c) of figure A.1 analogously shows the local triangulation distances. Both local distance mappings show a similar high (orange to red) value near the notch at high scales, as well as at corners. However, both distance values have been scaled within the same range (the lower range of the Chamfer distance was chosen) in order to demonstrate the difference in their behaviour. From that it becomes apparent that the triangulation distance is superior to the Chamfer distance in terms of yielding a better impact of the deepness of the protrusion or notch due to the much larger local distance value at higher scales.

The distance mappings are performed using a *logarithmic* lookup table to increase the visual impact of smaller values, while the local curvature mapping is based on a *linear* lookup table [Foley *et al.*, 1990].

7.6 Summary

This chapter has presented a new and effective method for multi-scale active shape description, applying the multi-scale active contour model to perform *fine-to-coarse* active shape evolution, or *coarse-to-fine* active shape focusing. Both techniques give rise to a multi-scale shape stack,

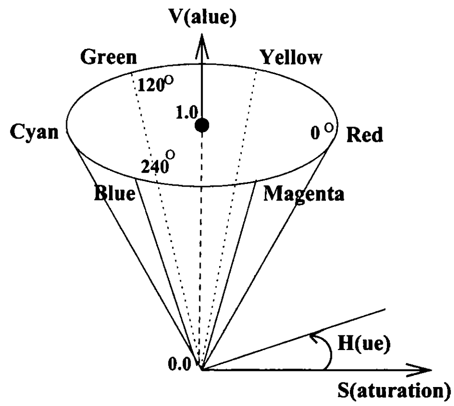


Figure 7.7: HSV colour model. The cylindrical coordinates correspond to hue, saturation, and value.

which can be suitably visualized as a $2\frac{1}{2}D$ structure for a single planar shape, or for a volumetric set of planar shapes as either a sequence of $2\frac{1}{2}D$ structures or a sequence of $3D$ structures. Scale-space notations in terms of the scale-related dimensionalities of the image scale-space, the shape tracking techniques in scale-space, and the resulting shape stacks have been introduced. The technique of multi-scale active shape description is based on the investigation of a multi-scale shape stack, as well as on a suitably chosen set of standard shape descriptors. The novelty of this approach is to investigate a planar shape or a volumetric set of planar shapes at various scale levels, rather than using a traditional single-scale approach. The applicability of all techniques has been demonstrated using the *notched rectangle* test image as an example, of which global and mean values across scales, as well as local shape mappings onto the shape stacks where shown to provide higher level shape information, which is not readily available from the known underlying shape. This chapter has also discussed reasonable scale sampling strategies with respect to the possible displacement of edges between adjacent scale levels, an issue which can be resolved by adjusting the local search space of the multi-scale active contour optimization adequately. Active shape evolution and focusing have been identified as dual techniques, with the advantage of active shape focusing over evolution of needing no *ground truth*, since it is performing implicit segmentation on the fly. Furthermore, the proposed techniques have been put into relation with existing techniques for single-scale and multi-scale contour analysis, and their higher level of applicability has been motivated. The following two chapters will test the presented techniques on fractal and other synthetic images (chapter 8), and medical applications (chapter 9).

Chapter 8

Application to Fractal and Other Synthetic Images

LE PETIT PRINCE LE REGARDA LONGTEMPS:
– TU ES UNE DRÔLE DE BÊTE, LUI DIT-IL ENFIN, MINCE COMME UN DOIGT...
– MAIS JE SUIS PLUS PUISSANT QUE LE DOIGT D'UN ROI, DIT LE SERPENT.

THE LITTLE PRINCE GAZED AT HIM FOR A LONG TIME.
"YOU ARE A FUNNY ANIMAL," HE SAID AT LAST. "YOU ARE NO THICKER THAN A FINGER..."
"BUT I AM MORE POWERFUL THAN THE FINGER OF A KING," SAID THE SNAKE.

Le Petit Prince, Antoine de Saint-Exupéry.

In this chapter two main types of synthetic objects will be investigated in order to test and demonstrate the applicability of the proposed multi-scale active shape description technique on objects with high curvature parts and protrusions, and fractal objects. All synthetic objects are silhouette-shaped, to allow for comparison with binary classic curve evolution techniques, and for analysis with respect to a ground truth. The application of active shape evolution and focusing to a true fractal structure will show that the proposed techniques preserve fractal characteristics, and provide a meaningful relationship between fractal resolution level and image scale.

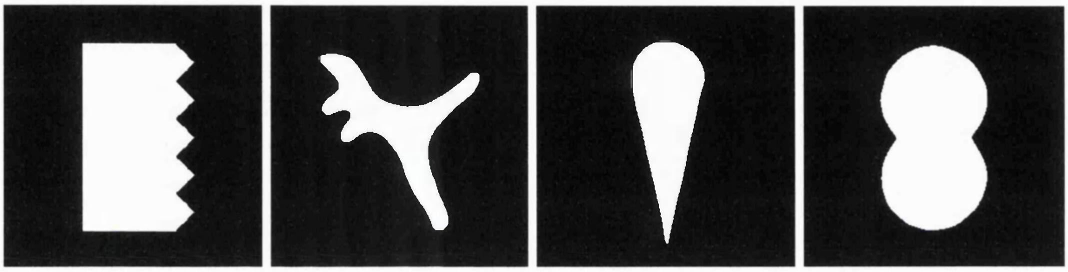
Table 8.1 lists the optimization parameters used for the active shape evolution and focusing techniques for all synthetic images, with the default parameters chosen as given in table 6.5 (chapter 6, page 141). Active shape evolution and focusing are performed according to algorithms 7.1 and 7.2, respectively, using the *greedy* algorithm for optimization. No preprocessing like thresholding or histogram equalization is required for any of the synthetic images.

8.1 Application to Strongly Curved Objects and Objects with Protrusions

Protruded objects are in general hard to describe in terms of quantitative shape description. Their high number of curvature changes poses difficulties for classic active contour models which by

Parameter	Description	Value
ϵ_1	Upper contour scale stabilizer	0.8
ϵ_2	Lower contour scale stabilizer	3.5
$M \times M$	*Greedy* search space	7×7

Table 8.1: Optimization parameters for active shape evolution and focusing of the synthetic images. Remaining default parameters can be found in table 6.5.



(a) (b) (c) (d)

Figure 8.1: Synthetic test images containing strongly curved objects and objects with protrusions. (a) *Saw-toothed rectangle*. (b) *Kangaroo*. (c) *Teardrop*. (d) *Blobs*.

their nature minimize the internal elasticity and curvature energy terms. One example for multi-scale active shape description of an object with a deep protrusion has already been demonstrated in the previous chapter in terms of the *notched rectangle* image. Figure 8.1 shows four more synthetic test images, which will be referred to as *saw-toothed rectangle*, *kangaroo*, *teardrop*, and *blobs*. They have been selected for their following properties: the *saw-toothed rectangle* image contains high curvature parts and corners, as well as a combination of global rectangular shape and specific and detailed saw-toothed edges; the *kangaroo* image has also high curvature parts, is widely elongated and of very low elasticity; the *teardrop* image contains a very sharp peak and is therefore complementary to the *notched rectangle* test image, and additionally its spatial width continually increases from the tip of its peak to its circular, blob-like end; finally, the *blobs* image consists of two large circular structures which are connected at a small subpart, yielding inward bound high curvature peaks. All objects, including the *notched rectangle*, are frequently used in scale-space research, e.g. in [Morse, 1994; Whitaker, 1994b], since they demonstrate that single-scale description is often not sufficient to capture all important shape characteristics.

Figure 8.2 shows the ground truths and the initial models which are used for the active shape evolution and focusing processes of the synthetic images in figure 8.1. A total of $n = 32$ scale samples has been chosen, with $\sigma_0 = 1$ and $\sigma_{31} = 32$, with the exception of the *kangaroo* image, for which $n = 20$ samples are used, with the scale ranging from $\sigma_0 = 1$ to $\sigma_{19} = 20$. This particular image needs to be processed with a smaller upper scale due to its significantly lower overall spatial width, which also leads to a decrease of the necessary number of scale samples in

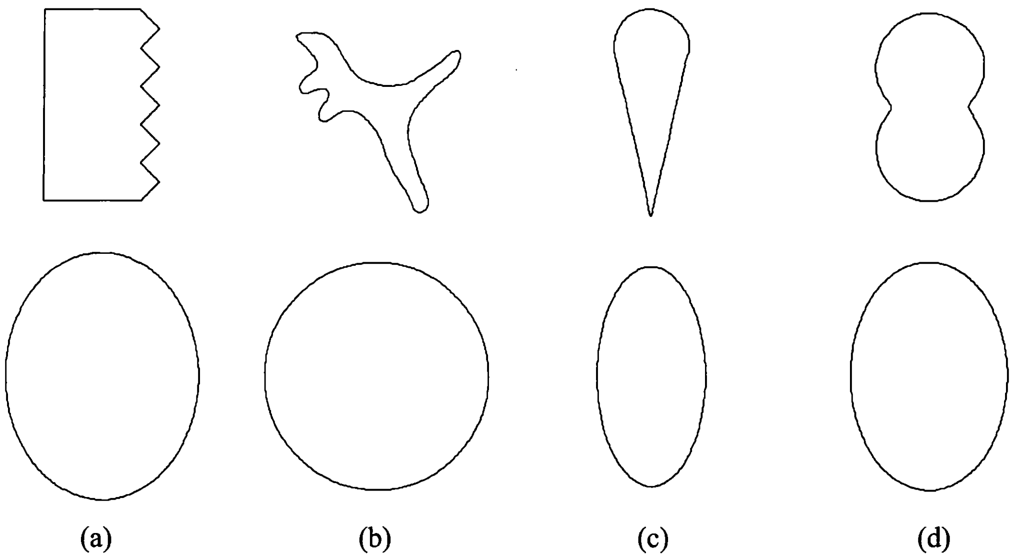


Figure 8.2: Ground truths (upper row) and initial models (lower row) for the respective synthetic test images of figure 8.1 (a)-(d). The size of the ellipse-shaped initial models was chosen to enclose the shapes of interest.

the sense of Bergholm [Bergholm, 1987] (see chapter 7, section 7.5).

8.1.1 Qualitative Description

Figures 8.3 and 8.4 show samples of the active shape evolution and focusing results, respectively. As for the *notched rectangle* example in the previous chapter, similar results for both techniques have been obtained for all synthetic shapes. From a qualitative visual inspection, the following observations can be made for the evolution and focusing of each synthetic shape:

Saw-toothed rectangle: At the highest scale, both evolution and focusing yield an ellipse, which slowly focuses down or has evolved from a rectangular shape. At intermediate scales, the saw-teeth start appearing, first in a smooth wave form, and then becoming more and more prominent. At the lowest scale, the sharp peaks of the saw-teeth are recovered. Both active shape evolution and focusing are therefore adequate techniques to reveal the two main characteristics of this particular shape, in terms of global (higher scale) overall shape and specific (lower-scale) detail, which single-scale techniques would not be capable of.

Kangaroo: As stated above, the kangaroo is evolved and focused in a lower scale range than the other synthetic examples. Recall from the reviewed techniques of *blobs* and *cores* (chapter 3, sections 3.4.3.1 and 3.4.3.3, respectively) that scale-selection plays an important role in feature detection. At different scales, different object properties can be investigated whose spatial size is directly related to the chosen scale. In other words, if one chooses the upper scale too small, the overall object shape cannot be recovered due to local minima of the

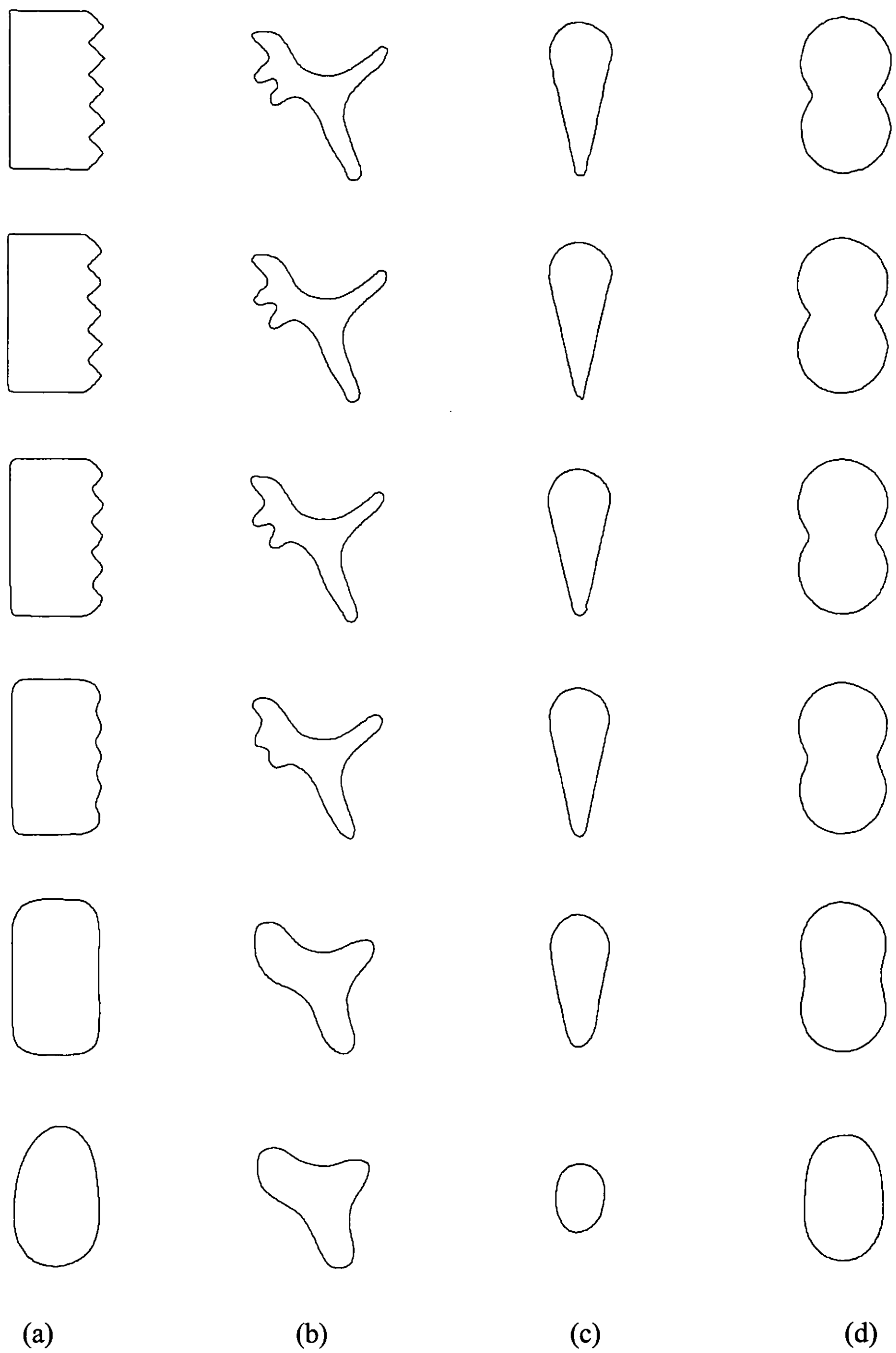


Figure 8.3: Active shape evolution results for synthetic test images containing strongly curved objects and objects with protrusions. The illustrated scale samples in columns (a), (c), and (d) are (from top to bottom): $\sigma = 1, 2.18712, 4.27752, 8.36588, 16.3618, 32$ with $\sigma_{n-1} = 32, \sigma_0 = 1, n = 32$. Column (b) was obtained in a slightly smaller scale range (from top to bottom): $\sigma = 1, 2.19975, 4.13309, 7.76563, 7.76563, 17.0826, 20$ with $\sigma_{n-1} = 20, \sigma_0 = 1, n = 20$. See text for further details.

multi-scale energy function of the active contour model, which are caused by fine details and smaller scale structures. If one chooses the upper scale too high, the object smoothes

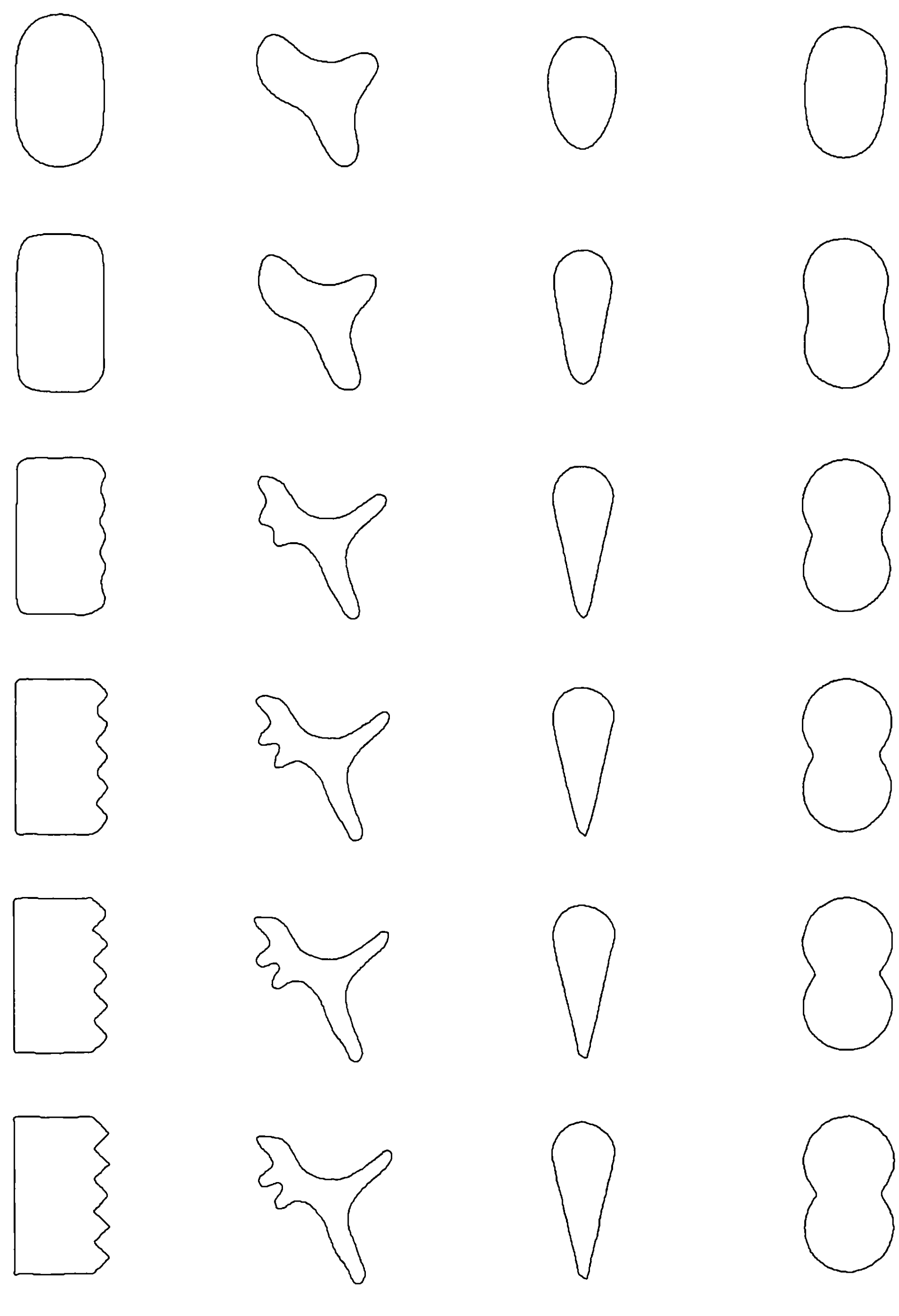


Figure 8.4: Active shape focusing results for synthetic test images containing strongly curved objects and objects with protrusions. The illustrated scale samples in columns (a), (c), and (d) are (from top to bottom): $\sigma = 32, 16.3618, 8.36588, 4.27752, 2.18712, 1$ with $\sigma_{n-1} = 32, \sigma_0 = 1, n = 32$. Column (b) was obtained in a slightly smaller scale range (from top to bottom): $\sigma = 20, 17.0826, 7.76563, 4.13309, 2.19975, 1$ with $\sigma_{n-1} = 20, \sigma_0 = 1, n = 20$. See text for further details.

out completely. If scale selection is performed with care, however, both active shape evolution and focusing are able to reveal the overall shape with three main elongations at higher scales, and recover smaller substructures (i.e. the legs, front paws, and tail) at their *adequate scales* corresponding to their spatial widths. Consequently, for a too high upper scale the techniques are only able to recover the main body of the *kangaroo* as substructures are *blocked* by high gradient values and ridges along the body or completely disappear, even though they reappear at smaller scales. Another problem with this particular image lies in the elasticity minimizing nature of the multi-scale active contour model: it is "cheaper" in terms of energy costs to make a short cut along the main body rather than following very "expensive" shape outlines of very high elasticity. The adaptive resampling technique, as was shown in chapter 6, partially but not completely eliminates this behaviour, as it is also restricted by the scale-related contour scale. Protrusions or elongations whose widths are smaller or only slightly larger than the respective contour and image scales can therefore only be tracked with difficulties, as will be further discussed in the next example.

Teardrop: The results for the *teardrop* show that the blob-shaped top of the structured is maintained throughout all scale levels, while the peak of the *teardrop* disappears at higher scales almost completely, and is continuously recovered at smaller scales. Two main problems can be observed for this particular example: The results of active shape evolution and focusing at the highest scale differ, and yield a more blob-like structure for shape evolution, and a more elongated, slightly larger structure for shape focusing. The second problem is that the peak of the *teardrop* is not completely recovered at the lowest scale with either technique. Both problems can be explained by the multi-scale adaptive resampling technique: As mentioned earlier, adaptive resampling uses a certain tolerance span for the enforcement of the contour scale, as well as lower contour scale limit. The first point may lead to a variation of the contour scale $\tilde{\varsigma}_i$ at scale level σ_i between $[\tilde{\varsigma}_i - \Delta\tilde{\varsigma}_i; \tilde{\varsigma}_i + \Delta\tilde{\varsigma}_i]$ (see algorithm 6.1, chapter 6), where $\Delta\tilde{\varsigma}_i$ is chosen to be in direct relation to the contour scale. In this case, the choice of $\Delta\tilde{\varsigma}_i = \frac{1}{2}\tilde{\varsigma}_i$ leads to a variation of the highest contour scale between $[32 - 16; 32 + 16]$. This may cause a significant difference in the local sampling density at high scales, which is the case in this particular example. Choosing a smaller tolerance span may eliminate this problem, but compromises the local flexibility of the model, as was discussed in the concept of strict uniform sampling having no tolerance span at all (see chapter 6, section 6.1.2.2). Alternatively, a lower upper scale limit may be chosen, leading to similar results for either technique. This can be seen when cutting the shape stacks in columns (c) of figures 8.3 and 8.4 below or above scale level $\sigma_i = 16.3618$, for example. This approach can also be justified by the relatively small overall size of the *teardrop*. The second problem of the inadequate tracking of the peak of the *teardrop* lies in the fixed lower contour scale, chosen to be $\tilde{\varsigma}_{min} = 3.5$. This leads to a relatively large minimum snaxel spacing of the contour model even at small scales, and consequently to an inability to track elongations or protrusions whose entrance is narrower. This problem can be solved by choosing a lower minimum contour scale, e.g. $\tilde{\varsigma}_{min} = 2.0$, in which case the size of the search space needs to be decreased to 3×3 to avoid overlapping search spaces. This naturally leads to the necessity of increasing the number of scale samples, in order to restrict the maximum shift of edges to one pixel only.

Blobs: The two circular blobs in this image are recovered throughout all scale levels in either process, while the peak-like concavities are continuously recovered at decreasing scales, similarly to the peak of the *teardrop*. The overall, ellipse-like shape becomes therefore apparent at the highest scale level, while the specific detail of the inward peaks only appears at intermediate and fine scales. Both active shape evolution and focusing yield very similar and robust multi-scale shape hierarchies of this particular image.

In the following section, the obtained multi-scale hierarchical shape stacks of figures 8.3 and 8.4 will be quantified with the methods presented in section 7.4 of the previous chapter.

8.1.2 Quantitative Description

In analogy to the example of multi-scale active shape description in the previous chapter, the multi-scale hierarchical shape stacks of the synthetic test images in this chapter can be investigated using different quantitative shape measurements across scales. Figures 8.5 and 8.6 illustrate the global planar shape metrics of table 6.5 (chapter 6, page 141) plotted against scale, and tables 8.2 and 8.3 list the corresponding mean shape metrics and slopes with respect to scale. At a first glance, one can perceive two things: first, the shape metrics of all four shapes are distinctively different from each other (with the exception for the Hausdorff distance measurement, whose limited ability of yielding similar values for similar shapes has been already pointed out and demonstrated in the previous chapter, but which has been included for completeness). Second, the shape metrics for each shape are very similar if obtained via active shape evolution and focusing, supporting the results of the previous chapter and the observation in the qualitative description made above that active shape evolution and focusing yield approximately the same results and are therefore dual techniques. The individual quantitative inspection of the shapes obtained by either shape evolution or focusing leads to the following conclusions:

Saw-toothed rectangle: When comparing the shape metrics with the ones obtained for the *notched rectangle* in the previous chapter, a number of similarities can be observed because these two shapes have the highest resemblance in the chosen test set. For intermediate to higher scales, the rectangular shape characteristics are recovered. Similar mean and slope measurements for area, perimeter, and compactness, with a small offset due to the larger overall size of the *notched rectangle* are obtained. The larger mean RMS triangulation dis-

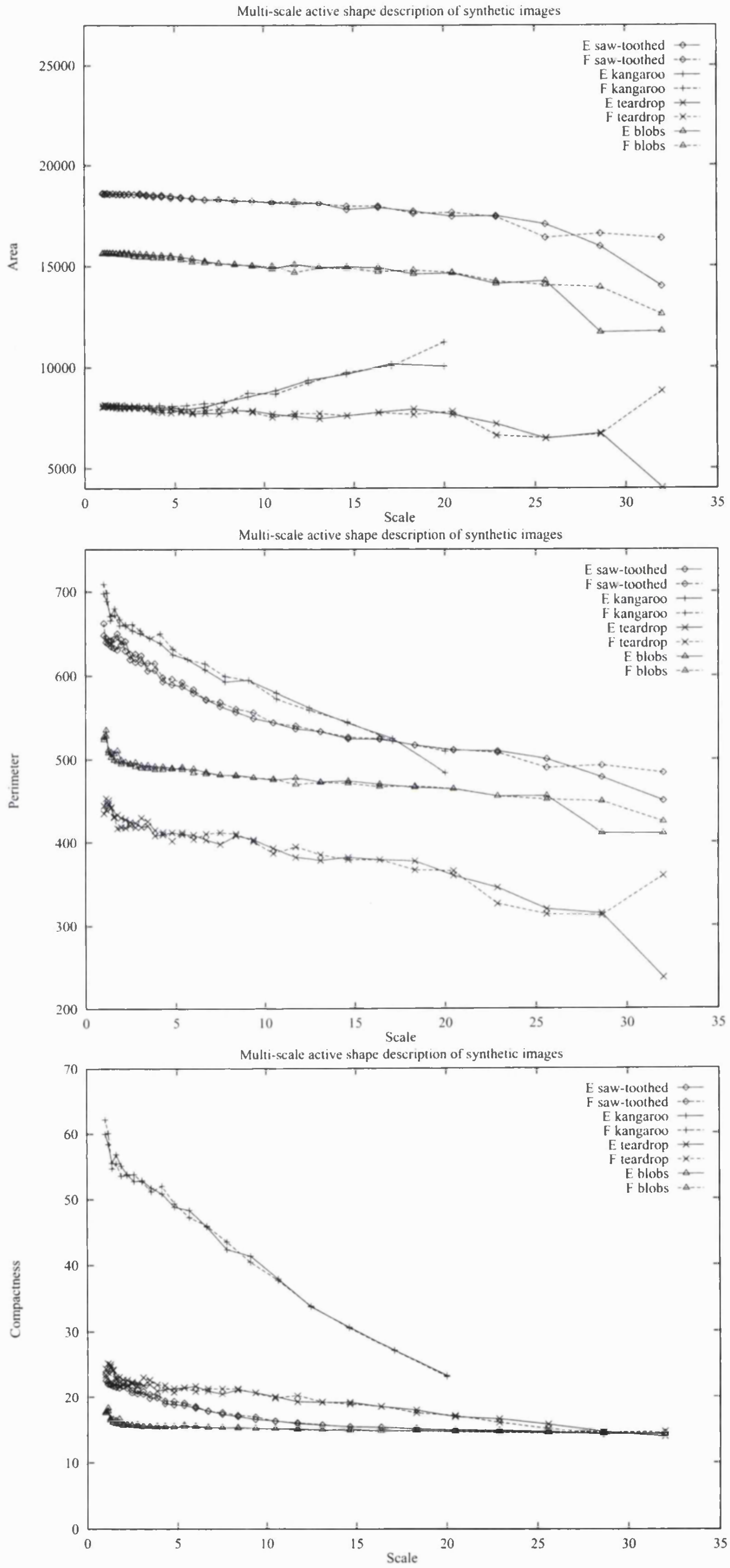


Figure 8.5: Area, perimeter and compactness of the synthetic images across scales for active shape evolution (E) and focusing (F).

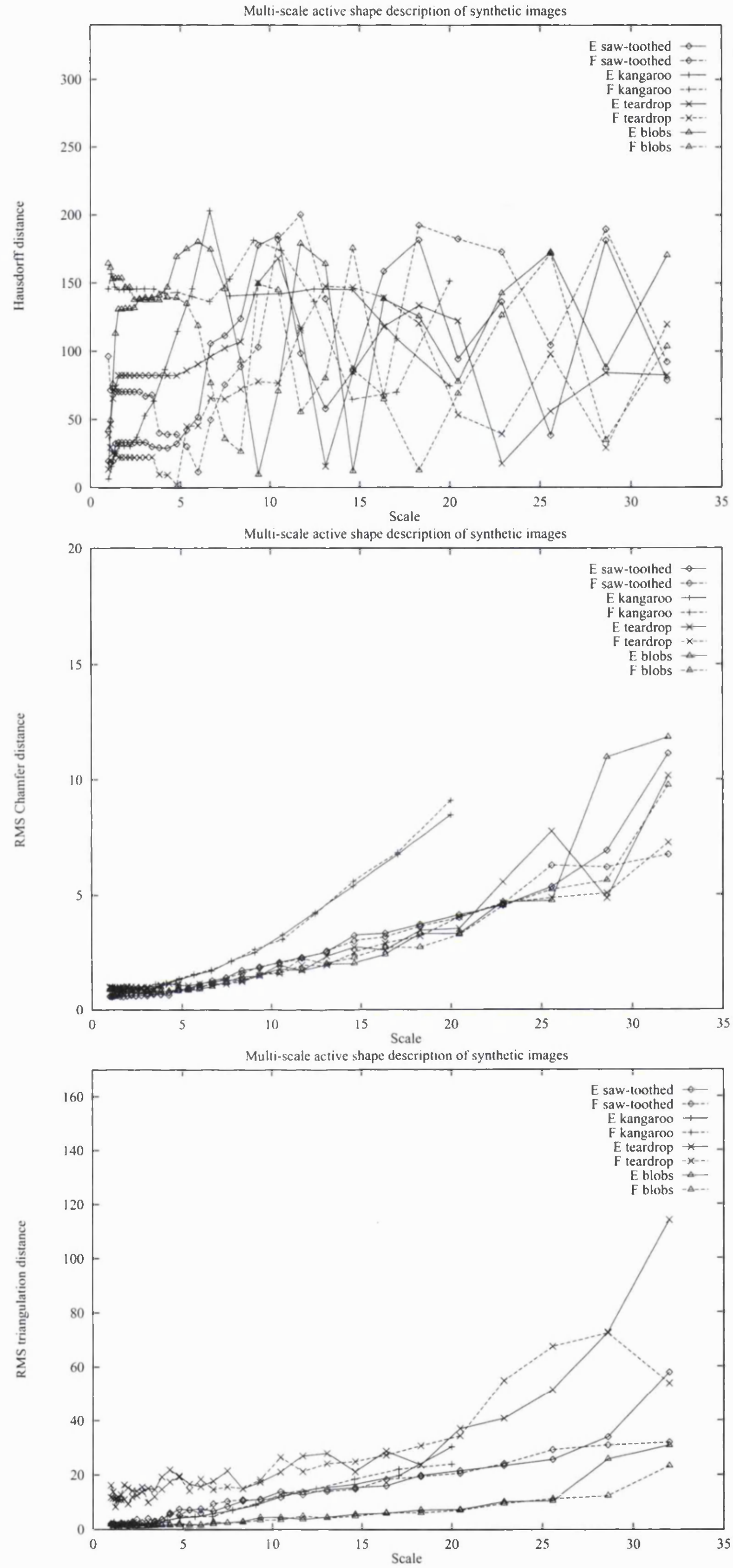


Figure 8.6: Hausdorff, RMS Chamfer, and RMS triangulation distances of the synthetic images across scales for active shape evolution (E) and focusing (F) with respect to the known shape.

Image	Method	$\overline{A}$	$\overline{P}$	$\overline{C}$	$\overline{\text{dist}_H}$	$\overline{\text{dist}_{C_{RMS}}}$	$\overline{\text{dist}_{T_{RMS}}}$
saw-toothed	E	18062	576.350	18.4658	73.5816	1.44561	26.58880
	F	18138	578.688	18.5400	92.3722	1.30706	24.53830
kangaroo	E	8488	619.300	46.3591	87.1255	2.06117	13.11710
	F	8582	622.225	46.3983	140.8110	2.12174	13.10450
teardrop	E	7676	396.891	20.5609	85.5625	1.82166	39.43280
	F	7798	400.503	20.6393	55.5206	1.62188	35.93470
blobs	E	15042	483.553	15.5556	122.8810	1.88198	8.58081
	F	15077	483.022	15.4790	118.2030	1.69480	7.16850

Table 8.2: Mean descriptors for active shape evolution (E) and focusing (F) of the synthetic images with respect to scale. The size measurements for the mean area perimeter, as well as the mean Hausdorff, RMS Chamfer, and RMS triangulation distance measurements are in pixel units, while the mean compactness is dimensionless.

Image	Method	$\frac{\Delta A}{\Delta\sigma}$	$\frac{\Delta P}{\Delta\sigma}$	$\frac{\Delta C}{\Delta\sigma}$	$\frac{\Delta\overline{\text{dist}_{C_{RMS}}}}{\Delta\sigma}$	$\frac{\Delta\overline{\text{dist}_{T_{RMS}}}}{\Delta\sigma}$
saw-toothed	E	-93.2691	-6.07053	-0.292346	0.189365	3.05724
	F	-68.6520	-5.72710	-0.294595	0.155573	3.04641
kangaroo	E	118.3960	-10.04900	-1.885900	0.343516	5.29156
	F	139.7460	-9.73863	-1.905110	0.354570	5.29158
teardrop	E	-68.7714	-4.74626	-0.302381	0.194024	3.05724
	F	-25.1534	-3.77179	-0.310750	0.151193	3.04642
blobs	E	-98.2331	-2.75860	-0.074620	0.238537	3.05724
	F	-69.8549	-2.21904	-0.070505	0.193869	3.04641

Table 8.3: Slope measurements for active shape evolution (E) and focusing (F) of the synthetic images with respect to scale. The slopes are estimated using linear regression on the graphs plotted in figures 8.5 and 8.6.

tance measurement for the *saw-toothed rectangle* may at first seem incorrect, since it appears to be more rectangular shaped throughout all scales. However, the *notched rectangle* is rectangular shaped everywhere except for a small part (the entrance to the notch), whereas one whole side of the *saw-toothed rectangle* deviates from the rectangular shape, though in smaller distance values. The most prominent difference between the two shapes is the continuous increase of area for decreasing scales for the *saw-toothed rectangle*, in contrast to the large local area minimum due to the deep protrusion of the *notched rectangle*.

Kangaroo: This shape is the only shape of the test set to have a decreasing area for decreasing scales, yielding a negative area slope measurement. At the same time, it has the steepest increase and highest overall perimeter, yielding the highest compactness and steepest increase for decreasing scales. This interesting aspect can be explained by the high convolutedness of the *kangaroo*. At high scales, short cuts between the substructures are made (as was observed in the qualitative description in section 8.1.1), yielding higher area measurements at larger scales, while shortening the overall perimeter. The area measurements for active shape evolution and focusing differ slightly at the highest scale due to the aforementioned problems of selecting the upper scale, leading to a lower slope value for the evolution result. Finally, the RMS Chamfer deviation of the kangaroo shape from its ground truth has the overall largest value from the test set, in contrast to the RMS triangulation distance which yields higher values for the *teardrop* which will be further discussed below.

Teardrop: The shape metrics of the *teardrop* shape show differing area, perimeter and RMS triangulation distance measurements at the highest scale, supporting the qualitative results of section 8.1.1. The very small size of the shape obtained via evolution at the highest scale is reflected by much lower perimeter and area measurements, and by a higher distance to the ground truth in comparison to the respective values for the focusing result. It can also be noted that this higher deviation from the reference shape is not revealed by the RMS Chamfer distance measurement due to its nature of measuring the closest rather than corresponding distance. Finally, due to the inadequate tracking of the peak at the lowest scales, the triangulation distance metric yields the highest values across scales.

Blobs: In contrast to the *teardrop*, this shape is the overall closest to its ground truth in terms of the mean and plotted triangulation distance values. This is due to the low compactness (or high roundness) of the *blobs*. The tracking of the inward peaks at the lower scales can be observed from the sharp increase of the perimeter. At high scales, the same problems in terms of a deviation between focusing and evolution occur, leading to different estimates in the slope measurements, but not affecting the mean values.

In the following section, the visualization and mapping of local shape metrics onto the multi-scale shape stacks of the synthetic test shapes will be presented.

8.1.3 Visualization

Figure 8.7 illustrates the *fine-to-coarse* surface rendering of the $2\frac{1}{2}D$ multi-scale shape stacks of the synthetic images obtained via active shape evolution and focusing. Slices of these stacks were shown in figures 8.3 and 8.4 above. Local curvature, Chamfer and triangulation distance mappings onto these stacks are illustrated in appendix A in figures A.2-A.5, with colour lookup

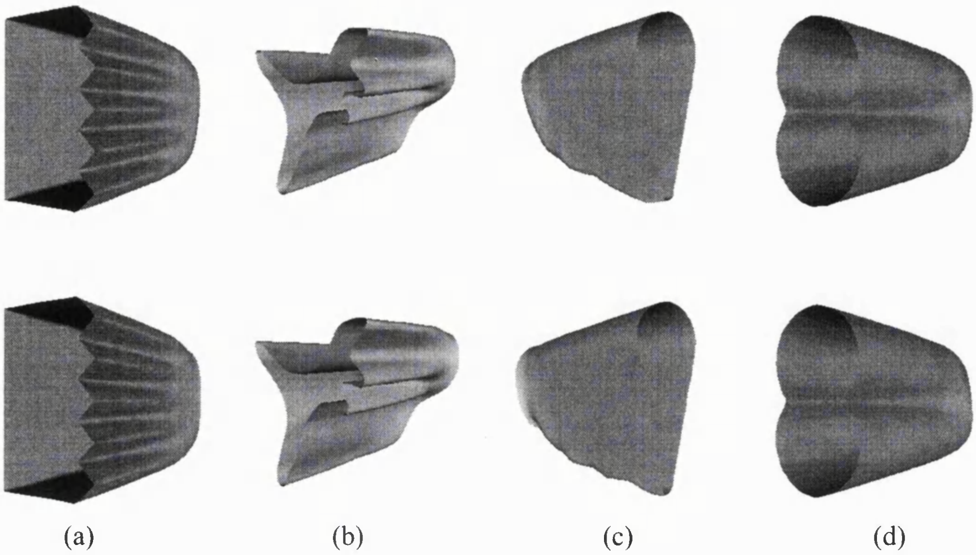


Figure 8.7: Interpolated shaded surface renderings of the hierarchical shape stack of the synthetic images obtained via active shape evolution (upper row) and active shape focusing (lower row). Surface rendering is performed in a *fine-to-coarse* view. See text for further details.

tables as presented in section 7.5 in the previous chapter. Regarding the local curvature mappings illustrated in column (a) of these figures, the saw-teeth and corners of the *saw-tooth rectangle*, the elongated substructures of the *kangaroo*, the peak of the *teardrop*, as well as the inward peaks of the *blobs* can be well perceived on both types of shape stacks, in terms of outward (mapped in green) and inward (mapped in red) prominent bending behaviour. The Chamfer and triangulation distance mappings, in columns (b) and (c), respectively, provide a good visual impact of the shape flow in either direction in scale-space. The poor tracking of the peak of the teardrop can only be perceived from the triangulation distance mapping, yielding a higher corresponding distance.

The following section applies active shape evolution and focusing to a true fractal structure of increasing degrees of self-similarity, in order to investigate the correspondence between the fractal resolution levels and image scales.

8.2 Application to Fractal Objects

Section 2.3.3 (chapter 2) has introduced the theory of fractals, whose main characteristic is their self-similarity, or the geometric interpretation of their fragmentation and irregularities. Shapes other than truly circular, ellipse-shaped, or polyhedral are often difficult to describe in mathematical terms, yet some of them show a particular space filling behaviour (and were originally described as *monster curves* or simply as *pathological*). The fractal dimension describes the degree in which planar curves of Euclidean dimension $D = 1$ fill out their embedding space of Euclidean dimension $D = 2$. For true fractals, an analytical value for their fractal dimension

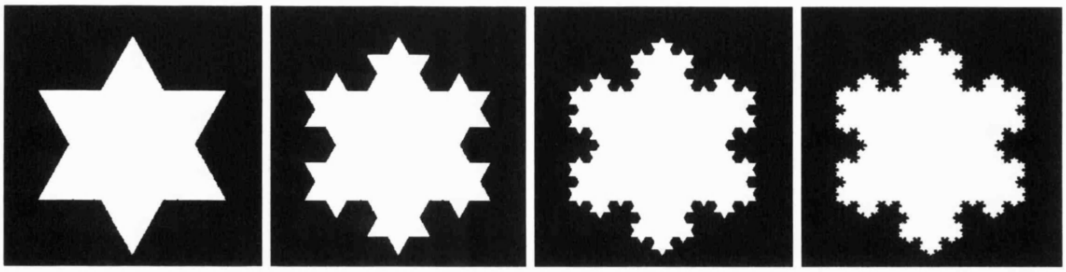


Figure 8.8: Von Koch curve (*snowflake*) test image series for increasing fractal resolution levels r. These images were obtained from the analytic curves of figure 2.10 (b) (chapter 2, page 46).

can be derived from their iterative fractal generation scheme, or by investigating the structure at each of its fractal generation or *resolution* levels. The term resolution is probably more appropriate, since *zooming* into a fractal structure recovers similarly structured subparts. This property is called *scale invariance*, as over all fractal scales the same fragmentation can be observed. Obviously, the topics of fractal resolution and scale-space theory (chapter 3) are strongly related, as they both give rise to multi-scale shape analysis. A curve which is not a true fractal structure (which is most often the case) can still exhibit *statistical scale invariance*, a feature which is often explored for classification purposes.

An example of a fractal generation has been shown in figure 2.10 (a), with resulting fractal curves of different fractal generations (or resolutions) shown in figure 2.10 (b). The resulting set of fractal shapes, depicted as synthetic silhouette images in figure 8.8, are in the following investigated using the proposed technique of multi-scale active shape description. The same energy and optimization parameters as for the other synthetic images above are used, with a chosen scale range of $\sigma \in [1; 32]$ for active shape evolution and $\sigma \in [32; 1]$ for active shape focusing, respectively, with $n = 32$ scale samples.

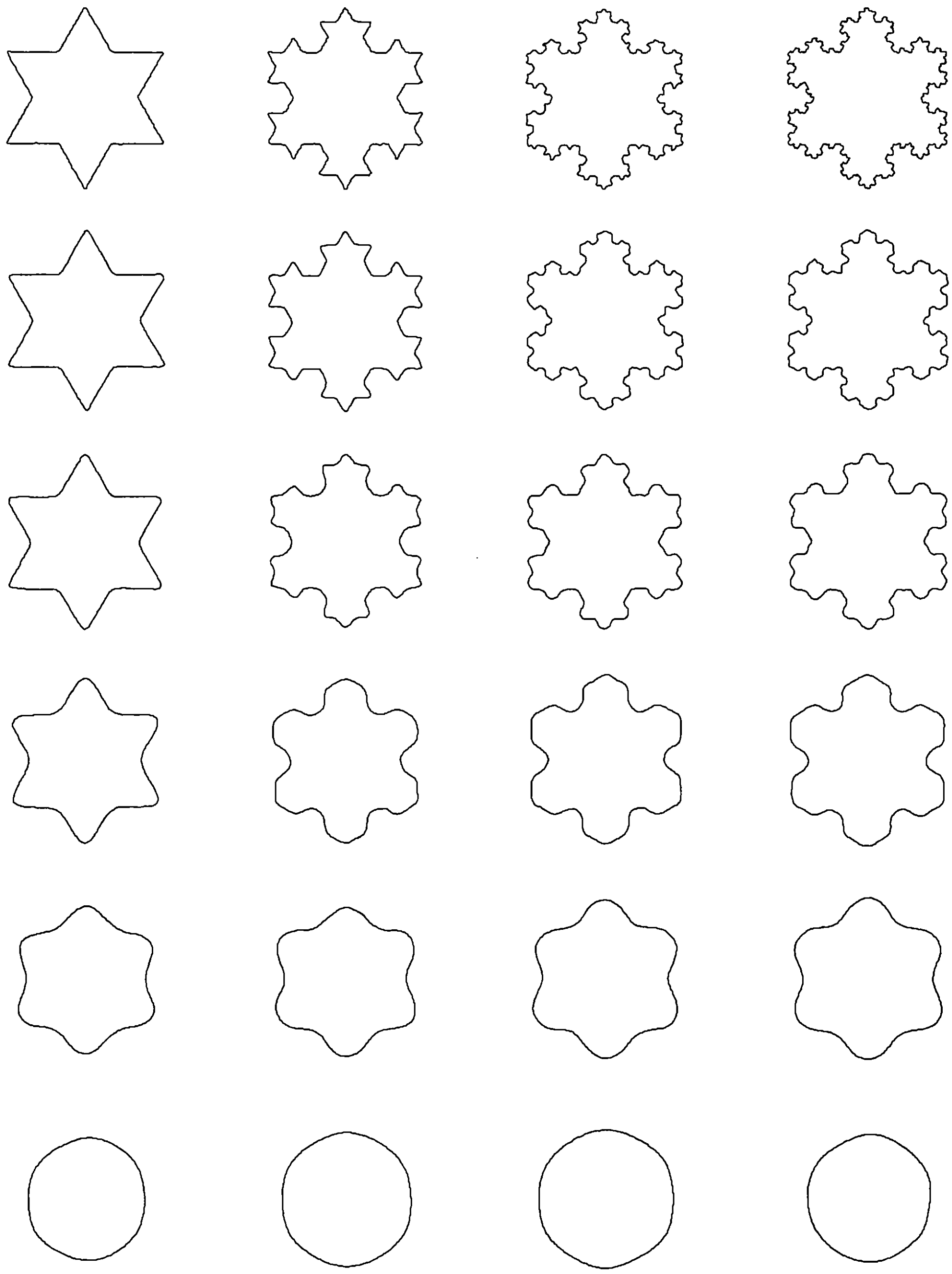
8.2.1 Qualitative Description

Figures 8.9 and 8.10 illustrate samples of the active shape evolution and focusing results for the different fractal resolution levels of the von Koch curve, yielding similar shapes at corresponding fractal resolutions and image scales for either technique. The most interesting aspect of the illustrated samples is the observation that between the highest two illustrated scale levels $\sigma_i \in [16.3618; 32]$, all fractal resolution curves for both techniques look almost identical. Between the scale levels $\sigma_i \in [4.27752; 16.3618]$, the fractal curves at resolution levels $r = 2, \cdots, 4$ still look identical for either technique, while the lowest fractal resolution curve approaches already its final star-shaped result. Between scale levels $\sigma_i \in [2.18712; 4.2775]$, the two higher fractal resolution curves with $r = 3, 4$ look still identical, while the next lower fractal resolution curve with $r = 2$ approaches its final, once substructured, star-like shape. Finally, the higher fractal resolution curves with $r = 3, 4$ diverge at the lowest scale level when approaching their final shapes of differently deep substructures. From this observation the following conclusion can be drawn: active shape focusing and evolution, when applied to true fractal structures, recover common substructures between different fractal resolution levels in a range of image scales. It is also important to note that in contrast to the hierarchical Fourier reconstruction of the von Koch curve in figure 2.7 (c) (see chapter 2, page 38), the topology is maintained. Additionally, the classic multi-scale contour representation of the same curve in figure 3.6 (c) (see chapter 3, page 71), shows almost identical shapes across scales. This is in contrast to the observation in section 7.5 that active shape evolution and focusing cause less shrinkage at higher scales, and can be explained by the global, circular shape of the fractal curves at high scales: the diffusion or heat equation eventually shrinks any shape into an ellipse or circle.

While the evolution results for fractal resolution $r = 4$ yields an adequate result at the lowest scale level, the focusing result for the same fractal resolution and image scale does not recover all fine details correctly, and seems to perform worse than the next lower fractal resolution result. This effect can be partially explained with the minimum contour stabilizer $\epsilon_2 = 3.5$ which prevents the multi-scale active contour model to be attracted from the outside into small concavities. The same applies for the adaptive resampling itself, which prevents a completely continuous focusing at small scales when removing too densely located snaxels. This does not affect the reverse evolution process, as the model is positioned on the ground truth right from the beginning. The other reason for the poor result of this particular curve of fractal resolution $r = 4$ is the so-called *partial pixel effect*, which describes image pixels at boundaries of structures that have a fuzzy object membership. In other words, the fine substructures of the ground truth fractal snowflake are of sub-pixel accuracy, which is lost by the creation of the corresponding test images, and which cannot be recovered as it is below the *inner scale*, and which leads to a merging of fine substructures.

8.2.2 Quantitative Description

Figures 8.11 and 8.12 illustrate the global planar shape descriptors of table 6.5 across scales for the fractal curves obtained via active shape evolution and focusing, respectively, and tables 8.4 and 8.5 list the respective mean and slope values with respect to scale. Again, for each fractal resolution curve, the results obtained from either process are very similar. The increasing area, perimeter, and compactness values for decreasing scales are ranked according to the fractal resolution level of the curves, with the higher fractal resolution curves having the largest values, which can be also seen from the mean values. The distance measurements (except the Hausdorff distance which is again only included for completeness) almost continuously decrease for decreasing scales when approaching their respective ground truths, and are of similar mean val-



(a) (b) (c) (d)

Figure 8.9: Active shape evolution results for the von Koch curves of increasing fractal resolution (resolutions $1 \cdots 4$ for columns (a)-(d)). The illustrated scale samples are (from top to bottom): $\sigma = 1, 2.18712, 4.27752, 8.36588, 16.3618, 32$ with $\sigma_{n-1} = 32, \sigma_0 = 1, n = 32$. See text for further details.

ues and slopes across scales. The triangulation distance shows a better monotonous behaviour of constant slope values (which are sightly lower for the focusing results), supporting the observation that the curves evolve and focus in a uniform manner. Another interesting behaviour can be observed when examining the area measurements across scales: they have a number of local

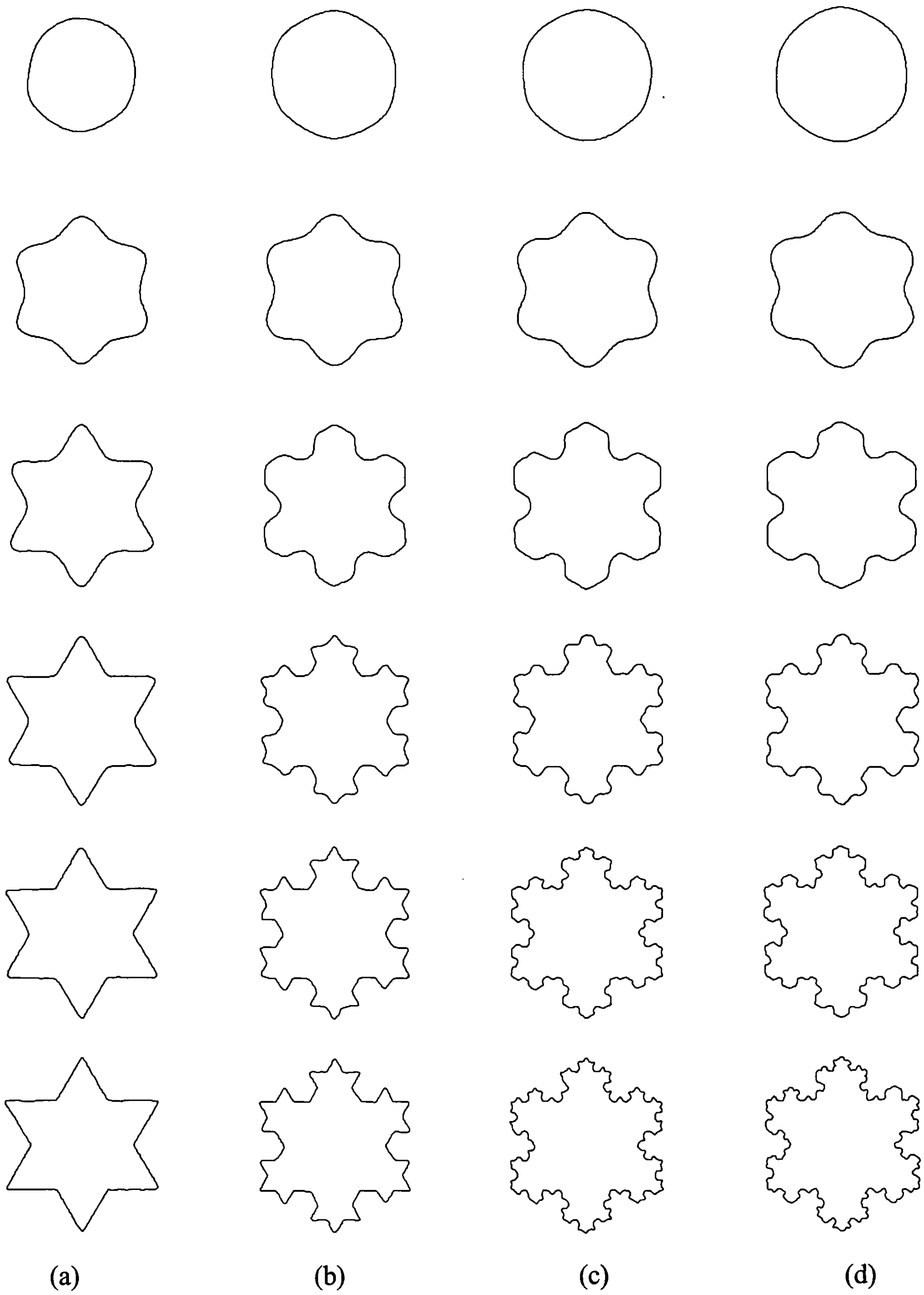


Figure 8.10: Active shape focusing results for the von Koch curves of increasing fractal resolution (resolutions $1 \cdots 4$ for columns (a)-(d)). The illustrated scale samples are (from top to bottom): $\sigma = 32, 16.3618, 8.36588, 4.27752, 2.18712, 1$ with $\sigma_{n-1} = 32, \sigma_0 = 1, n = 32$. See text for further details.

minima at intermediate and lower image scales which seem to be related to the fractal resolution level with an offset of one, except for the highest fractal resolution curves, which have like the next lower resolution curves two instead of three local minima. This can be motivated by the qualitative description in the previous section that common fractal substructures in form of more

r	Method	$\overline{A}$	$\overline{P}$	$\overline{C}$	$\overline{\text{dist}_H}$	$\overline{\text{dist}_{C_{RMS}}}$	$\overline{\text{dist}_{T_{RMS}}}$
1	E	19550	653.747	22.0995	63.9091	1.78018	24.8871
	F	19516	650.841	21.9132	105.5240	1.83869	25.6482
2	E	21803	778.294	28.8145	77.8134	2.34868	28.4989
	F	21886	777.922	28.6441	92.3041	2.31065	27.7656
3	E	22986	879.584	36.6219	84.7512	1.98855	28.5715
	F	22847	875.631	36.4191	75.3641	2.03132	29.7114
4	E	23518	900.719	38.3319	82.6819	1.70532	35.2772
	F	23524	887.325	36.5119	100.8640	1.73622	35.1066

Table 8.4: Mean descriptors for active shape evolution (E) and focusing (F) of the von Koch curves at fractal resolution levels $r = 1 \dots 4$ with respect to scale. The same remarks as in table 8.2 apply.

r	Method	$\frac{\Delta A}{\Delta \sigma}$	$\frac{\Delta P}{\Delta \sigma}$	$\frac{\Delta C}{\Delta \sigma}$	$\frac{\Delta \text{dist}_{C_{RMS}}}{\Delta \sigma}$	$\frac{\Delta \text{dist}_{T_{RMS}}}{\Delta \sigma}$
1	E	-95.1515	-10.2941	-0.568904	0.218378	3.26885
	F	-116.1380	-10.4147	-0.553382	0.220597	3.47370
2	E	-82.5631	-18.0766	-1.155720	0.261976	3.26886
	F	-70.6867	-17.6376	-1.129150	0.247998	3.47370
3	E	-63.7362	-26.2651	-1.936470	0.174442	3.26886
	F	-81.8739	-26.3637	-1.919430	0.183403	3.47370
4	E	-101.6820	-28.4650	-2.140650	0.168913	3.26885
	F	-87.8724	-26.9242	-1.948580	0.170178	3.47369

Table 8.5: Slope measurements for active shape evolution (E) and focusing (F) of the von Koch curves at fractal resolution levels $r = 1 \dots 4$ with respect to scale. The slopes are estimated using linear regression on the graphs plotted in figures 8.11 and 8.12.

detailed convexities are recovered in certain image scale ranges, which are cut off at higher scale levels by locally convex hulls. The most important observation, however, is the exponential increase in perimeter and compactness for decreasing scales, along with the ranking in terms of the size of the perimeter and compactness values over all scales with respect to the fractal resolution level of the curves. This is due to the fractal nature of the ground truth fractal curves: at each fractal resolution level, the perimeter increases by a factor $\frac{4}{3}$ (see section 2.3.3, chapter 2). Their enclosed area increases as well, though in a much smaller proportion, as new triangular structures are padded at each side of the fractal curves to yield the next higher resolution curve. Therefore the compactness of the fractal curves increases for higher fractal resolutions. This observation gives rise to a new *multi-scale fractal shape measurement*, which will be presented in section 8.2.4 after the investigation based on the visualization of the multi-scale shape stacks.

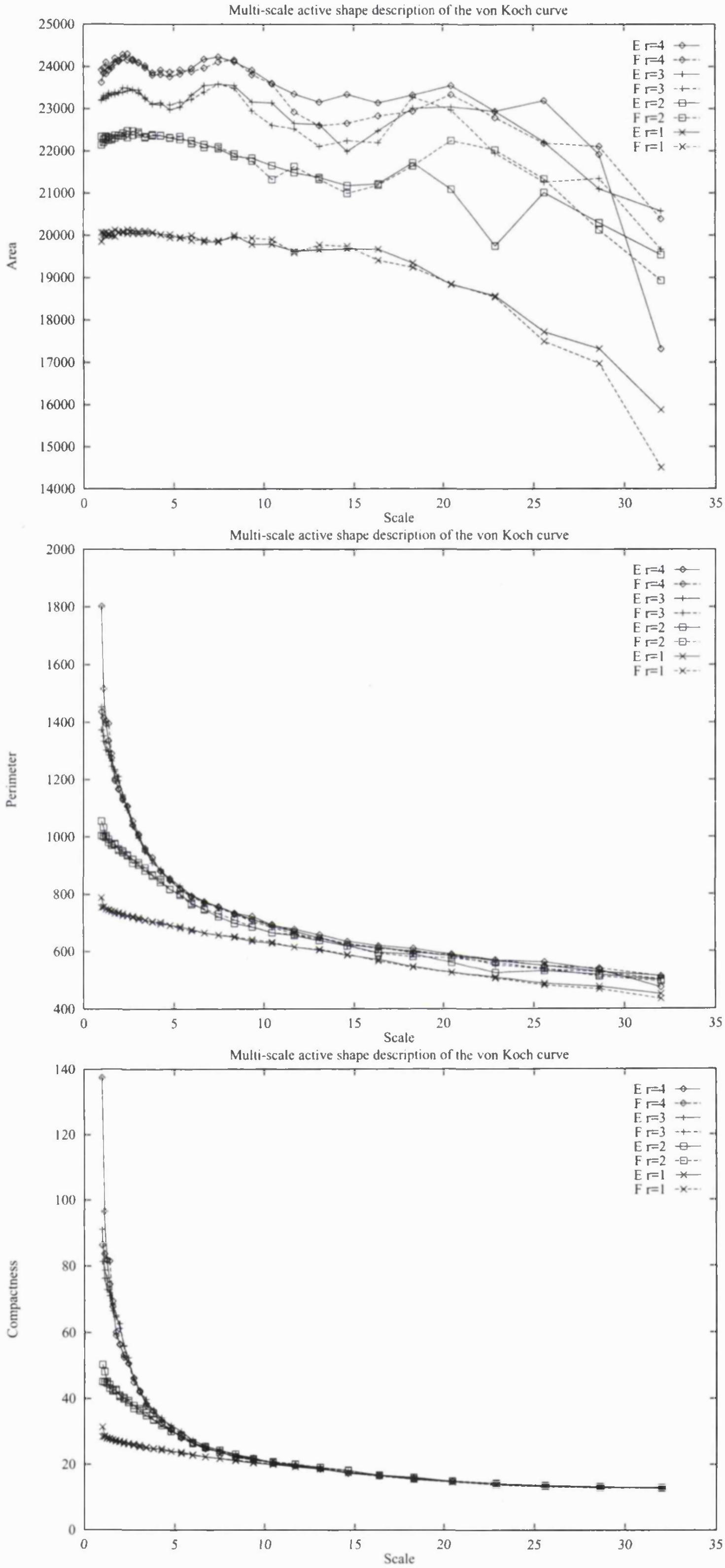


Figure 8.11: Area, perimeter and compactness of the von Koch curves across scales for active shape evolution (E) and focusing (F).

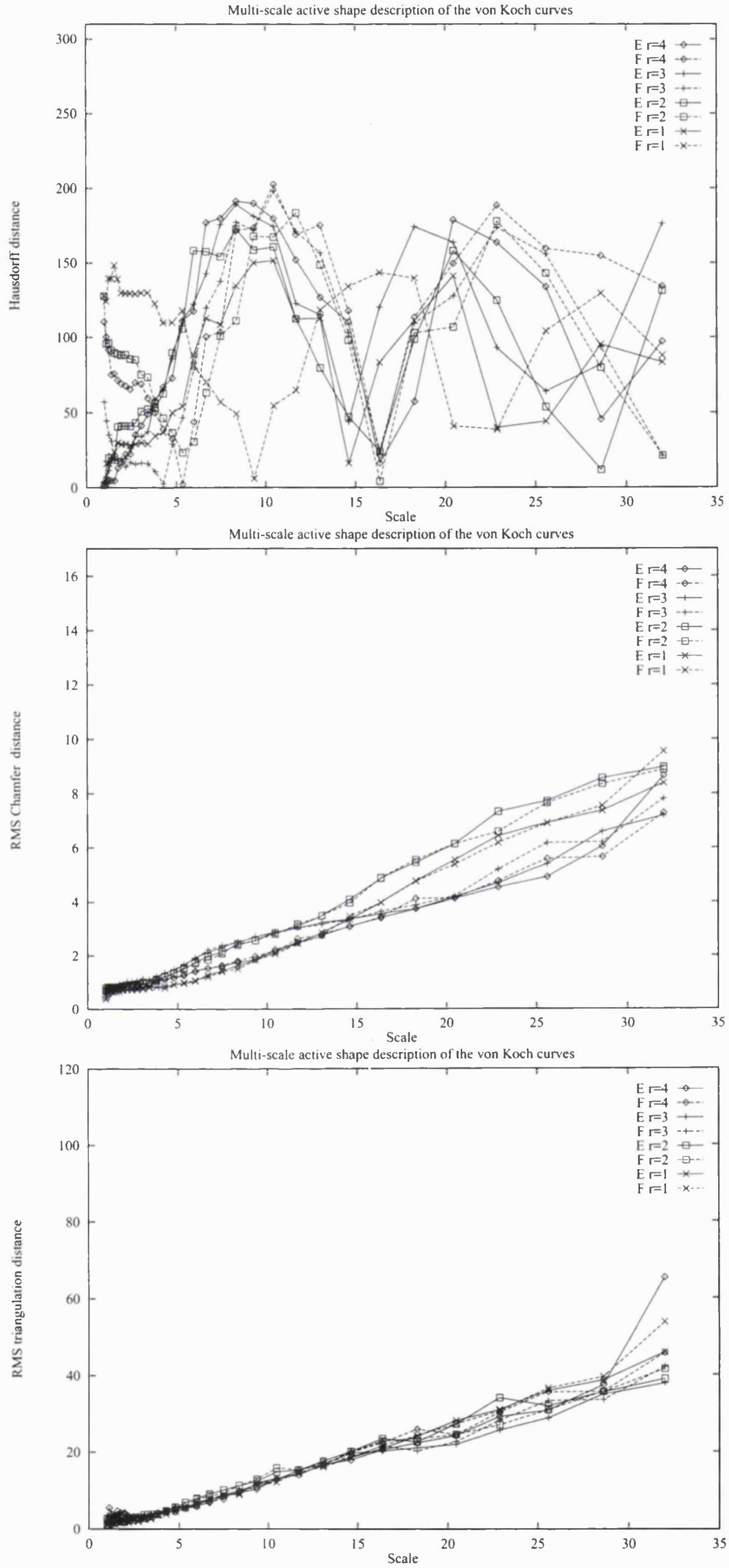
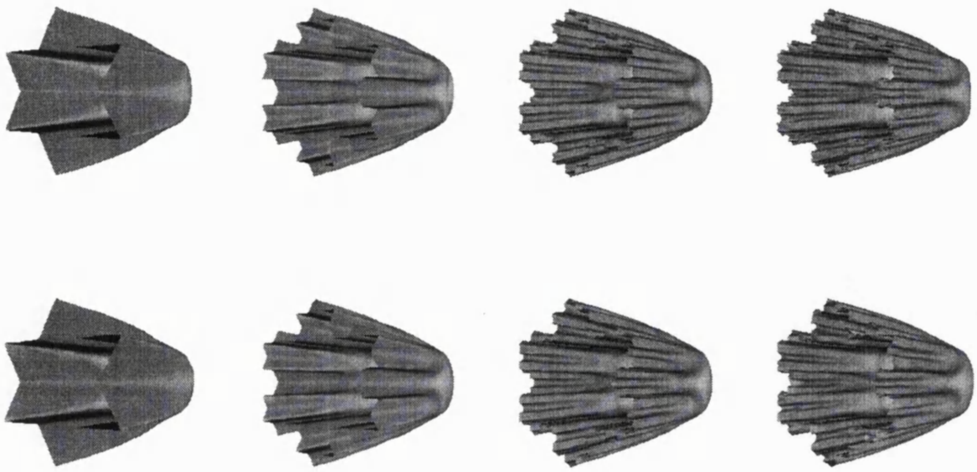


Figure 8.12: Hausdorff, RMS Chamfer, and RMS triangulation distances of the von Koch curves across scales for active shape evolution (E) and focusing (F) with respect to the known shape.



(a) (b) (c) (d)

Figure 8.13: Interpolated shaded surface renderings of the hierarchical shape stack of the von Koch curves obtained via active shape evolution (upper row) and active shape focusing (lower row). Surface rendering is performed in a *fine-to-coarse* view. See text for further details.

8.2.3 Visualization

Figure 8.13 illustrates the surface renderings of the multi-scale hierarchical shape stacks of the von Koch curves obtained via the active shape evolution and focusing techniques, respectively. Selected slices through these stacks have been shown in figures 8.9 and 8.10. Again, both techniques yield a similar visual impact. The shape flow from a circle over similar fractal states at higher and intermediate image scales can be well perceived. Additionally, one can observe the positions in scale when the fractal shapes diverge to their respective final snowflake shapes of different self-similarity depths. Colour mapping of local shape metrics on the hierarchical shape stack obtained via active shape focusing (the dual mapping of active shape evolution is not depicted here) is illustrated in figure A.6 in appendix A. Most strikingly, the curvature mappings in column (a) visualize the fine-scale snowflake structures in a very regular and prominent way, enabling to grasp the depth of the self-similarity of each fractal curve. The Chamfer distance mappings in column (b) illustrate the short cuts taken in these measurements, yielding very low distance measures (illustrated as no colour or slight yellow colour) at higher scales. The triangulation distance measurements in column (c) show more continuous behaviour, in accordance with the smooth perimeter increase and RMS triangulation distance decrease for decreasing scales observed in figures 8.11 and 8.12, with slightly higher values (in terms of hue) at the positions at intermediate and higher image scales where new substructures are formed at lower scales.

8.2.4 Multi-Scale Fractal Shape Measurement

In section 8.2.1 it was observed that similar fractal substructures for the von Koch curves of increasing fractal resolution are revealed in certain ranges of image scales. This was further quantified in section 8.2.2 by the exponential increase of perimeter and compactness for increasing scales in figure 8.11, as well as by the local minima of the area changes across scale, and the common fractal self-similarity depths were illustrated by the renderings of the multi-scale shape stacks in the previous section. As was mentioned in section 2.3.3, the fractal dimension of the von Koch curve can be computed by the increase of perimeter per fractal resolution level. The fractal relationship of area, perimeter and compactness with respect to the image scale can be further investigated in a logarithmic plot.

Figure 8.14 reveals linear relationships between the logarithmic area, perimeter, and compactness measurements with respect to scale (note that all measurements need to be scaled by the image scale at which they are obtained in order to take the scale-dependent self-similarity into account). For each measurement M, which is either the area, perimeter or compactness descriptor, the negative slope $-s_M = -\frac{\Delta \log M}{\Delta \log \sigma}$ can now be estimated by linear regression of the respective logarithmic plots:

$$\log\left(\frac{M(\sigma_i)}{\sigma_i}\right) = s_M \cdot \log(\sigma_i) \tag{8.1}$$

This slope estimation across scales has to be done carefully, as an adequate image scale selection needs to be performed. This is due to the inner and outer scales of the fractal shapes under investigation, as well as the inner and outer image and contour scales. As all fractal structures have been evolved and focused down correctly, the upper image and associated contour scale $\sigma_{n-1} = 32$ is of adequate size. In terms of the overall fractal object sizes and their spatial widths it could have even be chosen larger. The initiator of the ground truth fractal shapes is of side length $l_0 = 61.584$, and the side length of each higher resolution is computed by dividing the next lower fractal resolution length by three. Therefore each fractal shape is limited by its outer scale l_0, and its inner scale l_r, where r corresponds to the fractal resolution level or depth of the self-similarity. Additionally, the inner image scale (given by the pixel size) as well as the inner contour scale (given by the minimum distance $\varsigma_{min} = 3.5$), further restrict the range of scales. Table 8.6 lists the obtained values for all multi-scale fractal resolution curves obtained via active shape evolution and focusing, as well as the image scale ranges in which the slope estimation has been performed. From that it becomes apparent that for the higher three fractal resolution curves, both active shape evolution and focusing recover the fractal dimension of $D_{fractal} = \frac{\log(4)}{\log(3)} \approx 1.26$ by the perimeter slope measurement. The reason why the lowest fractal resolution curve recovers a too high dimensional value lies in the very limited scale range, which is effectively only in between $[20.52800; 32]$ due to the upper image and contour scale limit. The logarithmic slope estimations for area and com-

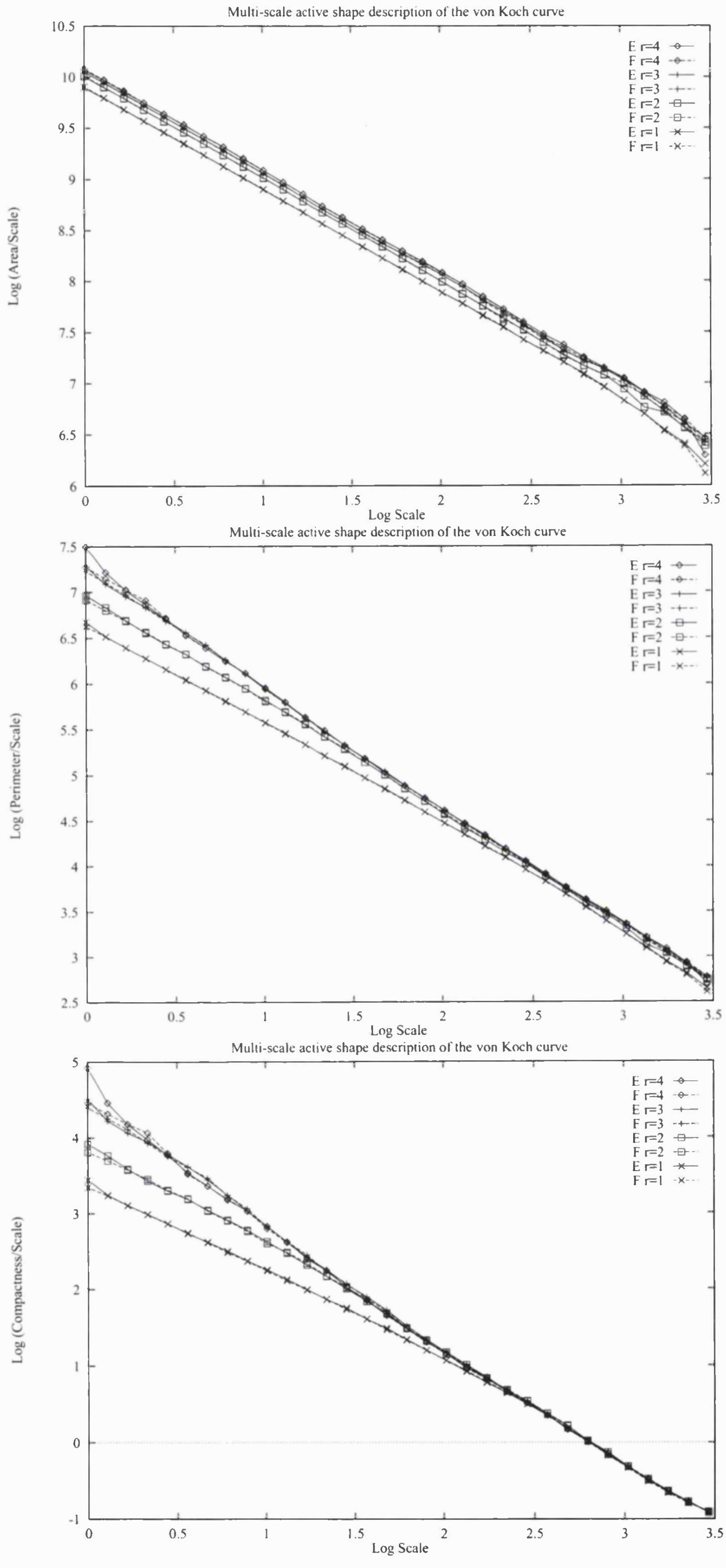


Figure 8.14: Logarithmic plot of scaled area, scaled perimeter, and scaled compactness across scales of the von Koch curves for active shape evolution (E) and focusing (E).

r	Method	Side Length	σ_{min}	$-\frac{\Delta \log(\frac{A}{\sigma})}{\Delta \log(\sigma)}$	$-\frac{\Delta \log(\frac{P}{\sigma})}{\Delta \log(\sigma)}$	$-\frac{\Delta \log(\frac{C}{\sigma})}{\Delta \log(\sigma)}$
1	E	61.58400	20.528000	1.454770	1.34966	1.24456
	F			1.716760	1.44066	1.16453
2	E	20.52800	5.132000	1.061280	1.25725	1.45323
	F			1.049860	1.25320	1.45653
3	E	6.84267	3.500000	1.037090	1.26669	1.49629
	F			1.049690	1.27096	1.49222
4	E	2.28089	3.500000	1.064870	1.26881	1.47275
	F			1.051530	1.26214	1.47277

Table 8.6: Logarithmic slope estimations for active shape evolution (E) and focusing (F) of the von Koch curves at fractal resolution levels $r = 1 \ldots 4$ with respect to scale. The negative slopes are estimated using linear regression on the graphs plotted in figures 8.14 within the scale ranges $[\sigma_{min}; \sigma_{n-1} = 32]$.

pactness with respect to scale show also a constant behaviour for the fractal resolution curves of $r > 1$, yielding negative slope values around between $[1.03; 1.06]$ for the area measurements, and between $[1.45; 1.49]$ for the compactness measurements. The reason for the underestimation of the logarithmic area slope lies in the non-continuous increase of the enclosed area of the fractal curves for decreasing scales, as finer scale concavities are recovered as convexities at higher scale, leading to local minima in correspondence to the degree or depth of self-similarity. This naturally affects the compactness slope measurement which depends on both perimeter and area measurements as well. The logarithmic slope estimation of the perimeter changes across scales obtained by active shape evolution or focusing correctly recovers the fractal dimension of the von Koch curve at different fractal resolution levels, based on a similar process to the composition or decomposition of a fractal curve outline into finer detailed substructures.

8.3 Summary

This chapter has tested the technique for multi-scale active shape description based on active shape evolution and active shape focusing on a set of synthetic test images. Different types of complex shape characteristics have been presented, as well as a synthetic set of true fractal shapes of increasing levels of fractal self-similarity. The silhouette-shaped images have been well tracked in both directions of the image scale-spaces, leading to qualitatively, quantitatively and visually similar results at the same scale levels for both active shape evolution and focusing. All levels of the resulting multi-scale shape stacks have been described with global planar shape metrics, as well as with mean and slope measurements across scales, leading to distinctively different values for the different types of shapes. The extracted shape measurements correspond to the expected behaviour of the different object shapes across scale, and demonstrate the potential of the framework for multi-scale active shape description to provide shape measures for the degree of "convolutedness" or smoothness, and for the scale location of protruded parts.

The application to strongly curved objects of varying compactness and objects with minor and major protrusions has demonstrated the applicability of multi-scale active shape description to a large variety of shapes. Simultaneously, some of the problems encountered with adequate image and contour scale selection with respect to the overall object size and tracking of substructures with small entrances (so-called *bottlenecks*) have been shown, and solutions have been suggested. The application of multi-scale active shape description to a true fractal structure suggests an intrinsic relationship between fractal resolution and image scale in terms of perimeter changes across scales. In particular, the slope estimation in a log-log plot of scaled perimeter versus image scale within an adequate range of scales has allowed to recover the fractal dimension of this particular structure at different fractal resolution levels. Although the fractal dimension estimation based on active shape evolution and focusing has not been evaluated on a larger set of fractal structures for its correctness, it provides an intuitive multi-scale shape measurement which reduces the large amount of measurements to a single meaningful number by organizing the scale information in a structural way.

While this chapter has been restricted to synthetic binary shapes of known ground truth, the following chapter tests the presented techniques on medical image data for three different medical applications. In absence of a ground truth, active shape focusing, which has been shown to be a dual technique to active shape evolution, is used for the construction of the medical shape stacks.

Chapter 9

Application to Medical Images

LE PETIT PRINCE EUT UN SOURIRE:
– TU N'EST PAS BIEN PUISSANT... TU N'AS MÊME PAS DE PATTES... TU NE PEUX MÊME PAS VOYAGER...
– JE PUIS T'EMPORTER PLUS LOIN QU'UN NAVIRE, DIT LE SERPENT.

THE LITTLE PRINCE SMILED.
"YOU ARE NOT VERY POWERFUL. YOU HAVEN'T EVEN ANY FEET. YOU CANNOT EVEN TRAVEL..."
"I CAN CARRY YOU FARTHER THAN ANY SHIP COULD TAKE YOU," SAID THE SNAKE.

Le Petit Prince, Antoine de Saint-Exupéry.

In this chapter, the presented framework for multi-scale active shape description will be tested on MRI data for three different types of clinical problems in order to demonstrate its functionality:

- Epilepsy, manifesting as local deformations and dysgenesis of the grey matter,
- Multiple Sclerosis (MS), manifesting as spinal cord atrophy, and
- neonatal data, manifesting as cortical impairment for early born children.

Control data are included for all cases to provide a comparison between normal shape variability and abnormal shape deformation. As the goal of this dissertation is to show the applicability of the framework rather than performing a clinical group study, the number of image data sets is comparatively small. Table 9.1 summarizes the information of the medical data used.

9.1 Pre-Processing, Initialization, and Optimization

In medical imaging, several degradations of image quality can occur. *Movement artefacts* are caused by movements of the patient in the scanner during image acquisition, and lead to discontinuous or unsharp object contours. This is especially the case with MRI, where acquisition times are relatively long. Partial volume effects are caused by sampling during image acquisition, yielding image *voxels* containing a mixture of different tissue types. Each type contributes with its own intensity value, and the resulting intensity is a weighted combination of the contributing

Image data information	Epileptic data	Spinal cord data	Neonatal data
Number of data sets	6	8 (4 scan/rescans)	2
Type of deformation	Cortical dysgenesis	Cord atrophy	Cortical impairment
Pixel size	0.9375×0.9375	0.4883×0.4883	0.0586×0.0586
Slice thickness	1.5	2.9297	5.0
Slice plane	Coronal	Reformatted	Transversal
Slice size	256×256	512×512	256×256
Region of interest	-	64×64	-
Linear zooming factor	-	4	-
Number of slices	124	10 (5 relevant)	10
MR type	T1	T2	T2

Table 9.1: Medical image data information. The size of the slices and regions of interest is given in pixel units, while the actual pixel size and slice thickness are given in mm^2 and mm, respectively. T1 and T2 refer to the weighting or relaxation times of the MRI acquisition.

intensity values. This effect might lead to inaccurate estimations of object size and location. *Intensity variations* within an image slice as well as between images slices can occur for various reasons; in MRI, for example, they are caused by nonuniform magnetic field strengths, and can be corrected using thresholding and contrast enhancement techniques. More sophisticated intensity corrections are based on a thin-plate spline interpolation of the intensity surface based on a given set of landmark points [Zijdenbos *et al.*, 1994]. For slice-by-slice acquisition techniques, each image plane is acquired with a certain *slice thickness*. Thinner slices increase the accuracy of measurements, but simultaneously decrease the *signal-to-noise-ratio*. Careful pre-processing is required, as otherwise the objects under investigation cannot be properly delineated from the surrounding tissues and other structures. A combination of the following items helps to achieve a proper separation, which are further discussed below:

1. Thresholding and intensity cutting
2. Contrast enhancement
3. Morphological processing
4. Scale selection
5. Edge information

As a first step, background noise and meninges (membranes enclosing brain and spinal cord) can be removed by *thresholding*, where the size of the threshold may vary between the data sets and is dependent on the image acquisition. In order to avoid strong edge responses from neighbouring

Pre-processing parameter	Epileptic data	Spinal cord data	Neonatal data
Intensity scaling range	[0; 255]		
Lower intensity threshold	$10 - 30$	5	$10 - 30$
Upper intensity threshold	255	255	-
Intensity cutting threshold	100 (upper 24 slices)	-	$110 - 115$
Image gradient threshold	10%		
Contrast enhancement	Histogram equalization		
Contrast	Light shape on dark background		
Morphological operator	Opening 5×5	-	Opening 3×3

Table 9.2: Medical image pre-processing parameters.

tissues with high water content such as soft tissue or the eyes, an upper threshold can be used for *intensity cutting* of very high intensity values by setting them to the lower threshold. This is usually only necessary for whole brain volumes in the upper slices containing the eyes, and otherwise only in a few slices. *Contrast enhancement* using histogram equalization for the redistribution of the thresholded intensities increases the contrast between the different tissue types. Small isolated structures blocking the entrance to protrusions can be removed using *morphological processing* techniques, e.g. by morphological *opening*. Morphological processing has already proven to be very useful for the extraction of the human brain, in particular for grey matter and white matter extraction in MRI [Kapur *et al.*, 1996]. For the actual focusing process, adequate *scale selection* needs to be performed with respect to the field of view. Additionally, if the object of interest is not isolated, but closely surrounded by other structures, large scales cause blurring across the edges of both the object and the other structures, possibly leading to a premature merging of the edges. Finally, the optimization of the multi-scale active contour model not only depends on the edge strength, but also on directional *edge information*. In chapter 6, directional tuning was presented which can be used to recover objects of specific contrast to the background, which is defined by the sign of the image energy terms, including the isophote image curvature, in equations 6.16, 6.22, and 6.23, respectively. Additionally, the normalization of the gradient allows to threshold small spurious edge values falling below a certain percentage of the overall range independently of the size or range of the original image values. Table 9.2 lists the image pre-processing parameters used for the image data for the different medical applications.

Different strategies are employed to obtain the initial models: For the application to epileptic data, an initial, circular shaped model is created for a single slice for each data set using manual interaction, followed by shape propagation as presented in chapter 7. The neonatal data has less spatial correspondence between neighbouring slices due to the much larger slice thickness, leading to inaccurate propagation results. Therefore each slice is individually initialized with an ellipse-

Optimization parameter	Epileptic data	Spinal cord data	Neonatal data
$\alpha_{elasticity}$ for propagation	0.001	-	-
$\alpha_{elasticity}$ for focusing	0.00001		
Iterations for propagation	10-20	-	-
Iterations for focusing	20	10	10
Initial scale	$\sigma_{initial} = 8$	$\sigma_{initial} = 16 \pm 8$	$\sigma_{initial} = 8$
Final scale	$\sigma_{final} = 1$		
Number of scale samples	16		
Greedy search space	5×5	3×3	3×3

Table 9.3: Medical image optimization parameters. The fixed default parameters can be found in table 6.5 (chapter 6, page 141).

shaped model, which is acceptable due to the limited number of available slices. The spinal cord data, though being a good candidate for shape propagation, is initialized in a slightly different manner. The cord has an ellipse-shaped outline, varying in overall size and elongation between patients and controls, as well as across slices. The detection of scale-space extrema or *blobs*, discussed in chapter 3, has been adapted to locate the maximum scale response for each image slice in order to determine the centre and size of the initial ellipse-shaped models, a technique which will be presented in the context of this application.

In absence of a ground truth, the multi-scale shape stacks for the different types of medical image data are obtained via active shape focusing, using the lowest scale result of each image slice as the reference shape. Table 9.3 lists the optimization parameters used which will be discussed in the following, and the default parameters can be found in table 6.5 (chapter 6, page 141). For the given medical data sets, lower upper scale limits than for the synthetic data in the previous chapter have been chosen in order to reflect a smaller field of view in terms of object sizes and finer detailed substructures, and to avoid interaction with neighbouring structures. Still, the chosen upper scale limits are large enough to capture the global shape outline, and in the case of the epileptic data allow for adequate propagation through the whole brain volume. The number of samples is adjusted to the upper image scale, while the lower image scale is set to one pixel unit. Therefore a denser sampling than for the synthetic data is used, but less samples are needed. The lower scale range for the focusing process implies that smaller and therefore more efficient search spaces can be used, and are adjusted to the expected complexity of the objects. In the following, general guidelines for the application of multi-scale active shape description to volumetric data will be given.

9.2 Application to Volumetric Brain Data in MRI

The cortical surface of the brain can be formulated as a set of contours, one for each image slice $L(x, y, z_k)$. This allows for the computation of two different types of multi-scale shape stacks, viz. $3D$ and $3\frac{1}{4}D$ stacks which are based on image scale-spaces of dimensionality $3\frac{1}{2}D$ and $4D$, respectively, and are consequently obtained from $2\frac{1}{2}D$ or $2\frac{3}{4}D$ active shape focusing. Note that either type of shape stack can be organized as a set of $2\frac{1}{2}D$ stacks, one for each image slice, or as a $3\frac{1}{2}D$ stack, with one volumetric instance per scale level. The latter is preferred due to the better visual inspection using volumetric techniques, but simple reformatting allows to investigate the stacks of the individual slices without the need of performing another focusing process. Both $2\frac{1}{2}D$ or $2\frac{3}{4}D$ active shape focusing have the same computational complexity as they are both based on tracking a shape through a volumetric image scale-space; however, they differ in the computational complexity of the underlying scale-space computation. The slice-by-slice scale-space computation is computationally more efficient, but does not take the correlation between neighbouring slices into account which generally leads to more *local* results than the $4D$ scale-space approach. Only linear scale-spaces are considered here, but extensions to nonlinear scale-spaces will be discussed in chapter 10. The application of multi-scale active shape description to volumetric brain images requires the consideration of several issues, like the initialization of all slices, the choice of the underlying image scale-space dimensionality, and the structural organization of the resulting multi-scale shape stacks. These issues are intrinsically related to each other, and will be discussed in the following.

In theory, each slice of the volumetric data sets can be initialized manually with a coarse estimate, but demand on expert time is very high, and user dependence should be avoided as much as possible. Using the concept of shape propagation as discussed in chapter 7, section 7.3.1, allows to reduce manual interaction to the initialization of a single slice per volumetric data set. However, shape propagation in MRI needs to be performed with consideration, as due to the slice thickness structures may suddenly appear or disappear from one slice to another. In order to avoid a trapping of the multi-scale active contour model in a local energy minimum caused by the sudden appearance of a substructure, the initialization of an initial circular shaped model is performed on an intermediate slice where the brain is most circular shaped and of maximum expansion, which is usually between slices 40 and 50. The elasticity energy term of the multi-scale active contour model enables to contract the model in lack of appropriate external image forces, allowing the model to recover when substructures like the brain stem disappear or shrink back. For this reason a larger weighting term $\alpha_{elasticity}$ is chosen for the shape propagation than for the following focusing process (see table 9.3). Shape propagation is performed at the upper chosen scale level, and is therefore also affected by the dimensionality of the underlying image scale-space. In a $3\frac{1}{2}D$ image scale-space, correspondence between neighbouring slices is lower than in a true $4D$

scale-space, where blurring is performed additionally in the slice direction. Therefore the scale-space dimensionality influences the stability of the propagation process. For a following focusing process, the scale-space dimensionality also indicates how much focus is on local properties and shape characteristics in individual slices. The $2\frac{1}{2}D$ or $3\frac{1}{2}D$ organization of the resulting multi-scale shape stacks (whose dimensionalities depend on the respective scale-space dimensionalities) is chosen to inspect individual brain slices, or the overall global shape of the volumetric brain structure. For either shape stack organization, an underlying $4D$ image scale-space takes the volumetric nature of the data and the associated anisotropic voxel size into account. At large image scales, this influence is very high, while it vanishes at lower image scales, and the $4D$ image scale-space approaches the $3\frac{1}{2}D$ image scale-space. The main difference of both image scale-spaces therefore lies in higher to intermediate scales. This leads to solutions of different locality in terms of detailed structures which are not inherent in the neighbouring slices, or which are only enhanced by the neighbouring slices, but not very prominent in the image slice under investigation. Consequently, a $4D$ image scale-space leads to a more global $3\frac{1}{4}D$ stack due to its averaged nature, where finer scale details may not be adequately tracked. A $3D$ stack which is based on a $3\frac{1}{2}D$ image scale-space captures all finer scale details in the individual slices, but lacks the knowledge of the more general shape. This might endanger the globality of the solution, and leads to the potential trapping of the multi-scale active contour model in local minima.

For multi-scale active shape description of volumetric images, volumetric rather than planar shape metrics which have been described in table 7.2, chapter 7, need to be computed. Both volume and surface area can be obtained from a triangulation of the sets of contours $\{\tilde{\mathbf{v}}_{z_k}(s;\sigma_i)\}$ at each scale level σ_i as was described in section 7.4. However, it was found that due to poor results of standard triangulation techniques for very complex volumetric shapes these measurements were of inadequate quality. Therefore, a more pragmatic approach has been taken to derive the volumetric measurements: perimeter and area measurements for each image slice and at each scale level are calculated and summed across slices considering the anisotropic voxel size and slice thickness. Volumetric compactness, indicating the deviation from sphereness, can then be computed using equation 7.2 (chapter 7, page 153). Similarly, the relative shape distance measurements with respect to the lowest scale result for each slice (since no ground truth is available) can be computed for each slice and scale level separately, and then averaged across all slices.

Before moving on to the results, it should be stressed that the aim of this dissertation is to test multi-scale active shape description as a new methodology on different types of medical data. Segmentation, though being an *implicit* byproduct of the active shape focusing process, is not the primary goal. Instead, all intermediate scale results are regarded as being of the same importance for concise and comprehensive shape investigation. In the following, the application of active shape description on epileptic data, spinal cord data, and neonatal data is presented.

9.3 Application to Patients with Epilepsy

Epilepsy is the most common major neurological disorder with a minimum prevalence of approximately 0.5%. There are two major types of epilepsy, viz. idiopathic epilepsy and symptomatic epilepsy. The first type is presumed to be caused by genetic factors and can most of the time be controlled with adequate drug treatment. The second type, however, usually develops as a result of structural abnormalities of the brain, and can be identified with high-resolution MRI in up to 60% of the cases, which allows for adequate planning of epilepsy surgery in order to remove the identified abnormal parts of the brain. In 70% of the cases above eventually undergoing surgical treatment, the patients become completely seizure free. A variation of symptomatic epilepsy is cryptogenic epilepsy, comprising all cases where a structural cause is suspected but not found despite of extensive investigation. Post mortem studies have suggested that cortical dysgenesis in form of subtle disruptions of the gyral/sulcal pattern is the most common structural cause in patients with epilepsy. A correspondence between subtle structural abnormalities like gyral thickening and deepening has been established, reflecting the strong link between cortical cytoarchitecture and epilepsy [Meencke, 1994]. This link has also been revealed based on the analysis of MRI in terms of the complexity of the white matter/cortical interface characterized by abnormal fractal dimensions [Cook *et al.*, 1995; Free *et al.*, 1997]. However, these methods currently require time-consuming grey/white matter segmentation and are based on the definition of a somewhat arbitrary set of ROI's. Studies have shown that abnormalities can be detected by careful visualisation in 3D that were not visible in 2D images, that correlated with the electronic seizure pattern [Sisodiya *et al.*, 1996]. Similarly, abnormal volumes and volume ratios can be detected in regions of the brain which appear normal on visual inspection [Sisodiya *et al.*, 1995]. Research has also recently been conducted in finding local intrinsic compactness measures based on $2D$ folded, triangulated brain surface patches and geodesic measurements [Castellano Smith *et al.*, 1997]. The multi-scale character of the presented methodology for active shape description approaches the problem of detecting structural differences in volumetric brain data from a slightly different angle. Though comparisons between the six cases will be made, the emphasis is on *intra-shape* rather than *inter-shape* comparison in terms of volume and surface changes across scale.

Figures 9.1 and 9.2 show scale samples of a $3D$ and a $3\frac{1}{4}D$ shape stack, respectively, for an intermediate slice of each data set for the application to epilepsy, showing the differences in the locality or globality of the solutions when based on a $3\frac{1}{2}D$ and $4D$ image scale-space. These slices also represent the position in the brain volumes at which the shape propagation process is initiated in order to obtain the initial estimates for each image slice. The circular shaped initial model is superimposed in black on the highest scale images for either scale-space dimensionality. Each brain data set is initialized with a circular model only slightly varying in overall size and location, as the data sets are not registered. For each brain the same associated initial model is taken for the two different scale-space approaches. The optimization result for each scale level is superimposed in white. Either shape stack consists of 16 samples of which only four are shown in figures 9.1 and 9.2, and has been computed within the scale range of $\sigma_i \in [1; 8]$ with a minimum contour scale of $\varsigma_{min} = 3.5$ pixel units.

Figures 9.3-9.8 illustrate the volumetric lowest scale levels of the obtained $3D$ and $3\frac{1}{4}D$ shape stacks, obtained by propagating the models in figures 9.1 and 9.2 through the whole brain volume, followed by performing active shape focusing in each individual image slice. The figures have been obtained by volume rendering of the extracted brains, where the extraction is performed by deriving and stacking the enclosed areas of the optimized sets of models $\{\mathbf{v}_{z_k}(s;\sigma_0)\}$. Volume rendering is performed as iso-surface rendering using the Visualization Toolkit [Schroeder *et al.*, 1997]. Iso-surface rendering, in contrast to other volume rendering techniques like maximum intensity projection or traditional *ray-tracing* (often also called *ray-casting*), renders volume surfaces of constant intensity values, and, if applied to an extracted volume, is similar to a *stop-at-first-voxel* method [Levoy, 1989]. This implies that the rendering result corresponds closely to the true extracted volume, and performs no additional visual improvement.

The following sections will summarize the qualitative and quantitative description of the stacks in terms of volumetric measurements for the detection of structural differences between patients and controls, as well as their visualization.

Figure 9.1: Samples of the active shape focusing results for the epileptic data obtained in a $3\frac{1}{2}D$ image scale-space. From top to bottom: patient 1, patient 2, patient 3, patient 4, control 1, control 2. Columns (a) $\sigma = 8$ (b) $\sigma = 4$ (c) $\sigma = 2$ (d) $\sigma = 1$. Columns (a) also contain the initial model superimposed in black.

(a) (b) (c) (d)

Figure 9.2: Samples of the active shape focusing results for the epileptic data obtained in a $4D$ image scale-space. From top to bottom: patient 1, patient 2, patient 3, patient 4, control 1, control 2. Columns (a) $\sigma = 8$ (b) $\sigma = 4$ (c) $\sigma = 2$ (d) $\sigma = 1$. Columns (a) also contain the initial model superimposed in black.

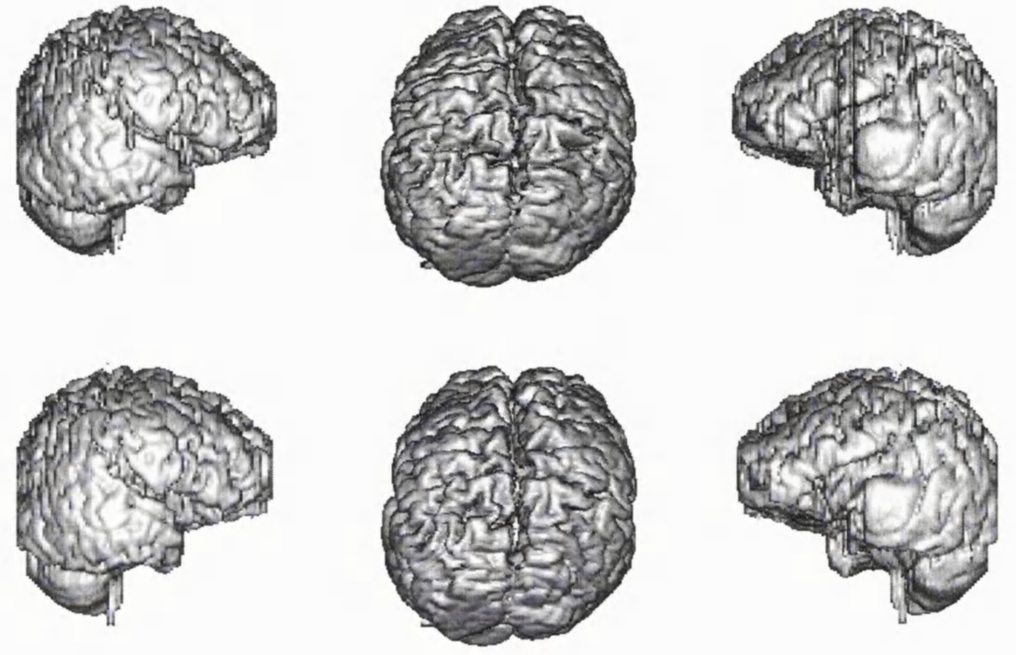


Figure 9.3: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of patient 1 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

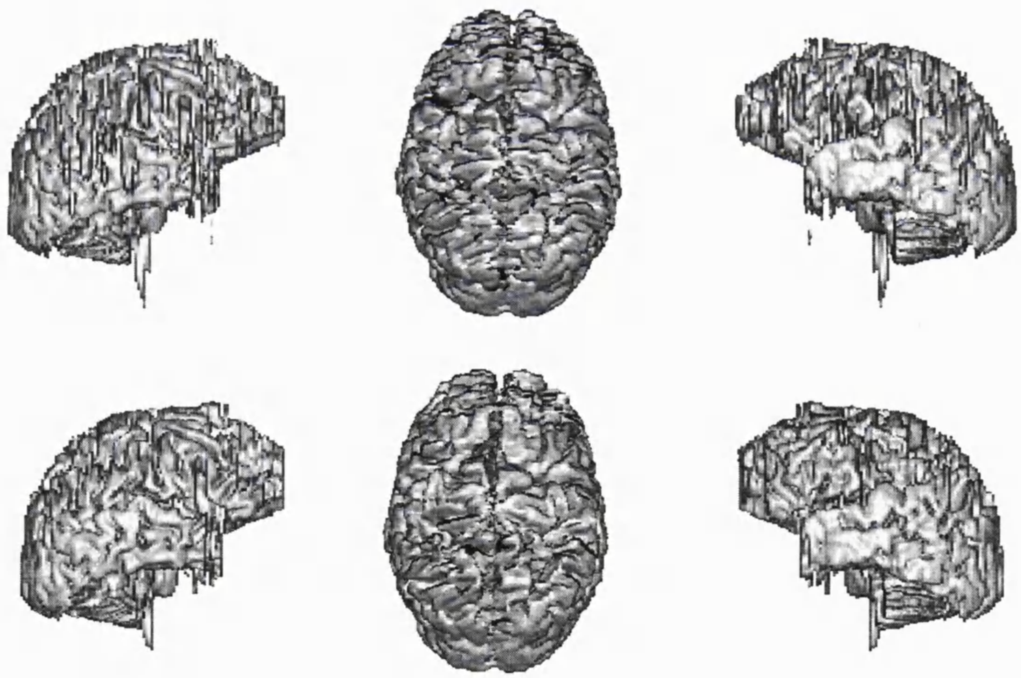


Figure 9.4: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of patient 2 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

Figure 9.5: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of patient 3 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

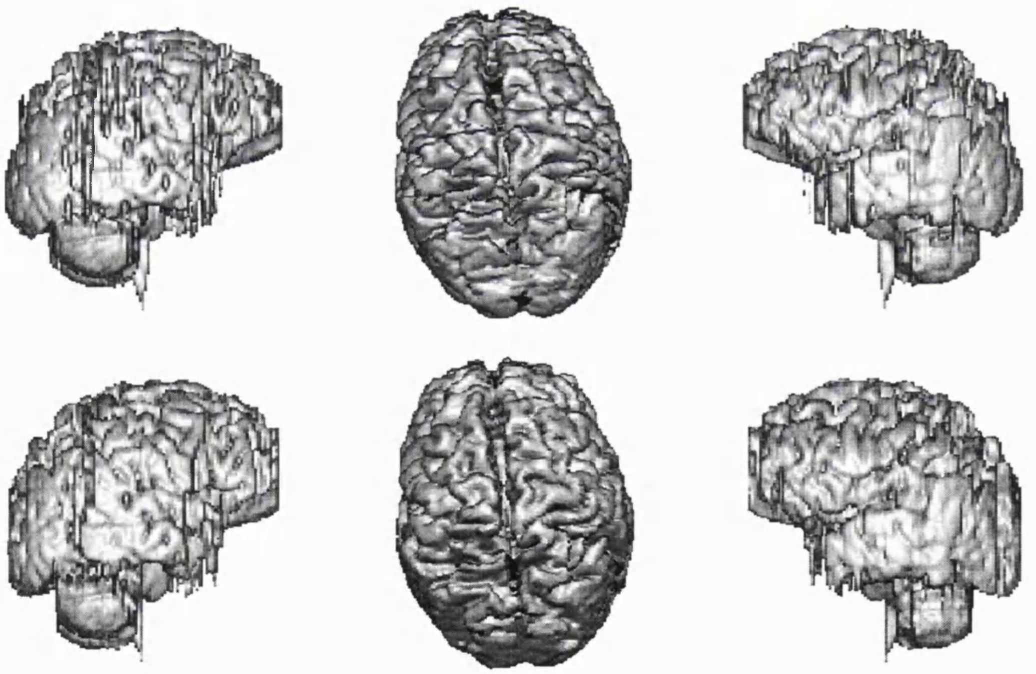


Figure 9.6: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of patient 4 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

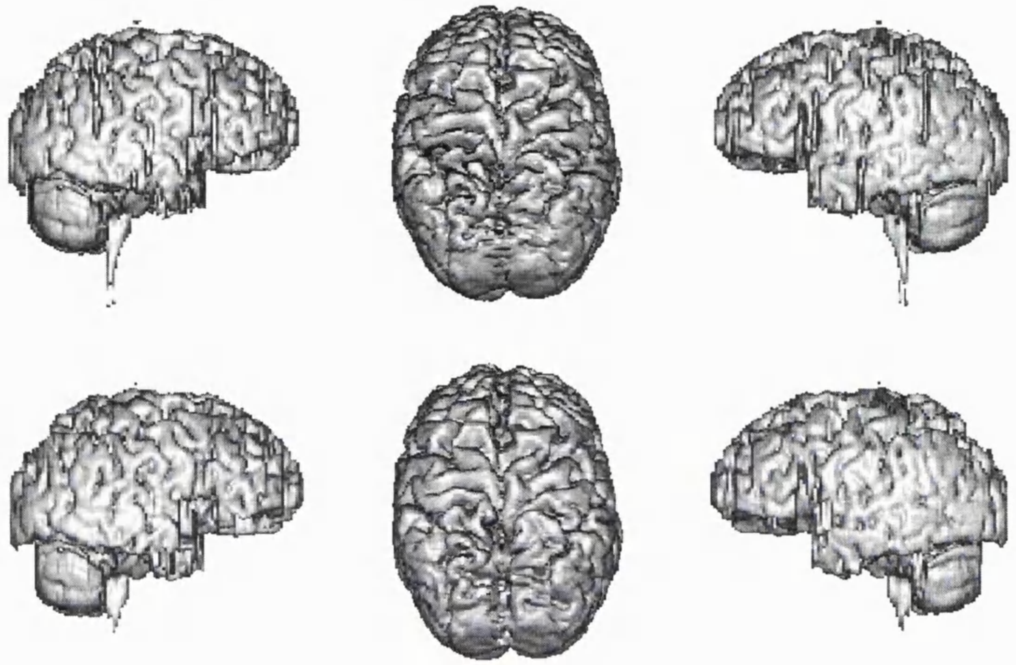


Figure 9.7: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of control 1 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

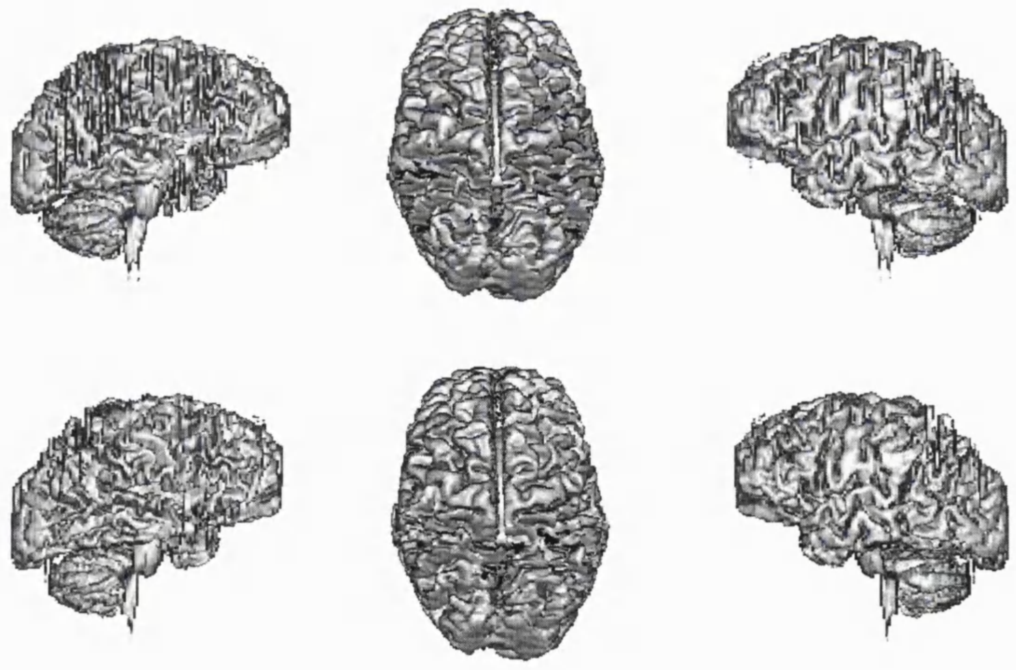


Figure 9.8: Right, top, and left views (from left to right) of the iso-surface rendered lowest scale results of the brain image of control 2 obtained via active shape focusing in a $3\frac{1}{2}D$ (top row) and $4D$ (bottom row) image scale-space.

9.3.1 Qualitative Description

From a first qualitative visual inspection, two observations can be made: first, despite the considerable differences due to structural abnormalities (which are more obvious for patients 1 and 2, and only subtle for patients 3 and 4) as well as natural shape variability between the individual cases, all slices have been well focused down using only a very coarse initial model. This was achieved by incorporating scale-space continuity into the active contour model in order to enlargen the capture region of prominent shape features, making the model apparently robust and less dependent on the initialization. Also, the curvature matching process enables the extraction of highly convoluted shapes and shapes of high local curvature like sulci and gyri of the cortex. The second observation lies in the different solutions with respect to the underlying image scale-space. At the highest scale level, the optimization results based on the $4D$ image scale-space appear to be much smoother. At intermediate to lower scales, they are more capable of capturing the overall global shape outline in terms of the separation of the brain from the skull, rather than following into small concavities and protrusions. The $3\frac{1}{2}D$ image scale-space based optimization results are more capable of tracking structures of fine detail, unless prevented by *bottlenecks*, structures below the minimum contour scale, or structures which are only partially contained in the image slice. Note that the latter is caused by the *partial volume effect*, and by topological changes, since the brain can not always be represented as a stack of planar shapes.

Using a higher order image scale-space can help to overcome this problem, which becomes more obvious when comparing the results of patient 4 and control 1 in either scale-space. Small details which are missed out in the $3\frac{1}{2}D$ image scale-space in figure 9.1 can be tracked in the $4D$ scale-space as shown in figure 9.2 due to the blurring in the z-direction. The optimal solution to solve large topological problems, for example the separation of the brain into the frontal lobes at higher image slices, can only be obtained by using a surface- rather than a contour-based model. Advantages and disadvantages of such an approach will be discussed in chapter 10. Another prominent difference arising from the image scale-spaces of different dimensionalities can be seen for patient 3, where the result does not correspond to the desired solution (see figure 9.1). Not the overall brain outline is extracted, but the model is partially trapped at one side of the brain despite using an adequate initial model. This is due to small fragmentations and fine details which do not provide a clear brain outline at the highest scale level, as well as intra-slice image intensity and contrast variations. In figure 9.2, however, the overall brain outline is enhanced across slices and can be well tracked, but on the cost of local detail in terms of small protrusions whose locations vary from slice to slice.

It must be stressed that the illustrated slices have not been especially selected for their good but rather their average quality, showing all occurring positive and negative aspects of active shape focusing in volumetric brain images. The proposed technique for active shape focusing has empirically been found to be very robust in extracting a large variety of complex brain contours in a slice-by-slice fashion. Additionally, shape propagation at the highest scale level has proven to be a very stable technique to obtain initial models in both scale-space dimensionalities, and is therefore an adequate technique to automate the initialization and following focusing process.

In contrast to the previous single-slice scale samples structured as $2\frac{1}{2}D$ stacks, figures 9.3-9.8 visualize the lowest scale level of the $3D$ and a $3\frac{1}{4}D$ shape stacks organized as $3\frac{1}{2}D$ (volumetric) stacks in right, top, and left view, and simultaneously show the power as well as some encountered limitations of the proposed active shape focusing technique for volumetric images. The different effects of the underlying image scale-space dimensionality can also be seen very clearly. The illustrated results, though not perfect segmentations of the brain for reasons discussed below, demonstrate an adequate tracking behaviour of the brain volumes through image scale-space in terms of their main shape outline. The top views of the illustrated volumetric brain shapes show very detailed tracking results for either scale-space, with the exception of patient 3, where the sulci are tracked in finer detail in the $3\frac{1}{2}D$ scale-space, but tend to be smoothed out in the $4D$ scale-space. In all cases except for patient 1, tracking problems towards the back of the brain (the lower part of the top views) occur due to changes in object size, contrast and intensity variations across and within slices. Experimentation has shown that manually adjusting the intensity distribution and scale range in these individual slices can help to improve the shape extraction results and yield locally better results, but cannot be generalized to all slices. More sophisticated intensity correction methods are expected to improve the results globally, but are not a topic of this dissertation. The automatic adjustment of the individual scale ranges is a non-trivial task which needs further research in terms of adequate scale selection of very complex and convoluted volumetric structures. For practical reasons, in this dissertation all image, scale, and optimization parameters are kept fixed and are chosen as general as possible, yielding as a result the best compromise between all parameters, instead of individually fine-tuning each image slice. The only parameter that varies between the individual brain data sets is the level of the lower threshold due to the overall different intensity ranges caused by different image acquisition parameters as well as artefacts and patient or control dependent characteristics, but is kept constant within each data set. The side views in figures 9.3-9.8 illustrate the good tracking of the brain stem for which the classic active contour model usually fails due to its elasticity and curvature minimizing nature. The curvature matching process developed in chapter 6, along with the adaptive sampling strategy enables the multi-scale active contour model to extract the brain stem adequately, as the model's curvature is adjusted to the underlying high curvature at the peak of the stem, and the sampling strategy allows for local flexibility. From the side views two more observations can be made: first, the slice-by-slice nature of the actual shape extraction in terms of active shape focusing becomes very apparent when using a $3\frac{1}{2}D$ image scale-space, as slicing becomes visible at brain parts of high convolution, and where neighbouring slices differ very much. Brains with smooth surfaces, however, are hardly affected (e.g. patient 1). Using a $4D$ image scale-space greatly reduces the slicing effect due to the incorporation of image information of adjacent planes.

This section has qualitatively investigated two different types of multi-scale shape stacks obtained from volumetric brain images, viz. stacks of dimensionality $3D$ and $3\frac{1}{4}D$ which are based on linear image scale-spaces of dimensionality $3\frac{1}{2}D$ and $4D$, respectively. Active shape focusing, based on either image scale-space, is a slice-by-slice, $2\frac{1}{2}D$ or $2\frac{3}{4}D$ multi-scale technique for *implicit* volumetric segmentation or shape regularization, which formulates the brain volume as a set of planar shapes. The dimensionality of the underlying image scale-space corresponds to the desired locality or globality of the solution within the individual slices and across the volume, which may to be adjusted from case to case. A good solution for this ambivalence may be a top-down approach which first acquires a more general, $3\frac{1}{4}D$ multi-scale shape stack organized as a $3\frac{1}{2}D$ structure in order to investigate more global volumetric properties, followed by reformatting this stack into a set of $2\frac{1}{2}D$ stacks in order to further locate specific shape properties. Finally, individual slices of higher interest can be processed on the basis of an efficient $3D$ scale-space (or a $3\frac{1}{2}D$ scale-space for the whole volume), if necessary with optimally adjusted parameters. In the following, the obtained $3D$ and $3\frac{1}{4}D$ multi-scale shape stacks for all six data sets will be quantitatively described on the basis of a $3\frac{1}{2}D$ stack organization.

9.3.2 Quantitative Description

As for the description of synthetic shapes in the previous chapter, mean measurements, slope measurements, and logarithmic slope measurements are computed from the $3D$ and $3\frac{1}{4}D$ multi-scale shape stacks of all volumetric brain data sets. Additionally, global and relative shape measurements are plotted across scale. In contrast to the description in synthetic images, however, the shape stacks are organized as $3\frac{1}{2}D$ structures, or as volumetric rather than planar instances in scale-space, requiring the computation of volumetric rather than planar shape quantifiers as described earlier in section 9.2.

Figures 9.9 and 9.10 illustrate the global and relative shape measurements of all data sets across scale. In addition to the results obtained from using $4D$ image scale-spaces, the corresponding results from using $3\frac{1}{2}D$ scale-spaces are shown in dashed lines. The first impression from these plots is that all shape changes are very smooth and uniform with respect to scale. Volume, surface area and compactness (see figure 9.9) increase continuously for decreasing scales, and differ for the different scale-space dimensionalities in terms of being of lower value for the $4D$ scale-space results. The distance measurements (see figure 9.10) decrease continuously when approaching the reference volumes, and are of lower values for the $3\frac{1}{2}D$ scale-space.

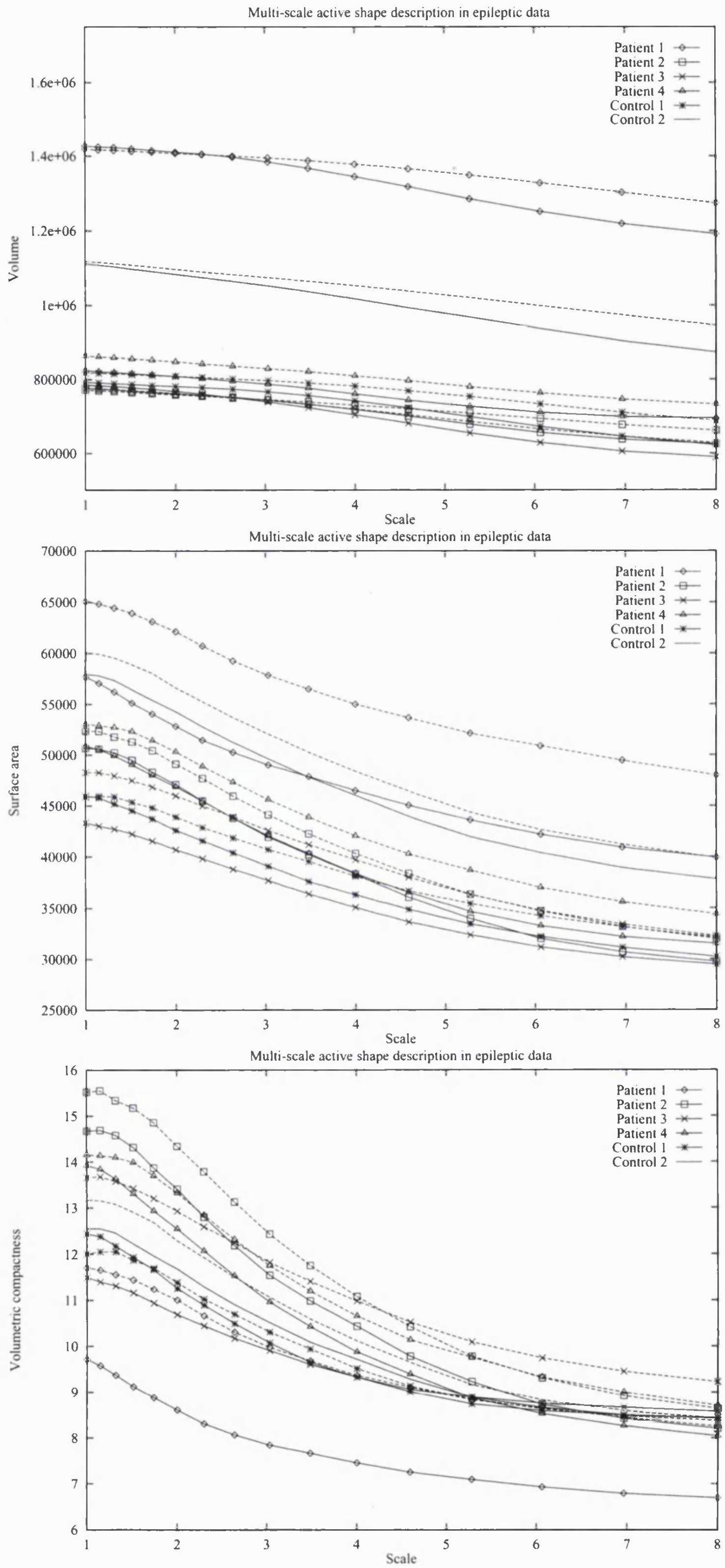


Figure 9.9: Volumetric global measurements for the application to epileptic data across scales. Dashed lines are the corresponding $3\frac{1}{2}D$ scale-space results.

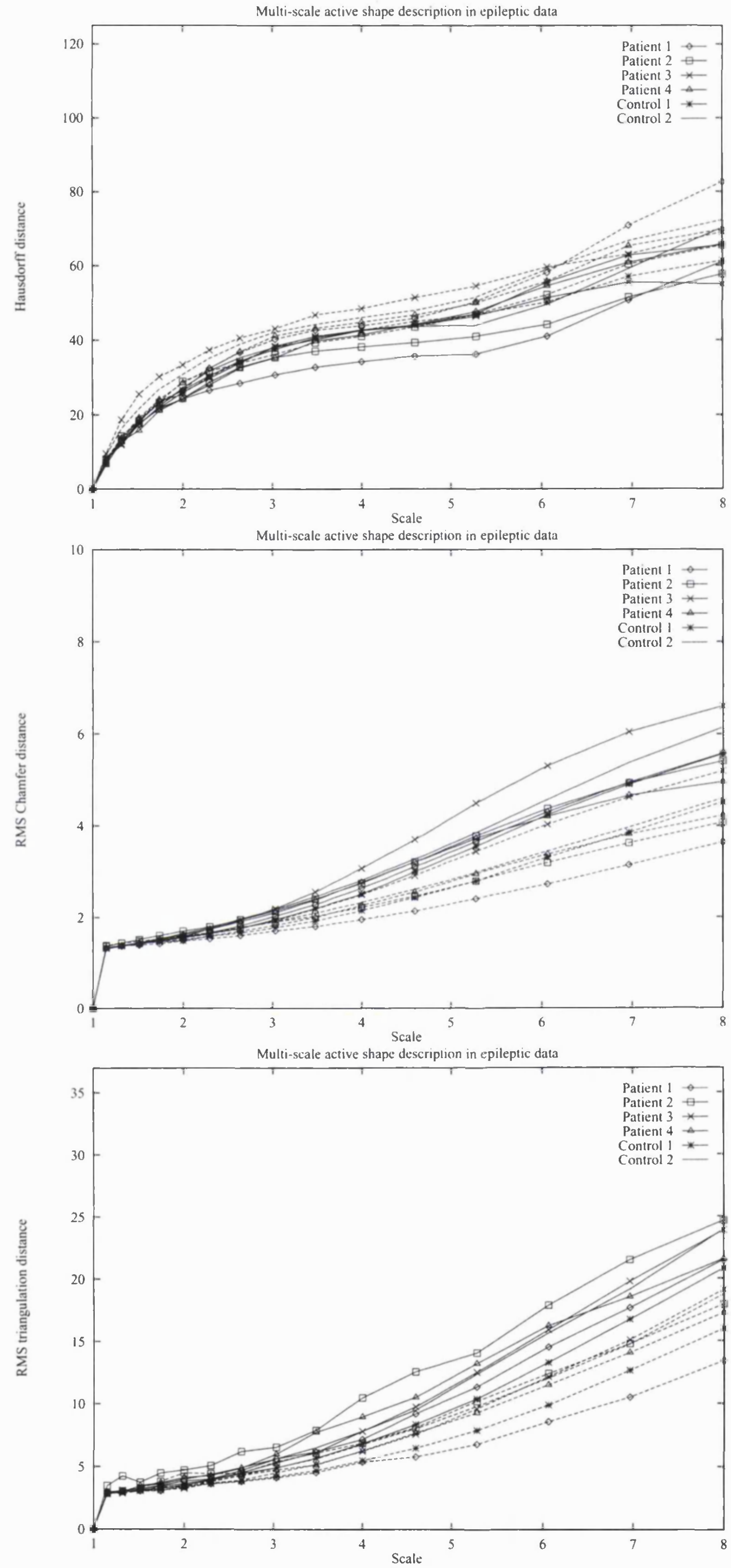


Figure 9.10: Volumetric relative distance measurements for the application to epileptic data across scales. Dashed lines are the corresponding $3\frac{1}{2}D$ scale-space results.

From both figures it becomes obvious that the shape measurements for the brain of patient 1 are distinctively different from all others in terms of the absolute values; it has the overall largest volume and perimeter, while at the same time it is of lowest compactness. Ranking next in overall volume and surface area size is control 2, which however has a much higher compactness due to a smaller surface to volume ratio. All other brains are of similar volume and much smaller, but between them quite similar surface area. It is important to note, however, that the absolute volume and surface area measurements can only be assigned much weight to if the brain data sets were appropriately registered. Since this is not the case in this study, the ratio between surface area and volume, given by the volumetric compactness as a dimensionless quantity, appears to be a more adequate shape descriptor, in addition to the volumetric slope measurements. For example, the surface area and hence the compactness of the brain of patient 1 only moderately increases for decreasing scales in comparison to the other sets. This corresponds to the intuition that this brain is very smooth and not very convoluted, which can also be observed from its overall rather low deviation from its reference volume, as well as from its lowest scale volumetric visualization in figure 9.3. The surface area measurements show a steeper increase for patients 2 and 4 than for all other cases, yielding therefore the highest and steepest compactness measurements for decreasing scales. Patient 3 has a very similar slope behaviour as patient 1, and is of next higher volumetric compactness, although much smaller in overall size (its overall low complexity becomes also obvious from figure 9.5). Finally, control 1 has a slightly steeper increase in surface area than patient 3, leading to a similar compactness as control 2. Global volumetric measurements across scale were found to provide a potentially discriminating measure for the investigated cases, by categorizing them into shapes of very low (patient 1 and 3), intermediate (controls), and high complexity (patient 2 and 4). The distance measurements show less discrepancy between the different cases, and are probably not very reliable quantifiers due to their highly averaged nature.

Differences between the $4D$ and corresponding $3\frac{1}{2}D$ volumetric and distance measurements are only subtle. For the volumetric quantifiers, the $3\frac{1}{2}D$ results are overall slightly higher than the corresponding $4D$ results due to the shrinking caused by blurring in the slice direction for the $4D$ image scale-space computation. The volume quantification differs at higher scales much more than at lower scales where it approaches similar values for both scale-space dimensionalities. The surface areas are characterized by an almost constant offset across scales, as are the volumetric compactness measurements with respect to the scale-space dimensionality. This observation indicates that the two scale-spaces allow to focus to local detail at a different speed, which may be quantified by the slope behaviour of the shape measurements across scale. The volume, which is a cubic quantifier, therefore diverges more at higher scales for the different scale-space dimensionalities than the surface area, which is only a squared quantifier. Similarly, the divergence of the distance metrics is higher at larger scales.

Descriptor	D_S	Patient 1	Patient 2	Patient 3	Patient 4	Control 1	Control 2
$\overline{V}$	$3\frac{1}{2}D$	1378990	736506	729106	818556	782444	1061740
	$4D$	1355440	726831	719606	775231	744394	1031910
$\overline{S}$	$3\frac{1}{2}D$	57935.8	43892.2	42001.9	45407.2	40409.6	51709.1
	$4D$	49363.7	41791.8	37398.4	42093.5	39046.3	49364.4
$\overline{C_V}$	$3\frac{1}{2}D$	10.1153	12.5005	11.7831	11.8161	10.3744	11.0700
	$4D$	8.0825	11.7393	10.0231	11.1332	10.3462	10.6047
$\overline{\text{dist}_H}$	$3\frac{1}{2}D$	37.0958	33.9456	39.4931	36.2594	33.4636	38.0819
	$4D$	28.8987	30.6944	32.7501	33.1949	34.3458	33.9084
$\overline{\text{dist}_{C_{RMS}}}$	$3\frac{1}{2}D$	1.85335	2.07291	2.33769	2.08636	2.06739	2.16179
	$4D$	2.44865	2.52236	2.80329	2.43206	2.40502	2.60738
$\overline{\text{dist}_{T_{RMS}}}$	$3\frac{1}{2}D$	5.11664	6.75600	6.38444	6.21958	5.62902	6.52806
	$4D$	7.40156	9.21337	7.85925	8.05681	6.93419	7.92694
$\frac{\Delta V}{\Delta\sigma}$	$3\frac{1}{2}D$	-33.5561	-26.3761	-37.3994	-32.6485	-30.6479	-40.8492
	$4D$	-60.2180	-39.1730	-51.3569	-34.9656	-41.9904	-58.2717
$\frac{\Delta S}{\Delta\sigma}$	$3\frac{1}{2}D$	-4.41602	-5.48491	-4.25802	-5.02314	-3.70060	-5.37020
	$4D$	-4.40825	-5.65955	-3.66259	-5.18233	-4.11427	-5.34478
$\frac{\Delta C_V}{\Delta\sigma}$	$3\frac{1}{2}D$	-0.00092	-0.00190	-0.00121	-0.00150	-0.00103	-0.00131
	$4D$	-0.00074	-0.00177	-0.00077	-0.00156	-0.00107	-0.00114
$\frac{\Delta\overline{\text{dist}_{C_{RMS}}}}{\Delta\sigma}$	$3\frac{1}{2}D$	0.00061	0.00072	0.00103	0.00080	0.00082	0.00085
	$4D$	0.00113	0.00110	0.00144	0.00104	0.00112	0.00125
$\frac{\Delta\overline{\text{dist}_{T_{RMS}}}}{\Delta\sigma}$	$3\frac{1}{2}D$	0.00245	0.00363	0.00384	0.00348	0.00305	0.00376
	$4D$	0.00450	0.00537	0.00511	0.00479	0.00426	0.00498
$-\frac{\Delta\log(V/\sigma)}{\Delta\log(\sigma)}$	$3\frac{1}{2}D$	1.25317	1.48122	1.47336	1.43696	1.37877	1.40222
	$4D$	1.46507	1.61402	1.42091	1.50348	1.35827	1.39900
$-\frac{\Delta\log(S/\sigma)}{\Delta\log(\sigma)}$	$3\frac{1}{2}D$	1.30477	1.55210	1.56080	1.52128	1.42031	1.44744
	$4D$	1.50634	1.68556	1.45544	1.54548	1.38357	1.44990
$-\frac{\Delta\log(C_V/\sigma)}{\Delta\log(\sigma)}$	$3\frac{1}{2}D$	1.18626	1.31582	1.20854	1.26126	1.20429	1.24319
	$4D$	1.19127	1.31561	1.16010	1.29342	1.21732	1.22235

Table 9.4: Mean, slope, and logarithmic slope descriptors for active shape focusing of the epileptic data with respect to scale. D_S indicates the underlying scale-space dimensionality.

Table 9.4 lists the quantitative mean and slope measurements for all brain data sets taken across scales, computed for the $3\frac{1}{2}D$ and $4D$ image scale-spaces. The volumetric mean measurements quantify the observed divergence between the different scale-space dimensionalities by yielding smaller values for the $4D$ image scale-space results. In particular, the volumetric compactness yields overall lowest results for patient 1, and overall highest values for patient 2. The other cases rank as observed from the plots above, with a single exception of patient 3: it was already observed in figure 9.5 that the different scale-space dimensionalities yield distinctively different results for this particular case, yielding a much higher volumetric compactness in $3\frac{1}{2}D$ than in $4D$, where it is only slightly higher than for patient 1. The mean distance measurements do not show a clear demarcation between the cases.

The volumetric slope measurements yield negative values for all cases, indicating an increase of the values for decreasing scales. For the volume slopes, significant differences arise from the different scale-space dimensionalities, yielding a much steeper increase for the $4D$ scale-space measurements. As mentioned before, this is due to the different speed in the respective focusing processes; in a true $4D$ scale-space, shapes are much smoother and smaller due to the smoothing and averaging in the slice direction, while at lower scales, the scale-spaces provide more similar results. The ranking in the increase of volume for decreasing scales is the following: in the $3\frac{1}{2}D$ case, the volume of the brain of control 2 has the steepest increase, closely followed by patient 3. Patients 2 and 4 and have the lowest increase, and the remaining two cases have very similar values. In the $4D$ image scale-space, however, patient 1 yields the highest slope, whereas the remaining cases have a similar behaviour as in the $3\frac{1}{2}D$ scale-space. Closer examination of the actual focusing results of patient 1 reveals that deep sulci protrusions are tracked at a high level of scale in the $3\frac{1}{2}D$ scale-space, whereas they are only subsequently located at intermediate to low scales in the $4D$ scale-space, leading to more prominent changes in volume. The increase of surface area for decreasing scales shows steepest behaviour for patient 2, closely followed by control 2 and patient 4. This naturally implies steepest increase of volumetric compactness for patients 2 and 4, while the two controls have an intermediate, and patients 1 and 3 lowest increase. Finally, the relative distance slope measurements for the investigated cases show no significant differences.

Figure 9.9 also illustrates an exponential behaviour of the volumetric global measurements. Recall from the previous chapter that the fractal dimension of a true fractal structure, the von Koch curve, has been recovered by analysing the logarithmic behaviour of the perimeter measurement across scales, where this measurement was obtained via active shape evolution or focusing. For structures which are not true fractals, but which may nonetheless show statistical self-similarity, a statistical estimate of such a *fractality* of the shape can be obtained in a similar fashion. For volumetric shapes, however, such an estimate can be computed either in a slice by slice fashion, yielding fractal measurements for each individual slice, or on the whole volume. In the latter case, the perimeter is replaced by the surface area (which is approximated by a summation over the perimeter across slices), yielding an *average* fractal measure. Similarly, planar area measurements can be replaced by the volume measurements, and dimensionless planar compactness by volumetric compactness.

The last three rows of table 9.4 list the obtained volumetric fractal measurements, describing the negative slopes of the logarithmic plots of volume, surface area, and volumetric compactness across scale. In contrast to the restricted scale range in the estimation for the von Koch curves in chapter 8, the entire scale range was used for the linear regressions as no prior knowledge about inner and outer scale of the brain volumes is available. Additionally, it was argued in the previous chapter that below the inner or minimum contour scale no further shape information is available for the von Koch curves. However, for grey-scale images of much finer detail this does not entirely hold; it can be observed from the plots of the volumetric measurements in figure 9.9 that there is still potential of getting attracted to details below the contour scale due to the approximating nature of the underlying continuous B-spline representation of the multi-scale active contour model, which even allows to model finer details which are located between the discrete *snaxels*. The estimated statistical fractal volumetric measurements for the different brain data sets in the application to epilepsy are mainly characterized by comparatively low values for the controls for the fractal volume and surface area measurements. In particular, the fractal surface measurement obtained from linear regression over the logarithmic surface area across scale lies for the controls roughly between $[1.38; 1.45]$, while from the results of patients 2 to 4 values between $[1.46; 1.69]$ are recovered. A similar, but slightly clearer demarcation can be found by the fractal volume measurements: the values for the controls range between $[1.36; 1.40]$, while for patients 2, 3, and 4 they lie in between $[1.42; 1.61]$. For both measurements, highest values are obtained from the results of patient 2, closely followed by patient 4, while for the controls, the values of control 2 are slightly higher than for control 1. Patient 1 only fits into the values for the other patients for the $4D$ scale-space results, but has the overall lowest absolute values in the $3\frac{1}{2}D$ results. This different behaviour may be explained by the observation made above that in the $3\frac{1}{2}D$ scale-space for patient 1, deep protrusions are tracked, which actually lead to an overestimation in surface area, and an underestimation in volume. Finally, the fractal volumetric compactness values across scales yield lowest values for patients 1 and 3 (between $[1.16; 1.20]$), intermediate values for the controls (between $[1.20; 1.24]$), and overall highest values for patients 2 and 4 (between $[1.26; 1.32]$), yielding a similar categorization as given above for the description based on global volumetric shape metrics across scale. It needs to be emphasized that the obtained fractal measurements are not volumetric (which would yield dimensionalities between 2 and 3), but planar, as they are based on average planar measurements across slices.

This section has quantitatively investigated volumetric global and relative distance metrics from $3D$ and $3\frac{1}{4}D$ multi-scale shape stacks obtained in $3\frac{1}{2}D$ and $4D$ image scale-spaces for the application to epilepsy. For the investigated cases, global volumetric measurements in terms of dimensionless volumetric compactness and slope measurements across scale have shown a potential for providing demarcating measures into three different categories, namely brains of overall smooth, average, and very convoluted shape, corresponding to the intuitive qualitative description in section 9.3.1. The included control data were found to be of intermediate category in terms of their complexity. The multi-scale fractal volume and surface measurements showed a potential for a different categorization, namely into data of low fractality (for the control data), and high fractality (for the patient data). While high fractality and compactness have been observed for data of high fragmentation and complexity, low fractality and intermediate compactness coincided more with data of average complexity and shape variability in the investigated cases. Data of low compactness has still been classified as being of high fractality, giving rise to the observation that the multi-scale fractal and volumetric compactness measurements are intrinsically, but not directly related, as they express different shape properties. Volumetric compactness measures the degree of *sphereness*, while the fractal surface and volume measurements indicate the degree of self-similarity. For structural abnormalities inherent in smooth shapes (as is the case for gyral thickening), and very fragmented shapes (as is the case for sudden disruptions of the normal gyral pattern), standard volumetric techniques are probably not capable of categorizing cases properly. The multi-scale fractal measurements investigate the brain surface and volume with respect to self-similarity under rescaling and therefore offer additional statistical shape information which may have a potential of distinguishing between normals and abnormals. It is of course not possible to generalize this empirical observation on the basis of the quantitative description of only a few data sets. A larger group study would be needed to further test this observation for its clinical correctness. In particular, the use of a brain atlas containing volumetric intensity and geometric information as presented in [Collins *et al.*, 1995] could be used to define and compare mappings of individual brains of high shape variability to such an atlas.

9.3.3 Visualization

In addition to the planar scale samples illustrated in figures 9.1 and 9.2, and the lowest scale volumetric iso-surface renderings shown in figures 9.3-9.8, qualitative local shape metrics can be mapped onto surface renderings of the obtained $3D$ and $3\frac{1}{4}D$ multi-scale shape stacks, as well as on multi-scale shape stacks of individual slices which can be obtained by reformatting the volumetric shape stacks organized as single $3\frac{1}{2}D$ stacks into a sequence of $2\frac{1}{2}D$ stacks.

In appendix A, figures A.7-A.12 illustrate the reformatted $2\frac{1}{2}D$ multi-scale shape stacks associated with figures 9.1 and 9.2, where the results obtained from $3\frac{1}{2}D$ and $4D$ image scale-spaces are both shown in order to illustrate the difference in the locality of the results which was already discussed in the qualitative description in section 9.3.1. In particular, the local curvature mapping visualizes the number and deepness of tracked sulci (characterized by inward curvature or red colour mapping) and gyri (characterized by outward curvature or green colour mapping) of the brains, along with the triangulation distances to the lowest scale shape, which is a more valid metric for $2\frac{1}{2}D$ stacks than across image slices as discussed earlier. The slice samples for patients 1 and 3 show overall very smooth outlines of low curvature and low relative distances, while for patients 2 and 4, disruptions and irregularities in the shape flow can be perceived. Both controls give a more regularly shaped visual impact in terms of more symmetric curvature behaviour.

For the visualization and qualitative mapping of local shape metrics for the $3D$ and $3\frac{1}{4}D$ multi-scale shape stacks, figures A.13-A.18 in appendix A illustrate four scale samples in form of volumetric instances of the brains obtained in image scale-spaces of dimensionality $3\frac{1}{2}D$ and $4D$. In contrast to the iso-surface rendering of the corresponding volumetric lowest scale results in figures 9.3-9.8, the volumetric shape stacks in the appendix were computed and displayed by triangulation of the focused contours of each scale level to volumetric structures (similar to the visualization of the $2\frac{1}{2}D$ stacks of planar shapes used in this dissertation). Several items, apart from the local shape mapping, are important: first, the difference in locality of the results with respect to the chosen scale-space dimensionality becomes very apparent. Especially at the highest scale, the $4D$ scale-space samples look very similar, while general structural differences are already apparent for the $3\frac{1}{2}D$ results, in particular deep protrusions of patient 1 which were already discussed above, in addition to apparently irregular shape patterns for patients 3 and 4. Second, volumetric visualization enables the inspection of brain images in regard to their structural complexity, such as high fragmentation parts occurring for patients 2 and 4, or overall smoothness occurring for patients 1 and 3. Third, the inadequacy of the triangulation of highly complex shapes becomes apparent when comparing the lowest scale volumetric brain visualizations obtained by triangulation in figures A.13-A.18 with the top views of the iso-surface rendered focusing results obtained at the lowest scale level in figures 9.3-9.8, which are of superior visual impact and accuracy. Local shape mapping, however, is in principle unaffected by the quality of the triangulation, except for the triangulation distance itself. Especially the local curvature mapping allows to visualize the location of sulci and gyri, as well as their absence characterized by large surface parts of low curvature, and their degree of symmetry. Additionally, the scale location at which highly curved structures become visible can be used as an indication for the spatial width of the structure. For example, larger sulci for the individual brains are tracked at higher scale levels, while smaller sulci and gyri are detected at intermediate to low image scales. The local Chamfer distance mapping enables to visualize parts of the brain surface of high smoothness in terms of partially very low distance values across scales, e.g. for patients 1 and 3.

9.3.4 Summary of Application to Patients With Epilepsy

In this section multi-scale active shape description was applied to brain volumes of patients with epilepsy and to controls. Qualitative and quantitative shape measurements across scale were performed, and resulting shape stacks were visualized with colour mapping of local shape information which enable the inspection of the brain shapes at varying degrees of detail. Further and more localized investigation can be carried out by reformatting the volumetric brain stacks into $2\frac{1}{2}D$ stacks for the individual image slices.

On the limited evidence available in this study, the hypothesis could be formed that global volumetric measurements across scale are capable of discriminating between normal and abnormal shapes. One such feature is the dimensionless volumetric compactness and its slope across scale, indicating the degree of *sphereness*, which have shown potential to categorize the investigated cases into shapes of very low, intermediate or average, and high complexity. Additionally, the fractal volumetric measurements correspond to intuition of shape regularity or irregularity in terms of statistical self-similarity under rescaling. They are in contrast to volumetric fractal dimension estimation like in [Free *et al.*, 1997] obtained on a quasi-planar basis by averaging across image slices, and may prove to be an indicator of high fragmentation and complexity, as well as of average complexity and normal shape variability, and therefore may be equally capable of distinguishing between normals and abnormals. A study based on a larger number of cases would be required to further test and confirm the formulated hypothesis. Such a study, however, is beyond the scope of this dissertation, which focuses on formulating new multi-scale methodologies for shape description.

9.4 Application to Patients with Multiple Sclerosis

Spinal cord atrophy implicating axonal loss in diameter of the cord and decrease in cord area has been observed in the development of disability for patients with multiple sclerosis. Recent studies performed at the Institute of Neurology at Queen Square, London [Losseff *et al.*, 1996b; Losseff *et al.*, 1996a], have shown that demyelination due to lesions additionally contributes to atrophy by changing the shape of the cord, and have demonstrated that there is a strong correlation between spinal cord atrophy and clinical disability. Spinal cord data is acquired using $3D$ MRI techniques. The number of contiguous slices containing useful information about the cord is very limited, and the data needs to be reformatted in the perpendicular direction to the spinal cord for each slice. In order to compare normal and pathology of the cord in terms of geometric shape, usually a segmentation of the MR image data is performed on the few slices available. Due to the large partial volume effect in this data, however, the appearance of the boundary between the cord and cerebral spinal fluid (CSF) is very blurred, and is commonly assumed by clinicians to lie at intensity values halfway between the values of the cord and the CSF. A standard segmentation technique for the spinal cord requires therefore the manual outlining of a region of interest (ROI) around the cord (called *inner* ROI) in the top slice, followed by outlining an ROI around the cord and CSF space (*outer* ROI) in the same slice in order to derive mean intensity values for CSF and the cord. The mean intensity value for CSF, $\overline{I}_{CSF}$, is then computed using the areas A and mean intensities $\overline{I}$ of the inner and outer ROIs [Losseff *et al.*, 1996b]:

$$\overline{I}_{CSF} = \frac{\left(\overline{I}_{outer} \cdot A_{outer}\right) - \left(\overline{I}_{inner} \cdot A_{inner}\right)}{A_{outer} - A_{inner}} \tag{9.1}$$

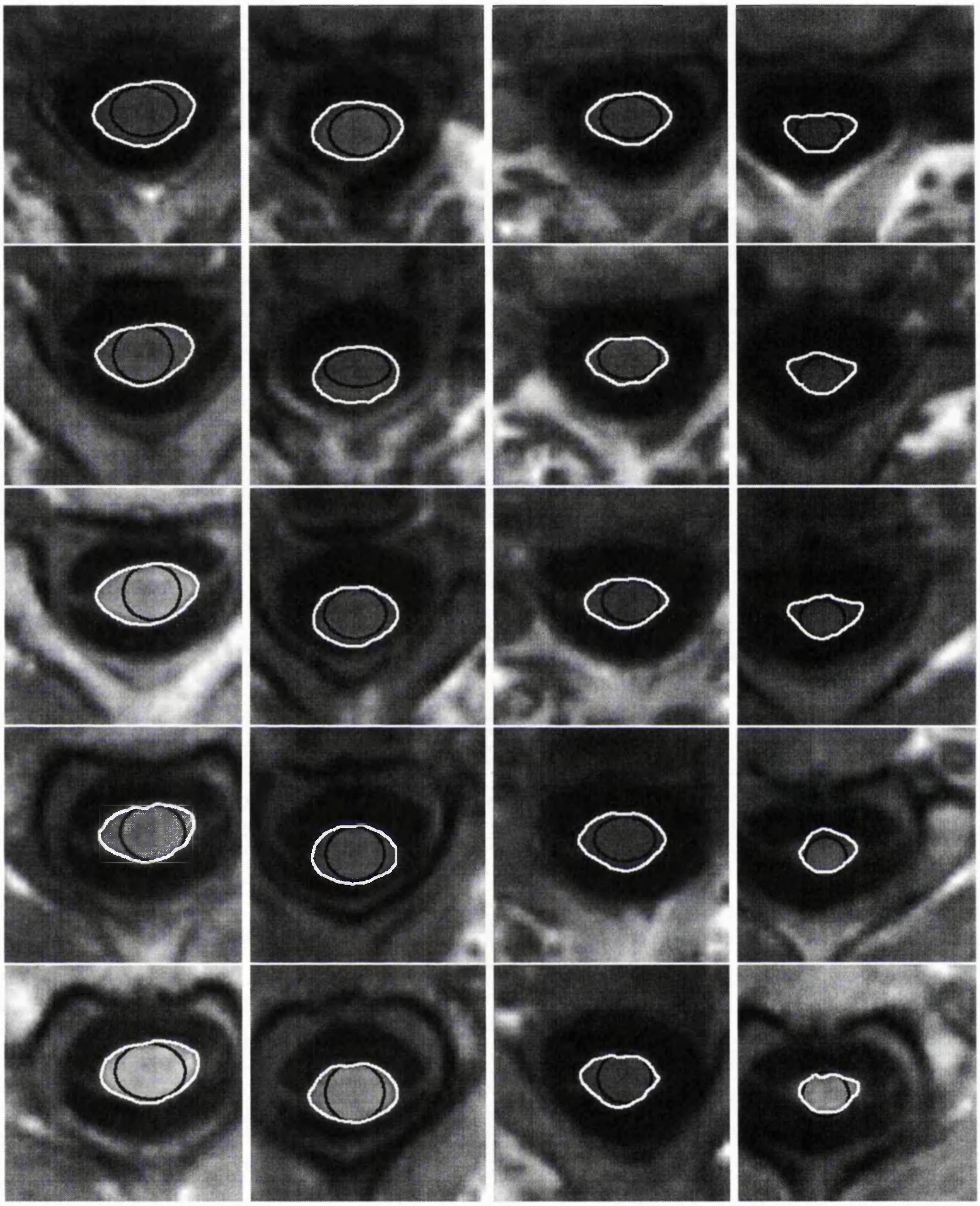
Deriving the boundary intensity as the mean of the cord and CSF intensity thus enables to automate the boundary detection in the five relevant axial slices, with resultant cross sectional area measurements, where the area mean across slices is used as a quantitative measure of atrophy.

The technique described above suffers from several problems: first, it is based on the image intensity distribution which may vary between patients as well as over time. Second, the semi-automated delineating of the cord boundary is still based on manual outlining of the inner and outer ROIs. An overestimation of the inner ROI decreases the mean intensity of the cord, and leads to a higher mean CSF intensity, causing a shift of the cord boundary in the following contouring process and leading to an overestimation of the cord area. Similarly, an underestimation increases the mean cord intensity, and leads to a subsequent underestimation of the area. It was argued in [Losseff *et al.*, 1996b] that if such an over- or underestimation is performed consistently in a clinical study, a natural regression to the mean is produced. Overestimation has been observed especially for normal sized cords, which is mainly attributed to the partial volume effect. Finally, demyelination of the cord manifests less in decrease of area, but in reduction in axonal diameter. Further study on how the pathology in the spinal cord changes the shape of the cord thus becomes interesting in view of ascertaining where and how the pathology happens. It is therefore very important to gather all available shape information which might prove to be useful for the analysis of normal and abnormal cord shapes in order to derive a quantification to what extent, or in what pattern (or the lack of it) spinal cord changes in MS occur.

The concept of multi-scale active shape description increases the amount of otherwise limited shape information, while at the same time organizes the multi-scale metrics in terms of their behaviour across scale. The application of this technique to spinal cord data involves minimal preprocessing (see table 9.2), as well as a semi-automated technique for initialization. As mentioned above, initialization is not performed using shape propagation, but using multi-scale blob detection which will be explained in the following in more detail. To demonstrate the functionality of multi-scale active shape description in this application, four scans with rescans over a time period are tested, including two controls and two patients suffering from MS. Figures 9.11 and 9.12 illustrate the time series of the upper five relevant image slices of the case, with the initial highest scale and final lowest scale models superimposed in black and white, respectively.

9.4.1 Initialization

Section 3.4.3 has introduced the concept of multi-scale blob detection by locating the normalized scale-space extrema of the normalized scale-space Laplacian. Constructing a dense scale-space for a given range of scales allows to compute the scale-space Laplacian whose absolute value is given by $\|\sigma^2 \Delta L(x, y, z_k; \sigma)\|$. Using this technique for the application to spinal cord data makes use of the observation that the spinal cord is a blob-like, ellipse-shaped structure of varying size in terms of its major axis (or diameter) and its minor axis, as well as its location. These parameters vary between cases, and may also need adjustment across slices as well as over time due to different acquisition parameters of a rescan.



(a) (b) (c) (d)

Figure 9.11: Upper five slices for spinal cord MR data sets of two controls (columns (a) and (b)) and two patients (columns (c) and (d)) from a first scan. The initial model is superimposed in black, the final, lowest scale model is superimposed in white.

In this dissertation, a semi-automated multi-scale blob detection has been developed for the initialization of the multi-scale active contour model in spinal cord data. For each slice of each data set, a dense $3D$ scale-space of 16 samples ranging between $\sigma \in [12; 28]$ (see table 9.2) is constructed, and the normalized Laplacian is computed as an appropriate *blob* measure. For the top slice of each data set, limited user interaction is required in order to derive the potential location of

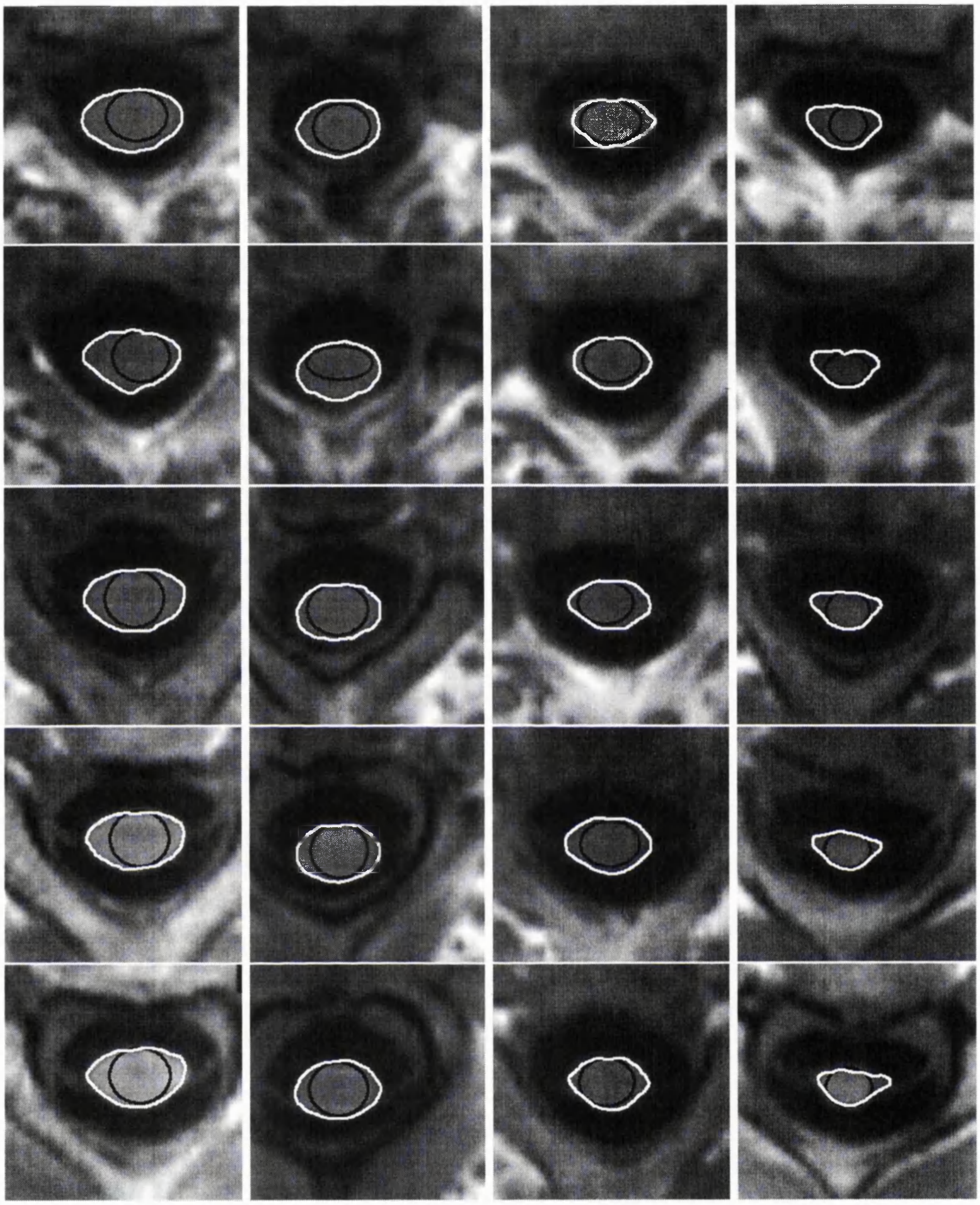


Figure 9.12: Upper five slices for spinal cord MR data sets of two controls (columns (a) and (b)) and two patients (columns (c) and (d)) from a rescan. The initial model is superimposed in black, the final, lowest scale model is superimposed in white.

the centre of the cord. This is performed by specifying a *seed point* anywhere inside the cord. Experimentation has shown that the following blob detection behaves robustly towards this manual interaction. A local scale-space extremum is then found by a simple tracking algorithm which is shown in algorithm 9.1. Local tracking within a neighbourhood whose size is chosen with respect to the image scale is preferred over global detection of scale-space extrema to improve compu-

---

// Initialize seed point $\mathbf{p}_{seed}$ for top slice, or take centre point $\mathbf{p}_{center}$ from previous slice, and set maximum scale response max to zero

// Find local maximum scale response of normalized Laplacian

for $i = 0$ **to** $n-1$ **do**

    for $y = y_{seed} - \sigma_i$ **to** $y_{seed} + \sigma_i$ **do**
        for $x = x_{seed} - \sigma_i$ **to** $x_{seed} + \sigma_i$ **do**
            if $\|\tilde{\Delta}L(x, y, \sigma_i)\| >$ max **then**
                set $\sigma_{blob} = \sigma_i$
                set $\mathbf{p}_{centre} = (x, y)$
                set max $= \|\tilde{\Delta}L(\mathbf{p}_{centre}; \sigma_{blob})\|$
        end for
    end for

end for

// Ellipse detection starting from $\mathbf{p}_{centre}$ in image slice $L_{\sigma_{blob}}$ (rotation is omitted for simplicity)

set $a_{max} = \sigma_{blob}$

set $b_{max} = \sigma_{blob}$

set $max = Boundariness(Ellipse(\mathbf{p}_{center}, a_{max}, b_{max})$

for $a = \frac{\sigma_{blob}}{2}$ **to** $2 \cdot \sigma_{blob}$ **do**

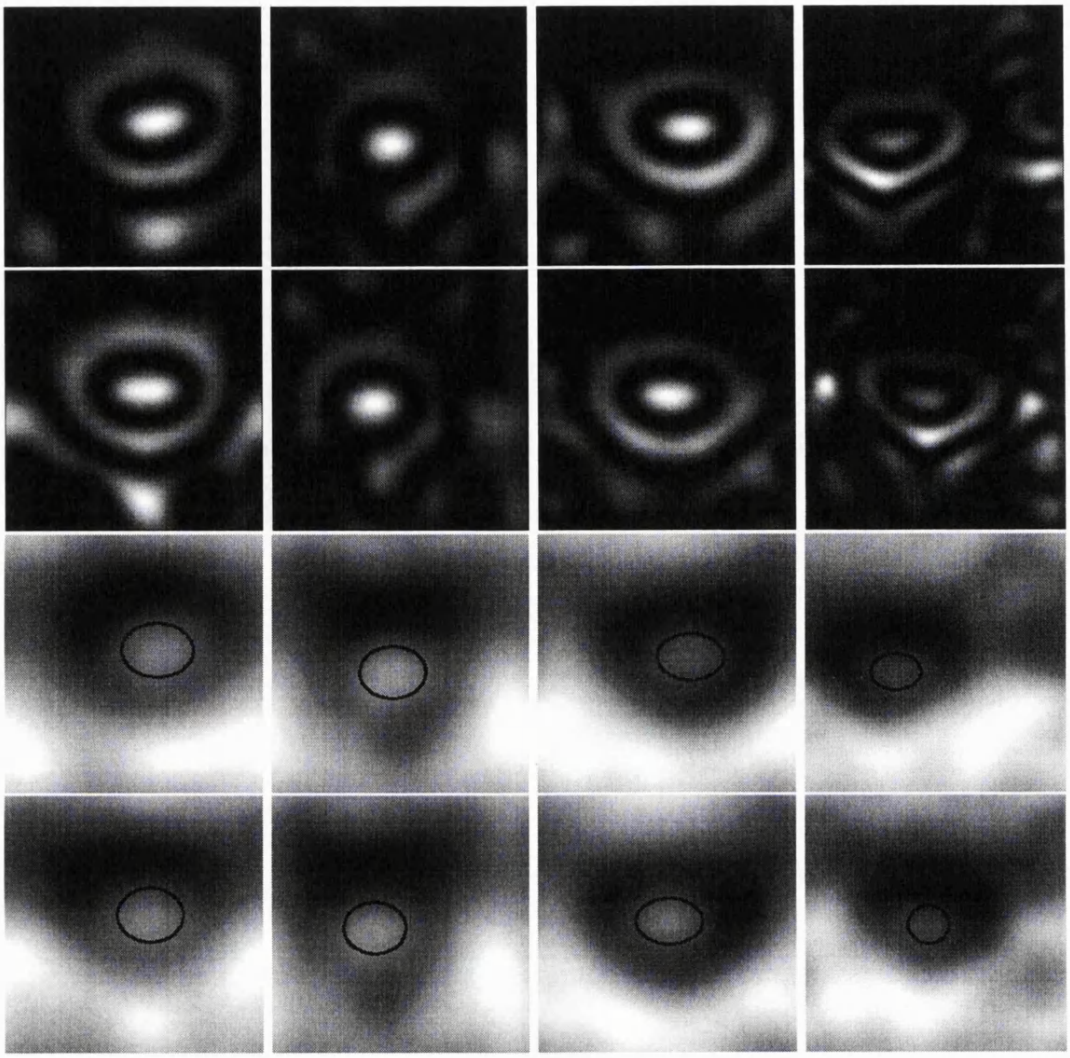
    for $b = \frac{\sigma_{blob}}{2}$ **to** $2 \cdot \sigma_{blob}$ **do**
        set $value = Boundariness(Ellipse(\mathbf{p}_{center}, a, b))$
        if $value > max$ **do**
            set $a_{max} = a$
            set $b_{max} = b$
            set $max = value$
    end for

end for

Algorithm 9.1: Algorithm for maximum scale response blob detection in spinal cord data. $\|\tilde{\Delta}L(x, y, \sigma_i)\|$ corresponds to the absolute normalized scale-space Laplacian at scale level σ_i, and the boundariness is the integrated directional edge response along the ellipse. The estimated final ellipse is located at σ_{blob} with centre p_{centre} and the maximum response values for major and minor axis a_{max} and b_{max}.

tational efficiency, and to avoid ambiguity for the case that the desired local extremum does not correspond to the global one (as the image might contain other scale-space extrema outside of the cord). The tracking is performed by taking the blob whose location is determined by its centre, and whose radius is given by the scale of maximum response, as an initial ellipse. Rotating, as well as changing the length of the major and minor axis allows to integrate the gradient magnitude along the ellipse boundary, and to find a best fitting.

Having obtained an initial ellipse for each top slice, the centre of the ellipse is propagated to the next lower slice as a new seed point, and the scale-space tracking and ellipse fitting process is



(a) (b) (c) (d)

Figure 9.13: Scale-space Laplacian at maximum response scale, and initial ellipses at maximum response scale for spinal cord data. The upper two rows correspond to the normalized Laplacians of the upper rows of figures 9.11 and 9.12, respectively, and the lower two rows show the correspondingly smoothed image slice with the initial ellipses superimposed.

repeated. As a result, not only an initial ellipse shaped model for each individual image slice is obtained, but also the individual maximum response scales as starting scales for the following active shape focusing process, along with the resulting ellipse parameters indicating overall object width in terms of their major and minor axes. Figures 9.11 and 9.12 illustrate the found initial models superimposed on black. Figure 9.13 illustrates for the top slices the absolute scale-space Laplacians and correspondingly smoothed images at the scale where the maximum scale-response was found, along with the superimposed located ellipses. From these images it becomes clear that in some cases several scale-space extrema are present, motivating the need for local tracking.

Table 9.5 lists the ranges for the maximum response scales for all spinal cord data sets. These

Data	Time	Maximum response scale	Major axis	Minor axis
Control 1	$t = 1$	[25.01; 26.46]	[31; 38]	[27; 30]
	$t = 2$	[23.64; 26.46]	[31; 33]	[27; 31]
Control 2	$t = 1$	[21.11; 25.01]	[32; 36]	[20; 32]
	$t = 2$	[21.11; 25.01]	[31; 36]	[18; 30]
Patient 1	$t = 1$	[21.11; 22.34]	[29; 33]	[23; 27]
	$t = 2$	[21.11; 22.34]	[30; 33]	[23; 26]
Patient 2	$t = 1$	[15.92; 18.86]	[22; 25]	[18; 24]
	$t = 2$	[15.04; 17.82]	[20; 25]	[17; 21]

Table 9.5: Maximum response scales for the spinal cord data sets, and resulting ellipse parameters. A linear scale-space with 16 scale samples with $\sigma \in [12; 28]$ was computed, and the scale-space extrema were tracked in a local neighbourhood of the normalized absolute scale-space Laplacian $\tilde{\Delta}L$. The ellipse major and minor axes parameters were found by efficient local search.

values along with the ellipse parameters provide potentially valuable shape information about the overall size and elongation of the spinal cord for the different cases. A decrease in overall size over time for all data can be observed which is probably due to different acquisition parameters of the scans and rescans. The maximum response scale and ellipse parameters are significantly lower for the data acquired of patient 2, with $\sigma_{blob} \in [15.04; 18.86]$ in comparison to the controls which range in $\sigma_{blob} \in [23.64; 26.46]$ (control 1) and $\sigma_{blob} \in [21.11; 25.00]$ (control 2), and patient 1 with $\sigma_{blob} \in [21.11; 23.33]$. Patient 1 and control 2 yield very similar maximum response scales, but differ in the elongation of the initial ellipse shape models, which are for patient 1 more circular shaped, i.e. with similar values for minor and major axis, than for control 2.

In the following, the results obtained for the application of multi-scale active shape description to the investigated spinal cord data will be presented, based on active shape focusing of the initial ellipse shape models obtained by multi-scale blob detection and tracking as described above. Each image slice is separately focused down in a $3D$ image scale-space which is sampled between the individual maximum response scale and pixel unit scale, and the number of samples is derived from the maximum response scale.

9.4.2 Description and Visualization

In order to qualitatively describe the active shape focusing results, figures 9.11 and 9.12 can be visually inspected. The final, lowest focusing results are superimposed in white on the individual slices. In general, focusing has been performed adequately despite the lack of a clear border of the cord rim, separating the cord from surrounding CSF. This is due to the ability of the multi-scale active contour model to locate *subjective* contours, as in lack of adequate edge forces the model is constrained by the elasticity term and by the adjustment to the underlying image curvature.

For patient 2 in the first scan in figure 9.11, however, the lower two slices cannot be properly delineated due to dymelination at the elongated parts of the cord. This is additionally enhanced by the comparatively low image scale derived as the maximum response scale, which affects the locality of the solution. As was discussed in the previous section, the located cord boundaries vary considerably between the four cases. Control 1 has the largest overall cross-sectional cord size, while patient 2 has the lowest (and least ellipse-shaped) size. Patient 1 and control 2 yield rather similar sizes, but it can be observed that the cord rim of patient 1 is less ellipse shaped. This indicates that size and diameter measurements are probably not sufficient quantification methods to reliably detect spinal cord atrophy.

In order to quantify the shape information obtained in the active shape focusing process, figures 9.14 and 9.15 illustrate the change of planar global shape metrics in terms of area, perimeter, and planar compactness across scales for all relevant slices of the spinal cord data at scan time $t = 1$ (first scan) and $t = 2$ (rescan), respectively, and figures 9.16 and 9.17 show the relative distance measurements to the respective lowest scale results (or reference shapes) in absence of a ground truth. The first two figures show a ranking of the largest area measurements in the order of controls 1 and 2, closely followed by patient 1, and, with a considerable difference, patient 2. A similar ranking can be observed from the perimeter measurements across scale. Both measurements almost continuously increase for decreasing scales due to the low complexity of the shape under investigation. It can be seen, however, that for some slices of control 1 which are less ellipse shaped, but sharper bended at the positions of maximum elongation, and most of the slices of patient 1 and some slices of patient 2, a slight decrease in area occurs at smaller scales, indicating a local deviation from the global ellipse shape. The compactness across scales yields small values for all data sets at higher scales due to the overall ellipse shape of the cord rim. At finer scales, however, it steeply increases for patient 2, and a little less for some slices of patient 1. Recall that the ratio between perimeter and volume in terms of the compactness measures the degree of *roundness* (yielding minimal values for circular shapes), and therefore has a potential of quantifying atrophy in addition to absolute area and perimeter measurements. Over the time series, only few differences can be seen. The most prominent observation is the steeper increase of area and perimeter for patient 2, leading also to a steeper increase of compactness. This might be an indication for a deterioration of the cord atrophy for patient 2, as all other measurements of the controls and patient 1 remain stable. This deterioration is not necessarily visible from qualitative inspection or the lowest scale area measurements. The relative distance measurements in terms of Chamfer and triangulation distance from the respective reference shapes yield largest distance values for the controls at higher scales, and continuous decreases for decreasing scales. This is due to the large upper scales in these cases, where the shrinking effect of the Gaussian blurring needs to be compensated, but the ellipse shape of the reference model is already captured glob-

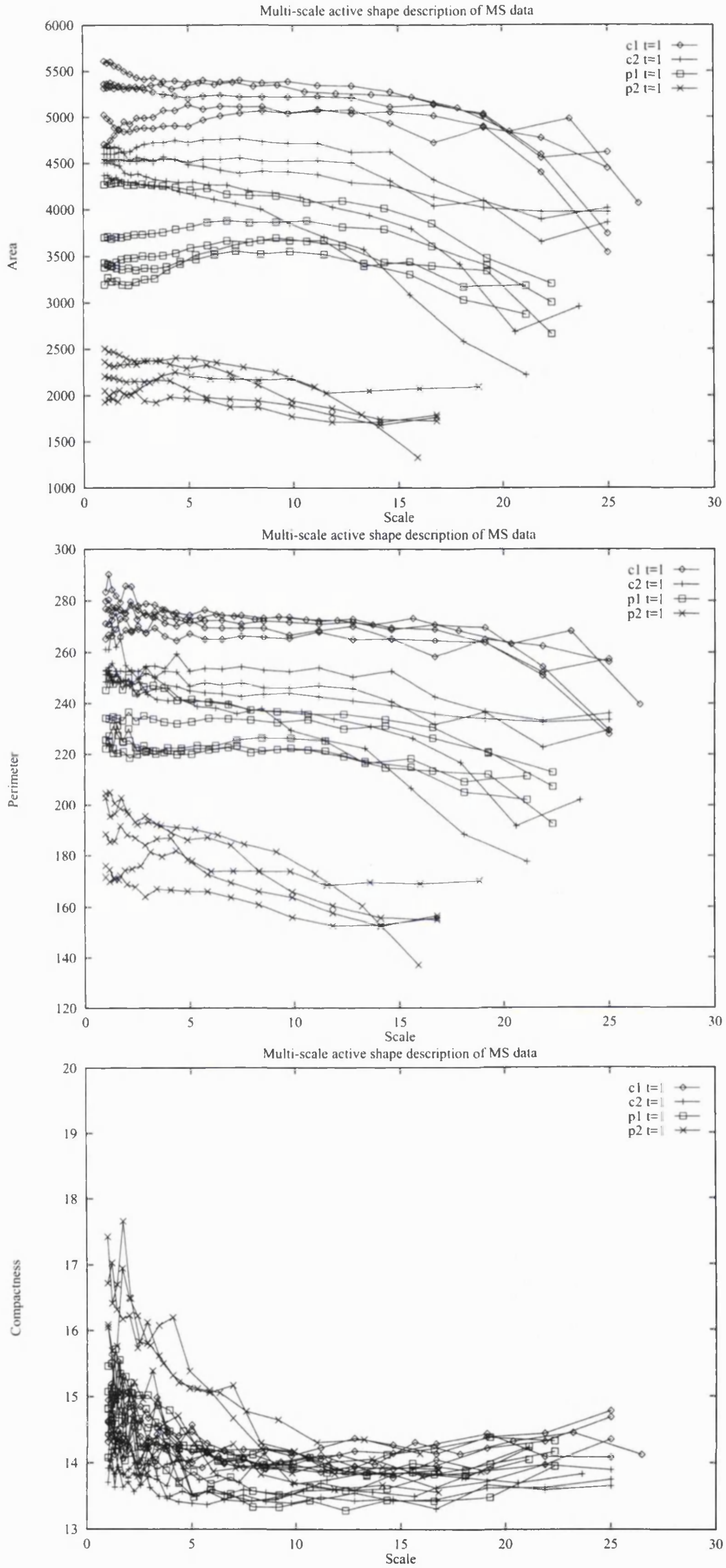


Figure 9.14: Area, perimeter and compactness of the spinal cord images of figure 9.11 across scales for active shape focusing.

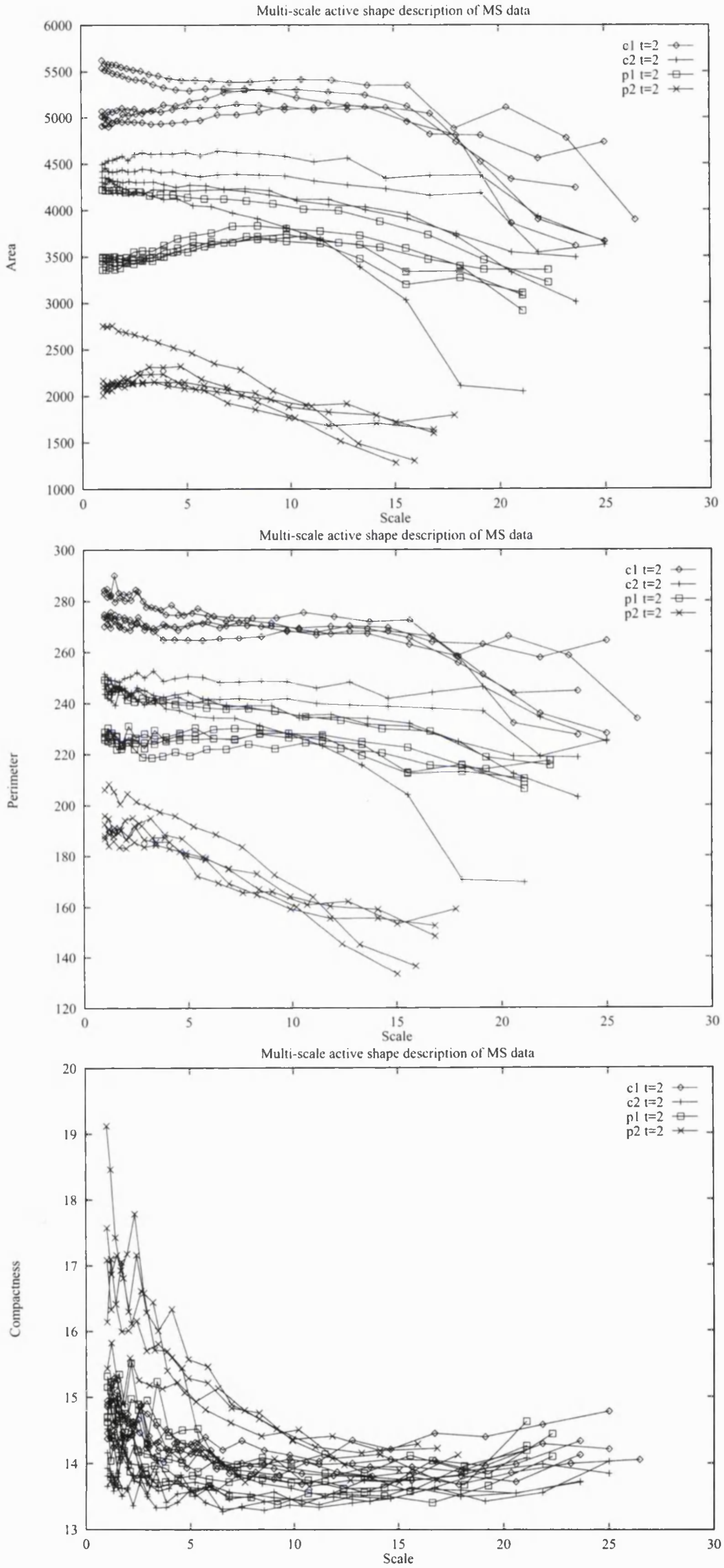


Figure 9.15: Area, perimeter and compactness of the spinal cord images of figure 9.12 across scales for active shape focusing.

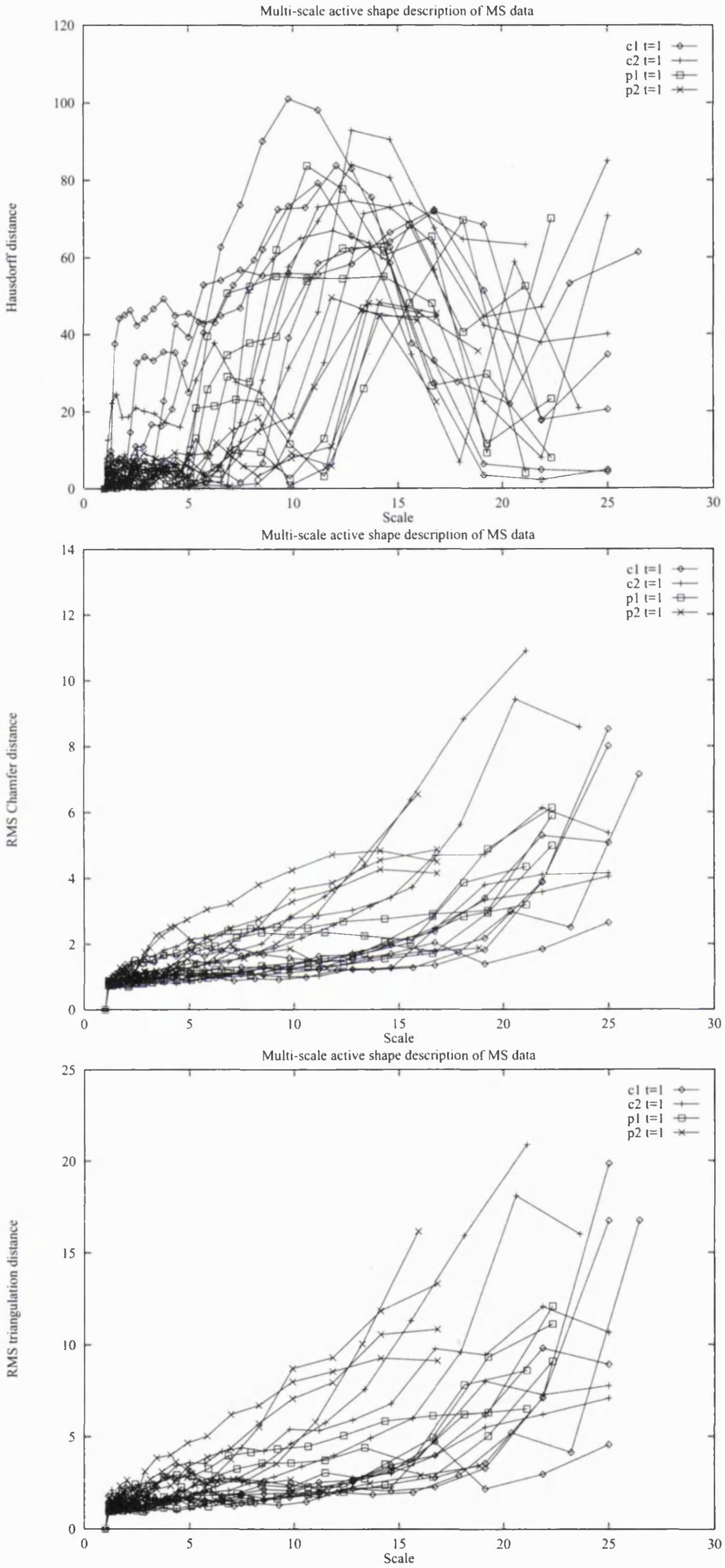


Figure 9.16: Hausdorff, RMS Chamfer, and RMS triangulation distances of the spinal cord images of figure 9.11 across scales for active shape focusing.

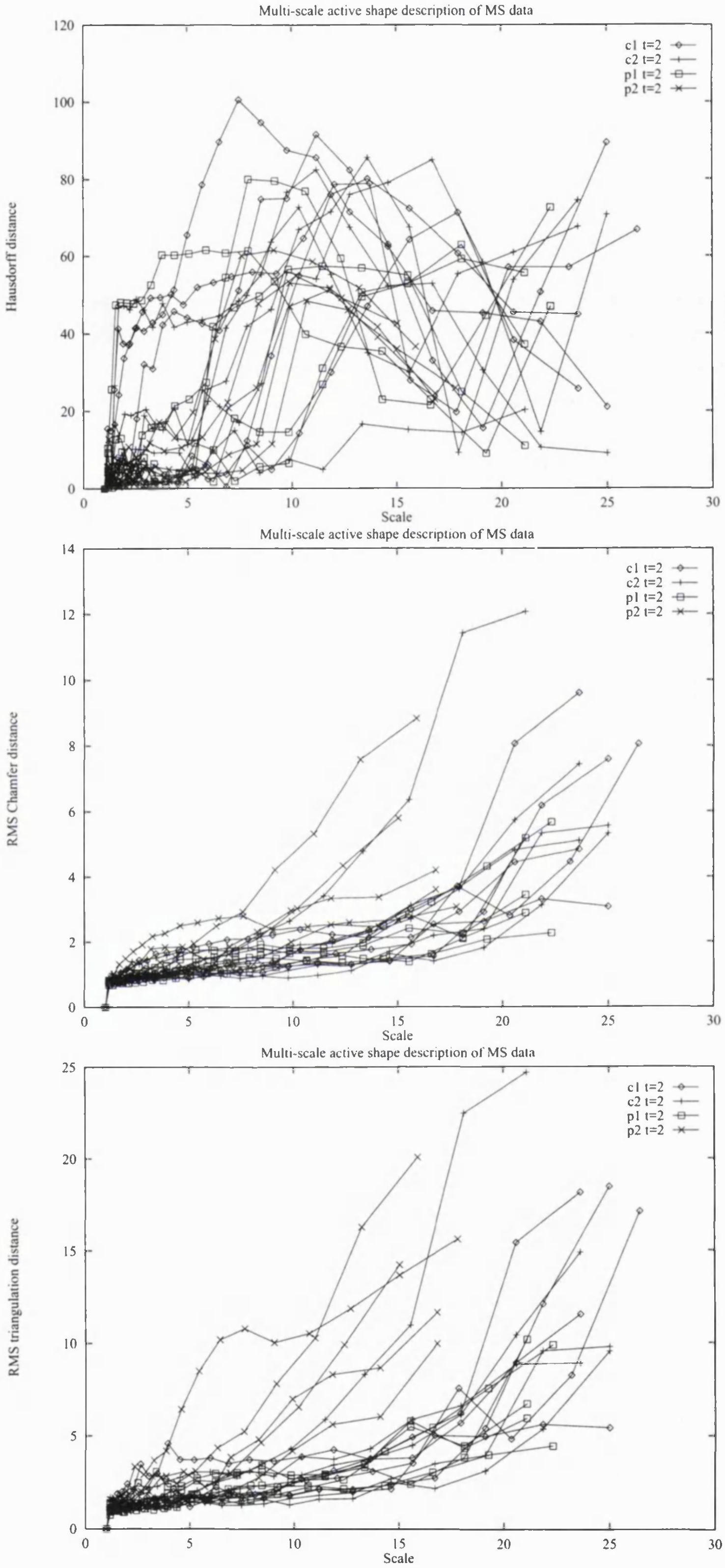


Figure 9.17: Hausdorff, RMS Chamfer, and RMS triangulation distances of the spinal cord images of figure 9.12 across scales for active shape focusing.

Descriptor	Time	Control 1	Control 2	Patient 1	Patient 2
$\overline{A}$	t=1	5109.624	4299.678	3611.464	2101.968
	t=2	5095.718	4171.872	3627.066	2095.194
$\overline{P}$	t=1	272.3230	243.2956	226.7558	177.2806
	t=2	269.6038	238.5722	226.5784	179.5996
$\overline{C}$	t=1	14.33150	13.81168	14.28408	14.97674
	t=2	14.28716	13.68962	14.18658	15.46552
$\overline{\mathrm{dist}}_H$	t=1	31.273720	23.406560	18.459320	10.136548
	t=2	33.770840	26.524732	24.989280	16.609780
$\overline{\mathrm{dist}}_{C_{RMS}}$	t=1	1.233290	1.727138	1.294718	1.565138
	t=2	1.416562	1.565230	1.150386	1.579266
$\overline{\mathrm{dist}}_{T_{RMS}}$	t=1	3.288424	4.752166	3.426866	4.780852
	t=2	3.957822	4.298010	3.040460	5.696468
$\frac{\Delta A}{\Delta\sigma}$	t=1	-26.40479	-50.31766	-16.21226	-31.11008
	t=2	-34.69116	-47.27934	-12.61000	-50.16094
$\frac{\Delta P}{\Delta\sigma}$	t=1	-0.893604	-1.692082	-0.937893	-2.218987
	t=2	-1.228322	-1.538757	-0.725929	-3.353634
$\frac{\Delta C}{\Delta\sigma}$	t=1	-0.018089	-0.018651	-0.054276	-0.148879
	t=2	-0.028838	-0.008630	-0.041649	-0.203738
$\frac{\Delta \overline{\mathrm{dist}}_{C_{RMS}}}{\Delta\sigma}$	t=1	0.120126	0.238270	0.137741	0.205899
	t=2	0.164168	0.224702	0.109957	0.244555
$\frac{\Delta \overline{\mathrm{dist}}_{T_{RMS}}}{\Delta\sigma}$	t=1	0.433275	0.493087	0.551592	0.754617
	t=2	0.443217	0.498059	0.559648	0.789252
$-\frac{\Delta\log(\frac{A}{\sigma})}{\Delta\log(\sigma)}$	t=1	1.076210	1.164550	1.084070	1.179000
	t=2	1.099510	1.164820	1.074130	1.288140
$-\frac{\Delta\log(\frac{P}{\sigma})}{\Delta\log(\sigma)}$	t=1	1.037490	1.080940	1.046790	1.120530
	t=2	1.051650	1.079020	1.039100	1.181810
$-\frac{\Delta\log(\frac{C}{\sigma})}{\Delta\log(\sigma)}$	t=1	0.998760	0.997330	1.009510	1.062050
	t=2	1.003790	0.993222	1.004080	1.075480

Table 9.6: Mean, slope, and logarithmic slope descriptors for active shape focusing of the spinal cord images with respect to scale. All values are averaged over the upper five slices, and the logarithmic slopes are estimated within the scale ranges $[\sigma_{min} = \varsigma_{min} = 3.5; \sigma_{blob}]$, where the respective values for σ_{blob} for each image slice are listed in table 9.5.

ally at high image scales. Distance measurements for patient 1, where the focusing process starts from a lower scale, yield therefore rather small values, but remain in the first scan overall higher at intermediate scales than for the controls. In the second scan, however, they appear to be lower across scales. Finally, the distance measurements for patient 2 are in their scale range the overall highest and have the steepest decrease for decreasing scales. The distances become even higher in the rescan, supporting the observation above that the cord rim of patient 2 is least ellipse shaped and of highest complexity, with a possible deterioration over time. Finally, the Hausdorff distance is not found to be a useful measure as it is here closely related to the overall size of the cord data.

Table 9.6 lists the mean and slope values of all data sets averaged across scales as well as across the individual slices for both scanning times. The mean area and perimeter measurements are constant over time, but show a clear demarcation between patients and controls, and also between the two patients, possibly categorizing them into subtle and apparent cord atrophy. The mean compactness has a much higher value for patient 2, as well as the observed deterioration over time. The other compactness values remain stable, showing lowest values for control 2 (with the most circular shaped cord), but similar values for control 1 and patient 1. The mean distance measurements are very much influenced by the high values for the controls obtained in the highest scale range, for which no corresponding data for the patients is available. It is assumed that these measurements should better be performed in a similar scale range, e.g. for $\sigma \in [1; 16]$ for which data of all cases exists. However, the triangulation distance for patient 2 is still overall highest despite the lack of information at higher scales, indicating the largest overall deviation from an ellipse shape, and the mean Hausdorff distance is lowest for the patients due to their smaller size, leading to lower worst mismatches. The slope measurements show mainly shape changes for patient 2 over time, as well as overall steepest increase of perimeter and compactness, and steepest decrease of the deviation from the lowest scale results for the same patient. The change of compactness is next lower for patient 1, followed with an offset by control 1 and control 2. In contrast to the slope of compactness, the slope measurements of area, perimeter, and Chamfer distance are not demarcating measurements for the investigated cases, as they highly depend on the overall size, individual scale ranges, and overall ellipse shape. For example, the Chamfer distance performs the closest distance measurements with respect to an underlying ellipse shape, missing out more subtle disruptions in the cord pattern. Again, it is possible that adjusting the range of scales appropriately may lead to more demarcating results for these quantifiers.

Finally, multi-scale planar fractal measurements with respect to area, perimeter, and compactness have been performed, which are listed in the last three rows of table 9.6. All measurements indicate a much lower level of self-similarity in comparison to the volumetric brain measurements of section 9.3, and are very close to 1 (note that the fractal dimension of a perfect ellipse is exactly 1, and lower values than that are theoretically not possible and therefore caused by small numerical problems). The fractal measurements indicate a progressive deterioration in terms of increase of fractality for patient 2, with overall highest values for all measurements. The fractal area and perimeter measurements show no clear demarcation for the remaining cases which have all much more ellipse shaped cords, but the fractal compactness divides the investigated cases quite clearly.

Visualization of the resulting $2\frac{1}{2}D$ multi-scale shape stacks is performed via triangulation of all focusing results of each individual image slice to the reference shape. Local curvature mapping, illustrated in figures A.19 and A.20 in appendix A for the relevant image slices at both scanning times, visualize effectively parts of high outward curvature (mapped in green) at the most elongated parts of the cord for the patients, and show more uniform bending behaviour for the controls.

The Chamfer distance mappings, shown in figures A.21 and A.22, show at the same location large local distances for patient 2, as well overall large distances for control 2 (which is especially the case at higher scales and could be observed in figures 9.16 and 9.17, respectively). The triangulation distance (figures A.23 and A.24) has similar mapping results for patient 1, but additionally shows several parts of high local distances for patient 2, which may be attributed to local lesions. The colour mapping onto the hierarchical shape stacks for spinal cord data shows that the global quantification in terms of mean and slope measurements, as well as multi-scale fractal quantification, can be well complemented with local shape information, which in this case has the advantage of giving a visual impression of the deviation from the initial ellipse shape for each individual image slice, rather than averaging across scales and slices.

9.4.3 Summary of Application to Patients With Multiple Sclerosis

In this section, multi-scale active shape description has been applied to spinal cord data. The shape description process, which is embedded into the active shape focusing process, requires in comparison to standard methodologies no prior knowledge of intensity distribution of the cord, or manual outlining of regions of interest, and is almost user independent. Visualization of local shape information, in addition to global quantification, allows to further explore spinal cord parts with respect to atrophy. A stable initialization process has been applied which yields as a byproduct an adequate scale selection scheme, as well as ellipse parameters which can be used for a first qualitative shape description. This initialization can be further improved by using directional edge information for ellipse detection [Lindeberg, 1994], as well as by tracking the local maxima in a finer sampled scale direction [Rueckert *et al.*, 1997]. $4D$ scale-space techniques have not been considered due to the reformatted nature of the image slices.

As for the application to epilepsy in section 9.3, the number of data sets used for the application to spinal cord data was very limited, as the aim of this work is to demonstrate the functionality of the framework for multi-scale active shape description. On the basis of this study, however, it could be hypothesised that the maximum response scales and resulting initial ellipse parameters, along with area and perimeter mean measurements, compactness slopes, relative distances and fractal compactness measures may have a potential of quantifying the degree of spinal cord atrophy occurring for patients with MS. Also, the monitoring over time using these shape features may allow to detect progressive deterioration in some cases. A larger study would be needed to support this hypothesis, and to establish the clinical relevance of the multi-scale shape measurements, as well as their potential of demarcating atrophy from normal shape variation, and of monitoring disease progress over time.

9.5 Application to Neonatal Data

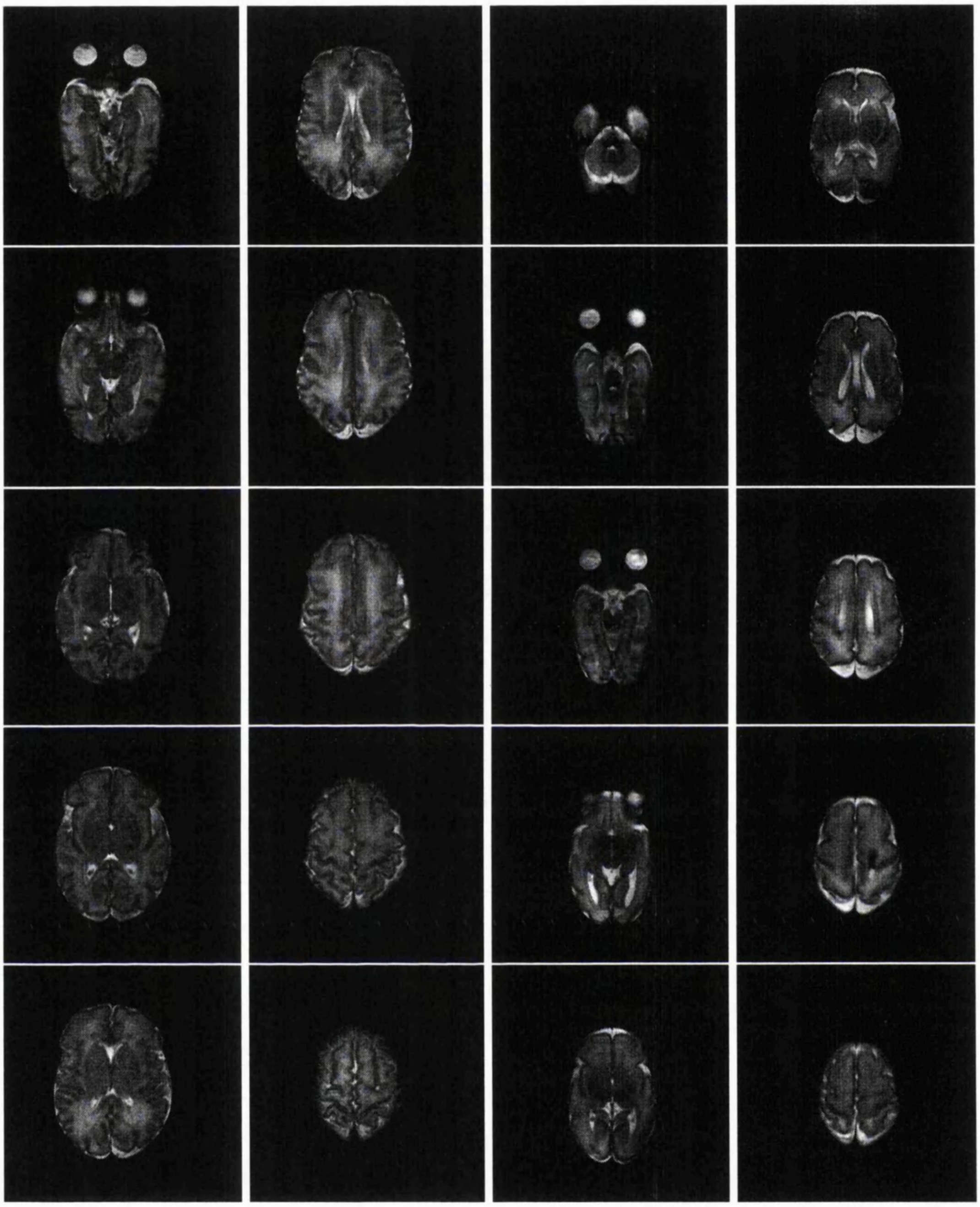
The developing human brain starts as an organ with a relatively smooth outline. As development proceeds, an increase in the volume of the grey matter in the brain is accommodated by an increase in cortical complexity with the development of convolutions in the brain surface which become fully developed gyri. This process occurs during gestation and in the first months of life and may be delayed by brain injury during gestation or at birth. Therefore, a quantitative measure of cortical complexity may provide a means of monitoring the rate of normal and abnormal brain development in order to provide a prognostic measure of the eventual degree of cerebral impairment in brain-injured neonates. Multi-slice T2 weighted MRI data sets of large slice thickness allow to assess the clinical outcome of brain-injured infants. Shape analysis measurements have already been shown to provide a quantitative measure of brain myelination during the first post-natal months of life [Thornton *et al.*, 1997], but will be further investigated using the developed techniques of this dissertation.

Figure 9.18 shows the neonatal data sets for a term brain (or a control) and a premature which show significant difference in area and complexity of the cortex outline. Due to the small overall size of either brain, only 10 contiguous slice are available. Each slice is of considerable slice thickness and has a very small pixel size (see table 9.1). The large slice thickness implies that the correlation between the individual slices is rather low. Therefore, as for the spinal cord application above, active shape focusing is performed using individual $3D$ image scale spaces associated with each image slice. Moreover, each of the 10 slices for the two data sets is manually initialized with a coarse ellipse shaped model. Due to the T2 weighting of the images, CSF appears as high intensity values. There are two possible choices for delineating the grey matter boundary, namely by adjusting the contrast parameter to detect dark grey matter shape on light CSF background, or by thresholding the high CSF values. The latter was favoured as the CSF is only visible at some parts of the grey matter boundary. The pre-processing parameters, including the threshold for intensity cutting for CSF, are listed in table 9.2. In the following, the results for multi-scale active shape description based on active shape focusing of the individual image slices will be presented.

9.5.1 Description and Visualization

Figure 9.19 shows the surface rendering results of the individual $2\frac{1}{2}D$ multi-scale shape stacks obtained for each image slice of the term and premature data sets. The difference in size, as well as the steeper tracking of sulci and gyri, and their number, has a large visual impact. The brain of the premature case seems visually much smoother, and does not seem to change very much for decreasing scales.

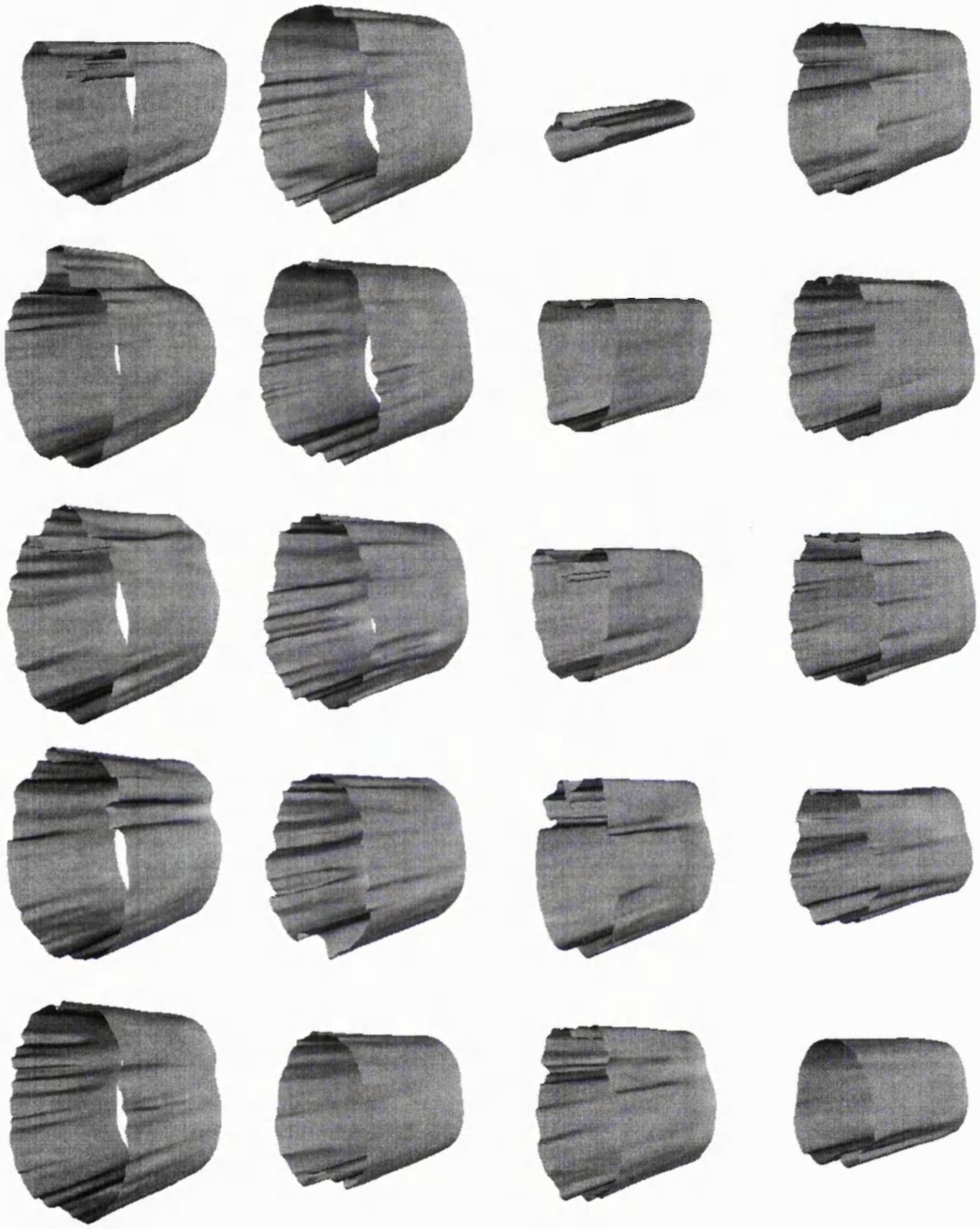
Figures 9.20 and 9.21 illustrate the global planar shape measurements as well as the distance



(a) (b) (c) (d)

Figure 9.18: Brain MR data set of a term case (columns (a) and (b)) and a premature case (columns (c) and d)).

measurements from the lowest scale reference shapes for all slices of both data sets across scale. While the term case is of overall higher area and perimeter, both cases are of surprisingly similar compactness, and even showing overall slightly higher values for the premature case. This could not be expected by the visualization of the data sets and the corresponding shape stacks which show opposite compactness behaviour, and will be further discussed below. Additionally, both cases have a similar continuous decrease in relative distance measurements for decreasing



(a) (b) (c) (d)

Figure 9.19: Interpolated shaded surface renderings of the hierarchical shape stack of the neonatal data obtained via active shape focusing. Surface rendering is performed in a *fine-to-coarse* view. See text for further details.

scale. The global planar measurements, however, show another behaviour over scale: the area for the term case is continuously decreasing rather than increasing for lower scales, while the premature case shows more constant area measurements across scale. Regarding the perimeter measurement, the term case shows a similar constant increase for all slices, while the premature case diverges in the different slices.

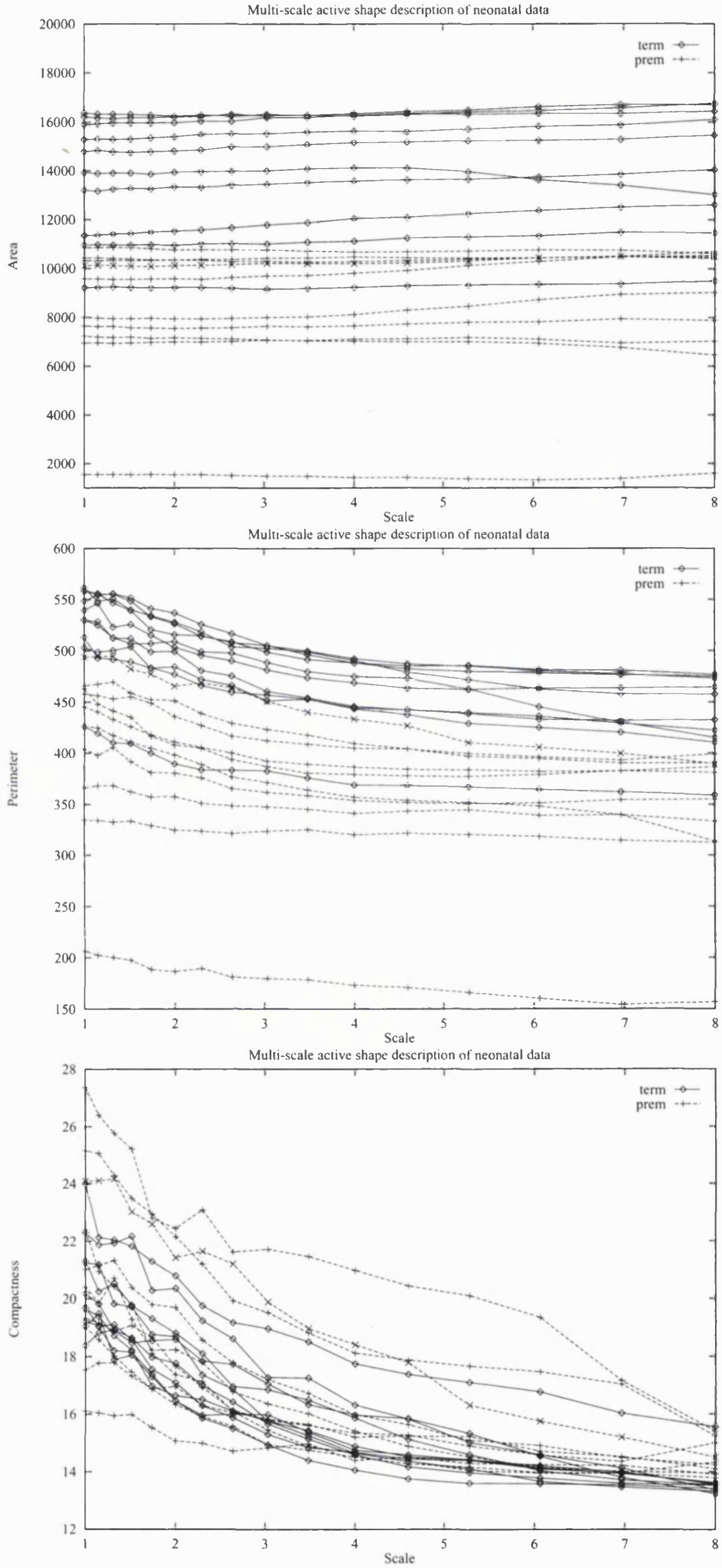


Figure 9.20: Area, perimeter and compactness of the neonatal data for active shape focusing.

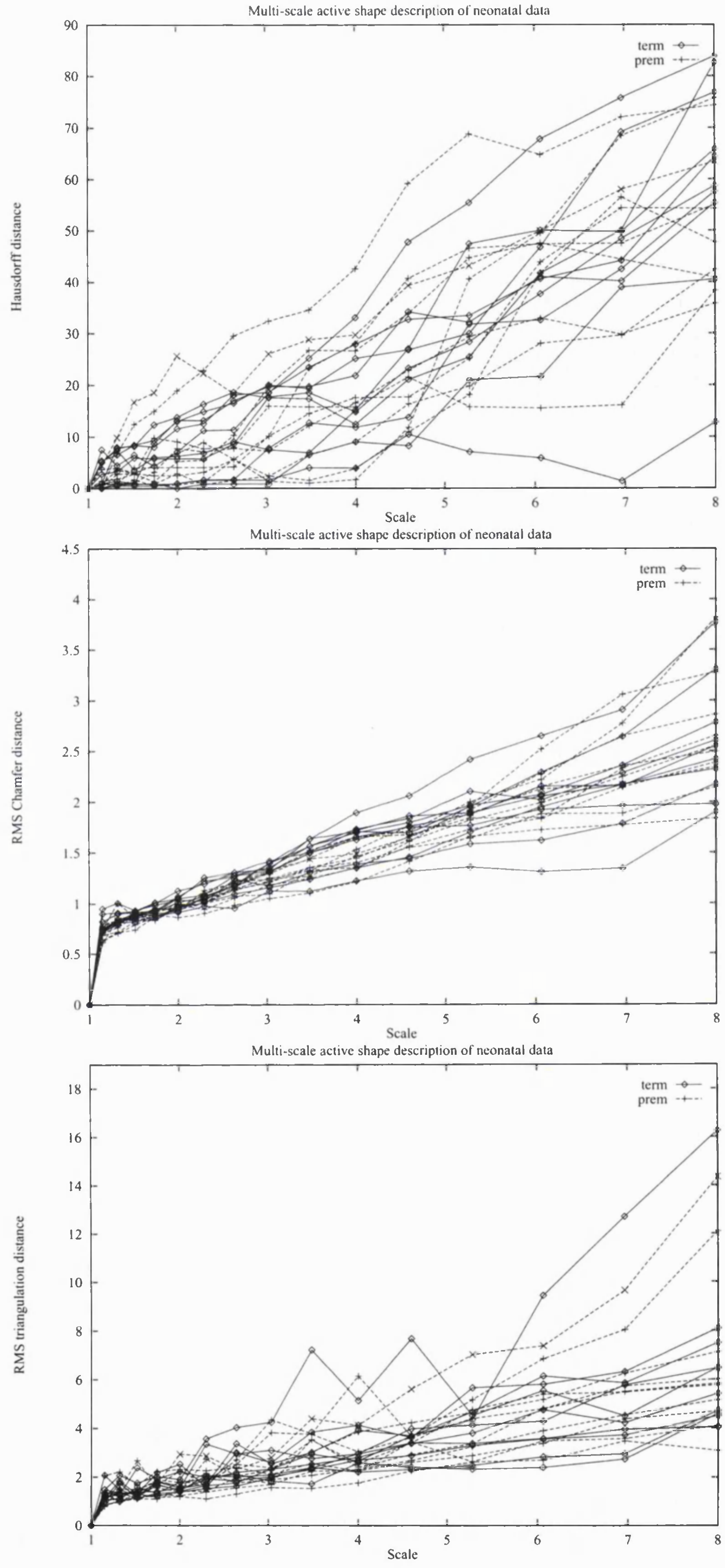


Figure 9.21: Hausdorff, RMS Chamfer, and RMS triangulation distances of neonatal data across scales for active shape focusing with respect to the known shape.

Descriptor	Term case	Premature case
$\overline{A}$	13897.300000	8307.130000
$\overline{P}$	480.217000	371.484000
$\overline{C}$	16.871400	17.715400
$\overline{\text{dist}_H}$	18.087900	18.472700
$\overline{\text{dist}_{C_{RMS}}}$	1.095820	1.057680
$\overline{\text{dist}_{T_{RMS}}}$	3.710300	3.686050
$\frac{\Delta A}{\Delta\sigma}$	77.360300	37.011500
$\frac{\Delta P}{\Delta\sigma}$	-12.733500	-9.397120
$\frac{\Delta C}{\Delta\sigma}$	-0.982766	-0.965931
$\frac{\Delta \overline{\text{dist}_{C_{RMS}}}}{\Delta\sigma}$	0.240428	-0.255215
$\frac{\Delta \overline{\text{dist}_{T_{RMS}}}}{\Delta\sigma}$	1.256820	1.256820
$-\frac{\Delta\log(\frac{A}{\sigma})}{\Delta\log(\sigma)}$	0.975164	0.977868
$-\frac{\Delta\log(\frac{P}{\sigma})}{\Delta\log(\sigma)}$	1.080950	1.072310
$-\frac{\Delta\log(\frac{C}{\sigma})}{\Delta\log(\sigma)}$	1.186740	1.166750

Table 9.7: Mean, slope, and logarithmic slope descriptors for active shape focusing of the neonatal data with respect to scale. The values were averaged across all slices, and the logarithmic slope estimations were performed in the scale range $[\sigma_{min} = \tilde{\varsigma}_{min} = 2.5; \sigma_{max} = 8]$.

Table 9.7 summarizes the mean, slope, and multi-scale fractal measurements averaged across all slices of each data set. The most apparent difference lies in the much larger mean area and perimeter measurements of the term case, which also manifests as higher negative slopes. All other shape quantifications, however, show very similar values, including the compactness related descriptors. As mentioned for the observation of similar compactness across scales, this is a quite unexpected result, as visually significantly different degrees of cortical complexity are visually observed. This result can be explained, however, by the decreasing area behaviour for the term case with simultaneous increase of perimeter, as opposed to rather constant area and perimeter behaviour for the premature case. This leads to similar values of compactness for both cases, but which are caused by different structural characteristics of the underlying shape. Recall from the previous chapter, that for the *kangaroo* test image a similar behaviour as for the term case could be observed, and that loss in area with simultaneous increase of perimeter for decreasing scales implies the tracking of deep structures, e.g. convolutions and protrusions. The overall much larger size of the brain of the term case leads to a small compactness value, but if the two brains were of equal size, a problem which might be solved by registering the brains. However, especially the difference in cross-sectional area is an important quantitative and diagnostic measure which probably should not be adjusted.

The visualization of local shape metrics onto the $2\frac{1}{2}D$ multi-scale shape stacks shown in figure 9.19 is given in appendix A. In particular, the local curvature mapping (figure A.25) shows the higher number of gyri and sulci, as well as the steeper tracking of the cortical structure for the term case, providing a good visual impression of the degree of potential cortical impairment for the premature case. The colour distance mappings, shown in figures A.26 and A.27, respectively, also locally illustrate the existence and deepness of sulci and gyri, in contrast to the global quantification measures arising from these descriptors given above.

9.5.2 Summary of Application to Neonatal Data

The application of multi-scale active shape description in this section supports the shape analysis in [Thornton *et al.*, 1997] in terms of demarcation via area and perimeter measurements across scale, but it failed in detecting the different degrees in cortical complexity for reasons given above. In particular, the compactness was found to be not a useful measure in this application, as an increasing perimeter for tracking deep structures is made at the cost of loss in area. Similar fractal measures between the two cases, however, may indicate that the shapes are statistically similar, but in different ranges of scale, similar to the von Koch curves at different fractal resolution levels in the previous chapter. In other words, the premature shape may be a simplified version of the more complex term shape. This empirical observation would require a large group study for verification.

9.6 Summary

This chapter has investigated the application of multi-scale active shape description in medical imaging on the basis of active shape focusing. Three different medical problems have been addressed. For the first application, six brain volumes for patients with epilepsy and controls have been successfully focused down in scale-spaces of different dimensionalities, yielding results differing in the locality of the solution. The main advantage of applying multi-scale active shape description to such data is that it is an almost automated method, with the only user interaction being the initialization of a single slice with a circular model for all data sets. It requires no prior brain segmentation, which in clinical practice is either performed manually or with semi-automated, but still time consuming methods. The adding of the extra scale dimension was found to increase the amount of available shape information by quantifying the speed of the active shape focusing process with respect to the underlying structural complexity of the investigated brain data sets. It is hypothesized that for a larger clinical study, several of the derived shape features, in particular the volumetric compactness and fractal surface measurements, could have a potential of classifying brain shapes into different degrees of cortical dysgenesis for patients with epilepsy, and into normal shape variability. The second application to patients with Multiple Sclerosis has shown some potential of distinguishing between normal shape variability, and spinal cord atrophy. In particular, the automatic scale selection scheme and initialization provide an attractive alternative to current manual delineating of the cord rim, providing several additional shape measures of potential quantitative use regarding the overall cord size, elongation and diameter. A study based on a larger body of data would be required to further test these observations. Finally, the application to neonatal data has supported recent clinical research results in finding different behaviour in area and perimeter measurements by quantifying the mean and slope measurements of perimeter and enclosed area across scale. However, the very apparent differences in cortical complexity could not be quantified, but only visualized with colour mapping rendering techniques. Again, an application to a larger set of data, including rescans, would be needed to evaluate the multi-scale shape descriptors for this application.

The description in medical imaging has been performed on a comparatively small number of data sets. These data sets were selected in order to represent a large variability in structural deformations in the different medical applications, and to give an impression of the functionality and applicability of the developed framework. The presented results can therefore not be generalized to give statistically representative measurements and solutions to clinical problems. However, they were shown to yield a large set of higher level shape measurements, some of which might prove to have a potential to be demarcating shape measurements in medical imaging if evaluated in large group studies, as well as in studies over time.

Limitations of the application of multi-scale active shape description to medical imaging in clinical neurology are mainly caused by the $2D$ nature of the shape extraction in volumetric images. Although the use of higher-order image scale-spaces allows to capture correlation between neighbouring image slices, topological changes between slices as well as partial volume effects and large slice thickness in some of the data can only be processed in $3D$ image slice scale-spaces, or using a higher order, volumetric model. Following this approach, an extended set of shape descriptors for the description of brain surfaces can be added to the planar techniques used in this work. The following chapter will discuss several potential extensions of the presented framework in order to address the aforementioned problems, including true $3D$ shape extraction, description, and the use of non-linear scale-spaces.

Chapter 10

Outlook and Conclusions

– ON N'EST JAMAIS CONTENT LÀ OÙ ON EST, DIT L'AIGUILLEUR.
"NO ONE IS EVER SATISFIED WHERE HE IS," SAID THE SWITCHMAN.
Le Petit Prince, Antoine de Saint-Exupéry.

This chapter will first focus on current and future work of the presented methodology of multi-scale active shape description in medical imaging, and will present some promising approaches in improving and extending this methodology while simultaneously analysing its limitations. After the outlook section, the main contributions of this dissertation will be briefly summarized and discussed, and more general future directions and conclusions will be made.

10.1 Outlook

In order to formulate possible extensions, one first needs to consider and investigate the limitations of the developed methodology. Multi-scale active shape description, as presented in this dissertation, is based on implicit shape regularization and description of planar shapes in a linear image scale-space. Volumetric images are described in a slice-by-slice fashion under the assumption that a volumetric shape can be represented as a stack of planar shapes. $4D$ image scale-spaces in combination with shape propagation through the image volume have been employed in order to achieve a volumetric correlation between planar shapes obtained from adjacent image slices. This approach has been shown to be very robust, yet still has failed in some cases due to several problems inherent in medical imaging in general, and in clinical neurology in particular.

The human brain is arguably one of the most complex shapes of the human body, and cannot always be modelled by concatenating a set of planar shapes. This is due to two things: First, the brain is not of simple topology, but of several constituent subparts. Second, even at topologically simple parts of the brain, there is the chance of structures disappearing temporarily from one image plane, only to partially reappear in another. The most likely part of the brain to show such behaviour is the brain stem, which is slightly curved, but it can also be observed in the complex cortex pattern. The partial volume effect leads to a possibly inaccurate classification of image voxels not only within the image plane, but also in the slice direction of the volume. For high slice thickness, like for the neonatal data investigated in the previous chapter, there is only little spatial correspondence between adjacent image slices. In these cases, $4D$ image scale-spaces and automatic initialization methods like shape propagation cannot be used, leading to the loss of volumetric information. Even if a $4D$ image scale-space can be employed, the resulting planar shapes of adjacent slices can greatly differ. Additionally, the use of a linear scale-space reduces valuable edge information at higher scales, and yields a premature merging and shrinkage of objects. Several non-linear diffusion techniques have been reviewed in chapter 3 which may help to reduce this effect, but except for the classic edge-affected diffusion, no straight-forward $4D$ extensions exists. Finally, the planar nature of the multi-scale active contour model, even when resulting in a concatenated set of planar shapes in scale-space, only gives rise to planar shape description, and volumetric metrics can only be approximated. However, the investigation of volumetric surface structure in terms of principal surface curvature directions, as well as ridge and crease structure would provide valuable additional volumetric shape information, but is beyond the scope of this dissertation.

These limitations naturally lead to two major possible and desirable extensions of multi-scale active shape description, namely the development of a *multi-scale active surface model* which also involves the extension to volumetric internal and external energy terms as well as the investigation of true volumetric surface structure, and the use of *non-linear scale-spaces* which involves the extension of existing ones to higher dimensions. These topics will be addressed in sections 10.1.1 and 10.1.2, respectively. Furthermore, two more general future topics will be discussed: Section 10.1.3 will present possible approaches to investigate parallels between the developed planar techniques of active shape evolution and focusing, and classic multi-scale contour evolution schemes, and section 10.1.4 will show possible extensions of the multi-scale active contour model to incorporate diffusion flow information, as well as the concept of a true multi-scale rather than *fine-to-coarse* or *coarse-to-fine* approach, with an outlook to a shape-based diffusion process.

10.1.1 Extension to Multi-Scale Active Surface Models

The multi-scale active contour model developed in this dissertation is a $2D$ model with respect to spatial dimensions, with an additional scale dimension. Using shape propagation and a $4D$ image scale-space helps to improve the correlation between neighbouring image slices in volumetric images, but an extension to a true volumetric image might be in some cases necessary in order to capture the topological aspects of the shape under investigation adequately. A general formulation of such a model has been given in chapter 4, yet its implementation in practice poses several difficulties that will need to be resolved. More precisely, the following items need to be considered:

- Surface representation.
- Volumetric image features.
- Volumetric optimization.
- Surface description.

10.1.1.1 Surface Representation

There are two main surface representation techniques for deformable models, namely *explicit* surface representation via a spline surface or a parametric model, and *implicit* surface representation based on iso-surfaces of the image. Explicit surface representation requires the formulation of a suitable interconnection between the points forming the actual surface model. For example, the logical extension of the planar multi-scale active contour model developed in this dissertation would be a $3D$ spline surface with continuous first-order and piecewise continuous second-order derivatives in two dimensions. A B-spline surface can be represented by two arc length parameters and three coordinate functions, i.e. as $\mathbf{v}(s,r) = (x(s,r), y(s,r), z(s,r))$. The B-spline patches are then represented as

$$\begin{aligned} x(s,r) &= \mathbf{S}\cdot\mathbf{M}_{Bs}\cdot\mathbf{G}_{Bs_x}\cdot\mathbf{M}_{Bs}^T\cdot\mathbf{R}^T \\ y(s,r) &= \mathbf{S}\cdot\mathbf{M}_{Bs}\cdot\mathbf{G}_{Bs_y}\cdot\mathbf{M}_{Bs}^T\cdot\mathbf{R}^T \\ z(s,r) &= \mathbf{S}\cdot\mathbf{M}_{Bs}\cdot\mathbf{G}_{Bs_z}\cdot\mathbf{M}_{Bs}^T\cdot\mathbf{R}^T \end{aligned} \tag{10.1}$$

in analogy to C^2 continuous two-dimensional B-spline curves [Foley *et al.*, 1990]. Here, the geometry matrix $\mathbf{G}$ consists of x, y, and z coefficients of 16 instead of 4 control points, $\mathbf{S}$ and $\mathbf{R}$ denote the two row vectors $\mathbf{S} = [s^3\ s^2\ s\ 1]$ and $\mathbf{R} = [r^3\ r^2\ r\ 1]$, and $\mathbf{M}_{Bs}$ is the standard B-spline basis matrix. The advantage of such a representation is given by the associated analytic internal smoothness constraints. The elasticity energy of the surface is computed by

$$\mathcal{E}_{elasticity}(\mathbf{v}(s,r)) = \int_0^1\int_0^1 (\mathbf{v}_s^2(s,r) + \mathbf{v}_r^2(s,r))\ \mathrm{d}s\ \mathrm{d}r \tag{10.2}$$

In [Cohen *et al.*, 1992] it was argued that the surface bending energy is analogously given by

$$\mathcal{E}_{bending}(\mathbf{v}(s,r)) = \int_0^1\int_0^1 (\mathbf{v}_{ss}^2(s,r) + 2\mathbf{v}_s(s,r)\mathbf{v}_r(s,r) + \mathbf{v}_{rr}^2(s,r))\ \mathrm{d}s\ \mathrm{d}r \tag{10.3}$$

For a $3D$ extension of the $2D$ curvature matching process developed in this dissertation, however, the computation of the actual surface curvature is more complex. Given the unit surface normal vector $\mathbf{n}(s,r) = \frac{v_s(s,r)\wedge v_r(s,r)}{\|v_s(s,r)\wedge v_r(s,r)\|}$ at any point of the surface, and the tangential plane defined by the vectors $\mathbf{v}_s(s,r)$ and $\mathbf{v}_r(s,r)$, the principal curvatures and directions are the eigenvalues and eigenvectors solving the general eigensystem resulting from the matrix expressed in terms of the first and second fundamental forms [Spivak, 1979]:

$$\frac{1}{\mathbf{E}\mathbf{G}-\mathbf{F}^2}\begin{pmatrix} \mathbf{E} & -\mathbf{F} \\ -\mathbf{F} & \mathbf{G} \end{pmatrix}\begin{pmatrix} \mathbf{e} & \mathbf{f} \\ \mathbf{f} & \mathbf{g} \end{pmatrix} \tag{10.4}$$

where $\mathbf{E} = (\mathbf{v}_s, \mathbf{v}_s)$, $\mathbf{F} = (\mathbf{v}_s, \mathbf{v}_r)$, $\mathbf{G} = (\mathbf{v}_r, \mathbf{v}_r)$, $\mathbf{e} = (\mathbf{n}, \mathbf{v}_{ss})$, $\mathbf{f} = (\mathbf{n}, \mathbf{v}_{sr})$, and $\mathbf{g} = (\mathbf{n}, \mathbf{v}_{rr})$. Note that when investigating only the intermediate neighbourhood of $\mathbf{v}(s,r)$ which is the origin of the Euclidean 3-space fixed by $(\mathbf{v}_s(s,r), \mathbf{v}_r(s,r), \mathbf{n})$, this expression simplifies, as the first fundamental form matrix in equation 10.4 becomes the identity, and the denominator can be dropped in the expression for the normal and the second fundamental form [Koenderink, 1990]. This yields the second-order part of the Taylor expansion for $\mathbf{v}(s,r)$, or the Hessian whose eigenvalues and eigenvectors constitute the principal curvatures and directions. An attempt for a suitable solution of how to base a curvature matching process on these measurements will be given below in the concept of volumetric image terms.

Another explicit surface representation is given by a parametric representation, e.g. in form of parametric surfaces [Székely *et al.*, 1996] or deformable superquadrics [Bardinet *et al.*, 1996a]. The former are based on a Fourier decomposition of the surface, where the advantages of a more compact representation with only few parameters are outweighed by the globality of the representation (also recall from chapter 2 that the preservation of topology may be endangered using Fourier truncation techniques). The latter are based on super-ellipsoids, which provide a very smooth surface representation and a high potential for global refinement fitting, but usually require a two-stage process, involving a prior coarse segmentation of the shape under investigation, followed by further approximation based on the super-ellipsoid. Moreover, superquadrics have so far only been used on data containing rather simple, high contrast shapes in cardiac imaging, making it hard to judge how well they would perform on more complex data like brain MRI. Assuming that a coarse parametric representation can be obtained at a very high level of image scale by some other process, however, refinement can presumably be extended to a parametric active shape focusing process. The disadvantage of using parametric models is that they are global techniques, which makes the extraction of local shape properties, e.g. as needed for a curvature matching process, not easily applicable. The main drawback of explicit surface representation techniques in general lies in their inability to accommodate for the simultaneous segmentation of multiple objects contained in an image, e.g. the inner and outer rim of the spinal cord. Additionally, a considerable theoretical as well as computational effort need to be performed in terms of accessing internal shape information (as addressed above), and solving newly imposed optimization problems, as will be discussed below.

Chapter 4 has presented several implicit surface representation schemes which are in general topologically more flexible in terms of being able to dynamically split into several objects, or to merge into a single one. This is achieved by their representation as iso-surfaces contained in the volumetric image, which are affected in their flow using a non-linear image diffusion framework. Their major drawback is that *a priori* knowledge cannot be easily accommodated for (only in terms of the expected object size and the associated relevant scale range), and that the convergence needs to be controlled by an additional external stopping criterion (as otherwise the objects in the scene are eventually completely smoothed out). Moreover, the lack of image independent intrinsic smoothness constraints deteriorates the performance of such an implicit approach for image degradations. This however dramatically restricts the number of possible applications of such an approach to rather simple shapes of closed shape outline, and the application to more complex problems in medical imaging is very limited [Niessen, 1997].

10.1.1.2 Volumetric Image Features

Given that the computation of internal constraints is either performed analytically, e.g. as discussed above for a spline surface, or implicitly, based on differential iso-surface properties, formulations of the volumetric image energy terms need to be made. Considering the volumetric embedding of an active surface model, the implementation of the $3D$ image gradient L_w and directional tuning is straightforward, as the surface normal is known and can be adjusted to the gradient direction given by ∇L (recall from chapter 4 that for an implicit formulation, the image gradient direction is equivalent to the iso-surface normal direction, and is used as an inflation force). In $2D$, the ridges of the gradient magnitude have been used in this dissertation. In $3D$, the ridge computation becomes more complex, as it is are based on the local extrema of the principal curvatures of $\|L_w\|$ (see chapter 3). Therefore, in order to find a suitable extension from $2D$ ridge lines to $3D$ ridge surfaces rather than lines, a choice needs to be made. For example, the fuzzy volumetric ridgeness operator presented in [Maintz, 1996] can be used. Originally, this ridgeness has been computed as a volumetric extension of the isophote image curvature $-\frac{L_{vv}}{L_w}$, which is based on the principal image curvatures and the corresponding principal directions. This concept can analogously be applied to derive a ridgeness measure of the volumetric gradient magnitude. However, higher order differentiation (e.g. third order differential measurements for the ridges of $\|L_w\|$) tends to be very unstable and sensitive to noise, and therefore it might be more appropriate and computationally more efficient to compute the zero-crossings of the $3D$ Laplacian ΔL of the image intensity, followed by a $3D$ Chamfer distance transform. In any rate, the computation of the $3D$ isophote image curvature of the image L, or the ridgeness of the image gradient magnitude L_w, can be generalized by computing the ridgeness of any image feature F by removing the superfluous third spatial dimension by rotation of the local coordinate system spanned by the x and y axes chosen perpendicular to the gradient direction. Formulating fuzzy ridge points as the second directional derivative of F in a direction a, where a lies in the local tangent plane, and also points into the direction where the second derivative is minimal, is then equivalent of finding the maximum concavity (or minimum convexity) by solving

$$\min \frac{1}{\|a\|^2}(s \cdot \nabla)^2 F \tag{10.5}$$

under the constraint $\nabla L \cdot \mathbf{a} = 0, \mathbf{a} \neq \vec{0}$ [Maintz, 1996]. If this system of equations is maximized, the principal curvatures and principal directions are obtained. Note that ridges obtained from this framework take the volumetric image information into account, and are surfaces rather than lines.

It was pointed out above that in order to perform a suitable curvature matching process, the volumetric definitions of both active surface curvature and isophote image curvature need to be considered. Two different solutions exist for this problem: Either the principal curvatures and directions are matched directly, e.g. by directional tuning of the principal surface curvature vectors with the underlying isophote image curvature vectors in $3D$, or a similar concept as for the $3D$ isophote image curvature is performed for the active surface curvature. In the first case, the local principal surface curvatures $\kappa_1(s,r)$ and $\kappa_2(s,r)$ and their corresponding directions $\mathbf{e}_1(s,r)$ and $\mathbf{e}_2(s,r)$ can be derived as the eigenvalues and eigenvectors of the local surface Hessian $\mathbf{H}(s,r)$. Similarly, the principal curvatures of the image intensity at the surface location $\mathbf{v}(s,r)$ in the image space L are given by $\kappa_{image_1}(s,r) = -\frac{L_{uu}(s,r)}{L_w(s,r)}$ and $\kappa_{image_2}(s,r) = -\frac{L_{vv}(s,r)}{L_w(s,r)}$ in a local $3D$ gauge (u,v,w), with w pointing into the gradient direction, and u and v denoting the orthogonal tangent directions depending on the gauge conditions $L_u = L_v = L_{uv} = 0$ [Florack, 1993] (see also chapter 3). The computation of the principal directions is more complex. Similar to the ridgeness computation above, however, equation 10.5 can be solved, yielding the principal directions $\mathbf{e}_{image_1}(s,r)$ and $\mathbf{e}_{image_2}(s,r)$ of $\kappa_{image_1}(s,r)$ and $\kappa_{image_2}(s,r)$ lying within the local tangent plane. Curvature matching can then be performed as directional tuning:

$$\begin{aligned} E_{bending}(\mathbf{v}(s,r)) &= (\kappa_{image_1}(s,r) \pm \kappa_1(s,r))^2 \left(\frac{\mathbf{e}_{image_1}(s,r)}{\|\mathbf{e}_{image_1}(s,r)\|} \cdot \frac{\pm\mathbf{e}_1(s,r)}{\|\mathbf{e}_1(s,r)\|} \right)^m + \\ & (\kappa_{image_2}(s,r) \pm \kappa_2(s,r))^2 \left(\frac{\mathbf{e}_{image_2}(s,r)}{\|\mathbf{e}_{image_2}(s,r)\|} \cdot \frac{\pm\mathbf{e}_2(s,r)}{\|\mathbf{e}_2(s,r)\|} \right)^m \end{aligned} \tag{10.6}$$

where m is used for broadening or narrowing the tuning operator, and is typically chosen as $m = 2$. The choice of the sign depends on the chosen normal direction. If the second approach is adopted, a $3D$ surface curvature in form of a singular value has to be computed in analogy to the $3D$ image ridgeness measure, by replacing in equation 10.5 the image feature F by the local surface definition $\mathbf{v}(s,r)$. Denoting the thus obtained values for the active surface and the image by $\kappa(s,r)$ and $\kappa_{image}(s,r)$, curvature matching can be performed similarly to the $2D$ equivalent used for the multi-scale planar active contour model in this work (compare to equation 6.16, page 130):

$$E_{bending}(\mathbf{v}(s,r)) = (\kappa(s,r) \pm \kappa_{image}(s,r))^2 \tag{10.7}$$

The curvature matching as well as autonomous shape constraints need only be performed for explicit model representations, as they are inherent in the internal ones. Volumetric extensions of the image forces, however, need to be used for either representation.

10.1.1.3 Volumetric Optimization

Having decided on a suitable volumetric energy functional, the question of performing the actual optimization needs to be addressed. For an explicit representation, there are two possible approaches: The first, computationally more efficient and easier to be implemented one is a so-called slice-by-slice model, where the points of the surface model are constrained to associated image slices [Cohen *et al.*, 1992]. Optimization is then carried out within the respective image planes, but under consideration of $3D$ image and internal surface constraint forces. The second approach allows the spline surface points to move freely in the z direction as well, and is therefore probably much more useful in coping with topological changes, and partial disappearance and reappearance of structures as mentioned above. It is a true volumetric optimization method, rather than a sequential, slice-by-slice optimization, and therefore of much higher computational complexity. In either case, the *greedy* algorithm used in this thesis can be extended to optimize the continuous spline surface by allowing to move the surface control points in the planar neighbourhoods of their respective image slices, or in volumetric neighbourhoods, depending on the chosen deformation approach. However, finite element methods (FEM) as used in [Cohen and Cohen, 1993] might prove to be computationally more efficient.

Finally, in a scale-space setting, multi-scale sampling needs to be formulated in $3D$ in order to locally change the resolution of the volumetric model, which is computationally very complex. For example, if the model is based on an approximating spline surface, a normalized surface representation $\tilde{\mathbf{v}}(s, r)$ is required.

10.1.1.4 Surface Description

Multi-scale active surface description based on the topics formulated above involves an active surface focusing process of true $3D$ dimensionality (between $2\frac{3}{4}$ and $3D$ if the model is still constrained to the image planes, but affected by volumetric internal surface constraints) in a $4D$ image scale-space and associated volumetric scale-space image features. Active surface focusing gives rise to a true $3\frac{1}{2}D$ multi-scale shape stack, which is based on instances of a volumetric surface $\tilde{\mathbf{v}}(s, r; \sigma)$ in image scale-space. Description of the volumetric multi-scale shape stack involves the comparison of each stack level to the lowest scale *reference surface* in terms of suitable $3D$ distance metrics. The problem of finding volumetric point correspondences needs to be addressed in order to correctly register the surface models in scale-space. Surface features like the principal curvatures and directions (naturally leading to geometric ridge and crease surface structure) can be examined, and used for multi-scale registration. Examination of the surface flow in scale-space can therefore be applied for a complete volumetric shape description. Related work has been done in [Fidrich, 1997], where feature lines of iso-surfaces in scale-space have been directly extracted, allowing for the visualization of singular (umbilic) surface points as local principal curvature extrema, parabolic lines, crest lines and surfaces, ridges, and ravines. However, in many cases (and in particular in brain MRI), iso-surfaces cannot be directly extracted, and therefore an explicit surface segmentation in scale-space, like multi-scale active surface models, becomes necessary.

In the following, the application of non-linear image scale-space is addressed, which are expected to avoid the premature fusion of shapes due to their edge and other image feature preserving nature.

10.1.2 Extension to Non-Linear Scale-Spaces

The choice of the underlying image scale-spaces plays an important role [Niessen *et al.*, 1996], as they vary from shape shrinking [Florack *et al.*, 1994] to shape preserving. The most well-known scale-space generation schemes have been presented in chapter 3, but their appropriateness as an underlying image representation for a multi-resolution shape analysis needs to be further investigated. In particular, the efficient implementation of scale-space generators needs to be considered. In contrast to the linear Gaussian scale-space, non-linear scale-spaces need to be implemented by a finite differencing scheme. Recent advances provide efficient implicit schemes for scale-space computation which allows larger scale steps to be taken without sacrificing robustness of the underlying diffusion process. However, so far only linear diffusion and edge-affected diffusion by Perona and Malik [Perona and Malik, 1990] can be computed implicitly. Diffusion schemes which additionally incorporate the L_{vv} operator (e.g. the Euclidean shortening flow) cannot be extended to $3D$ in a straightforward way. This has been already discussed in the previous section in terms of the $3D$ extension of the $-\frac{L_{vv}}{L_w}$ operator, and similar solutions can be found.

In order to show the potential applicability of non-linear scale-spaces to multi-scale active shape description, figure 10.1 illustrates the final active shape focusing result using a $3\frac{1}{2}D$ scale-space based on the Euclidean Shortening flow for the brain of the epileptic data for patient 1 (for the linear results see chapter 9, section 9.3). Visually the difference to the linear scale-space (see figure 9.3) is small, but the measurements across scale listed in table 10.1 show a much larger mean volume across scale than for the linear case (see table 9.4), which is also more preserved during the focusing process. This can be perceived by the very small volume increase, and generally smaller Chamfer and triangulation distance deviations. Additionally, the surface area increases only with less than half the slope for decreasing scales compared to the linear case.

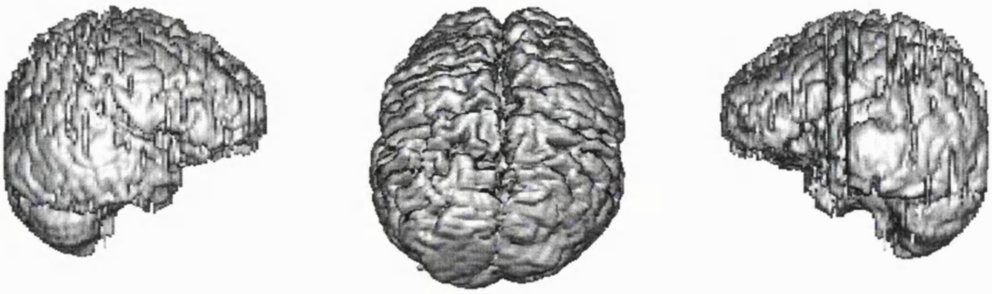


Figure 10.1: Right, top, and left views of the iso-surface rendered lowest scale results of the volumetric brain image of the epileptic data for patient 1 obtained via active shape focusing in a $3\frac{1}{2}D$ Euclidean shortening flow image scale-space. Active shape focusing was performed with $t_{max} = 32$, $t_{min} = 0.5$, $n = 16$ samples, $\Delta\tau = 0.1$, approximately corresponding to $\sigma_{max} = 8$ and $\sigma_{min} = 1$.

Figure 10.2 shows maximum intensity projections of the linear and several non-linear $3D$ scale-spaces of the intermediate slice of patient 1 (see also figure 9.1, page 197), demonstrating the visual difference in shape flow. Figure 10.3 illustrates the multi-scale shape stacks obtained from the scale-spaces in figure 10.2, where the same circular shape as in chapter 9 was used as an initial model. In fact, figure 10.2 (a) illustrates the rendered $3D$ linear scale-space shown in figure 9.1 for patient 1, and, figure 10.3 shows the corresponding surface rendered shape stack (colour rendering results of this particular linear stack can be found in appendix A, figure A.7). A preliminary conclusion of these visualizations is that from the non-linear schemes, the affine shortening flow and the Euclidean shortening flow are probably more suitable for active shape focusing than the other three schemes, which tend to preserve smaller structures over a larger set of scales. This observation leads to another aspect involved with non-linear diffusion schemes which is still an open topic in scale-space research - the investigation of appropriate scale-space sampling strategies. For the linear case, an exponential sampling scheme (as used in this work) is thought to be optimal, but the choice for non-linear schemes is more complex. In conjunction with this question is the variation of control parameters, such as regularization level, and edge threshold - the question referred to as the *scale-space recipe* by Whitaker [Whitaker, 1994b], as well as the correspondence between these parameters and diffusion times for the different schemes. This plays a particularly important role for the comparison of non-linear scale-spaces in terms of implicit segmentation abilities in a volumetric active shape focusing process. In order to judge the appropriateness of the non-linear schemes, resulting shapes need to be compared at adequate levels of image scale.

Descriptor	Value
$\overline{V}$	1403150
$\overline{S}$	57259.4
$\overline{C_V}$	9.773620
$\overline{\text{dist}_H}$	62.216900
$\overline{\text{dist}_{C_{RMS}}}$	1.745250
$\overline{\text{dist}_{T_{RMS}}}$	4.982440
$\frac{\Delta V}{\Delta\sigma}$	-9.257280
$\frac{\Delta S}{\Delta\sigma}$	-2.940210
$\frac{\Delta C_V}{\Delta\sigma}$	-0.000683
$\frac{\Delta \text{dist}_{C_{RMS}}}{\Delta\sigma}$	0.000262
$\frac{\Delta \text{dist}_{T_{RMS}}}{\Delta\sigma}$	0.001443
$-\frac{\Delta \log(V/\sigma)}{\Delta \log(\sigma)}$	1.116470
$-\frac{\Delta \log(S/\sigma)}{\Delta \log(\sigma)}$	1.163590
$-\frac{\Delta \log(C_V/\sigma)}{\Delta \log(\sigma)}$	1.137550

Table 10.1: Mean, slope, and logarithmic slope descriptors for active shape focusing based on the Euclidean shortening flow of the epileptic data for patient 1. Compare with values for patient 1 in table 9.4 on page 208 obtained from linear scale-spaces.

10.1.3 Comparison to Multi-Scale Contour Analysis

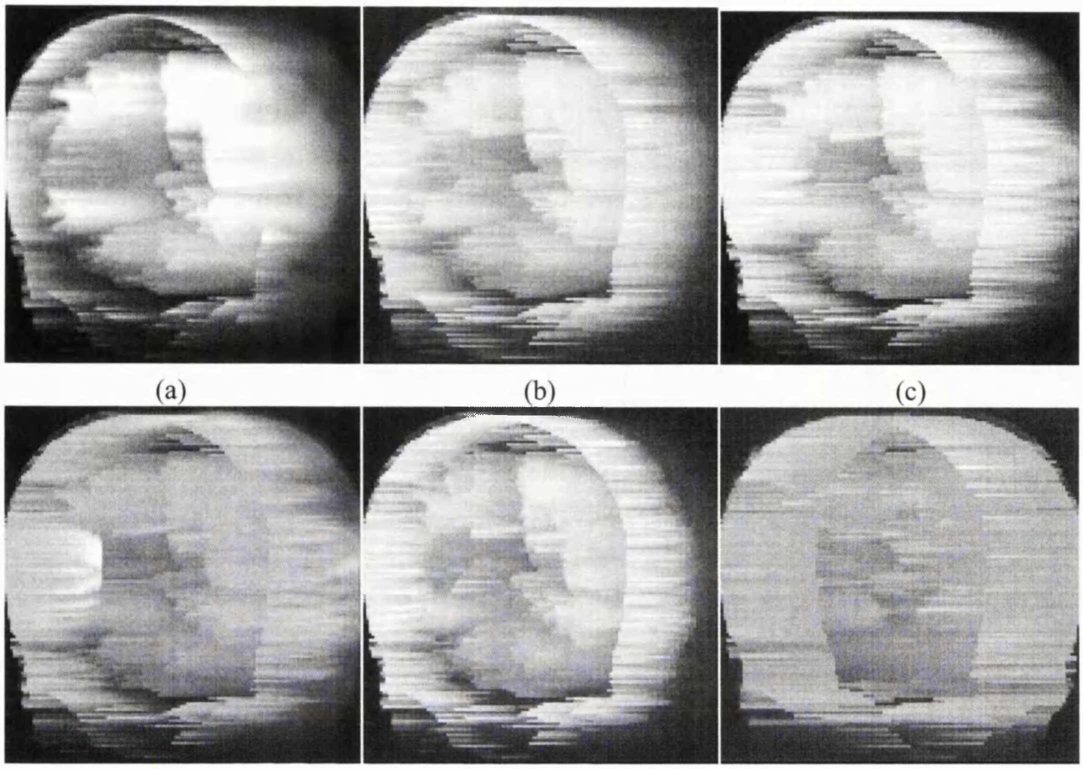
It was stated earlier that active shape evolution and its dual, active shape focusing, differ from classic contour evolution schemes in terms of the scale-space dimensionality, as they are embedded in an image scale-space rather than being based on a contour scale-space. Several issues regarding differences of the results can be further investigated, like differences in resulting *fingerprints* [Witkin, 1983] and *curvature scale-spaces* [Mokhtarian and Mackworth, 1986] (see chapter 3), as well as comparisons to theoretical expressions for the behaviour of area and perimeter (and hence compactness) with respect to scale. In particular, it was shown by [Gage and Hamilton, 1986] that the heat equation for a planar curve can be written as the system

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial s^2} \quad \text{and} \quad \frac{\partial y}{\partial t} = \frac{\partial^2 y}{\partial s^2} \tag{10.8}$$

from which the shrinkage of a planar curve due to the heat equation can be quantified in terms of the change of perimeter or length L and the enclosed area A:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -\int_0^L \kappa^2(s)\mathrm{d}(s) \quad \text{and} \quad \frac{\mathrm{d}A}{\mathrm{d}t} = \int_{x,y\in\Re^2} w(x,y)\mathrm{d}x\,\mathrm{d}y \tag{10.9}$$

where $w(x, y)$ is the winding number of the curve with respect to the point (x, y), and the rate of decrease in area is -2π times the rotation index of the curve. In general, a curve evolution embedded in an image scale-space will lead to a lower rate of shrinkage of the level set curves, than



(a) (b) (c)

(d) (e) (f)

Figure 10.2: Maximum intensity projections of different $3D$ linear and non-linear scale-spaces for patient 1 in a *fine-to-coarse* view. (a) Linear. (b) Affine shortening flow. (c) Euclidean shortening flow. (d) Entropy (reaction-diffusion). (e) Modified affine shortening flow. (f) Edge-affected (Perona and Malik). All scale spaces have been computed between $t_{min} = 0.5$ and $t_{max} = 32$, $n = 16$ samples, $\Delta\tau = 0.1$ (excepting for the linear and edge-affected diffusion, where $\Delta\tau = 0.25$).

a classic curve evolution scheme. Therefore active shape evolution and focusing can be regarded as non-linear curve evolution processes, with the advantage of the latter that it requires no ground truth.

10.1.4 Multi-Scale Shape-Driven Techniques

Directly arising from the multi-scale active contour model presented in this dissertation are two more potential extensions. The first one is based on the linear diffusion process which can be used to steer the deformation of the model between adjacent image scale levels. It is referred to as *active shape flow* in the following. The second topic is based on a true multi-scale approach, enabling the multi-scale active contour model to evolve freely in image scale-space, rather than tracking it through the scale-space in a slice-by-slice fashion. The result is a space curve, which forms the basis of an *active shape diffusion* process by reformatting the image scale-space with respect to the multi-scale contour model, and provides a hierarchical tool for image-based shape interpretation.

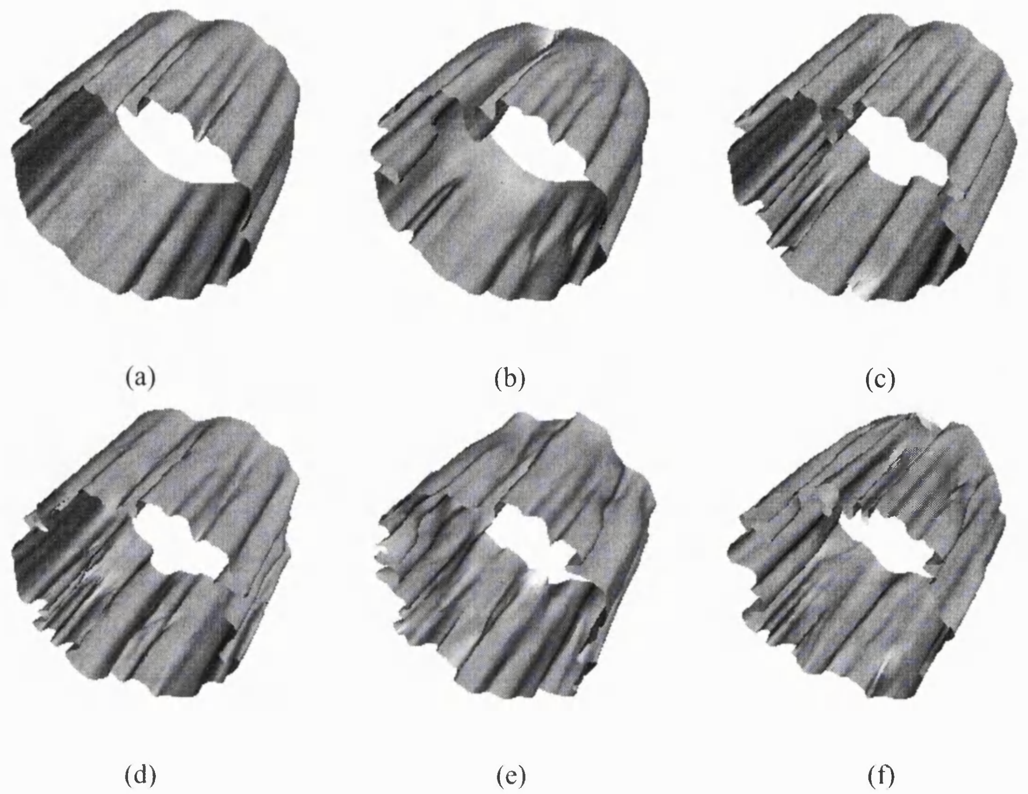


Figure 10.3: Interpolated shaded surface renderings of the hierarchical shape stacks for patient 1 of different $3D$ non-linear scale-spaces. (a) Linear. (b) Affine shortening flow. (c) Euclidean Shortening flow. (d) Entropy (reaction-diffusion). (e) Modified affine shortening flow. (f) Edge-affected (Perona and Malik). All scale spaces have been computed between $t_{min} = 2$ and $t_{max} = 128$, using $n = 16$ samples.

10.1.4.1 Active Shape Flow

When constructing a linear image scale-space, directional information of the intensity flow in scale-space can be obtained. More specifically, if the image at evolution time t_0 is known and a diffusion step of $\Delta\tau$ is performed, one has two images $L(\mathbf{x}; t_0)$ and $L(\mathbf{x}; t_0 + \Delta\tau)$ which are adjacent in scale-space and therefore differ only slightly. The difference between these two images in $2D$ is given by the *diffusion feature*. This means that at any point $\mathbf{x}$ in both blurred images, the local difference is given by the differences of the local intensity changes derived from the local weighted Laplacian of the less blurred image. The diffusion feature can then be formulated as velocities in x and y directions as

$$\begin{aligned} v_x(L_{i,j}) &= |L_{i,j} - L_{i-1,j}| - |L_{i,j} - L_{i+1,j}| \\ v_y(L_{i,j}) &= |L_{i,j} - L_{i,j-1}| - |L_{i,j} - L_{i,j+1}| \end{aligned} \quad (10.10)$$

where sub-indices denote the spatial location of the image pixels, and L denotes the intensity at that position. Then the velocity vector $\mathbf{f}(L_{i,j}; t_0) = (v_x(L_{i,j}), v_y(L_{i,j}))$ provides the local

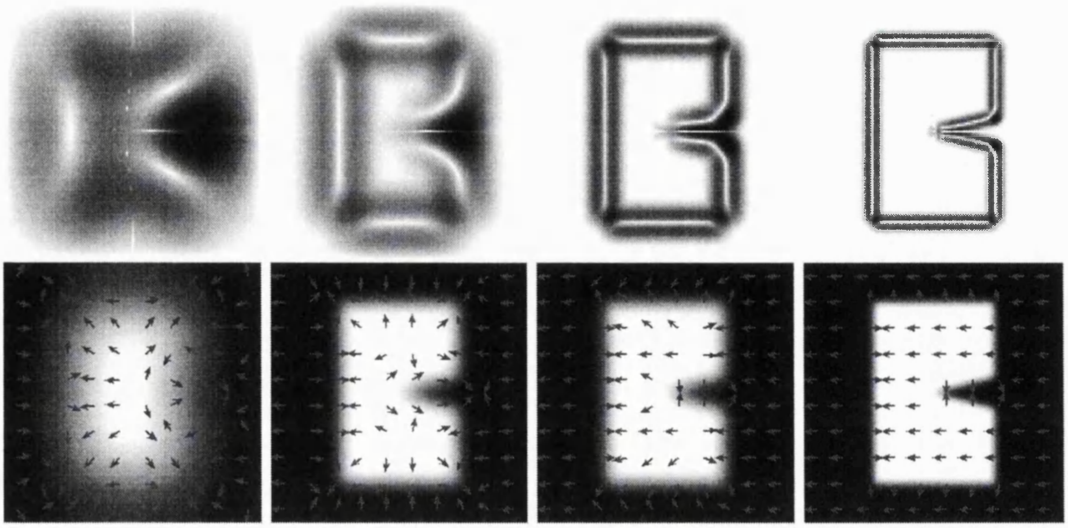


Figure 10.4: Linear diffusion magnitude (top) and direction (bottom) for the *notched rectangle* test image at different levels of scale. From left to right: $t = 32, 8, 2, 0.5$, with $\Delta\tau = 0.25$.

diffusion magnitude and *diffusion direction* as the absolute value and angle. In other words, given an image $L(\mathbf{x}; t_0)$, it is known in which direction and with which magnitude the image locally diffuses.

This knowledge can be applied as a data dependent, self-regularized local inflation or deflation force of a multi-scale active contour model, as it can locally push the model towards the direction of the flow of the underlying image feature. The diffusion can be formulated as an additional term in the energy function:

$$E_{flow}(\mathbf{v}(s); t_0) = \pm\, \alpha_{flow}\, \mathbf{f}(\mathbf{v}(s); t_0)^2 \left(\frac{\mathbf{f}(\mathbf{v}(s); t_0)}{\|\mathbf{f}(\mathbf{v}(s); t_0)\|} \cdot \mathbf{n}(s) \right)^m \tag{10.11}$$

where α_{flow} is a constant weighting parameter, m is used for broadening or narrowing the influence of this force, $\mathbf{n}(s)$ is the unit normal vector of the contour model, and the sign of the whole expression is chosen with respect to the chosen normal direction. This is equivalent to directional tuning of the model with respect to the diffusion flow. One should note, however, that in contrast to the classic balloon model presented in chapter 4, this flow force can simultaneously inflate or deflate the model at different parts, and is additionally using adaptive data-dependent diffusion flow magnitudes. This technique becomes particularly interesting in terms of tracking a multi-scale active contour model through image scale-space.

Figure 10.4 shows the linear diffusion magnitudes and directions for the *notched rectangle* test images at four different evolution times. In contrast to edge potentials (see figure 7.3, page 157), a strong image force based on the diffusion magnitude is achieved at high scales within the notch, as well as at corners. These are the positions which are most diffused, and therefore are usually the hardest to track in a scale-based scheme. A scheme only based on the flow of diffusion magnitude and direction, however, is only well defined in synthetic binary images, and even then will suffer from the problem that in homogeneous parts of an image the diffusion magnitude will be zero, and use of the direction is therefore not applicable. Integrating the energy force of equation 10.11 into the energy function used in this dissertation, and choosing α_{flow} adequately, e.g. $\alpha_{flow} = 1$, allows to push the model in homogeneous areas towards edges and adjust it to the underlying image curvature, while simultaneously improving its elastic behaviour. At more heterogeneous parts, i.e. near edges where diffusion has a stronger effect, directed attraction potential is provided.

Most recently, some related concepts have been introduced which will be briefly listed in the following. The *edge flow* scheme performs boundary detection and segmentation by predicting the direction of change in colour and texture at each image location at a given fixed scale [Ma and Manjunath, 1997]. Another approach is the fixed-scale *gradient vector flow* as an external force for snakes [Xu and Prince, 1997] which is computed as a diffusion of the image gradient vectors, and is therefore similar to the classic optic flow [Horn and Schunk, 1981].

10.1.4.2 Active Shape Diffusion

Until now, the levels of the underlying image scale-space have been investigated separately in a *fine-to-coarse* or *coarse-to-fine* manner, yielding a multi-scale stack of shapes, each with an adjusted inner scale. As mentioned in chapter 6, however, it might be more desirable to adjust the contour scale locally to the required level of detail. It was argued that at parts of low curvature, only a sparse sampling is needed, while at corners and complex structures, a high level of contour detail is needed. Obviously, contour and image detail should still be related, which gives rise to a space curve in scale-space with varying contour scale and local adjustment to the respective image scale level the contour passes through.

In order to develop such a technique, however, several topics need to be further investigated, namely a spline-based space curve representation, variable sampling strategies, scale-space geometry, and the effect of optimization in image scale-space. In order to formulate a multi-scale contour representation, an extra coordinate function is added to the previously planar B-spline representation, which does not refer to a spatial location, but to a natural scale location $\tilde{\sigma}(s)$:

$$\tilde{\mathbf{v}}(s) = (\tilde{x}(s), \tilde{y}(s), \tilde{\sigma}(s)) \qquad (10.12)$$

Note that the natural representation does not refer to a fixed-scale setting, but to locally variable scales. This implies that at each point of the multi-scale contour, the contour scale becomes variable as well, i.e. it is defined as $\tilde{\varsigma}(s)$. In chapter 6, three variable, but fixed-scale sampling strategies based on internal and external curvature properties have been presented, which were based on the adaptive sampling algorithm developed in this work. Similarly, adaptive variable sampling can use the underlying image scale as a distance metric. This leads naturally to the topic of scale-space geometry. The planar model in this dissertation is embedded in a fixed-scale Euclidean setting. Hence the distance between two points of the contour is given by the Euclidean distance of their natural representation. For a multi-scale setting, however, the scale-space geometry becomes Riemannian. Therefore the distance of two points on the space curve needs to be computed along their geodesic paths, and the criterion of whether a points needs to be removed or inserted is based on this Riemannian scale-space distance measurement. The actual insertion of a new point not only involves the interpolation of the spatial coordinates, but also the interpolation of the scale coordinate. The B-spline representation accommodates for that, as the natural representation ensures linear sampling in the natural scale direction, although the scale is actually exponentially sampled. Evolving a contour directly in scale-space rather than in a slice-by-slice fashion imposes the need of appropriate differentiation of the image potentials, as the scale dimension needs to be considered in the differentiation process. Moreover, the curvature of an arc length parameterized space curve is of slightly extended form [Bronstein and Semendjajew, 1989], with

$$\kappa(s) = \frac{((x_s^2 + y_s^2 + z_s^2)(x_{ss}^2 + y_{ss}^2 + z_{ss}^2) - (x_s x_{ss} + y_s y_{ss} + y_s y_{ss})^2)^{\frac{1}{2}}}{(x_s^2 + y_s^2 + z_s^2)^{\frac{3}{2}}} \tag{10.13}$$

and it is also affected by a torsion term:

$$\tau(s) = \frac{x_s(y_{ss}z_{sss} - z_{ss}y_{sss}) - y_s(x_{ss}z_{sss} - z_{ss}x_{sss}) + z_s(x_{ss}y_{sss} - y_{ss}x_{sss})}{x_s^2 + y_s^2 + z_s^2} \tag{10.14}$$

where expressions for arbitrary parameterizations exist. Additionally the principal normal vector $\mathbf{n}(s)$ is defined to be in the plane of the tangent vector $\mathbf{t}(s)$ which points toward the concave side of the curve, and a binormal vector is defined to be orthogonal to both. These three vectors form the so-called *Frenet frame*, with the inter-relationship $\mathbf{b}(s) = \mathbf{t}(s) \wedge \mathbf{n}(s)$. Torsion and curvature are then related by the *Frenet equations*:

$$\frac{\mathrm{d}\mathbf{t}(s)}{\mathrm{d}s} = \kappa(s)\mathbf{n}(s) \qquad \frac{\mathrm{d}\mathbf{n}(s)}{\mathrm{d}s} = -\kappa(s)\mathbf{t}(s) + \tau(s)\mathbf{b}(s) \qquad \frac{\mathrm{d}\mathbf{b}(s)}{\mathrm{d}s} = -\tau(s)\mathbf{b}(s) \tag{10.15}$$

Most recently, in [Mokhtarian, 1997b] a theoretical framework for a torsion scale-space (TSS) has been formulated, which is based on multi-scale space curves and provides an extension of the curvature scale-space presented in chapter 3 from planar curves to space curves. The main difference between the proposed concept above and the TSS is that the multi-scale contour representation above is still spatially planar, but becomes a space curve with respect to the scale of the image scale-space in which it is embedded, while for the TSS, the spatial contour representation is a true space curve, with an extra, but fixed-scale dimension.

For a basic demonstration of the concept of this true multi-scale model, linear scale-spaces for the *notched rectangle* and *teardrop* test images, whose active shape evolution and focusing results have been presented and discussed in chapter 7 and 8, respectively, have been computed. A

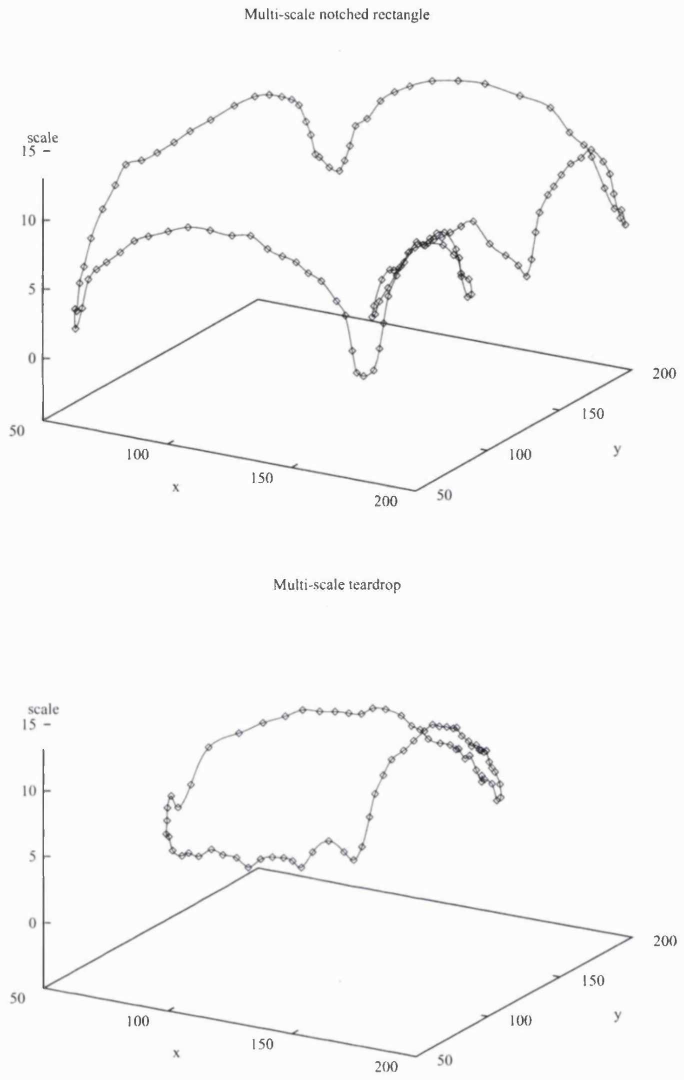


Figure 10.5: Multi-scale contour optimization results for the *notched rectangle* and *teardrop* test images.

coarse-to-fine approach, using the same ellipse-shaped initial models as for the previous active shape focusing with additional natural scale coordinates, was performed, by setting all natural scale coordinates to the highest natural scale level (i.e. to $i = n - 1$). The equivalent *fine-to-coarse* approach would be based on using the *ground truth* models, and setting the the scale coordinate to the lowest natural scale level (i.e. to $i = 0$). Extending the optimization strategy, for instance the presented multi-scale *greedy* algorithm, from a local planar search space to a local volumetric search space, allows to deform the multi-scale model spatially as well as with respect to natural scale. For a first simple approach, only the gradient magnitude (without directional tuning), and the elasticity given by $\mathbf{v}_s(s)$ have been taken into account.

Figure 10.5 shows the final multi-scale contours obtained for both test images which yielded good extraction results with respect to all global *and* local shape details, plotted against the local scale $\sigma(s)$ obtained from the natural scale coordinates $\tilde{\sigma}(s)$. For the *notched rectangle*, several observations can be made: At extended straight parts of the boundary, the local scale $\sigma(s)$ and associated contour scale $\varsigma(s)$ is rather high. At all corners, both scale measurements become rather low, and continuously increase in a smooth manner towards the straight parts. The peak of the notch is of lowest scale (see right hand side of the plot). Note that if all scale coordinates were set to $\tilde{\sigma}(s) = 0$, the accurate planar shape outline of the notched rectangle would be obtained. For the *teardrop*, the outward peak (towards the back of the plot) is also recovered at a very low scale level, while the straight sides of the teardrop are of overall highest scale, starting from the centre of the blob like end, followed by an intermediate scale level for the circular end (at the front of the plot).

The local scales obtained from the multi-scale contour optimization process therefore can be regarded as the *adequate scales*, as they correspond to the maxima of the underlying *scale-space signatures* of the gradient magnitude (only slightly regularized by the elasticity term). The resulting contour for both cases is equivalent to the *boundary at the scale of the core* (BASOC), as reviewed in chapter 3, but with the suggested scheme they are much easier to obtain, and are continuous by nature. This gives rise not only to a true multi-scale shape representation, but also to the investigation of the locally adequate scales as local shape descriptors. Note also that a multi-scale contour optimization results into a single contour rather than into a whole stack, and can be compared with slicing a multi-scale shape stack into a single meaningful plane (which is only planar though with respect to the spatial shape locations). This leads to another important aspect: Having obtained the most meaningful scales, or the most meaningful variable scale slice for a shape stack, this knowledge can be back-projected into the image. Similarly to the backprojection of a core to the boundary that contributed most to it, every image pixel can be investigated with respect to the multi-scale shape boundary, by reducing the image scale-space to the most meaningful scales with respect to the located shape. Naturally, direct knowledge of the adequate scales is only available at the image pixels underlying the shape boundary, but interpolation can be used to extend this knowledge over the whole image scale-space. A suitable interpolation is given by thin-plate splines (see chapter 2 for the theoretical background on this topic). Defining a thin-plate spline mapping function on the basis of the contour control points allows to model the global displacement field with respect to the displacement of the natural scale coordinates of the shape during the optimization process. Having obtained the displacement for each image pixel, the image scale-space can be sliced through these natural scale coordinates, yielding an *active shape diffusion* result. Figure 10.6 shows the resulting, sliced image scale-spaces for the *notched rectangle* and *teardrop* images, visualizing the interpolated scale levels with respect to the multi-

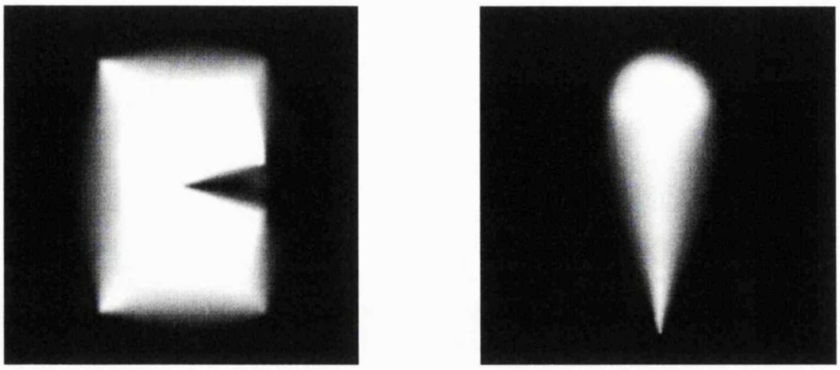


Figure 10.6: Active shape diffusion obtained from back-projection of the multi-scale contours illustrated in figure 10.5 into the image scale-space using thin-plate spline interpolation.

scale contour optimization processes.

This concept, though quite attractive at first sight, needs to be further investigated in terms of adjusting the internal and external image energy terms with respect to space curve geometry and the Riemannian geometry of the image scale-space. Also it needs to be appropriately compared with the concept of *cores* and the *boundary at the scale of the core*. The *active shape diffusion* process as described above then allows to investigate images with respect to the shapes that they contain, yielding additional valuable information about local object size and changes thereof. This scheme can be the starting point to true multi-scale image analysis, providing the extraction of locally adequate and meaningful shape properties.

10.2 Conclusions

In this dissertation a novel scale-based framework for automated shape description with an application to medical imaging has been developed. The primary purpose of this approach was to provide higher-level shape information than currently available, and to incorporate the shape description step into an implicit shape extraction process. In this way, and in contrast to traditional shape description methods in medical imaging, the developed approach does not rely on a prior (mostly manual) shape segmentation, but instead works directly on the grey-level MR images. Implicit shape extraction, however, is performed at multiple levels of image detail in order to extract shape information in a concise manner, capturing shape characteristics of varying locality without the loss of image context. Rather than increasing the amount of available shape information due to the extra scale dimension, shape characteristics have been structured with respect to their behaviour across image scales, from which novel multi-scale and fractal shape metrics have been derived.

Several important methodologies have been developed for this purpose:

1. A multi-scale active contour model, which is based on a multi-scale spline representation and differential invariants in scale-space.

2. Active shape evolution and focusing, which have been identified as dual techniques for tracking a shape in opposite directions of a linear image scale-space.

3. The concept of a multi-scale shape stack, which organizes the shapes obtained from either multi-scale tracking process into a shape hierarchy, and which can be visualized in order to inspect local shape information.

The first item not only yields a scale-based shape extraction tool, but also combines several important topics in multi-scale image processing, like the relation between scale-space density and contour resolution, multi-scale spline sampling strategies, as well as the adjustment of shapes to the underlying geometric image structure like isophote curvature properties. The second item is based on the application of the multi-scale active contour model to a *fine-to-coarse* or *coarse-to-fine* image scale-space. The former is similar to classic multi-scale contour evolution, as it yields subsequently blurred versions of the known *ground truth* of a shape. It differs from classic evolution in that it uses a higher order scale-space dimensionality, i.e. the fuzzy image-scale space in which the shape is embedded rather than the binary contour scale-space. The latter is similar to edge focusing in that it reverses the blurring process in order to derive a shape close to the (possibly unknown) *ground truth*. In contrast to edge focusing, however, active shape focusing comprises the extraction of all intermediate scale results, rather than discarding all but the result obtained at the final lowest scale. The third item directly arises from applying the multi-scale active contour model to active shape evolution or focusing, and is similar to classic scale-based shape sketches like scale-space fingerprints or the curvature scale-space in that it provides a structural organization and useful visual representation of the obtained shape hierarchy in scale-space. All items together form the theoretical framework for multi-scale active shape description as the main contribution of this dissertation.

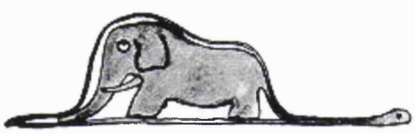
This dissertation also serves as a preliminary study of the applicability of the developed methodologies in clinical neurology for the description of complex structures like the human brain. Three different types of medical problems have been investigated, to which currently exist no standard solutions in the clinical environment. The application to patients with epilepsy has shown promising results in quantifying structural characteristics in the cortex pattern which may prove to be useful for distinguishing between abnormal patterns and normal shape variability in a larger clinical study. The application to patients with multiple sclerosis has additionally demonstrated a novel technique for automatic scale-based initialization. Finally, the application to neonates has illustrated the difference between visual shape complexity and shape compactness and fractal similarity measurements. The main clinical contributions of this work therefore lie in providing clinicians with a high-level framework for shape regularization, visualization and analysis with respect to scale, which automates the otherwise time-consuming task of shape extraction as much as possible, while avoiding inter- and intra-observer variability.

The use of scale continuity, the automated nature, and the high level of shape abstraction make the presented framework for multi-scale active shape description suitable for the application to larger scale clinical studies. It is hoped that this framework, along with the outlined future extensions, will provide a useful platform for shape analysis not only in clinical neurology, but also in other medical applications where shape deformations are observed and need to be analysed.

– LES HOMMES ONT OUBLIÉ CETTE VÉRITÉ, DIT LE RENARD. MAIS TU NE DOIS PAS L'OUBLIER. TU DEVIENS RESPONSABLE POUR TOUJOURS DE CE QUE TU A APPRIVOISÉ. TU ES RESPONSABLE DE TA ROSE...

"MEN HAVE FORGOTTEN THIS TRUTH," SAID THE FOX. "BUT YOU MUST NOT FORGET IT. YOU BECOME RESPONSIBLE, FOREVER, FOR WHAT YOU HAVE TAMED. YOU ARE RESPONSIBLE FOR YOUR ROSE..."

Le Petit Prince, Antoine de Saint-Exupéry.

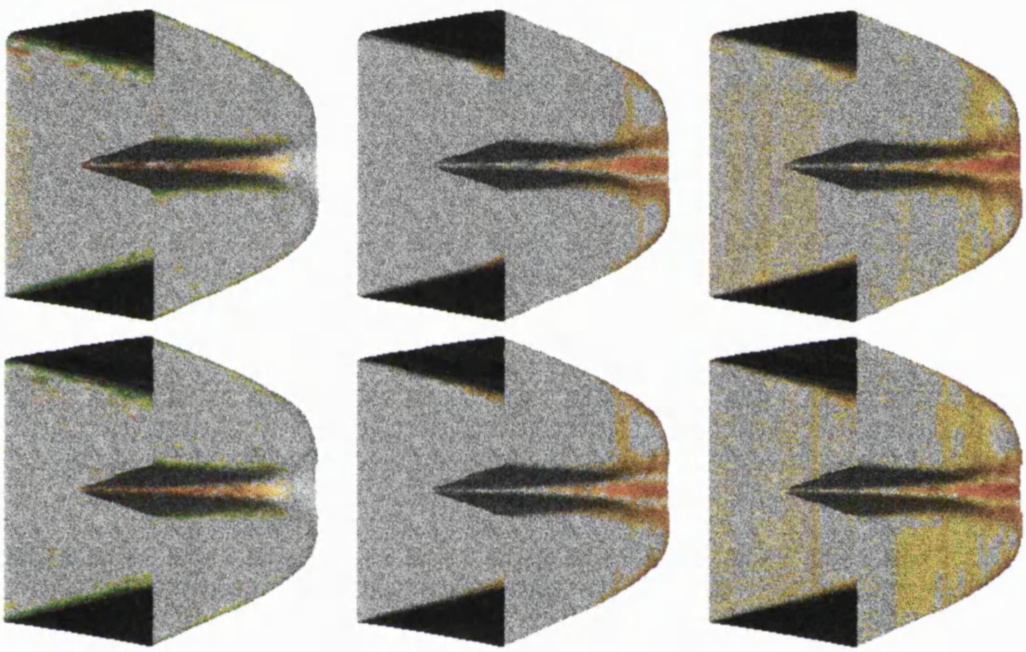


DESSIN N^o 2. – DRAWING NO. 2.

Le Petit Prince, Antoine de Saint-Exupéry.

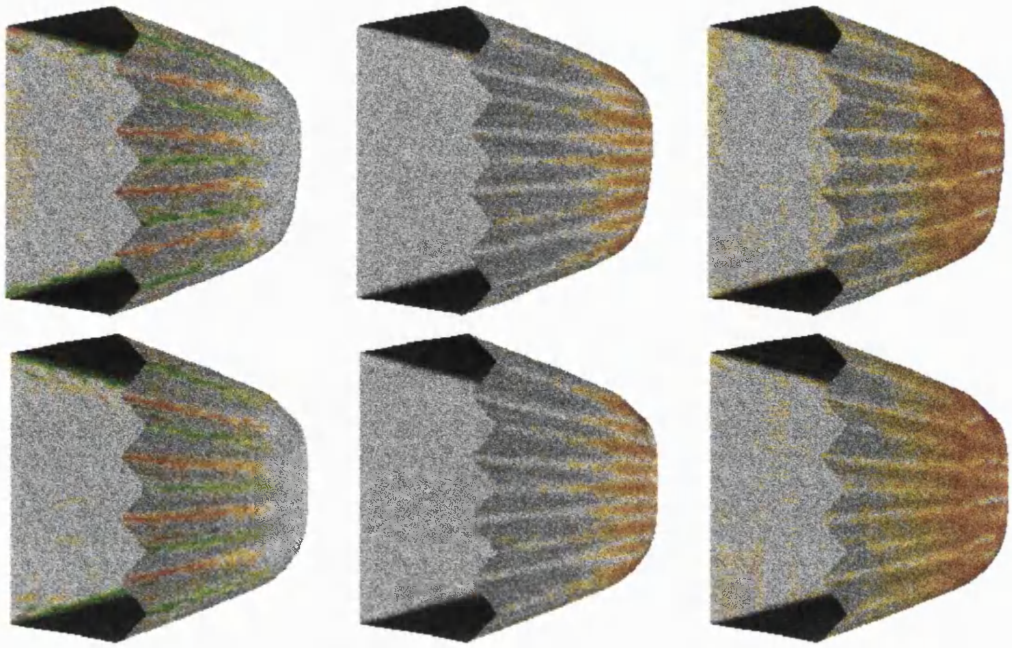
Appendix A

Colour Plates



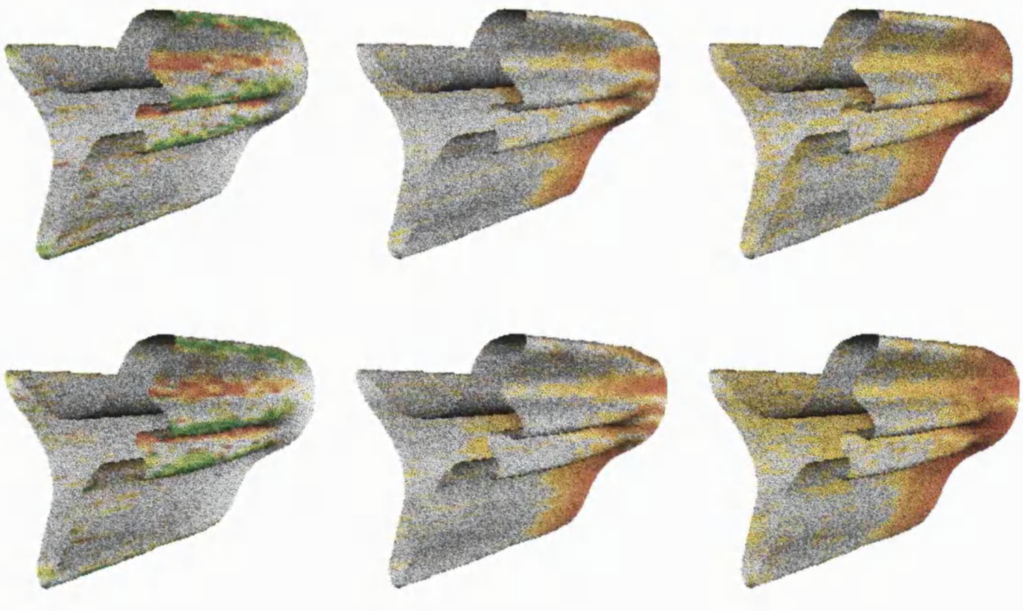
(a) (b) (c)

Figure A.1: Colour mapping of local shape descriptors onto the shape stack obtained via active shape evolution (upper row) and focusing (lower row) of the *notched rectangle* image. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



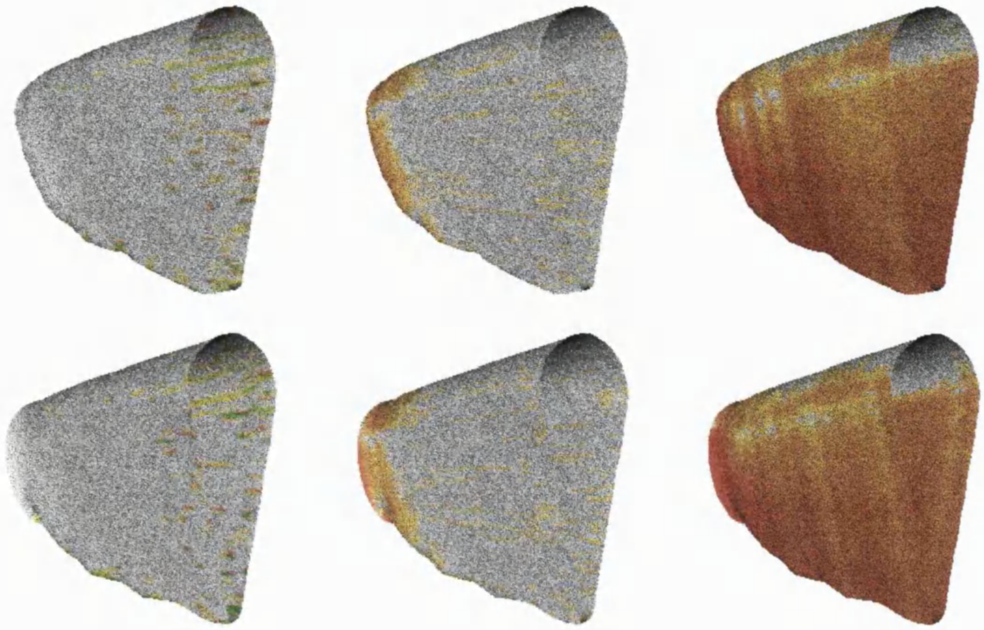
(a) (b) (c)

Figure A.2: Colour mapping of local shape descriptors onto the shape stack obtained via active shape evolution (upper row) and focusing (lower row) of the *saw-toothed rectangle* image. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



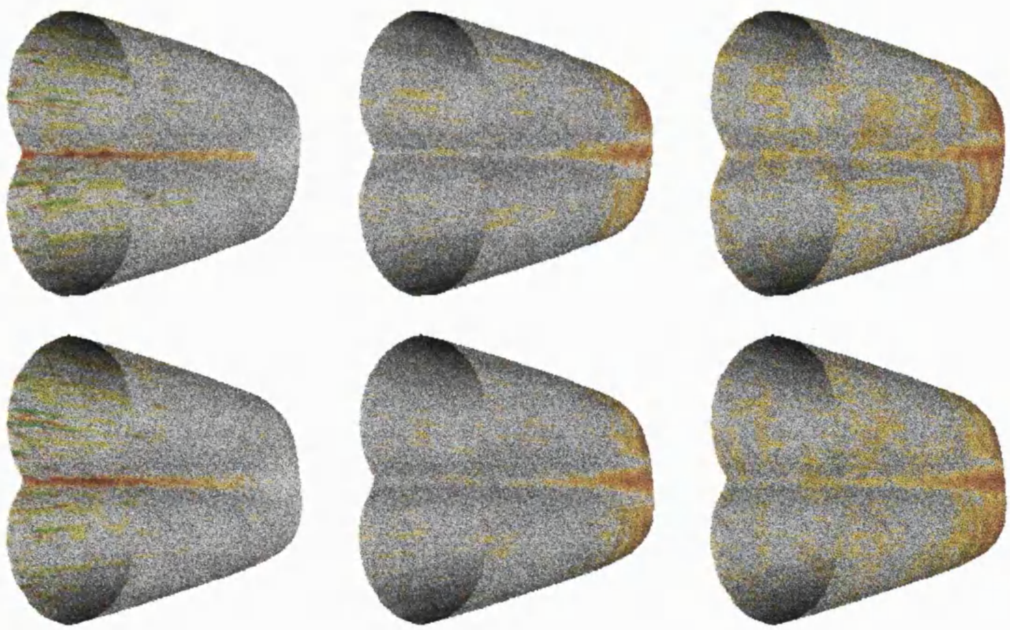
(a) (b) (c)

Figure A.3: Colour mapping of local shape descriptors on the shape stack obtained via active shape evolution (upper row) and focusing (lower row) of the *kangaroo* image. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



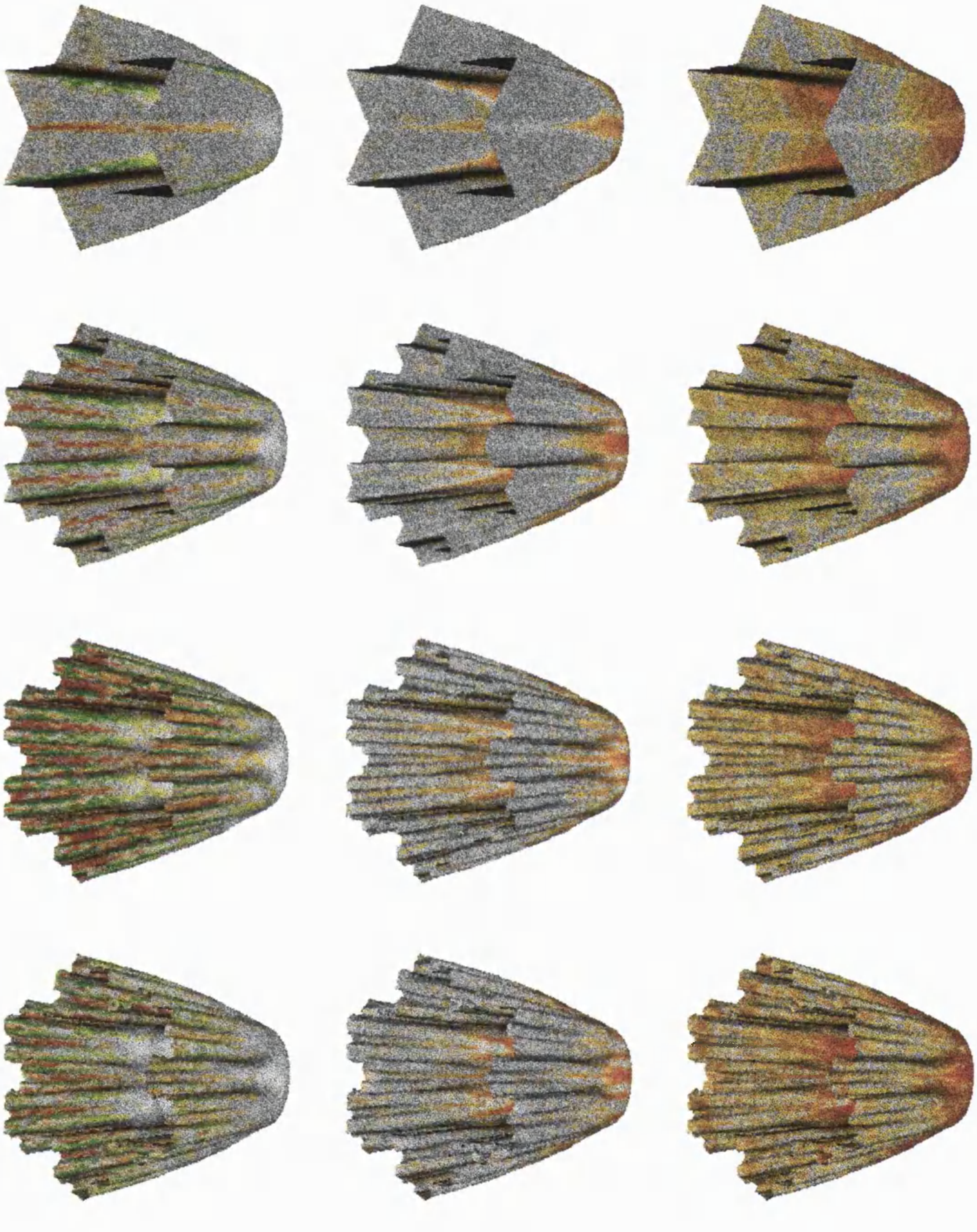
(a) (b) (c)

Figure A.4: Colour mapping of local shape descriptors on the shape stack obtained via active shape evolution (upper row) and focusing (lower row) of the *teardrop* image. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



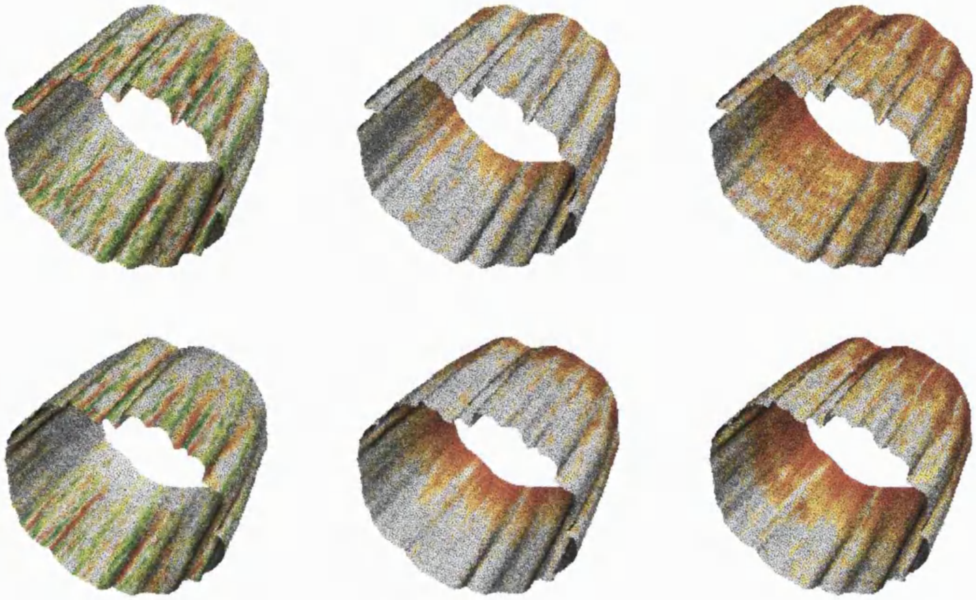
(a) (b) (c)

Figure A.5: Colour mapping of local shape descriptors on the shape stack obtained via active shape evolution (upper row) and focusing (lower row) of the *blobs* image. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



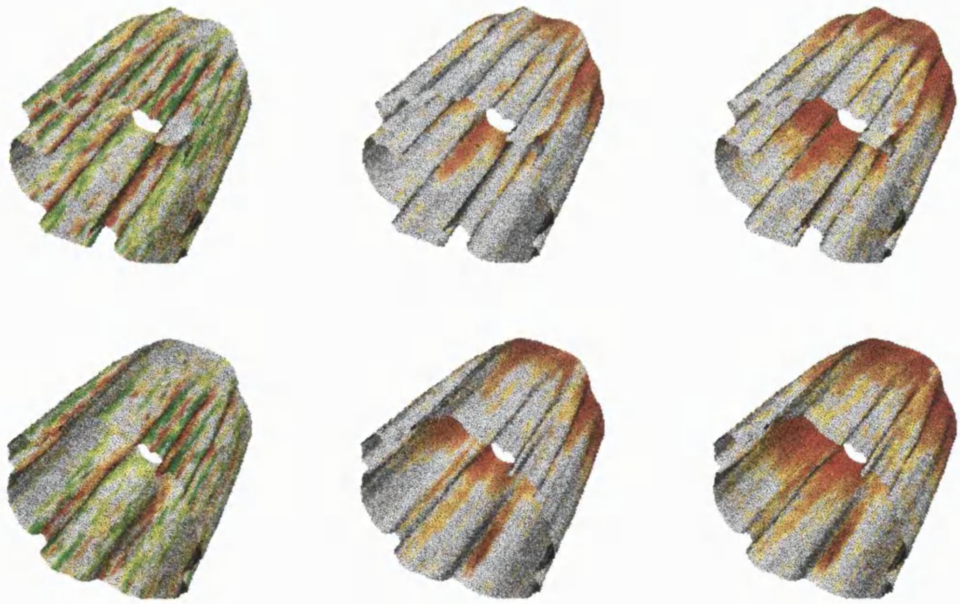
(a) (b) (c)

Figure A.6: Colour mapping of local shape descriptors onto the shape stack obtained via active shape focusing of the von Koch curve for increasing fractal generations (from top to bottom). Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



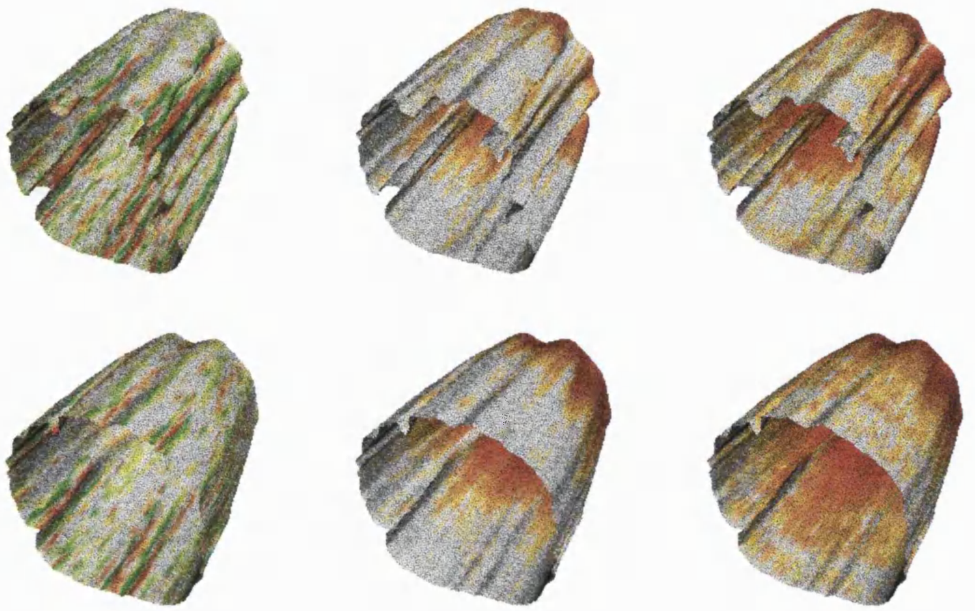
(a) (b) (c)

Figure A.7: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of patient 1. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



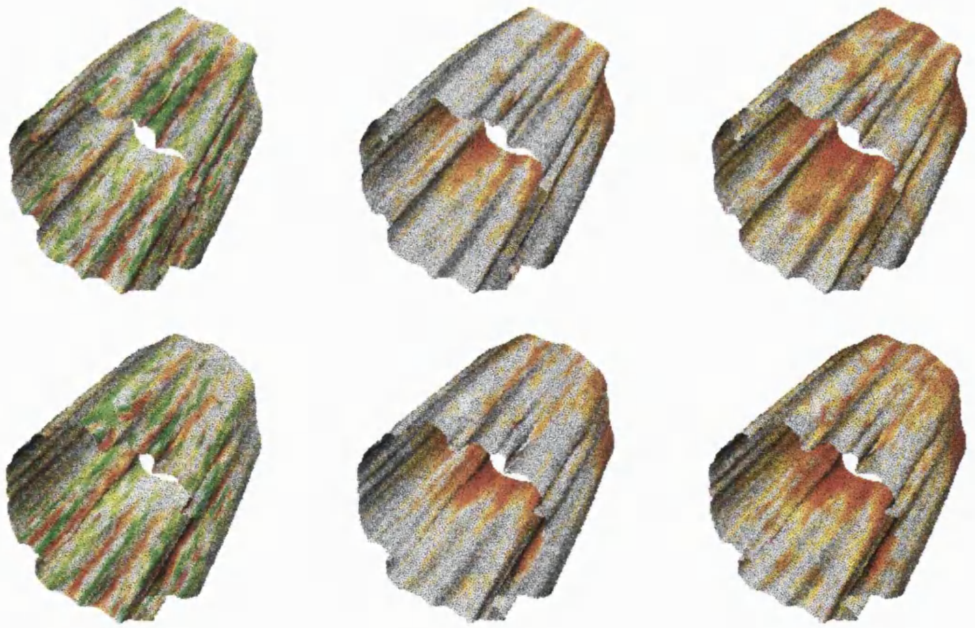
(a) (b) (c)

Figure A.8: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of patient 2. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



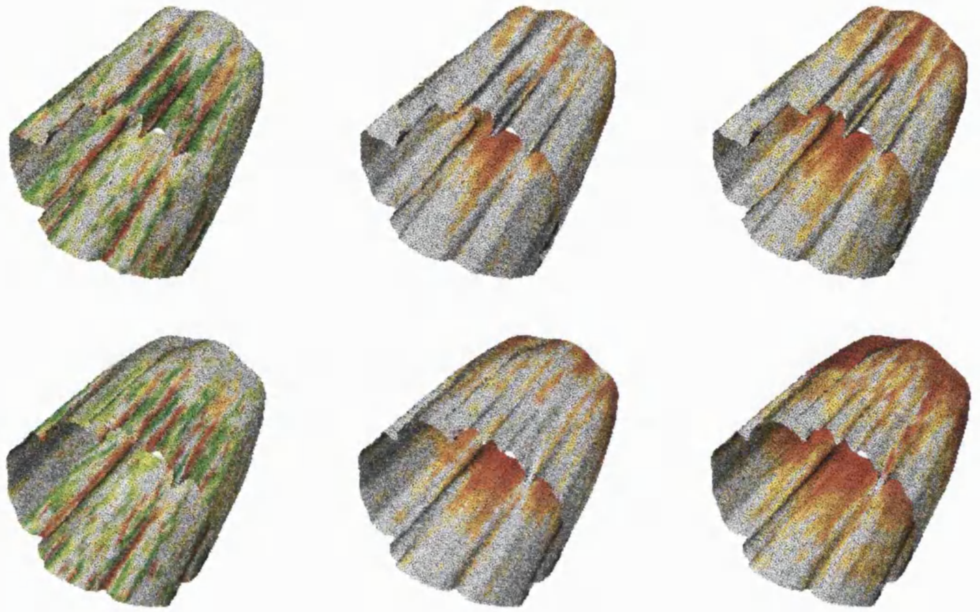
(a) (b) (c)

Figure A.9: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of patient 3. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



(a) (b) (c)

Figure A.10: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of patient 4. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.



(a) (b) (c)

Figure A.11: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of control 1. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.

(a) (b) (c)

Figure A.12: Colour mapping of the local shape descriptors onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ (top) and $4D$ (bottom) scale-space for the epileptic data of the intermediate slice of control 2. Columns (a) Curvature. (b) Chamfer distance. (c) Triangulation distance.

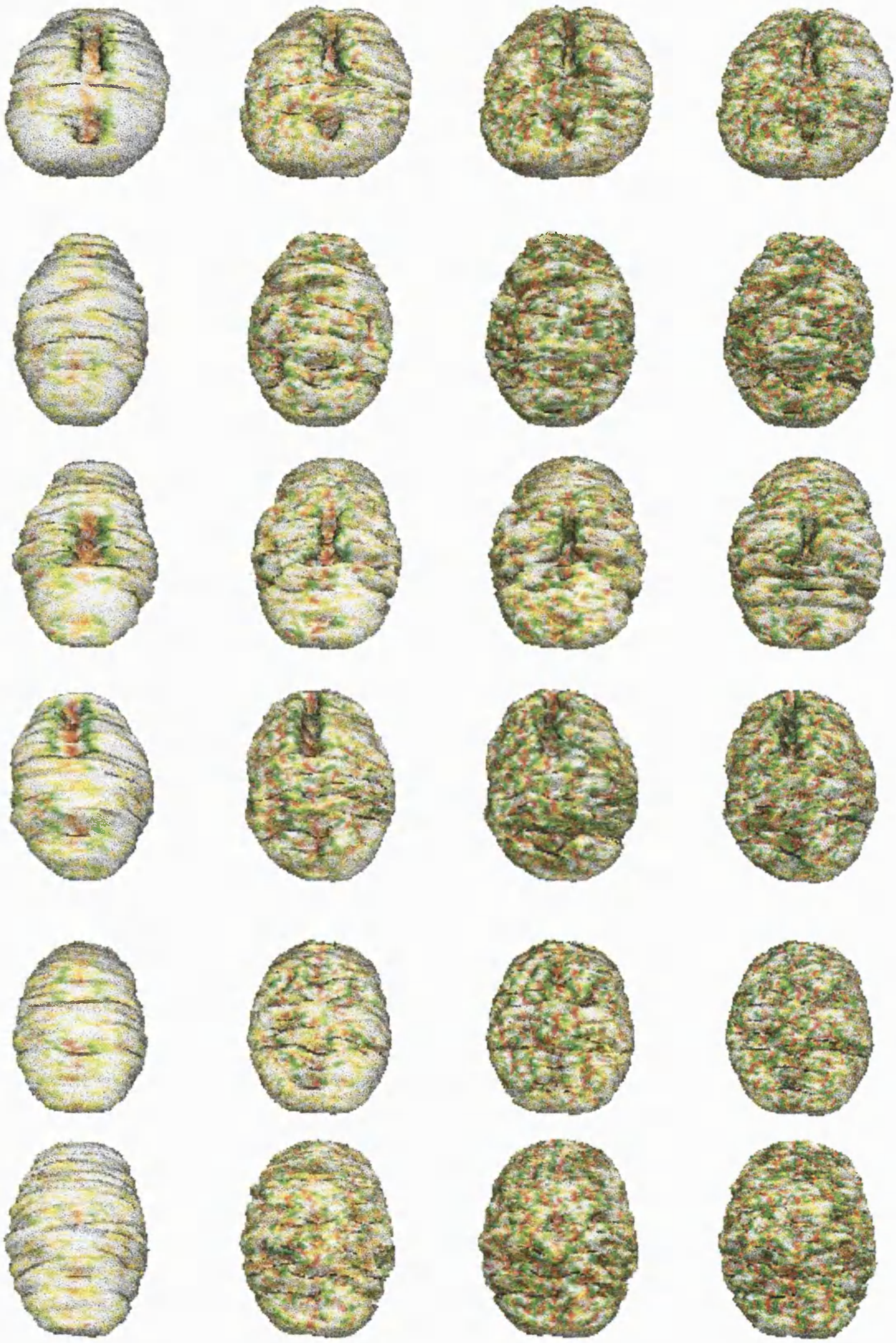


Figure A.13: Local curvature mapping onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

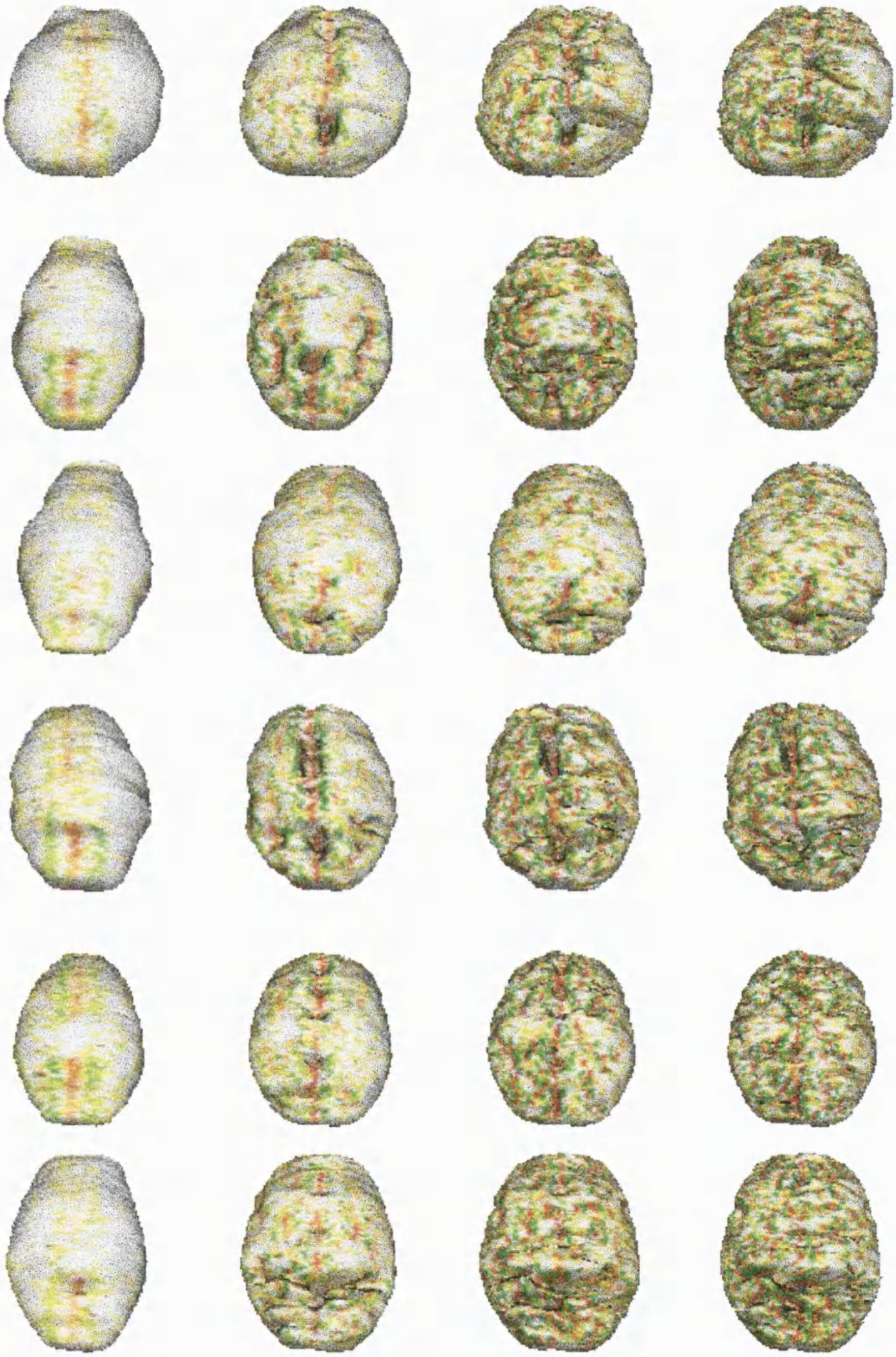


Figure A.14: Local curvature mapping onto the shape stacks obtained via active shape focusing in a $4D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

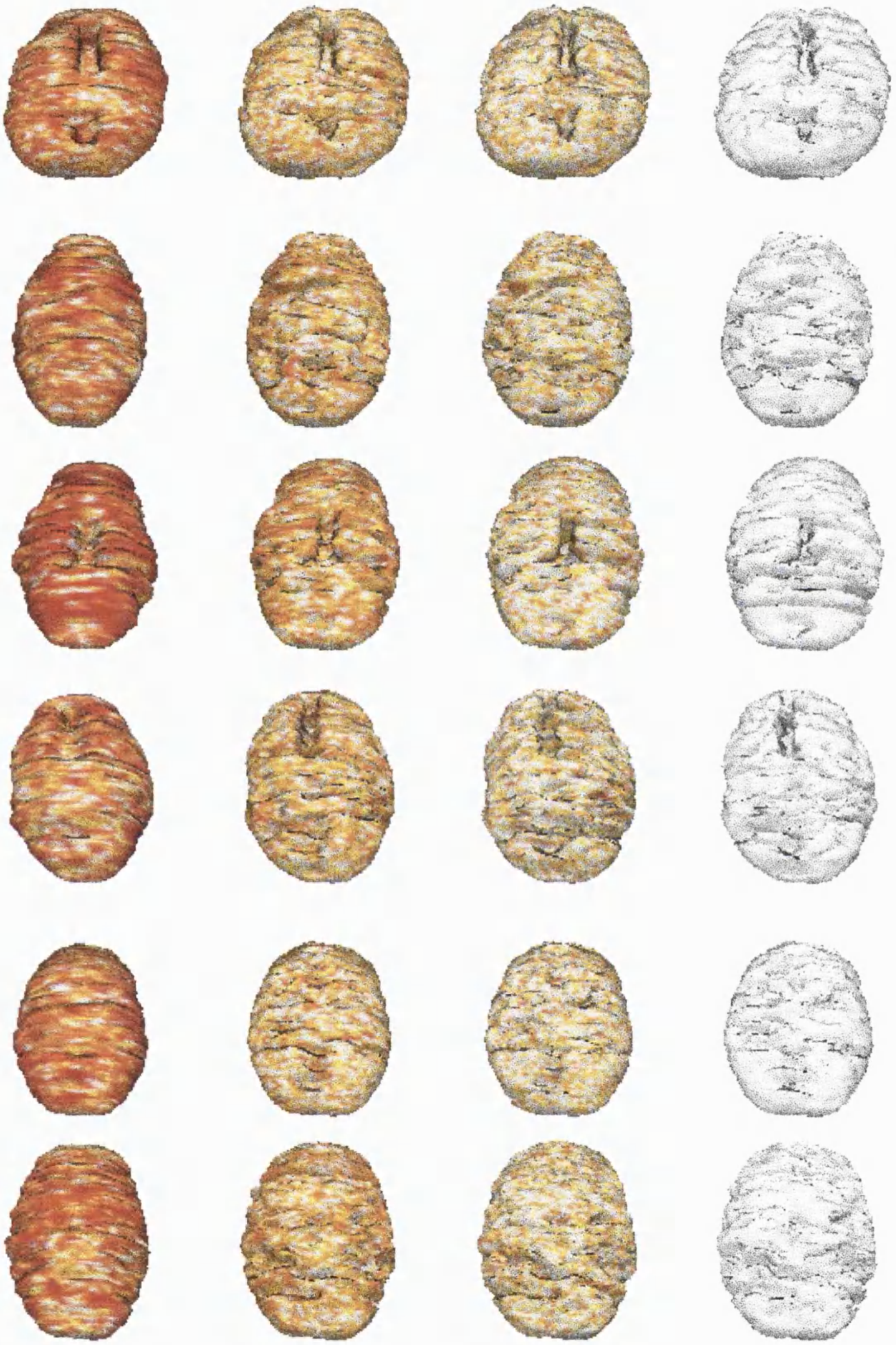


Figure A.15: Local Chamfer distance mapping onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

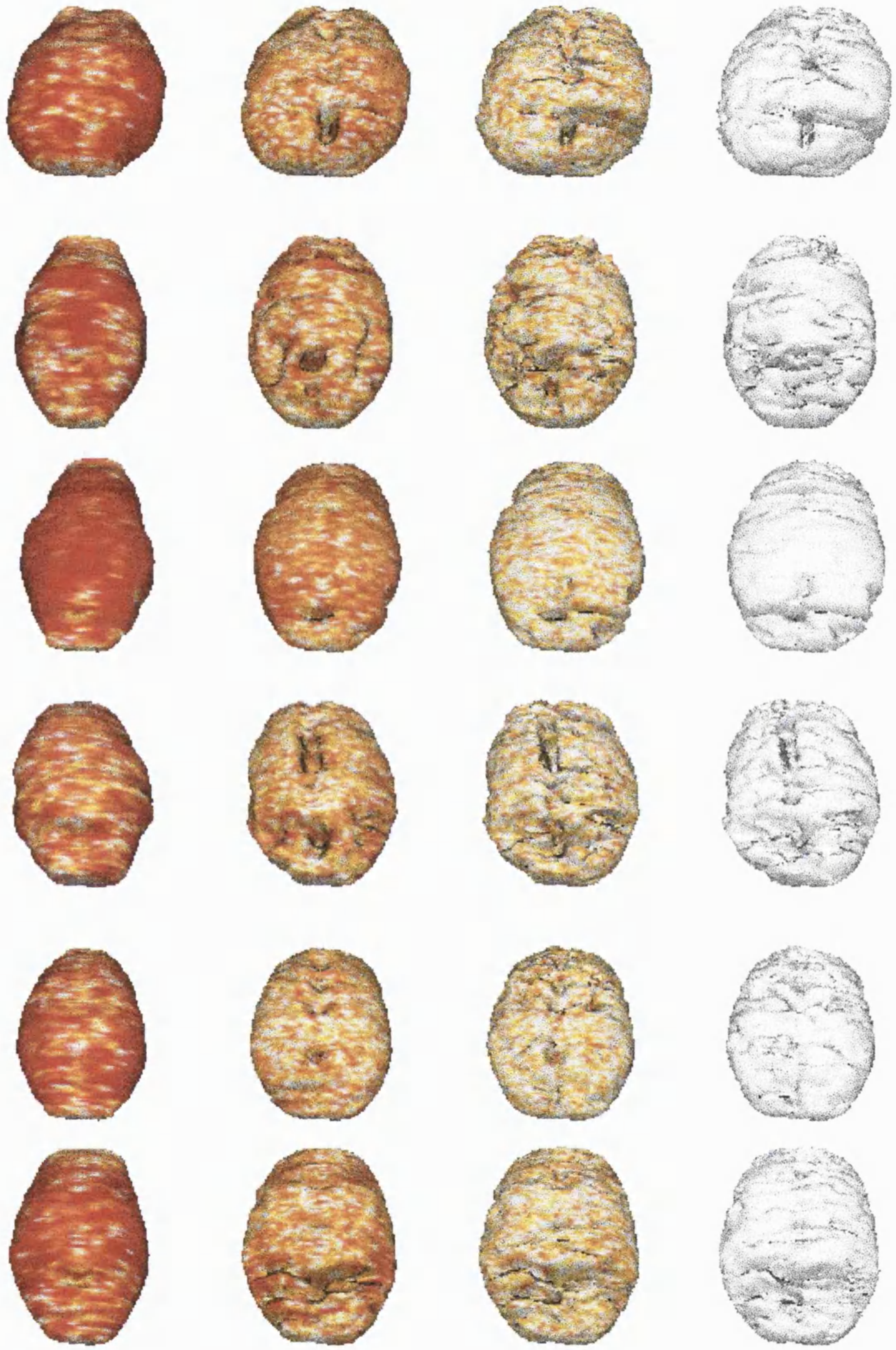


Figure A.16: Local Chamfer distance mapping onto the shape stacks obtained via active shape focusing in a $4D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

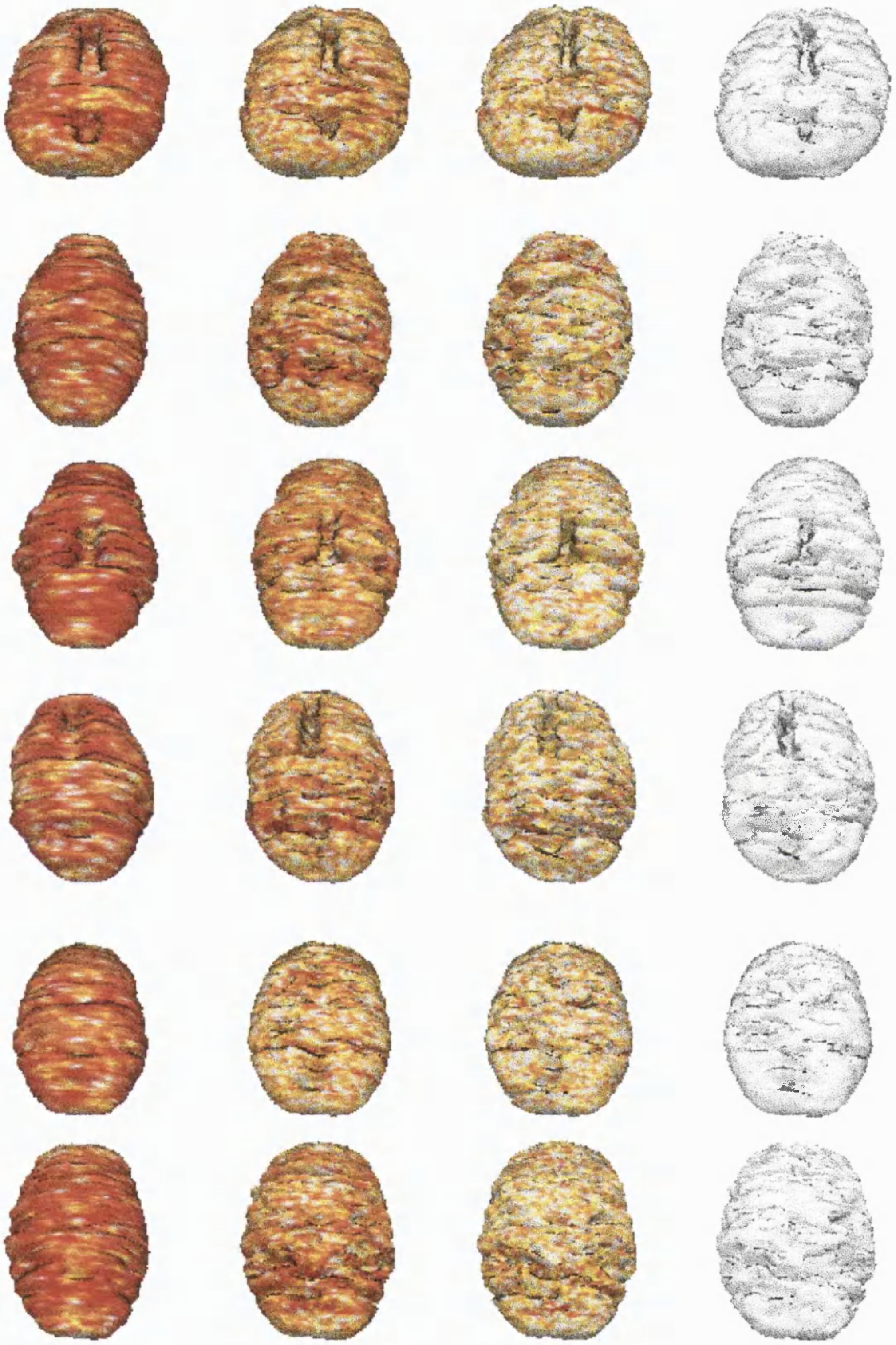


Figure A.17: Local triangulation distance mapping onto the shape stacks obtained via active shape focusing in a $3\frac{1}{2}D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

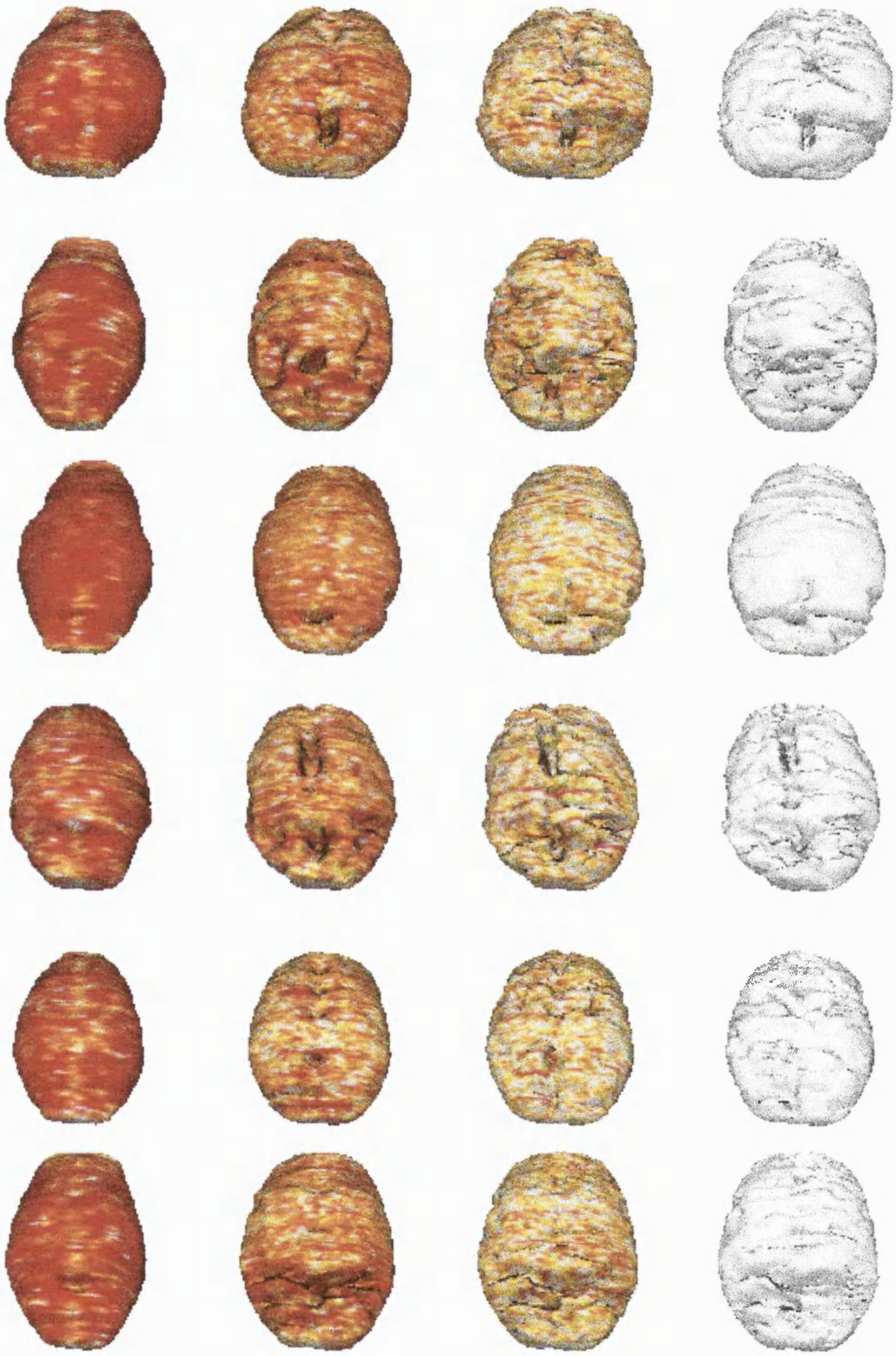
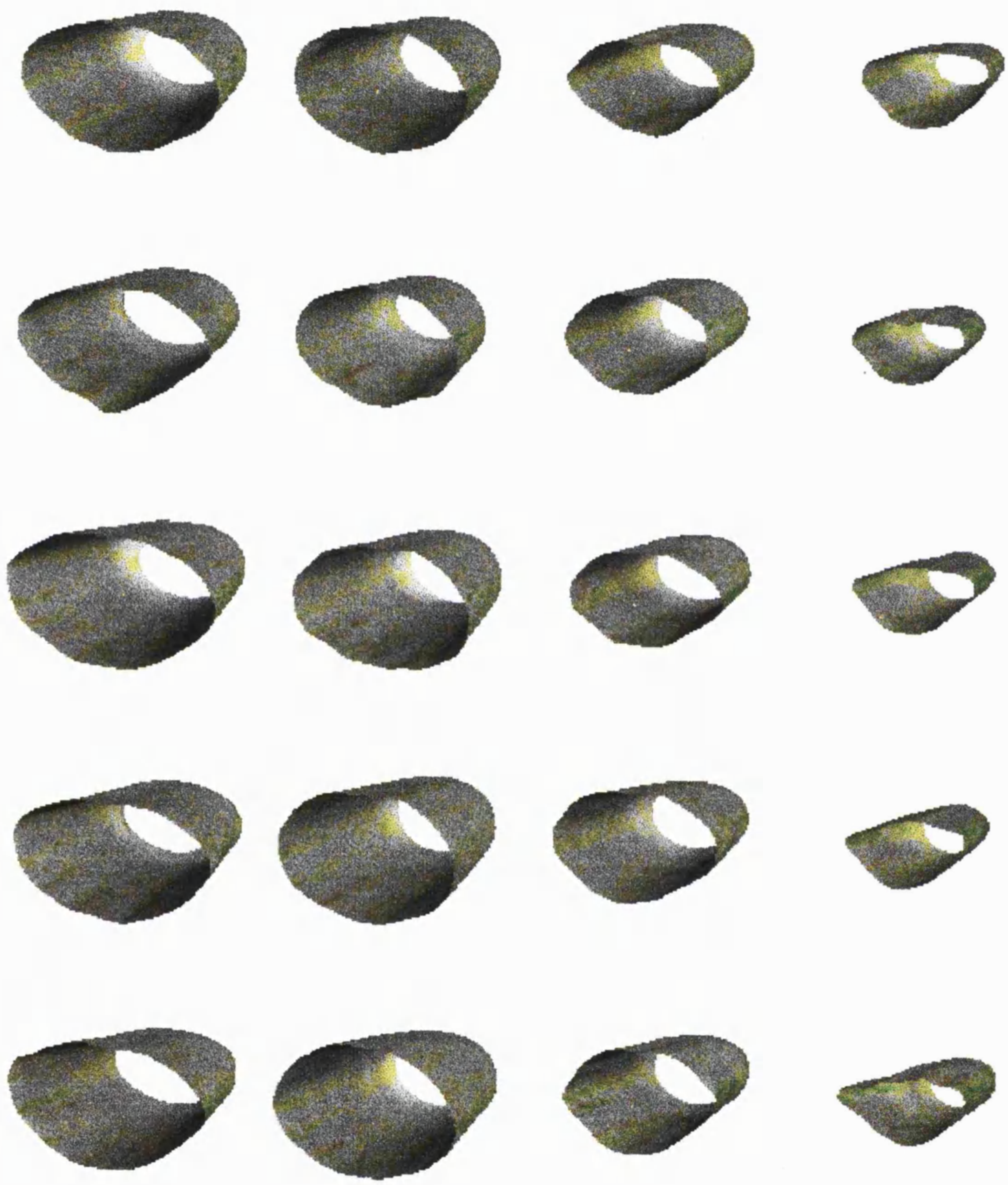


Figure A.18: Local triangulation distance mapping onto the shape stacks obtained via active shape focusing in a $4D$ scale-space for the epileptic data. From top to bottom: patient1, patient 2, patient 3, patient 4, control 1, control 2. Scale samples (from left to right): $\sigma = 8, 4, 2, 1$.

Figure A.19: Colour mapping of local shape descriptors onto the shape stack obtained via active shape focusing of the spinal cord data from the first scan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2.



(a) (b) (c) (d)

Figure A.20: Colour mapping of local shape descriptors onto the shape stack obtained via active shape focusing of the spinal cord data from the rescan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2.

(a) (b) (c) (d)

Figure A.21: Local Chamfer distance mapping onto the shape stack obtained via active shape focusing of the spinal cord data from the first scan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2.

(a) (b) (c) (d)

Figure A.22: Local Chamfer distance mapping onto the shape stack obtained via active shape focusing of the spinal cord data from the rescan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2.

(a) (b) (c) (d)

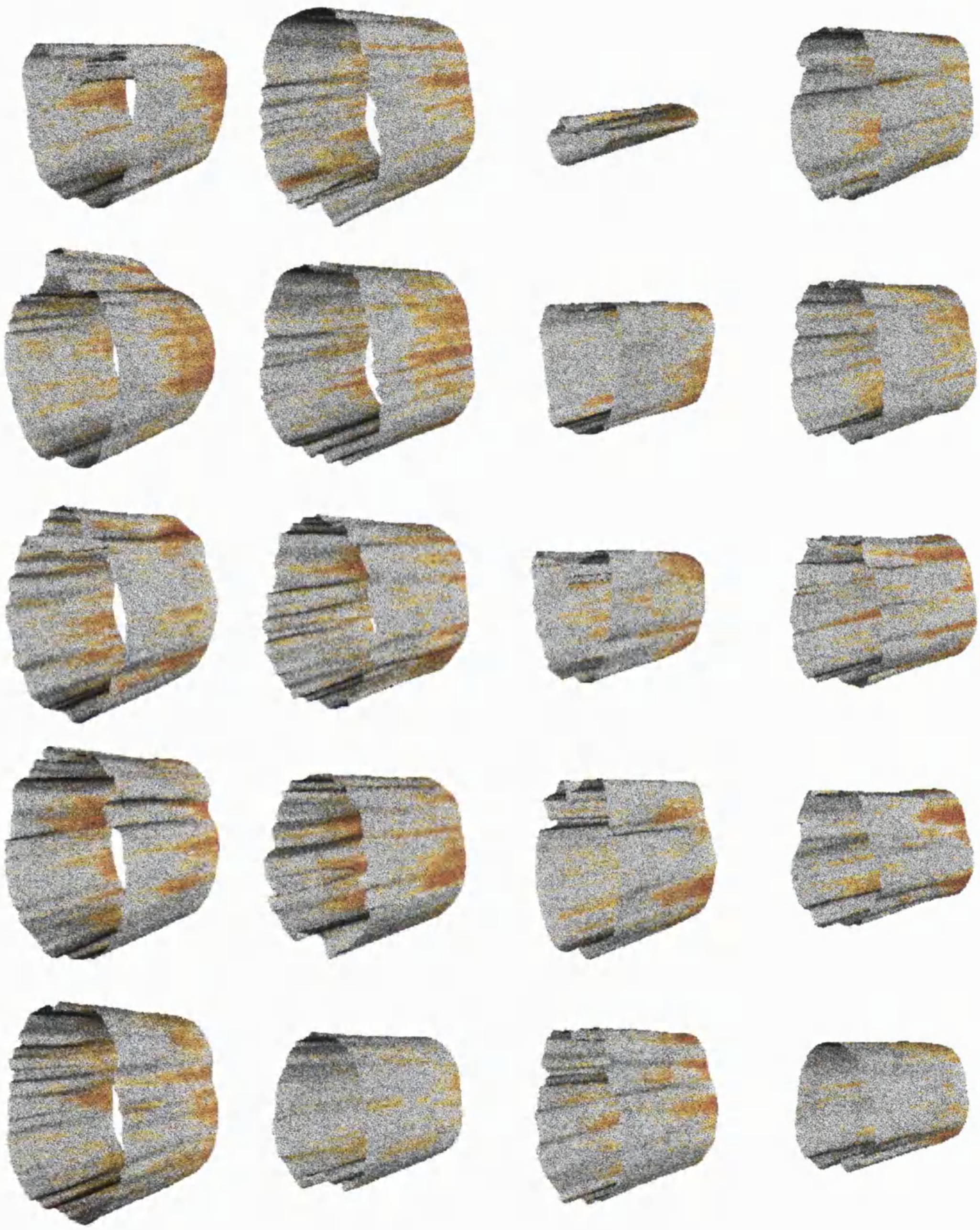
Figure A.23: Local triangulation distance mapping onto the shape stack obtained via active shape focusing of the spinal cord data from the first scan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2

(a) (b) (c) (d)

Figure A.24: Local triangulation distance mapping onto the shape stack obtained via active shape focusing of the spinal cord data from the rescan, for columns: (a) Control 1. (b) Control 2. (c) Patient 1. (d) Patient 2.

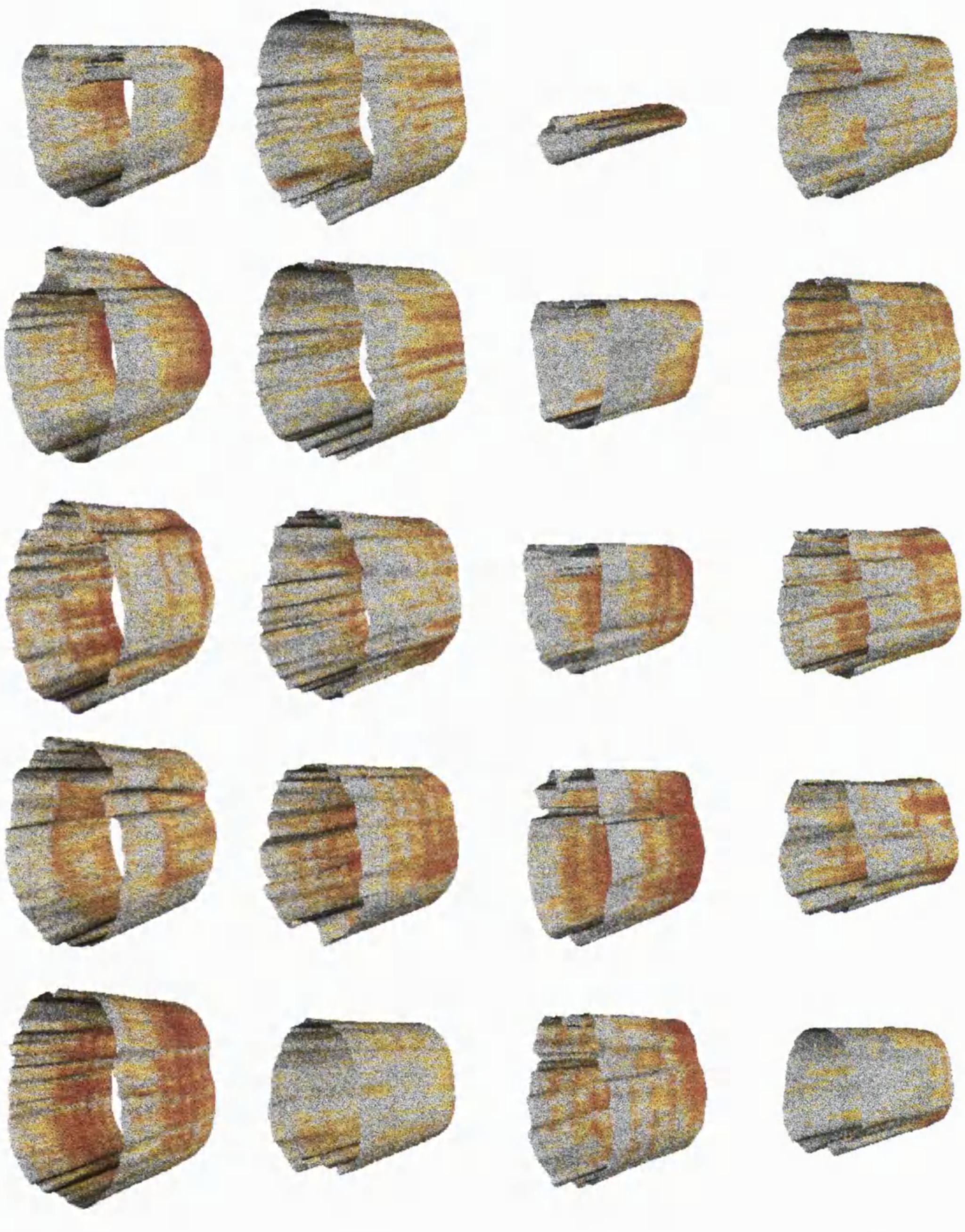
(a) (b) (c) (d)

Figure A.25: Colour mapping of the local curvature onto the shape stacks obtained via active shape focusing of the neonatal data. Columns (a)-(b): Term case. Columns (c)-(d): Premature case.



(a) (b) (c) (d)

Figure A.26: Colour mapping of the local Chamfer distance onto the shape stacks obtained via active shape focusing of the neonatal data. Columns (a)-(b): Term case. Columns (c)-(d): Premature case.



(a) (b) (c) (d)

Figure A.27: Colour mapping of the local triangulation distance onto the shape stacks obtained via active shape focusing of the neonatal data. Columns (a)-(b): Term case. Columns (c)-(d): Premature case.

Bibliography

[Abbasi *et al.*, 1997] S. Abbasi, F. Mokhtarian, and J. Kittler. Reliable classification of chrysanthemum leaves through curvature scale space. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space '97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 284–295. Springer Verlag, 1997.

[Aguadao *et al.*, 1996] A.S. Aguadao, E. Montiel, and M.S. Nixon. On using directional information for parameter space decomposition in ellipse detection. *Pattern Recognition*, 29(3):369–381, 1996.

[Akl and Toussaint, 1978] S.G. Akl and G. Toussaint. A fast convex hull algorithm. *Information Processing Letters*, 7:216–219, 1978.

[Alvarez *et al.*, 1992] L. Alvarez, P.-L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM Journal on Numerical Analysis*, 29(3):845–866, 1992.

[Amini *et al.*, 1990] A.A. Amini, T.E. Weymouth, and R.C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):855–867, 1990.

[Amini *et al.*, 1995a] A. Amini, R. Curwen, and J. Gore. Measurement of tag deformations with DP snakes and snake grids. In Y. Bizais, C. Barillot, and R. DiPaola, editors, *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, Computational Imaging and Vision, pages 363–364. Kluwer Academic Publishers, 1995.

[Amini *et al.*, 1995b] A.A. Amini, R.W. Curwen, R.T. Constable, and J.C. Gore. Dynamic programming snakes, snake grids, and smooth warps for analysis of 2D tissue deformations. Technical report, Departments of Diagnostic Radiology and Electrical Engineering, Yale University, 1995.

[Arridge and Simmons, 1997] S.R. Arridge and A. Simmons. Multispectral probabilistic diffusion using Bayesian classification. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space '97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 224–235. Springer Verlag, 1997.

[Asada and Brady, 1986] H. Asada and M. Brady. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):2–14, 1986.

[Babaud *et al.*, 1986] J. Babaud, A.P. Witkin, M. Baudin, and R.O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):26–33, 1986.

[Ballard, 1981] D.H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.

[Bangham *et al.*, 1996a] J.A. Bangham, R. Harvey, P.D. Ling, and R.V. Aldridge. Nonlinear scale-space from n-dimensional sieves. In B. Buxton and R. Cippola, editors, *Computer Vision – Proc. 4th European Conference on Computer Vision (ECCV'96)*, volume 1064 of *Lecture Notes in Computer Science*, pages 189–198. Springer Verlag, 1996.

[Bangham *et al.*, 1996b] J.A. Bangham, P.D. Ling, and R. Harvey. Nonlinear scale-space causality preserving filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:520–528, 1996.

[Bardinet *et al.*, 1996a] E. Bardinet, L. Cohen, and N. Ayache. Tracking and motion analysis of the left ventricle with deformable superquadrics. *Medical Image Analysis*, 1(2):129–150, 1996.

[Bardinet *et al.*, 1996b] E. Bardinet, L. Cohen, and N. Ayache. Tracking medical 3D data with a deformable parametric model. In B. Buxton and R. Cippola, editors, *Computer Vision – Proc. 4th European Conference on Computer Vision (ECCV'96)*, volume 1064 of *Lecture Notes in Computer Science*, pages 317–328. Springer Verlag, 1996.

[Barnsley, 1988] M. Barnsley. *Fractals Everywhere*. Academic Press, 1988.

[Bartels *et al.*, 1987] R. Bartels, J. Beatty, and B. Barsky. *An Introduction to Splines for Use in Computer Graphics and Geometric Modelling*. Morgan Kaufmann, 1987.

[Bartels *et al.*, 1994] K.A. Bartels, A.C. Bovik, and C.E. Griffin. Spatio-temporal tracking of material shape change via multi-dimensional splines. In *IEEE Workshop on Biomedical Image Analysis*, pages 110–116. IEEE Computer Society Press, 1994.

[Bascle and Deriche, 1992] B. Bascle and R. Deriche. Features extraction using parametric snakes. In *Proc. International Conference on Pattern Recognition ICPR'92*, volume III, Conference C: Image, Speech, and Signal Analysis, pages 659–662. IEEE Computer Society Press, 1992.

[Bascle and Deriche, 1993] B. Bascle and R. Deriche. Stereo matching, reconstruction and refinement of 3D curves using deformable contours. In *Proc. 4th International Conference on Computer Vision (ICCV'93)*, pages 421–430. IEEE Computer Society Press, 1993.

[Bascle and Deriche, 1995] B. Bascle and R. Deriche. Region tracking through image sequences. In *Proc. 5th International Conference on Computer Vision (ICCV'95)*, pages 302–307. IEEE Computer Society Press, 1995.

[Baumberg and Hogg, 1995] A. Baumberg and D. Hogg. An adaptive eigenshape model. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 87–96. BMVA, 1995.

[Bergholm, 1987] F. Bergholm. Edge focusing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(6):726–741, 1987.

[Berkner, 1997a] K. Berkner. Reconstruction of self-similar functions from scale-space. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space'97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 311–314. Springer Verlag, 1997.

[Berkner, 1997b] K. Berkner. A wavelet-based solution to the inverse problem for fractal interpolation functions. In J.L. Véhel, E. Lutton, and C. Tricot, editors, *Fractals in Engineering'97*, pages 81–92. Springer Verlag, 1997.

[Besag, 1986] J. Besag. On the statistical analysis of dirty pictures. *Journal Royal Statistical Society B*, 48(3):259–302, 1986.

[Besl and McKay, 1992] P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–256, 1992.

[Blum and Nagel, 1978] H. Blum and R.N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10:167–180, 1978.

[Blum, 1967] H. Blum. A transformation for extracting new descriptors of shape. In W. Wathen-Dunn, editor, *Models for the perception of speech and visual form*, pages 362–380. MIT Press, 1967.

[Blum, 1973] H. Blum. Biological shape and visual science (Part I). *Journal of Theoretical Biology*, 38:205–287, 1973.

[Bookstein and Green, 1992] F.L. Bookstein and W.D.K. Green. A feature space for edgels in images with landmarks. In D.C. Wilson and J.N. Wilson, editors, *Mathematical Methods in Medical Imaging 1992*, volume 1768, pages 228–247. SPIE, 1992.

[Bookstein, 1989] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[Bookstein, 1991a] F.L. Bookstein. *Morphometric tools for landmark data : geometry and biology*. Cambridge University Press, 1991.

[Bookstein, 1991b] F.L. Bookstein. Thin-plate splines and the atlas problem for biomedical images. In A.C.F. Colchester and D.J. Hawkes, editors, *Information Processing in Medical Imaging: Proc. 12th International Conference (IPMI'91)*, volume 511 of *Lecture Notes in Computer Science*, pages 326–342. Springer Verlag, 1991.

[Bookstein, 1996a] F.L. Bookstein. Applying landmark methods to biological outline data. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 59–70. Leeds University Press, 1996.

[Bookstein, 1996b] F.L. Bookstein. Visualizing group differences in outline shape: methods from biometrics and landmark points. In K.H. Höhne and R. Kikinis, editors, *Proc. 4th International Conference on Visualization in Biomedical Computing (VBC'96)*, volume 1131 of *Lecture Notes in Computer Science*, pages 405–410. Springer Verlag, 1996.

[Borgefors, 1986] G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 34:344–371, 1986.

[Borgefors, 1996] G. Borgefors. On digital image transforms in three dimensions. *Computer Vision and Image Understanding*, 64(3):368–376, 1996.

[Brammer and Bullmore, 1994] M. Brammer and E. Bullmore. Characterisation of cerebral grey/white matter boundaries in Magnetic Resonance Imaging scans using fractal dimension and radius of gyration. In K.V. Mardia and C.A. Gill, editors, *14th Leeds Annual Statistical Research (LASR) Workshop: Medical Imaging: Shape and Visualisation*, page 10. University of Leeds, Department of Statistics and CoMIR, 1994.

[Brandt and Algazi, 1992] J.W. Brandt and V.R. Algazi. Continuous skeleton computation by Voronoi diagram. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(3):329–338, 1992.

[Brechbühler *et al.*, 1995] C. Brechbühler, G. Gerig, and O. Kübler. Parameterization of closed surfaces for 3D shape description. *Computer Vision and Image Understanding*, 61(2):154–170, 1995.

[Brewer and Anderson, 1977] J.A. Brewer and D.C. Anderson. Visual interaction with Overhauser curves and surfaces. In *Proc. SIGGRAPH'77*, volume 11(2), pages 132–137. ACM, 1977.

[Bronstein and Semendjajew, 1989] I.N. Bronstein and K.A. Semendjajew. *Taschenbuch der Mathematik.* Verlag Harri Deutsch, 1989.

[Bullmore *et al.*, 1994] E. Bullmore, M. Brammer, I. Harvey, R. Presaud, R. Murray, and M. Ron. Fractal analysis of the boundary between white matter and cerebral cortex in magnetic resonance images: a controlled study of schizophrenic and manic-depressive patients. *Psychological Medicine*, 24(3):771–781, 1994.

[Bulpitt and Efford, 1994] A.J. Bulpitt and N.D. Efford. Using 3D deformable models to recognise anatomical structure. In *14th Leeds Annual Statistical Research (LASR) Workshop: Medical Imaging: Shape and Visualisation*, page 3. University of Leeds, Department of Statistics and CoMIR, 1994.

[Bulpitt and Efford, 1995] A.J. Bulpitt and N.D. Efford. An efficient 3D deformable model with self-optimising topology. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 37–46. BMVA, 1995.

[Burbeck and Pizer, 1994] C.A. Burbeck and S.M. Pizer. Object representation by cores: Identifying and representing primitive spatial regions. TR94-160, University of North Carolina at Chapel Hill, 1994.

[Canny, 1987] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1987.

[Caselles *et al.*, 1993] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1–31, 1993.

[Caselles *et al.*, 1997] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.

[Castellano Smith *et al.*, 1997] A.D. Castellano Smith, D.L.G. Hill, and D.H. Hawkes. Surface simplification and explicit surface shape measures applied to the human brain. In C.J. Taylor, J.A. Noble, and J.M. Brady, editors, *Proc. 1st Annual Conference on Medical Image Understanding and Analysis'97*, pages 181–184. BMVA, 1997.

[Catmull and Rom, 1974] E. Catmull and R. Rom. A class of local interpolating splines. In R. Barnhill and R. Riesenfeld, editors, *Computer Aided Geometric Design*, pages 317–326. Academic Press, 1974.

[Catté *et al.*, 1992] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.

[Chakraborty and Duncan, 1995] A. Chakraborty and J. Duncan. Integration of boundary finding and region-based segmentation using game theory. In Y. Bizais, C. Barillot, and R. DiPaola, editors, *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, Computational Imaging and Vision, pages 189–200. Kluwer Academic Publishers, 1995.

[Chakraborty *et al.*, 1996] A. Chakraborty, L.H. Staib, and J.S. Duncan. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Transactions on Medical Imaging*, 15(6):859–870, 1996.

[Chalana and Kim, 1996] V. Chalana and Y. Kim. A methodology for evaluation of image segmentation algorithms on medical images. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 178–189. SPIE, 1996.

[Chalana *et al.*, 1996] V. Chalana, D.T. Linker, D.R. Haynor, and Y. Kim. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 15(3):290–298, 1996.

[Cham and Cipolla, 1996] T.-J. Cham and R. Cipolla. Automated B-spline curve representation with MDL-based active contours. In R.B. Fisher and E. Trucco, editors, *Proc. 7th British Machine Vision Conference (BMVC'96)*, volume I, pages 363–372. BMVA, 1996.

[Chen and Chin, 1993] M.-H. Chen and R.T. Chin. Partial smoothing splines for noisy boundaries with corners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):12081216, 1993.

[Chen *et al.*, 1989] C.-C. Chen, J.S. DaPonte, and M.D. Fox. Fractal feature analysis and classification in medical images. *IEEE Transactions on Medical Imaging*, 8(2):133–142, 1989.

[Chen, 1989] C.L. Chen. Computing the convex hull of a simple polygon. *Pattern Recognition*, 22:561–565, 1989.

[Choy and Jin, 1996] M.M. Choy and J.S. Jin. Morphological image analysis of left-ventricular endocrdial borders in 2D echocardiograms. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 852–862. SPIE, 1996.

[Christiansen and Sederberg, 1978] H.N. Christiansen and T.W. Sederberg. Conversion of complex contour line definitions into polygonal element mosaics. In *Proc. SIGGRAPH'78*, volume 12(3), pages 187–192. ACM, 1978.

[Cipolla and Blake, 1992] R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *International Journal of Computer Vision*, 9(2):83–112, 1992.

[Cohen and Cohen, 1993] L.D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2D and 3D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131–1147, 1993.

[Cohen and Wang, 1994] F.S. Cohen and J.-Y. Wang. Part I: Modeling image curves using invariant 3-D object curve models - a path to 3-D recognition and shape estimation from image contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1):1–12, 1994.

[Cohen *et al.*, 1992] I. Cohen, L.D. Cohen, and N. Ayache. Using deformable surfaces to segment 3-D images and infer differential structures. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 56(2):242–263, 1992.

[Cohen, 1990] L.D. Cohen. Note on active contour models and balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2):211–218, 1990.

[Cohen, 1995] L.D. Cohen. *Méthodes variationelles pour le traîtement d'images.* Habilitation à diriger des recherches, Université Paris-IX Dauphine, U.F.R. Mathématiques de la décision, 1995.

[Collins *et al.*, 1995] D.L. Collins, A.C. Evans, C. Holmes, and T.M. Peters. Automatic 3D segmentation of neuro-anatomical structures from MRI. In Y. Bizais, C. Barillot, and R. Di-Paola, editors, *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, Computational Imaging and Vision, pages 139–152. Kluwer Academic Publishers, 1995.

[Cook *et al.*, 1995] M.J. Cook, S.L. Free, M.R.A. Manford, D.R. Fish, S.D. Shorvon, and J.M. Stevens. Fractal description of cerebral cortical patterns in frontal lobe epilepsy. *European Neurology*, 35:327–335, 1995.

[Cootes and Taylor, 1992] T.F. Cootes and C.J. Taylor. Active shape models - 'Smart Snakes'. In *Proc. 3rd British Machine Vision Conference (BMVC'92)*, pages 266–275. Springer Verlag, 1992.

[Cootes *et al.*, 1992a] T.F. Cootes, D.H. Cooper, C.J. Taylor, and J. Graham. Trainable method of parametric shape description. *Image and Vision Computing*, 10(5):289–294, 1992.

[Cootes *et al.*, 1992b] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. 3rd British Machine Vision Conference (BMVC'92)*, pages 9–18. Springer Verlag, 1992.

[Cootes *et al.*, 1993] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical imaging. In H.H. Barret and A.F. Gmitro, editors, *Information Processing in Medical Imaging: Proc. 13th International Conference (IPMI'93)*, volume 687 of *Lecture Notes in Computer Science*, pages 33–47. Springer Verlag, 1993.

[Cootes *et al.*, 1995] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[Cortelazzo *et al.*, 1996] G. Cortelazzo, G. Deretta, G.A. Mian, and P. Zamperoni. Normalized weighted Levensthein distance and triangle inequality in the context of similarity discrimination of bilevel images. *Pattern Recognition Letters*, 17:431–436, 1996.

[Danielsson, 1980] P.E. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

[Dean *et al.*, 1996] D. Dean, P. Buckley, F. Bookstein, J. Kamath, D. Kwon, L. Friedman, and C. Lys. Three dimensional MR-based morphometric comparison of schizophrenic and normal cerebral ventricles. In K.H. Höhne and R. Kikinis, editors, *Proc. 4th International Conference on Visualization in Biomedical Computing (VBC'96)*, volume 1131 of *Lecture Notes in Computer Science*, pages 363–373. Springer Verlag, 1996.

[Delingette *et al.*, 1992] H. Delingette, M. Hebert, and K. Ikeuchi. Shape representation and image segmentation using deformable surfaces. *Image and Vision Computing*, 10(3):132–144, 1992.

[Duncan *et al.*, 1991] J.S. Duncan, F.A. Lee, A.W.M. Smeulders, and B.L. Zaret. A bending energy model for measurement of cardiac shape deformity. *IEEE Transactions on Medical Imaging*, 10(3):307–320, 1991.

[Dyn *et al.*, 1983] N. Dyn, D. Levin, and S. Rippa. Surface interpolation and smoothing by thin plate splines. In C.K. Chui, L.L. Schumaker, and J.D. Ward, editors, *Proc. International Symposium on Approximation Thory IV*, pages 445–449. Academic Press, 1983.

[Eberly and Pizer, 1994] D. Eberly and S.M. Pizer. Ridge flow models for image segmentation. TR, University of North Carolina at Chapel Hill, 1994.

[Eberly *et al.*, 1993] D. Eberly, R. Gardner, B. Morse, and C. Scharlach. Ridges for image analysis. TR93-055, University of North Carolina at Chapel Hill, 1993.

[Eberly, 1994a] D.H. Eberly. *Geometric Methods for Analysis of Ridges in N-Dimensional Images*. PhD thesis, University of North Carolina at Chapel Hill, 1994.

[Eberly, 1994b] D.H. Eberly. The multiscale medial axis and its application in image registration. *Pattern Recognition Letters*, 15(5):445–452, 1994.

[Eberly, 1996] D.H. Eberly. *Ridges in Image and Data Analysis*, volume 7 of *Computational Imaging and Vision*. Kluwer Academic Publishers, 1996.

[Edwards *et al.*, 1995] P.J. Edwards, D.L.G. Hill, J.A. Little, V.A.S. Sahni, and D.H. Hawkes. Medical image registration incorporating deformations. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 2, pages 691–699. BMVA, 1995.

[Edwards *et al.*, 1997] P.J. Edwards, D.L.G. Hill, J.A. Little, and D.J. Hawkes. Deformation for image guided interventions using a three component tissue model. In J. Duncan and G. Gindi, editors, *Information Processing in Medical Imaging: Proc. 15th International Conference (IPMI'97)*, volume 1230 of *Lecture Notes in Computer Science*, pages 218–231. Springer Verlag, 1997.

[Evans and Spruck, 1991] L. Evans and J. Spruck. Motion of level sets by mean curvature I. *Journal of Differential Geometry*, 33:635–681, 1991.

[Farin, 1993] G. Farin. *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*. Computer Science and Scientific Computing. Academic Press, third edition, 1993.

[Fidrich, 1997] M. Fidrich. Following feature lines across scale. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space '97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 140–151. Springer Verlag, 1997.

[Florack *et al.*, 1992] L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, 10(6):376–388, 1992.

[Florack *et al.*, 1994] L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever. Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351, 1994.

[Florack, 1993] L.M.J. Florack. *The Syntactical Structure of Scalar Images*. PhD thesis, Utrecht University, The Netherlands, 1993.

[Foley *et al.*, 1990] J.D. Foley, A. van Damm, S.K. Feiner, and J.F. Hughes. *Computer Graphics - Principles and Practice*. Addison Wesley, second edition, 1990.

[Fortin *et al.*, 1992] C. Fortin, R. Kumaresan, W. Ohley, and S. Hoefer. Fractal dimension in the analysis of medical images. In B. Onaral and W. Ohley, editors, *Fractals and Scaling Theory*, volume 11(2), pages 65–71. IEEE Computer Society Press, 1992.

[Free *et al.*, 1997] S.L. Free, S.M. Sisodiya, M.J. Cook, D.R. Fish, and S.D. Shorvon. Three-dimensional fractal analysis of the white matter surface from magnetic resonance images of the human brain. *Cerebral Cortex*, 6(6):830–836, 1997.

[Freeman, 1974] H. Freeman. Computer processing of line drawing images. *Computer Surveys*, 6:57–98, 1974.

[Friedland and Adam, 1989] N. Friedland and D. Adam. Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing. *IEEE Transactions on Medical Imaging*, 8:344–353, 1989.

[Fritsch *et al.*, 1995] D.S. Fritsch, D. Eberly, S.M. Pizer, and M.J. McAuliffe. Stimulated cores and their applications in medical imaging. TR, University of North Carolina at Chapel Hill, Dept. of Radiation Oncology, 1995.

[Fritsch, 1994] D.S. Fritsch. Cores for image registration. TR, University of North Carolina at Chapel Hill, 1994.

[Gage and Hamilton, 1986] M. Gage and R.S. Hamilton. The heat equation shrinking convex plane curves. *Journal of Differential Geometry*, 23:69–96, 1986.

[Ge and Fitzpatrick, 1996] Y. Ge and J.M. Fitzpatrick. On the generation of skeletons from discrete Euclidean distance maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1955–1066, 1996.

[Geman and Geman, 1984] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[Gerig *et al.*, 1995] G. Gerig, G. Szekely, G. Israel, and M. Berger. Detection and characterization of unsharp blobs by curve evolution. In Y. Bizais, C. Barillot, and R. DiPaola, editors, *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, Computational Imaging and Vision, pages 165–176. Kluwer Academic Publishers, 1995.

[Goldberg, 1989] D.E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, 1989.

[Goldberger, 1992] A.L. Goldberger. Fractal mechanisms in the electrophysiology of the heart. In B. Onaral and W. Ohley, editors, *Fractals and Scaling Theory*, volume 11(2), pages 47–52. IEEE Computer Society Press, 1992.

[Goldman and Warren, 1993] R. Goldman and J. Warren. An extension of chaiken's algorithm to B-spline curves with knots in geometric progression. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 55(1):58–62, 1993.

[Graham and Yao, 1983] R.L. Graham and F.F. Yao. Finding the convex hull of a simple polygon. *Journal of Algorithms*, 4:324–331, 1983.

[Graham, 1972] R.L. Graham. An efficient algorithm for determining the convex hull of a planar set. *Information Processing Letters*, 1:132–133, 1972.

[Griffin and Colchester, 1994] L.D. Griffin and A.C.F. Colchester. The intrinsic geometry of the cerebral cortex. *Journal of Theoretical Biology*, 166:261–273, 1994.

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

[Griffin, 1995] L.D. Griffin. *Description of image structure*. PhD thesis, University of London, 1995.

[Grzeszczuk and Levin, 1994] R. Grzeszczuk and D.N. Levin. Brownian strings: segmenting images with stochastically deformable contours. In R.A. Robb, editor, *Proc. 3rd International Conference on Visualization in Biomedical Computing (VBC'94)*, volume 2359, pages 72–89. SPIE, 1994.

[Harvey *et al.*, 1997] R. Harvey, J.A. Bangham, and A. Bosson. Scale-space filters and their robustness. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space'97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 341–344. Springer Verlag, 1997.

[Heap and Hogg, 1995] T. Heap and D. Hogg. Automated pivot location for the cartesian-polar hybrid point distribution model. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 97–106. BMVA, 1995.

[Horn and Schunk, 1981] B.K.P. Horn and B.G. Schunk. Determining optic flow. *Artificial Intelligence*, 17:185–203, 1981.

[Hough, 1962] P.V.C. Hough. Method and means for recognizing complex patterns. U.S. Patent No. 3069654, 1962.

[Howell *et al.*, 1993] G.W. Howell, D.W. Fausett, and L.V. Fausett. Quasi-circular splines: A shape-preserving approximation. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 55(2):89–97, 1993.

[Hu, 1962] M.K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, IT-8:179–187, 1962.

[Hussain, 1988] Z. Hussain. A fast approximation to a convex hull. *Pattern Recognition Letters*, 8:289–294, 1988.

[Huttenlocher *et al.*, 1993] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[Ivins and Porrill, 1993a] J. Ivins and J. Porrill. Everything you always wanted to know about snakes (but were afraid to ask). AIVRU Technical Memo 86, University of Sheffield, 1993.

[Ivins and Porrill, 1993b] J. Ivins and J. Porrill. Statistical snakes: active region models. AIVRU Technical Memo 89, University of Sheffield, 1993.

[Ivins and Porrill, 1994a] J. Ivins and J. Porrill. Statistical Snakes: Active Region Models. In E.R. Hancock, editor, *Proc. 5th British Machine Vision Conference (BMVC'94)*, volume 1-2, pages 377–386. BMVA, 1994.

[Ivins and Porrill, 1994b] J. Ivins and J. Porrill. Statistical snakes II: mathematical notes on active region models. AIVRU Technical Memo 92, University of Sheffield, 1994.

[Jackway and Deriche, 1996] P.T. Jackway and M. Deriche. Scale-space properties of the multiscale morphological erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):38–51, 1996.

[Jackway, 1996] P.T. Jackway. Gradient watersheds in morphological scale-space. *IEEE Transactions on Image Processing*, 5(6):913–937, 1996.

[Kapur *et al.*, 1996] T. Kapur, W.E.L. Grimson, W.M. Wells III, and R. Kikinis. Segmentation of brain tissue from magnetic resoncance images. *Medical Image Analysis*, 1(2):109–128, 1996.

[Kass *et al.*, 1987a] M. Kass, A. Witkin, and D. Terzopoulos. Snakes - active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1987.

[Kass *et al.*, 1987b] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *Proc. 1st International Conference on Computer Vision (ICCV'87)*, pages 259–268. IEEE Computer Society Press, 1987.

[Keller *et al.*, 1989] J.M. Keller, S. Chen, and R.M. Crownover. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45:150–166, 1989.

[Kimia and Siddiqi, 1996] B.B. Kimia and K. Siddiqi. Geometric heat equation and nonlinear diffusion of shapes and images. *Computer Vision and Image Understanding*, 64(3):305–322, 1996.

[Kimia *et al.*, 1995] B.B. Kimia, A.R. Tannenbaum, and S.W. Zucker. Shapes, shocks, and deformations I: The components of two-dimensional shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.

[Koenderink and van Doorn, 1990] J.J. Koenderink and A.J. van Doorn. Receptive-field families. *Biological Cybernetics*, 63(4):291–297, 1990.

[Koenderink, 1984] J.J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[Koenderink, 1990] J.J. Koenderink. *Solid Shape*. The MIT Press, 1990.

[Koster *et al.*, 1996] A.S.E. Koster, K.L. Vincken, C. De Graaf, O.C. Zander, and M.A. Viergever. Heuristic linking models in multiscale image segmentation. *Computer Vision and Image Understanding*, 64(3):382–402, 1996.

[Koster, 1995] A.S.E. Koster. *Linking models for multiscale image segmentation.* PhD thesis, Utrecht University, The Netherlands, 1995.

[Leavers, 1992] V.F. Leavers. *Shape detection in computer vision using the Hough transform.* Springer Verlag, 1992.

[LeGoualher *et al.*, 1996] G. LeGoualher, C. Barillot, Y.J. Bizais, and J.-M. Scarabin. Three-dimensional segmentation of cortical sulci using active models. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 254–263. SPIE, 1996.

[Leitner *et al.*, 1990] F. Leitner, I. Marque, S. Lavallée, and P. Cinquin. Dynamic segmentation: finding the edge with snake splines. In *International Conference on Curves and Surfaces (ICCS).* Academic Press, 1990.

[Lepard and Robb, 1996] K.O. Lepard and R.A. Robb. Shape-based segmentation and characterization of biomedical images. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 232–242. SPIE, 1996.

[Lester and Arridge, 1997] H. Lester and S.R. Arridge. Summarising fluid registration by thin-plate spline warps with many landmarks. In C.J. Taylor, J.A. Noble, and J.M. Brady, editors, *Proc. 1st Annual Conference on Medical Image Understanding and Analysis '97*, pages 53–56. BMVA, 1997.

[Levoy, 1989] M. Levoy. *Display of surfaces from volume data.* PhD thesis, University of North Carolina at Chapel Hill, 1989.

[Lifshitz and Pizer, 1990] L.M. Lifshitz and S.M. Pizer. A multiresolution hierarchical approach to image segmentation based on intensity extrema. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(6):529–540, 1990.

[Lindeberg, 1993a] T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch - a method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318, 1993.

[Lindeberg, 1993b] T. Lindeberg. On scale selection for differential operators. In *8th Scandinavian Conference on Image Analysis*, pages 857–866, 1993.

[Lindeberg, 1994] T. Lindeberg. *Scale-space theory in computer vision*, volume 256 of *Engineering and Computer Science.* Kluwer Academic Publishers, 1994.

[Little *et al.*, 1996] J.A. Little, D.L.G. Hill, and D.H. Hawkes. Constraining rigid structures within a landmark based registration. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 170–177. Leeds University Press, 1996.

[Lobregt and Viergever, 1995] S. Lobregt and M.A. Viergever. A discrete dynamic contour model. *IEEE Transactions on Medical Imaging*, 14(1):12–24, 1995.

[Losseff *et al.*, 1996a] N.A. Losseff, W.I. MacDonald, D.H. Miller, and A.J. Thompson. Spinal-cord atrophy - a sensitive and reproducible measure of disease progression in Multiple-Sclerosis. *Neurology*, 46(2):70001, 1996.

[Losseff *et al.*, 1996b] N.A. Losseff, S.L. Webb, J.I. O'Riordan, R. Page, L. Wang, G.J. Barker, P.S. Tofts, W.I. MacDonald, D.H. Miller, and A.J. Thompson. Spinal cord atrophy and disability in Multiple-Sclerosis: a new reproducible and sensitive MRI method with potential to monitor disease progression. *Brain*, 119(3):701–708, 1996.

[Ma and Manjunath, 1997] W.Y. Ma and B.S. Manjunath. Edge flow: a framwework of boundary detection and image segmentation. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pages 744–749. IEEE Computer Society Press, 1997.

[Maes *et al.*, 1995] F. Maes, D. Vandermeulen, P. Suetens, and G. Marchal. Automatic image partitioning for generic object segmentation in medical images. In Y. Bizais, C. Barillot, and R. DiPaola, editors, *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI'95)*, Computational Imaging and Vision, pages 215–226. Kluwer Academic Publishers, 1995.

[Maintz *et al.*, 1996a] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Comparison of edge-based and ridge-based registration of CT and MR brain images. *Medical Image Analysis*, 1(2):151–161, 1996.

[Maintz *et al.*, 1996b] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):353–365, 1996.

[Maintz, 1996] J.B.A. Maintz. *Restrospective registration of tomographic images.* PhD thesis, Utrecht University, The Netherlands, 1996.

[Malladi *et al.*, 1995] R. Malladi, J.A. Sethian, and B.C. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):153–175, 1995.

[Malladi *et al.*, 1996] R. Malladi, R. Kimmel, D. Adalseinssohn, G. Sapiro, V. Caselles, and J.A. Sethian. A geometric approach to segmentation and analysis of 3D medical images. In *Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis MMBIA'96*. IEEE Computer Society Press, 1996.

[Mandelbrot, 1982] B.B. Mandelbrot. *The Fractal Geometry of Nature.* W.H. Freeman and Co., 1982.

[Maragos, 1996] P. Maragos. Differential morphology in image processing. *IEEE Transactions on Image Processing*, 5(6):822–937, 1996.

[Marr and Hildreth, 1980] D. Marr and E. Hildreth. A theory of edge detection. In *Royal Society*, volume B221, pages 187–217, 1980.

[Marr, 1982] D. Marr. *Vision: a computational investigation into the human representation and processing of visual information.* W.H. Freeman and Co., 1982.

[McAuliffe *et al.*, 1996] M.J. McAuliffe, D. Eberly, D.S. Fritsch, E.L. Chaney, and S.M. Pizer. Scale-space boundary evolution initialized by cores. In K.H. Höhne and R. Kikinis, editors, *Proc. 4th International Conference on Visualization in Biomedical Computing (VBC'96)*, volume 1131 of *Lecture Notes in Computer Science*, pages 173–182. Springer Verlag, 1996.

[McInerney and Terzopoulos, 1995a] T. McInerney and D. Terzopoulos. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *IEEE Transactions on Computerized Medical Imaging and Graphics*, 19(1):69–83, 1995.

[McInerney and Terzopoulos, 1995b] T. McInerney and D. Terzopoulos. Medical image analysis with topologically adaptive snakes. In *Proc. First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, 1995.

[McInerney and Terzopoulos, 1995c] T. McInerney and D. Terzopoulos. Topologically adaptable snakes. In *Proc. 5th International Conference on Computer Vision (ICCV'95)*, pages 840–845. IEEE Computer Society Press, 1995.

[McInerney and Terzopoulos, 1996] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[McQueen and Toussaint, 1985] M.M. McQueen and G.T. Toussaint. On the ultimate convex hull algorithm in practice. *Pattern Recognition Letters*, 3:29–34, 1985.

[Meencke, 1994] H.J. Meencke. Minimal developmental disturbances in epilepsy and MRI. In S.D. Shorvon et al., editor, *Advanced MR and Epilepsy*. Pergamon Press, 1994.

[Menet *et al.*, 1990a] S. Menet, P. Saint-Marc, and G. Medioni. Active contour models - overview, implementation and applications. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 194–199, 1990.

[Menet *et al.*, 1990b] S. Menet, P. Saint-Marc, and G. Medioni. B-snakes: implementation and application to stereo. In *DARPA Image Understanding Workshop*, 1990.

[Miller *et al.*, 1991] J.V. Miller, D.E. Breen, W.E. Lorensen, R.M. O'Bara, and M.J. Wozny. Geometrically deformed models: A method for extracting closed geometric models from volume data. In T.W. Sederberg, editor, *Proc. SIGGRAPH'91*, volume 25(4), pages 217–226. ACM, 1991.

[Mokhtarian and Mackworth, 1986] F. Mokhtarian and A.K. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):34–43, 1986.

[Mokhtarian and Mackworth, 1992] F. Mokhtarian and A.K. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):789–805, 1992.

[Mokhtarian *et al.*, 1996] F. Mokhtarian, S. Abbasi, and J. Kittler. Robust and effcient shape indexing through curvature scale space. In R.B. Fisher and E. Trucco, editors, *Proc. 7th British Machine Vision Conference (BMVC'96)*, volume I, pages 53–62. BMVA, 1996.

[Mokhtarian, 1995a] F. Mokhtarian. Convergence properties of curvature scale space representations. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 357–366. BMVA, 1995.

[Mokhtarian, 1995b] F. Mokhtarian. Silhouette-based isolated object recognition through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):539–544, 1995.

[Mokhtarian, 1996] F. Mokhtarian. Silhouette-based object recognition with occlusion through curvature scale space. In B. Buxton and R. Cippola, editors, *Computer Vision – Proc. 4th European Conference on Computer Vision (ECCV'96)*, volume 1064 of *Lecture Notes in Computer Science*, pages 566–578. Springer Verlag, 1996.

[Mokhtarian, 1997a] F. Mokhtarian. Multi-scale contour segmentation. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space'97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 296–307. Springer Verlag, 1997.

[Mokhtarian, 1997b] F. Mokhtarian. A theory of multiscale, torsion-based shape representation for space curves. *Computer Vision and Image Understanding*, 68(1):1–17, 1997.

[Morse, 1994] B.S. Morse. *Computation of Object Cores from Grey-Level Images*. PhD thesis, University of North Carolina at Chapel Hill, 1994.

[Nair and Saunders (Jr.), 1996] P.S. Nair and A.T. Saunders (Jr.). Hough based ellipse detection algorithm. *Pattern Recognition Letters*, 17:777–784, 1996.

[Najman and Schmitt, 1996] L. Najman and M. Schmitt. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12):1163–1173, 1996.

[Nastar and Ayache, 1994] C. Nastar and N. Ayache. Classification of nonrigid motion in 3D images using physics-based vibration analysis. In *IEEE Workshop on Biomedical Image Analysis*, pages 61–69. IEEE Computer Society Press, 1994.

[Niessen *et al.*, 1996] W.J. Niessen, K.L. Vincken, and M.A. Viergever. Comparison of multi-scale representations for a linking-based image segmentation model. In *Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis MMBIA '96*, pages 263–272. IEEE Computer Society Press, 1996.

[Niessen *et al.*, 1997] W.J. Niessen, B.M. ter Haar Romeny, L.M.J. Florack, and M.A. Viergever. A general framework for geometry-driven diffusion equations. *International Journal of Computer Vision*, 21(3):187–205, 1997.

[Niessen, 1997] W.J. Niessen. *Multiscale Medical Image Analysis*. PhD thesis, Utrecht University, The Netherlands, 1997.

[Ogniewicz, 1994] R.L. Ogniewicz. Skeleton-space: a multiscale shape description combining region and boundary information. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR '94)*, pages 746–751. IEEE Computer Society Press, 1994.

[O'Rourke, 1986] J. O'Rourke. The signature of a plane curve. *SIAM Journal on Computing*, 15:34–51, 1986.

[Papoulis, 1965] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill, 1965.

[Park and Lee, 1996] K.-R. Park and C.-N. Lee. Scale-space using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1121–1126, 1996.

[Parker *et al.*, 1998] G.J.M. Parker, J.A. Schnabel, and G.J. Barker. Identification of spatially random signal intensity using Approximate Entropy (ApEn) - a method for image noise reduction. In *Proc. 6th Meeting of the International Society for Magnetic Resonance in Medicine*, page 2083, 1998. To appear.

[Pavlidis, 1977] T. Pavlidis. Polygonal approximations by Newton's method. *IEEE Transactions on Computing*, C-26:800–807, 1977.

[Peleg *et al.*, 1994] S. Peleg, J. Naor, R. Hartley, and D. Avnir. Multiple resolution texture analysis and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(4):518–523, 1994.

[Penn and Loew, 1996] A.I. Penn and M.H. Loew. Estimating fractal dimension of medical images. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 840–851. SPIE, 1996.

[Pentland, 1984] A.P. Pentland. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):661–674, 1984.

[Perona and Malik, 1990] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[Persoon and Fu, 1986] E. Persoon and K.-S. Fu. Shape discrimination using fourier descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(3):388–397, 1986.

[Phillips and Rosenfeld, 1987] T. Y. Phillips and A. Rosenfeld. A method of curve partitioning using arc-chord distance. *Pattern Recognition Letters*, 5:285–288, 1987.

[Poon *et al.*, 1994a] C.S. Poon, M. Braun, R. Fahrig, and A. Ginige. A multi-feature integrated approach to MR image segmentation. In *Proc. Society of Magnetic Resonance (SMR '94)*, page 61, 1994.

[Poon *et al.*, 1994b] C.S. Poon, M. Braun, R. Fahrig, A. Ginige, and A. Dorrell. Segmentation of medical images using an active contour model incorporating region-based image features. In R.A. Robb, editor, *Proc. 3rd International Conference on Visualization in Biomedical Computing (VBC'94)*, volume 2359, pages 90–97. SPIE, 1994.

[Press *et al.*, 1992] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C - The Art of Scientific Programming.* Bridge, second edition, 1992.

[Prokop and Reeves, 1992] R.J. Prokop and A.P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 54(5):438–460, 1992.

[Rattarangsi and Chin, 1992] A. Rattarangsi and R.T. Chin. Scale-based detection of corners of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(4):430–449, 1992.

[Reinhardt and Higgins, 1996] J. M. Reinhardt and W. E. Higgins. Comparison between the morphological skeleton and morphological shape decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):951–957, 1996.

[Rougon and Prêteux, 1993a] N. Rougon and F. Prêteux. Directional adaptive deformable models for segmentation with application to 2D and 3D medical images. Technical report, Télécom Paris, Département Images, 1993.

[Rougon and Prêteux, 1993b] N.F. Rougon and F. Prêteux. Directional adaptive deformable models for segmentation with application to 2D and 3D medical images. In M.H. Loew, editor, *Proc. SPIE Medical Imaging 1993: Image Processing*, volume 1898, pages 193–207. SPIE, 1993.

[Rougon, 1993] N. Rougon. On mathematical foundations of local deformations analysis. In J.N. Wilson and D.C. Wilson, editors, *Proc. Mathematical Methods in Medical Imaging II*, volume 2035, pages 2–13. SPIE, 1993.

[Rueckert and Burger, 1995a] D. Rueckert and P. Burger. Contour fitting using an adaptive spline model. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 207–216. BMVA, 1995.

[Rueckert and Burger, 1995b] D. Rueckert and P. Burger. Contour fitting using stochastic and probabilistic relaxation for cine MR images. In H.U. Lemke, K. Inamura, C.C. Jaffe, and R. Felix, editors, *Proc. Computer Assisted Radiology (CAR'95)*, pages 137–142. Springer Verlag, 1995.

[Rueckert and Burger, 1997] D. Rueckert and P. Burger. Geometrically deformable templates for shape-based segmentation and tracking in cardiac MR images. In *Proc. International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR'97)*, Lecture Notes in Computer Science. Springer Verlag, 1997.

[Rueckert *et al.*, 1997] D. Rueckert, P. Burger, S. M. Forbat, R. D. Mohiaddin, and G. Z. Yang. Automatic tracking of the aorta in cardiovascular MR images using deformable models. *IEEE Transactions on Medical Imaging*, 16(5):581–590, 1997.

[Rueckert, 1993] D. Rueckert. Bildsegmentierung durch stochastisch optimierte Relaxation eines geometric deformable models. Diplomarbeit (MSc. thesis), Technische Universität Berlin, FG Computer Graphics, FB Computer Science, 1993.

[Ruff *et al.*, 1996] C.F. Ruff, A. Bhalerao, S.W. Hughes, T.J. D'Arcy, and D.J. Hawkes. The estimation of fetal organ volume using statistical shape analysis. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G. Farman, editors, *Proc. Computer Assisted Radiology (CAR'96)*, pages 280–285. Elsevier Publishers, 1996.

[Saint-Exupéry, 1946] A. de Saint-Exupéry. *Le Petit Prince*. Gallimard, 1946.

[Saint-Marc *et al.*, 1991] P. Saint-Marc, J.S. Chen, and G. Medioni. Adaptive smoothing - a general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):514–529, 1991.

[Saint-Marc *et al.*, 1993] P. Saint-Marc, H. Rom, and G. Medioni. B-spline contour representation and symmetry detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1191–1197, 1993.

[Sakar and Chaudhuri, 1992] N. Sakar and B.B. Chaudhuri. An efficient approach to estimate fractal dimension of textural images. *Pattern Recognition*, 25(9):1035–1041, 1992.

[Samadani, 1991] R. Samadani. Adaptive snakes: control of damping and material parameters. In C. Vemuri, editor, *Proc. SPIE on Geometric Models in Computer Vision*, volume 1570, pages 202–213. SPIE, 1991.

[Samarabandu *et al.*, 1993] J. Samarabandu, R. Acharya, E. Hausmann, and K. Allen. Analysis of bone X-rays using morphological fractals. *IEEE Transactions on Medical Imaging*, 12(3):466–470, 1993.

[Sanchez-Ortiz *et al.*, 1996a] G.I. Sanchez-Ortiz, D. Rueckert, and P. Burger. Knowledge-based anisotropic diffusion of vector-valued 4-dimensional cardiac MR images. In R.B. Fisher and E. Trucco, editors, *Proc. 7th British Machine Vision Conference (BMVC'96)*, volume 2, pages 605–614. BMVA, 1996.

[Sanchez-Ortiz *et al.*, 1996b] G.I. Sanchez-Ortiz, D. Rueckert, and P. Burger. Motion and deformation analysis of the heart using thin-plate splines and density and velocity encoded MR images. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 71–78. Leeds University Press, 1996.

[Sanchez-Ortiz *et al.*, 1997] G.I. Sanchez-Ortiz, D. Rueckert, and P. Burger. Knowledge-based anisotropic diffusion of cardiac cine MR images. In C.J. Taylor, J.A. Noble, and J.M. Brady, editors, *Proc. 1st Annual Conference on Medical Image Understanding and Analysis'97*, pages 89–96. BMVA, 1997.

[Sapiro and Ringach, 1996] G. Sapiro and D.L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5(11):1582–1586, 1996.

[Sapiro and Tannenbaum, 1993] G. Sapiro and A. Tannenbaum. Affine invariant scale-space. *International Journal of Computer Vision*, 11(1):25–44, 1993.

[Schnabel and Arridge, 1995] J.A. Schnabel and S.R. Arridge. Active contour models for shape description using multi-scale differential invariants. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 197–206. BMVA, 1995.

[Schnabel and Arridge, 1996a] J.A. Schnabel and S.R. Arridge. Hierarchical shape description of MR brain images based on active contour models and multi-scale differential invariants. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 36–43. Leeds University Press, 1996.

[Schnabel and Arridge, 1996b] J.A. Schnabel and S.R. Arridge. Multi-scale shape description of MR brain images using active contour models. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 596–606. SPIE, 1996.

[Schnabel and Arridge, 1996c] J.A. Schnabel and S.R. Arridge. A multi-scale shape description tool for 3D MR brain images. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G.

Farman, editors, *Proc. Computer Assisted Radiology (CAR'96)*, pages 292–297. Elsevier Publishers, 1996.

[Schnabel and Arridge, 1997a] J.A. Schnabel and S.R. Arridge. Active shape focusing. In C. Arcelli, L.P. Cordella, and G. Sanniti di Baja (eds.), editors, *Advances in Visual Form Analysis - Third International Workshop on Visual Form*, pages 492–501. World Scientific, 1997.

[Schnabel and Arridge, 1997b] J.A. Schnabel and S.R. Arridge. Active shape focusing. *Image and Vision Computing*, 1997. Special issue of the 3rd International Workshop on Visual Form, 28-30 May 1997, Capri, Italy. Invited paper for submission. Submitted.

[Schnabel and Arridge, 1997c] J.A. Schnabel and S.R. Arridge. Multi-scale active shape description. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space'97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 337–340. Springer Verlag, 1997.

[Schnabel and Arridge, 1997d] J.A. Schnabel and S.R. Arridge. Multi-scale shape description in medical imaging. In C.J. Taylor, J.A. Noble, and J.M. Brady, editors, *Proc. 1st Annual Conference on Medical Image Understanding and Analysis'97*, pages 141–144. BMVA, 1997.

[Schnabel *et al.*, 1995] J.A. Schnabel, L. Wang, D. Rueckert, and S.R. Arridge. Fractal feature analysis and classification of MS lesions. In H.U. Lemke, K. Inamura, C.C. Jaffe, and M.V. Vannier, editors, *Proc. Computer Assisted Radiology (CAR'95)*, pages 260–265. Springer Verlag, 1995.

[Schnabel *et al.*, 1996] J.A. Schnabel, L. Wang, and S.R. Arridge. Shape description of spinal cord atrophy in MS patients. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G. Farman, editors, *Proc. Computer Assisted Radiology (CAR'96)*, pages 286–291. Elsevier Publishers, 1996.

[Schnabel, 1993a] J.A. Schnabel. Die lokale fraktale Dimension als Segmentierungsmerkmal: Realisation und Evaluierung von regionen- und graph-orientierten kantenbasierten Verfahren. Diplomarbeit (MSc. thesis), FG Computer Graphics, FB Computer Science, Technische Universität Berlin, 1993.

[Schnabel, 1993b] J.A. Schnabel. Kantenverstärkung durch Auswertung der lokalen fraktalen Dimension eines Bildes. Studienarbeit (First thesis), FG Computer Graphics, FB Computer Science, Technische Universität Berlin, 1993.

[Schnabel, 1994] J.A. Schnabel. Edge detection using the local fractal dimension. In K.V. Mardia and C.A. Gill, editors, *14th Leeds Annual Statistical Research (LASR) Workshop: Medical Imaging: Shape and Visualisation*, pages 20–21. University of Leeds, Department of Statistics and CoMIR, 1994.

[Schnabel, 1995] J.A. Schnabel. Shape description methods for medical images. Technical report TR/95/12, Department of Computer Science, University College London, 1995.

[Schroeder *et al.*, 1997] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit.* Prentice Hall, second edition, 1997.

[Sclaroff, 1995] S.E. Sclaroff. *Modal matching: a method for describing, comparing, and manipulating digital signals.* PhD thesis, Massachusetts Institute of Technology, 1995.

[Sethian, 1996] J.A. Sethian. *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Materials Science.* Cambridge Monograph on Applied and Computational Mathematics. Cambridge University Press, 1996.

[Sheehy *et al.*, 1996] D.J. Sheehy, C.G. Armstrong, and D.J. Robinson. Shape description by medial surface construction. *IEEE Transactions on Visualization and Computer Graphics*, 2(1):62–72, 1996.

[Sherbrooke *et al.*, 1996] E.C. Sherbrooke, N.M. Patrikalakis, and E. Brisson. An algorithm for medial axis transform of 3D polyhedral solids. *IEEE Transactions on Visualization and Computer Graphics*, 2(1):45–59, 1996.

[Simmons, 1992] A. Simmons. *Segmentation of Neuroanatomy in Magnetic Resonance Images*. PhD thesis, Department of Medical Physics and Bioengineering, University College London, University of London, UK, 1992.

[Sisodiya *et al.*, 1995] S.M. Sisodiya, S.L. Free, J.M. Stevens, D.R. Fish, and S.D. Shorvon. The demonstration of gyral abnormalities in patients with cryptogenic partial epilepsy using three-dimensional mri. *Archives of Neurology*, 53:28–34, 1995.

[Sisodiya *et al.*, 1996] S.M. Sisodiya, J.M. Stevens, S.L. Free, D.R. Fish, and S.D. Shorvon. Widespread cerebral structural changes in patients with cortical dysgenesis and epilepsy. *Brain*, 118:1039–1050, 1996.

[Sjöberg and Bergholm, 1988] F. Sjöberg and F. Bergholm. Extraction of diffuse edges by edge focusing. *Pattern Recognition Letters*, 7(3):181–190, 1988.

[Sluzek, 1996] A. Sluzek. Identification and inspection of 2-D objects using new moment-based shape-descriptors. *Pattern Recognition Letters*, 16:687–697, 1996.

[Smith and Jain, 1982] S.P. Smith and A.K. Jain. Chord distribution for shape matching. *Computer Graphics and Image Processing*, 20(3):259–271, 1982.

[Spivak, 1979] M. Spivak. *Differential Geometry*, volume 3. Publish or Perish Inc., second edition, 1979.

[Staib and Duncan, 1992] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, 1992.

[Storvik, 1994] G. Storvik. A Bayesian-approach to dynamic contours through stochastic sampling and Simulated Annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):976–986, 1994.

[Székely *et al.*, 1996] G. Székely, A. Kelemen, C. Brechbühler, and G. Gerig. Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Medical Image Analysis*, 1(1):19–34, 1996.

[Szeliski and Lavallée, 1994] R. Szeliski and S. Lavallée. Matching 3-D anatomical surfaces with non-rigid deformations using octree-splines. In *IEEE Workshop on Biomedical Image Analysis*, pages 144–153. IEEE Computer Society Press, 1994.

[Tek and Kimia, 1994] H. Tek and B.B. Kimia. Shock-based reaction-diffusion bubbles for image segmentation. TR LEMS-138, Brown University, Providence, RI, 1994.

[Tek and Kimia, 1995a] H. Tek and B.B. Kimia. Automatic volumetric segmentation of three-dimensional images. In H.U. Lemke, K. Inamura, C.C. Jaffe, and M.V. Vannier, editors, *Proc. Computer Assisted Radiology (CAR'95)*, pages 164–170. Springer Verlag, 1995.

[Tek and Kimia, 1995b] H. Tek and B.B. Kimia. Image segmentation by reaction-diffusion bubbles. In *Proc. 5th International Conference on Computer Vision (ICCV'95)*, pages 156–162. IEEE Computer Society Press, 1995.

[Tek and Kimia, 1997] H. Tek and B.B. Kimia. Volumetric segmentation of medical images by three-dimensional bubbles. *Computer Vision and Image Understanding*, 65(2):246–258, 1997.

[ter Haar Romeny *et al.*, 1991] B.M. ter Haar Romeny, L.M.J. Florack, J.J. Koenderink, and M.A. Viergever. Scale-space: Its natural operators and differential invariants. In A.C.F. Colchester and D.J. Hawkes, editors, *Information Processing in Medical Imaging: Proc. 12th International Conference (IPMI'91)*, volume 511 of *Lecture Notes in Computer Science*, pages 239–255. Springer Verlag, 1991.

[ter Haar Romeny *et al.*, 1993] B.M. ter Haar Romeny, L.M.J. Florack, A.H. Salden, and M.A. Viergever. Higher order differential structure of images. In H.H. Barret and A.F. Gmitro, editors, *Information Processing in Medical Imaging: Proc. 13th International Conference (IPMI'93)*, volume 687 of *Lecture Notes in Computer Science*, pages 77–93. Springer Verlag, 1993.

[ter Haar Romeny, 1996] B.M. ter Haar Romeny. Introduction to scale-space theory: multiscale geometric image analysis. Tutorial VBC'96, Hamburg, Germany ICU-96-21, Utrecht University, The Netherlands, 1996.

[ter Haar Romeny, 1997] B.M. ter Haar Romeny. *Front-end vision and multiscale image analysis: Introduction to scale-space theory.* Computational Imaging and Vision. Kluwer Academic Publishers, 1997. In preparation.

[Terzopoulos, 1986a] D. Terzopoulos. Image analysis using multigrid relaxation methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):129–139, 1986.

[Terzopoulos, 1986b] D. Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424, 1986.

[Terzopoulos, 1996] D. Terzopoulos. Deformable models and the analysis of medical images. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 194–201. Leeds University Press, 1996.

[Thornton *et al.*, 1997] J. Thornton, P. Amess, J. Penrice, K. Chong, J.S. Wyatt, and R.J. Ordidge. Quantitative cerebral T2 relaxometry in human neonates at 2.4 Tesla: methodology and the effects of maturation. In *Proc. European Society of Magnetic Resonance in Medicine and Biology*, 1997.

[Toennies and Rueckert, 1994] K.D. Toennies and D. Rueckert. Image segmentation by stochastically relaxing contour fitting. In M.H. Loew, editor, *Proc. SPIE Medical Imaging 1994: Image Processing*, volume 2167, pages 18–27. SPIE, 1994.

[Toennies and Schnabel, 1994] K.D. Toennies and J.A. Schnabel. Edge detection using the local fractal dimension. In *IEEE Symposium on Computer-Based Medical Systems*, pages 34–39. IEEE Computer Society Press, 1994.

[Toussaint, 1985] G.T. Toussaint. A historical note on convex hull finding algorithms. *Pattern Recognition Letters*, 3:21–28, 1985.

[Unser *et al.*, 1993] M. Unser, A. Aldroubi, and M. Eden. The L(2) polynomial spline pyramid. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4):364–379, 1993.

[van den Boumgard and Smeulders, 1994] R. van den Boumgard and A. Smeulders. The morphological structure of images - the differential equations of morphological scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1101–1113, 1994.

[van den Boumgard, 1997] R. van den Boumgard. Affine invariant deformation curves - a tool for shape characterization. In C. Arcelli, L.P. Cordella, and G. Sanniti di Baja (eds.), editors, *Advances in Visual Form Analysis - Third International Workshop on Visual Form*, pages 594–603. World Scientific, 1997.

[van den Elsen *et al.*, 1995] P.A. van den Elsen, J.B. Maintz, E.-J.D. Pol, and M.A. Viergever. Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Transactions on Medical Imaging*, 14(2):384–396, 1995.

[Vincken *et al.*, 1996] K.L. Vincken, W.J. Niessen, and M.A. Viergever. Blurring strategies for image segmentation using a multiscale linking model. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR'96)*, volume Conference C: Image, Speech and Signal Analysis, pages 21–26. IEEE Computer Society Press, 1996.

[Vincken *et al.*, 1997] K.L. Vincken, A.S.E. Koster, and M.A. Viergever. Probabilistic multiscale image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):109–120, 1997.

[Vincken, 1995] K.L. Vincken. *Probabilistic multiscale image segmentation by the hyperstack.* PhD thesis, Utrecht University, The Netherlands, 1995.

[Voss, 1985] R.F. Voss. Random fractal forgeries. In R.A. Earnshaw, editor, *Fundamental Algorithms for Computer Graphics*, pages 805–836. Springer Verlag, 1985.

[Wahba, 1990] G. Wahba. *Spline models for observational data.* Society for Industrial and Applied Mathematics, 1990.

[Wang and Cohen, 1994] J.-Y. Wang and F.S. Cohen. Part II: 3-D object recognition and shape estimation from image contours using B-splines, shape invariant matching, and neural network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1):13–23, 1994.

[Wang *et al.*, 1996] M. Wang, J. Evans, L. Hassebrook, and C. Knapp. A multistage, optimal active contour model. *IEEE Transactions on Image Processing*, 5(11):1586–1591, 1996.

[Whitaker and Chen, 1994] R.T. Whitaker and D.T. Chen. Embedded active surfaces for volume visualization. In M.H. Loew, editor, *Proc. SPIE Medical Imaging 1994: Image Processing*, volume 2167, pages 340–352. SPIE, 1994.

[Whitaker and Gerig, 1994] R. Whitaker and G. Gerig. Vector-valued diffusion. In B.M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, chapter 4, pages 93–134. Kluwer Academic Publishers, 1994.

[Whitaker and Pizer, 1993] R.T. Whitaker and S.M. Pizer. A multiscale approach to nonuniform diffusion. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 57(1):99–110, 1993.

[Whitaker, 1993] R.T. Whitaker. Geometry-limited diffusion in the characterization of geometric patches of images. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 57(1):111–120, 1993.

[Whitaker, 1994a] R. Whitaker. Volumetric deformable models: active blobs. In R.A. Robb, editor, *Proc. 3rd International Conference on Visualization in Biomedical Computing (VBC'94)*, volume 2359, pages 121–134. SPIE, 1994.

[Whitaker, 1994b] R.T. Whitaker. *Geometry-Limited Diffusion.* PhD thesis, University of North Carolina at Chapel Hill, 1994.

[Williams and Shah, 1992] D.J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(1):14–26, 1992.

[Witkin *et al.*, 1987] A. Witkin, D. Terzopoulos, and M. Kass. Signal matching through scale space. *International Journal of Computer Vision*, 1(2):133–144, 1987.

[Witkin, 1983] A. Witkin. Scale–space filtering. In *Proc. International Joint Conference on Artificial Intelligence*, pages 1019–1022, 1983.

[Wright *et al.*, 1994] M.W. Wright, R. Cipolla, and P.J. Giblin. Skeletonisation using an extended Euclidean distance transform. Technical Report CUED/F-INFENG/TR 196, Department of Engineering, University of Cambridge, 1994.

[Wu and Wang, 1993] W.-Y. Wu and M.-J. Wang. Detecting the dominant points by the curvature–based polygonal approximation. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 55(2):79–88, 1993.

[Xu and Prince, 1997] C. Xu and J.L. Prince. Gradient vector flow: a new external forces for snakes. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 66–71. IEEE Computer Society Press, 1997.

[Yoo, 1996] T.S.-W. Yoo. *Image geometry through multiscale statistics*. PhD thesis, University of North Carolina at Chapel Hill, 1996.

[You and Jain, 1984] Z. You and A.K. Jain. Performance evaluation of shape matching via chord length distribution. *Computer Vision, Graphics, and Image Processing*, 28:185–198, 1984.

[Yuille and Poggio, 1986] A.L. Yuille and T.A. Poggio. Scaling theorems for zero-crossings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):15–25, 1986.

[Zhu and Yuille, 1996] S.C. Zhu and A.L. Yuille. Region competition: unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.

[Zijdenbos *et al.*, 1994] A.P. Zijdenbos, B.M. Dawant, R.A. Margolin, and A.C. Palmer. Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Transactions on Medical Imaging*, 13(4):716–724, 1994.

Publications

Journal Articles

- J.A. Schnabel and S.R. Arridge. Active shape focusing. *Image and Vision Computing*. Special issue of the 3rd International Workshop on Visual Form, 28-30 May 1997, Capri, Italy. Invited paper for submission. Submitted.
- G.J.M. Parker, J.A. Schnabel and G.J. Barker. Identification of spatially random signal intensity using Approximate Entropy (ApEn) - a method for image noise reduction. In preparation.

Articles in Conference Proceedings

- J.A. Schnabel and S.R. Arridge. Active shape focusing. In C. Arcelli, L.P. Cordella and G. Sanniti di Baja, editors, *Advances in Visual Form Analysis - Third International Workshop on Visual Form*, pages 492-501, Capri, Italy, 28-30 May 1997. World Scientific.
- J.A. Schnabel and S.R. Arridge. Hierarchical shape description of MR brain images based on active contour models and multi-scale differential invariants. In K.V. Mardia, C.A. Gill, and I.L. Dryden, editors, *16th Leeds Annual Statistical Research (LASR) Workshop: Image Fusion and Shape Variability Techniques*, pages 36–43. Leeds University Press, University of Leeds, Department of Statistics and CoMIR, 3-5 July 1996.
- J.A. Schnabel and S.R. Arridge. A multi-scale shape description tool for 3D MR brain images. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G. Farman, editors, *Proc. Computer Assisted Radiology (CAR'96)*, pages 292–297, Paris, France, 26-29 June 1996. Elsevier Publishers, The Netherlands.
- J.A. Schnabel, L. Wang, and S.R. Arridge. Shape description of spinal cord atrophy in MS patients. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G. Farman, editors, *Proc. Computer Assisted Radiology (CAR'96)*, pages 286–291, Paris, France, 26-29 June 1996. Elsevier Publishers, The Netherlands.
- J.A. Schnabel and S.R. Arridge. Multi-scale shape description of MR brain images using active contour models. In M.H. Loew and K.M. Hanson, editors, *Proc. SPIE Medical Imaging 1996: Image Processing*, volume 2710, pages 596–606, Newport Beach, California, USA, 1996. SPIE.
- J.A. Schnabel and S.R. Arridge. Active contour models for shape description using multiscale differential invariants. In D. Pycock, editor, *Proc. 6th British Machine Vision Conference (BMVC'95)*, volume 1, pages 197–206, Birmingham, UK, 11-14 September 1995. BMVA.
- J.A. Schnabel, L. Wang, D. Rueckert, and S.R. Arridge. Fractal feature analysis and classification of MS lesions. In H.U. Lemke, K. Inamura, C.C. Jaffe, and M.V. Vannier, editors,

Proc. Computer Assisted Radiology (CAR'95), pages 260–265, Berlin, Germany, 21-24 June 1995. Springer Verlag.

- K.D. Toennies and J.A. Schnabel. Edge detection using the local fractal dimension. In *IEEE Seventh Symposium on Computer–Based Medical Systems*, pages 34–39, Winston–Salem, North Carolina, 1994. IEEE Computer Society Press.

Short Papers and Abstracts in Conference Proceedings

- G.J.M. Parker, J.A. Schnabel and G.J. Barker. Identification of spatially random signal intensity using Approximate Entropy (ApEn) - a method for image noise reduction. In Proc. 6th Meeting of the International Society for Magnetic Resonance in Medicine, page 2083, Sydney, Australia, 18-24 April 1998. To appear.

- J.A. Schnabel and S.R. Arridge. Multi-scale shape description in medical imaging. In C.J. Taylor, J.A. Noble, and J.M. Brady, editors, *Proc. 1st Annual Conference on Medical Image Understanding and Analysis'97*, pages 141–144, University of Oxford, 7-8 July 1997. BMVA.

- J.A. Schnabel and S.R. Arridge. Multi-scale active shape description. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space'97 - 1st International Conference on Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 337–340, Utrecht, The Netherlands, 2-4 July 1997. Springer Verlag.

- J.A. Schnabel. Edge detection using the local fractal dimension. In K.V. Mardia and C.A. Gill, editors, *14th Leeds Annual Statistical Research (LASR) Workshop: Medical Imaging: Shape and Visualisation*, pages 20-21. Leeds University Press, University of Leeds, Department of Statistics and CoMIR, 11-12 July 1994.

Technical Reports and Theses

- J.A. Schnabel and S.R. Arridge. Scale-space implementation in the UK IUE project. Technical report TR/98/1, Department of Computer Science, University College London, Gower Street, UK - London WC1E 6BT, 1998.

- J.A. Schnabel, G. Bailey and S.R. Arridge. Volume rendering implementation in the UK IUE project. Technical report TR/98/2, Department of Computer Science, University College London, Gower Street, UK - London WC1E 6BT, 1998.

- J.A. Schnabel and S.R. Arridge. Snake implementation in the UK IUE project. Technical report TR/98/3, Department of Computer Science, University College London, Gower Street, UK - London WC1E 6BT, 1998.

- J.A. Schnabel. Shape description methods for medical images. Technical report TR/95/12, Department of Computer Science, University College London, Gower Street, UK - London WC1E 6BT, 1995.

- J.A. Schnabel. Die lokale fraktale Dimension als Segmentierungsmerkmal: Realisation und Evaluierung von regionen- und graph-orientierten kantenbasierten Verfahren. Diplomarbeit (MSc. thesis). FG Computer Graphics, FB Computer Science, Technische Universität Berlin, Franklinstr. 28/29, D-10587 Berlin. 1993.

- J.A. Schnabel. Kantenverstärkung durch Auswertung der lokalen fraktalen Dimension eines Bildes. Studienarbeit (First thesis). FG Computer Graphics, FB Computer Science, Technische Universität Berlin, Franklinstr. 28/29, D-10587 Berlin. 1993.

Glossary

$G(\mathbf{x};\sigma)$: Gaussian kernel	56
H : Hurst coefficient	47
$L(x,y)$: $2D$ image	144
$L(x,y,z)$: $3D$ image	144
$L(x,y,z;\sigma)$: $4D$ image scale-space	144
$L(x,y,z;\sigma_i)$: Sample of a $4D$ image scale-space	144
$L(x,y,z_k)$: Slice of a $3D$ image	144
$L(x,y,z_k;\sigma_i)$: Slice of a $4D$ image scale-space	144
$L(x,y;\sigma)$: $3D$ image scale-space	144
$L(x,y;\sigma_i)$: Sample of a $3D$ image scale-space	144
$L(\mathbf{x})$: *N-D* image	144
$L(\mathbf{x};\sigma)$: *(N+1)-D* image scale-space	144
$L(\mathbf{x};\sigma_i)$: Sample of an *(N+1)-D* image scale-space	144
$\Delta\tau$: Numerical diffusion time step	58
Γ : Graph representation of an image	67
δ_{ij} : Kronecker tensor	59
ϵ : Hidden scale	69
ϵ_{ij} : Lévi-Civita tensor	59
γ_n : Morphological closing operator	68
μ_{pq} : Cartesian moment of order $(p+q)$	38
ν_{pq} : Normalized central moments of order $(p+q)$	39
ω_n : Morphological opening operator	68
$\sigma(t)$: Scale recipe	64
σ_0 : Fixed-scale parameter	69
τ : Diffusion iteration step	58
$\tilde{\kappa}$: Normalized curvature	129
$\tilde{\sigma}$: Natural scale parameter	69
ς : Contour scale	114
$\{L_{z_k}(x,y;\sigma)\}$: $3\frac{1}{2}D$ or *slice-by-slice* image scale-space	144

$\{L_{z_k}(x, y; \sigma_i)\}$: sample of a $3\frac{1}{2}D$ image scale-space 144

$\{\tilde{\mathbf{v}}_{z_k}(s)\}$: Set of active contour models for slice-by-slice segmentation 145

$f(x, y)$: Implicit form of a curve 29

k: Conductance parameter 63

s : Arc length parameter 29

t : Diffusion time parameter, or evolution time, defined as $\Delta\tau \cdot \tau$ 57

$y(x)$: Explicit form of a curve 29

$\mathbf{v}(s)$: Parametric form of a curve 29

$\mathbf{v}(s, r)$: Parametric form of a surface 82

$\mathbf{v}_i$: i^{th} snaxel of a contour 82

$\mathbf{v}_s(s)$: First order curve derivative with respect to s 83

$\mathbf{v}_{ss}(s)$: Second order curve derivative with respect to s 83

$\mathcal{E}$: Continuous energy function 82

$\mathcal{E}^*$: Discrete energy function 82

$\overline{d}$: Mean distance between snaxels 83

BASOC : Boundary at the scale of the core 79

Compactness : Dimensionless quantity defined by $\frac{perimeter^2}{area}$ 45

CSF : Cerebral spinal fluid 213

CSS : Curvature scale-space 73

CT : Computer Tomography 20

DCP : Distance to the closest point algorithm 52

Diameter : Length of a medial axis 42

DP : Dynamic programming, a technique for solving variational problems 92

Edgel : Edge pixel in an image 40

fBm : Fractional Brownian motion 47

FEM : Finite element methods 242

FIF : Fractal interpolation function 48

FIM : Fractal interpolation with midpoints 48

FLAIR : Fluid Attenuated Inversion Recovery 21

Fractal dimension : Measure for self-similarity 46

GA : Genetic algorithm, a global optimization technique based on genetic operators 97

GDM : Geometrically Deformed Model 95

HMAT : Hough-like medial axis transform 78

HSV : Hue-Saturation-Value colour model 162

ICM : Iterated conditional modes, a probabilistic relaxation technique 95
ICP : Iterated closest point algorithm 53
IFS : Iterated function system (or fractal curve) 48

MDL : Minimum description length, a hierarchical technique for region grouping 77
MMA : Multi-scale medial axis, or *core* 78
MRF : Markov random field 94
MRI : Magnetic Resonance Imaging 20
MRT: Magnetic Resonance Tomography 20
MS : Multiple Sclerosis 21

PDM : Point distribution model 96
Perimeter : Length of a closed boundary 44
PET : Positron Emission Tomography 21
Pixel : Point in an image 23

RMS : Root-mean-squared error 131
ROI : Region of interest 213

SA : Simulated annealing, a stochastic relaxation technique 94
Snake : Active or deformable contour model for segmentation 81
Snaxel : Point of an active contour model 82
SPAMM : Spatial Amplitude Magnetization Modulation 21
SPECT : Single Photon Emission Computed Tomography 21

TSS : Torsion scale-space 250

Voxel : Volumetric image pixel 74

Index