

TEL AVIV UNIVERSITY

THE IBY AND ALADAR FLEISCHMAN FACULTY OF ENGINEERING

The Zandman-Slaner Graduate School of Engineering

Prior based Image Segmentation

By

Tamar Riklin-Raviv

THESIS SUBMITTED FOR THE DEGREE OF “DOCTOR OF PHILOSOPHY”
SUBMITTED TO THE SENATE OF TEL-AVIV UNIVERSITY

August 2007

TEL AVIV UNIVERSITY

THE IBY AND ALADAR FLEISCHMAN FACULTY OF ENGINEERING

The Zandman-Slaner Graduate School of Engineering

Prior based Image Segmentation

By

Tamar Riklin-Raviv

THESIS SUBMITTED FOR THE DEGREE OF “DOCTOR OF PHILOSOPHY”
SUBMITTED TO THE SENATE OF TEL-AVIV UNIVERSITY

This Research Work was Carried Out at Tel-Aviv University
The Faculty of Engineering

Under the Supervision of Prof. Nahum Kiryati

August 2007

This work was carried out under the supervision of

Prof. Nahum Kiryati

Acknowledgments

First and foremost I would like to express my deepest gratitude to my advisor Prof. Nahum Kiryati for his guidance and support throughout my PhD studies. The atmosphere in his group, his wide experience, his true scientific intuition and integrity significantly contributed to the success of this work and to me as a researcher.

I am deeply indebted to Dr. Nir Sochen for his invaluable contribution to this thesis. Nir introduced me to the world of calculus of variations and gave me a new and intriguing perspective on the challenging computer vision problems. I greatly appreciate his unlimited support and patience throughout this work.

I would like to thank Prof. Jonathan Nissanov, Dr. Smadar Gefen and Louise Bertrand from Drexel University. Their collaboration contributed to this thesis.

I would like to thank the present and the former VIA lab members at Tel-Aviv University for their friendship, encouragement and great tips - Amalia Potruch, Tomer Amiaz, Ruthy Katz, Leah Bar, Shay Rochel, Tal Darom, Anner Kushnir, Yan Ivanchenko. Special thanks to Nir Ben-Zadok with whom I had the pleasure to collaborate.

I would like to thank my dear and old friends from my days at the Hebrew University of Jerusalem - Amit Gruber and Jonathan Wexler for their long term friendship, encouragement and fruitful discussions.

I would also like to thank my friends and colleagues at Tel-Aviv University - Shiri Gordon, Sigal Trattner and Chen Sagiv for their friendship, empathy and fruitful con-

versations. Special thanks to Dr. Hayit Greenspan who contributed to my knowledge in medical image analysis.

I would like to acknowledge the A.M.N. foundation, the Yizhak and Chaya Weinstein Research Institute for Signal Processing and MUSCLE for the financial support.

I want to thank the School of Electrical Engineering at Tel-Aviv University that became my second home during my PhD studies, providing a friendly and supportive atmosphere.

Finally, I would like to express my love and gratitude to my dear husband, Alon Raviv, for his endless love, support and encouragement and to our lovely children, Gal and Omer for their joy and love.

Abstract

Object detection and segmentation can be facilitated by the availability of prior knowledge. This dissertation considers the incorporation of prior shape knowledge within a segmentation framework. The information about the expected shape of the object to extract is obtained from another image of the object. In the first part of the thesis we assume that the object boundaries in the prior image are known. The main challenge is accounting for projective transformations between the different object views. We address it by concurrent segmentation and registration processes. This is accomplished by the construction of a cost functional, where the dynamic variable is the object boundary represented by the zero level of a level-set function. The functional is optimized using calculus of variations.

Explicit shape prior is not always available. Consider the simultaneous segmentation of *two* object views. When neither of the images can be correctly segmented based on its edges and gray levels alone, the shape of the region extracted in either of them cannot be used as a reliable prior for the other. We therefore suggest an alternate minimization framework in which the evolving segmentation of each image provides a dynamic prior for the other. We call this process *mutual segmentation*.

When only a single image is given but the object taken is known to be symmetrical, the symmetry property forms a significant shape constraint and thus can be used to support segmentation. The third part of this thesis deals with the extraction of objects with either bilateral or rotational symmetry in the presence of perspective distortion. The key idea is the use of the *symmetrical counterpart* image obtained by a flip or rotation of the source image as another view of the object. The theoretical

foundation of the proposed method is a theorem, proven in this thesis, showing that symmetrical counterpart images are related by planar projective transformation.

The methods suggested are demonstrated on a variety of images that were taken in the presence of noise, shadows, occlusions or clutter. For each of the examples, accurate extraction of the object boundaries is shown together with the recovery of the planar projective transformation that relates the object views.

Some of the concepts developed are demonstrated on bio-medical applications. We show the delineation of uterine fibroids in MR images. Volumetric segmentation of mouse brain structures from histological data is also presented.

Contents

1	Introduction	1
1.1	Segmentation with prior knowledge	1
1.2	Variational approaches to segmentation	2
1.2.1	Level-set representation	4
1.2.2	Shape based segmentation	4
1.2.3	Two view geometry and contour correspondence	5
1.2.4	Implicit shape priors	5
1.3	Contributions and organization of the thesis	6
1.3.1	Level-set based segmentation	6
1.3.2	Segmentation with shape priors	7
1.3.3	Mutual segmentation	7
1.3.4	Segmentation with shape symmetry	8
1.3.5	Application: segmentation of mouse brain atlas	8
2	Level-set based Segmentation	10
2.1	Statistical set up and prior art	10
2.1.1	General principles	10
2.1.2	Region-based data term	12
2.1.3	Geodesic active contour: data part	13
2.1.4	Edge alignment term	15
2.1.5	Syntactic prior: geometry	15
2.1.6	Unified cost functional	17

2.2	Setting the weights of the energy terms	18
2.3	Example: Segmentation of uterine fibroids	19
2.3.1	General	19
2.3.2	Background and clinical motivation	20
2.3.3	Algorithm summary	21
2.3.4	Experimental results	21
2.4	Semantic prior: shape term	22
3	Segmentation with Shape Priors	25
3.1	Background and previous work	25
3.2	Variational framework	28
3.2.1	Chan-Vese two-phase model	28
3.3	Perspectivity	32
3.3.1	Cone of rays	32
3.3.2	Generalized cone	34
3.3.3	Formulation of the transformation	34
3.3.4	The six-parameter algorithm	37
3.4	From six to eight parameters	38
3.4.1	Planar projective homography	38
3.4.2	Implicit recovery of the homography	41
3.4.3	Explicit recovery of the homography	41
3.5	Objective functional	42
3.5.1	Representation of the prior shape	42
3.5.2	Recovery of the transformation	43
3.5.3	The eight-parameter algorithm	44
3.5.4	Optimization	45
3.6	Experimental results	46
3.7	Concluding remarks	51

4	Mutual Segmentation	60
4.1	Background and related works	60
4.2	Mutual segmentation with projectivity	62
4.2.1	Biased shape dissimilarity measure	63
4.2.2	Projective invariance	65
4.2.3	Unified cost functional	66
4.3	Implementation	66
4.3.1	Recovery of the transformation parameters	67
4.3.2	Algorithm	68
4.4	Experiments	69
4.5	Concluding remarks	71
5	Segmentation with Symmetry Constraint	79
5.1	Background and related works	79
5.2	Level sets framework and shape representation	83
5.2.1	Notation	83
5.2.2	Shape term	84
5.3	Symmetry and projectivity	86
5.3.1	Symmetry	86
5.3.2	Projectivity	87
5.3.3	Theoretical results	88
5.3.4	Examples	92
5.4	Symmetry based segmentation	95
5.4.1	Symmetry imperfection	95
5.4.2	Symmetry constraint	97
5.4.3	Biased shape dissimilarity measure	98
5.4.4	Recovery of the transformation	100
5.4.5	Unified segmentation functional	101
5.5	Implementation and further implications	101

5.5.1	Algorithm	101
5.5.2	Segmentation with multiple symmetrical counterparts	102
5.5.3	Segmentation with partial information	103
5.6	Experiments	104
5.7	Concluding remarks	106
6	Segmentation of Mouse Brain Atlas	110
6.1	Propagating distributions for the segmentation of mouse brain atlas .	110
6.2	Algorithm description	112
6.2.1	Level-set representation of the region of interest	112
6.2.2	The K-L divergence term	112
6.2.3	Unified cost functional	114
6.2.4	Slice by slice segmentation	114
6.3	Experiments	115
6.4	Concluding remarks	116
7	Discussion and Future Directions	120

List of Figures

2.1	Smooth approximation of the Heaviside function	12
2.2	The edge alignment constraint	16
2.3	MR image of uterine fibroids and the surrounding anatomy of the uterus	20
2.4	Initial level set function and the corresponding contour	21
2.5	MR images of various types of uterine fibroids and their segmentations	24
3.1	A comparison of the registration results using similarity, affine and perspective model	29
3.2	The cone of rays	33
3.3	A generalized cone	35
3.4	Segmentation results of a mannequin head image	39
3.5	The homography induced by a plane.	40
3.6	Failed segmenation and registration due to onvergence to a local minimum.	47
3.7	shape dissimilarity measure	47
3.8	Segmentation results of a synthetic image of the digit five	54
3.9	Segmentation results of a synthetic image	55
3.10	Segmentation results of a toy elephant image.	56
3.11	Segmentation results of the scissors image.	57
3.12	Segmention results of a bottle opener.	58
3.13	Segmentation results of Europe image taken from a satellite	59

4.1	The inherent labeling ambiguity demonstrated on two boot images.	62
4.2	A comparison of the mutual segmentation results using biased and unbiased dissimilarity measures.	73
4.3	Segmentation results of the hand images	74
4.4	Segmentation results of the boot images.	75
4.5	Segmentation results of the license plate images.	76
4.6	Segmentation results of the stop sign images.	77
4.7	Segmentation results of the chess images.	78
5.1	Shape transformation	83
5.2	The image of an approximately symmetrical object and its symmetrical counterpart.	83
5.3	An object with rotational symmetry and its symmetrical counterpart.	84
5.4	Examples of symmetrical object distorted by Euclidean and projective transformations.	94
5.5	Example of an object with rotational symmetry distorted by affine transformation.	95
5.6	Example of translational symmetry.	96
5.7	Segmentation results using biased and unbiased dissimilarity measures.	99
5.8	Segmentation results of a guitar.	104
5.9	Segmentation results of a butterfly image.	105
5.10	Segmentation results of a man-shadow image.	106
5.11	Segmentation results of a swan reflection image.	107
5.12	The extraction of a flower depicted with rotational symmetry.	109
6.1	Segmentation of the last slice in a 10 slice sequence.	117
6.2	Slice-by-slice segmentation captures the merge of two regions of interest	118
6.3	Slice-by-slice segmentation captures the split into two regions.	119

List of Tables

3.1	A comparison of the recovered and true transformation parameters for the digits image	49
3.2	A comparison of the recovered and true transformation parameters for the synthetic image.	50
3.3	A comparison of the homography matrix entries for the elephant image.	50
4.1	A comparison of the homography matrix entries	69
4.2	The means of the errors of the homography matrix entries	70
6.1	Correspondence between the slice-by-slice automatic segmentation and the expert segmentation.	116

List of Abbreviations

BU	Bottom Up
CT	Computed Tomography
GAC	Geodesic Active Contour
GMM	Gaussian Mixture Model
K-L	Kullback-Leibler
MBL	Mouse Brain Library
MR	Magnetic Resonance
MRgFU	Magnetic Resonance-guided Focused Ultrasound
M-S	Mumford Shah
MRF	Markov Random Field
PDE	Partial Differential Equation
PDF	Probability Density Function
ROI	Region of Interest
US	Ultrasound
VOI	Volume of Interest

Chapter 1

Introduction

1.1 Segmentation with prior knowledge

Segmentation can be considered as the task of partitioning the image into foreground object (or objects) and background regions. Since the concept of *object* is subjective and application dependent, segmentation in general is not well defined. Moreover, even when the object of interest is known, segmentation is challenging. This is so since in the imaging process the inherent properties of the inspected object may be corrupted due to noise, occlusion, illumination conditions and more. General *syntactical* assumptions such as continuity and smoothness of edges and homogeneity of semantic regions should thus be complemented with a-priori *semantical* information on the objects to be segmented.

The nature of prior knowledge varies, depending on the features that distinguish the object of interest from its surroundings. Color distribution, texture, motion and shape or their combinations are commonly used to characterize the regions of interest given the entire scene. The combination of such high level visual cues with the low level image features (edges and gray levels) yields a powerful segmentation framework [112]. Yet, there is a catch. Classically, the partitioning of the image into meaningful regions is a preliminary step towards image understanding. The next phase is identifying the segments based on their internal characteristics and interrelations. Only then, one may have an intelligent say on the ‘world’ depicted by the camera. The interleaving top-down and bottom-up processes seem to be an

incompatible flow.

The first argument to support this approach relies on recent research conducted in neurobiology and psychophysics. The hypothesis about symbiotic relations between high level cognitive vision and the low level visual processing is supported by exploration of interactions of various cortical visual processing pathways in human and animal visual systems, e.g. [39].

A possibly stronger claim stems from the fact that in most real-life applications at least some knowledge is available in advance. Consider for example the images acquired by medical imaging devices such as MR, CT, Ultrasound etc. There is much information on the patient (age, gender, medical history, physical conditions); on the imaged organ and tissue; on the expected pathologies and the normal, healthy states and also on the imaging modalities used and their limits. Medical imaging analysis problems are usually difficult but the solutions must be as reliable as possible. The reasoning is clear. When prior knowledge exists, why not using it?

Mathematical modeling of prior information and its incorporation in segmentation frameworks is an active field of research. One of the pioneering work in this domain was done by Borenstein and Ullman for top-down class-based segmentation using the statistics of image patches [5]. The main difficulty resides in combining the knowledge about the expected object appearance with the actual image data. This thesis suggests a variational approach to prior-based segmentation which allows a convenient way to integrate information from several sources in a coherent manner. We will therefore relate to variational methods in the forthcoming discussion.

1.2 Variational approaches to segmentation

Intensive image processing research in the last two decades raised up new insights about the nature of images. Modern segmentation approaches adopted these new concepts. The most influential observation is that an image can be approximated by a piecewise smooth function on \mathbb{R}^2 . The essence of this observation, made in 1988 by Mumford and Shah [70], is that a natural image is composed of homogeneous regions. The image segments and their delimiting boundaries obey the Gestalt principle of minimal description length. About a year later, Perona and Malik [79] suggested

the anisotropic diffusion for image denoising, based on the same concept of piecewise smoothness. Note that image denoising together with edge detection and integration were the objectives of the Mumford-Shah framework as well. These image processing tasks are in many senses complementary to image segmentation.

Partitioning of the image into semantic regions is equivalent to the detection of the closed contours that bound them. The key assumption behind the *region based* segmentation approaches is that meaningful image regions are homogeneous in terms of their color or texture. The Mumford-Shah functional was the origin of most region-based segmentation frameworks. Prominent methods include the works of Zhu and Yuille [128], Paragios and Deriche [78] and Chan and Vese [11].

Classical approaches to segmentation consider the correspondence of the image edges with the objects boundaries. Various image filters were suggested for edge detection [8]. Yet, the detection of the prominent image gradients had to be followed by a grouping process to obtain a coherent edge map that is compatible with image contours. The active contour methods originated by the *snakes* of Kass, Witkin and Terzopoulos [49] suggest an elegant solution to the problem of edge grouping. The main principle was the construction of a cost functional that imposes the alignment of the segmenting contour with the local maxima of the image gradients while maintaining its smoothness. The contour evolution was determined by partial differential equations (PDEs) derived from the first variation of the functional. The *classical snakes* method was elaborated by the introduction of the *balloon* term by Cohen and Cohen [13], that controls the direction of the contour propagation. The *geodesic snakes* by Caselles, Kimmel and Shapiro [10] and by Kichenassamy, Kumar, Olver, Tannenbaum and Yezzi [51] suggested a fundamental modification of the edge-based segmentation functional related to the anisotropic diffusion term of Perona and Malik [79].

The foundation of this thesis is a unified functional that is composed of edge-based terms together with a fidelity term based on the piecewise-smoothness assumption. The latter term is much in the spirit of the Chan-Vese level-set framework [11].

1.2.1 Level-set representation

The level-set methods introduced by Osher and Sethian [75] are numerical techniques to trace moving interfaces. Similar ideas have been proposed by Dervieux and Thomasset [25] a decade earlier. Applied to image segmentation problems, they allow an implicit and non-parametric representation of the propagating contour. The contour is the intersection of the level-set function with the $x - y$ (image) plane, or equivalently the zero level of a level-set function defined on the image plane. The level-set formulation enables automatic change in the contour topology (it can split or merge) and shows numerical robustness. Level-set functions have a special significance in this study, being used here also for the representation of the shape priors.

1.2.2 Shape based segmentation

Specific information on the segmented object, based on its known characteristics, is incorporated into the cost functional as additional terms. Model-based approaches embody common properties of the object class. Refer for example to [40], that extracts thin structures such as blood vessels, or to [92] that incorporates geometric information to segment road networks. When the object shape is specified, resemblance of the segmented object to the reference shape can also be imposed. The inclusion of prior shape knowledge within the variational framework for segmentation was pioneered by Leventon, Faugeras, Grimson and Wells [58, 57], Cremers, Kohlberger and Schnorr [19] and Rousson and Paragios [96]. Shape statistics seems to be an appropriate approach to deal with the variability in shape appearances. However, a comprehensive data-base of shapes is usually unavailable. Having only a single prior, previous works could only account for similarity transformations. Similarity transformations, however, rarely approximate satisfactorily the actual transformation between objects taken close to the camera. One of the main novelties of this thesis is the incorporation of the two-view geometry model in a variational framework for segmentation to account for planar projective transformation. This was enabled using a unique shape representation together with a novel dissimilarity measure between the object to segment and the prior shape.

1.2.3 Two view geometry and contour correspondence

The prior representations and the dissimilarity measure between the evolving segmentation and the prior representation are chosen to suit the two view geometry model. Consider two image planes f and f' . The rays that intersect the planes f and f' at points p and p' correspondingly, coincides at a single point P at the world plane. Fig. 3.5 demonstrate the idea. The points p and p' are related by planar projective homography. Assuming that a set of world points $\{P\}$ are coplanar, we can match their corresponding views by a single homography. Let C' denote the prior contour, such that $C' \in f'$ and $p' \in C'$. In the same manner, C denotes the contour being segmented, $C \in f$ and $p \in C$. An assumption made in this thesis is that the object contour points $P \in \mathbf{C}$ are approximately coplanar .

The core of the proposed methods is a process of concurrent segmentation and registration. By the end of the registration process the homography between the image to segment and the reference is recovered. The use of the planar projective model is made possible thanks to a special shape dissimilarity measure - introduced here. We construct a level-set function that defines the prior object. We then look for the coordinate transformation of the prior that best matches the object domain in the segmented image. We alternately look for the segmentation that corresponds the transformed prior. In the proposed formulation the level-set functions are not restricted to be sign-distance-functions. When a sign distance function undergoes an affine or a projective transformation it is no longer a sign distance function. This fact impaired previous approached and limited their ability to deal with transformations beyond similarity and scaling. The approach presented here can cope with more general transformations.

1.2.4 Implicit shape priors

The thesis proceeds to account for the cases where an explicit shape prior is not available. We consider the concurrent segmentation of two images of the same object in a framework termed here *mutual segmentation*. In this setting, the evolving contour in either of the images provides a dynamic prior for the other. The key idea is the definition of a biased dissimilarity term that incorporates the semantic knowledge

gained in the segmentation process of the image pair, accounting for excess or deficient parts in the estimated object shape.

We then consider the segmentation of symmetrical objects. The replicative form of the object induced by its symmetry is used as a shape constraint. A related idea was used in a work by Zabrodsky, Peleg and Avnir [125] for completion of occluded shapes using (rotational) symmetry. Our approach however is completely different from previous approaches to symmetry. Using the level-set representation, the symmetrical object shape is represented by a single entity and not as a collection of landmarks or feature points. The foundation of our symmetry-based approach is a coherent set of theoretical results that link between symmetry, two view geometry and level-set representation.

To the best of our knowledge, variational segmentation with implicit shape priors has never been considered before.

1.3 Contributions and organization of the thesis

Parts of the work presented here have been published or submitted to publication [3, 86, 84, 85, 87, 90, 83, 89, 91, 88]. The thesis includes seven chapters, where the main contributions are presented in Chapters 3 to 6.

1.3.1 Level-set based segmentation

Chapter 2 reviews some of the state-of-the-art concepts of level-set based segmentation. A minor contribution of this thesis is a reformulation of the alignment term proposed in [53] and [115]. We then construct a cost functional which is composed of region-based, edge-based and smoothness terms. In section 2.2 a new method for determining the weights of the energy terms is presented. The effectiveness of the bottom-up segmentation functional and the underlying gradient descent approach is demonstrated on a real medical problem - the delineation of uterine fibroids from MR images of the uterus. We conclude by giving the outline of the shape similarity measures to be described in details in the forthcoming chapters.

1.3.2 Segmentation with shape priors

In chapter 3 we present a method to facilitate segmentation given a single reference object. We assume that the prior shape is related to the shape being segmented by a planar projective transformation. Generalizing the Chan-Vese level set framework, we introduce a novel shape-similarity measure and embed the projective homography between the prior shape and the image to segment within a region-based segmentation functional. We first introduce the *unlevel-sets* prior-based segmentation method. The prior shape is represented by a *generalized cone* that its base has the shape of the prior object. The intersections of the cone with the image plane in different orientations and poses correspond to possible instances of the prior shape under perspective distortion and scaling. This technique has a nice intuitive geometric interpretation and a relatively simple mathematical formulation. This method, however, only handles perspective (a six-parameter transformation), and is applicable to the limited class of star-shaped objects. Nevertheless, it provides the conceptual basis for the method developed in this chapter. Next, we generalize the algorithm to eight-parameter projectivity model and remove the restriction to star-shaped objects. We demonstrate prior-based segmentation algorithms on a variety of images. The transformation parameters between the object views are recovered and verified when possible.

1.3.3 Mutual segmentation

In chapter 4 we present a variational approach for simultaneous segmentation of *two* images of the same object taken from different viewpoints. Due to noise, clutter and occlusions, neither of the images contains sufficient information for correct object-background partitioning. The evolving object contour in each image provides a dynamic prior for the segmentation of the other object view. We call this process *mutual segmentation*. The foundation of the proposed method is a unified level-set framework for region and edge based segmentation, associated with a shape similarity term. The suggested shape term incorporates the semantic knowledge gained in the segmentation process of the image pair, accounting for excess or deficient parts in the estimated object shape. Transformations, including planar projectivities, between the object views are accommodated by a registration process held concurrently with the

segmentation. The proposed segmentation algorithm is demonstrated on a variety of image pairs. The homography between each of the image pairs is estimated and its accuracy is evaluated.

1.3.4 Segmentation with shape symmetry

In chapter 5 we consider the extraction of objects with either bilateral or rotational symmetry in the presence of perspective distortion. Information on the symmetry axis of the object and the distorting transformation is obtained as a by-product of the segmentation process. The key idea is the use of a flip or rotation of the image to segment as another view of the object. We call this generated image the *symmetrical counterpart* image. The theoretical foundation of the proposed method is a theorem, proven in this thesis, showing that symmetrical counterpart images are related by planar projective homography. This homography is determined by the unknown planar projective transformation that distorts the object symmetry. The limits on the ability to recover the distorting projective transformation from this homography are stated. The homography that aligns the symmetrical counterpart images is recovered via a registration process carried out concurrently with the segmentation. Promising segmentation results of various images of approximately symmetrical objects are shown.

1.3.5 Application: segmentation of mouse brain atlas

In chapter 6 we present a method for segmentation of anatomical structures in the mouse brain from histological data. Segmentation is carried out slice-by-slice where the successful segmentation of one section provides a prior for the subsequent one. Intensities and spatial locations of the region of interest and the background are modeled by three-dimensional Gaussian mixtures. The key idea is a novel region-based energy term that considers the Kullback-Leibler (K-L) divergence between the probability density functions (PDFs) of the partitions. This information adaptively propagates across the sections. Segmentation is inferred by minimizing a cost functional that enforces the compatibility of the partitions with the corresponding models together with the alignment of the boundaries with the image gradients. The algorithm is demon-

strated on histological images of mouse brain. The segmentation results compare well with manual segmentation.

Chapter 2

Level-set based Segmentation

2.1 Statistical set up and prior art

2.1.1 General principles

Most segmentation approaches, despite their diversity, are motivated by the same Gestalt principles of perception, i.e. simplicity (minimum description length), similarity (homogeneity of semantic regions), continuity, proximity and closure. The image is then partitioned according to a subset of these rules using various mathematical tools. Commonly, segmentation is obtained by minimizing an objective functional either by discrete (graph-based) approaches such as Markov Random Fields or by the continuous formulation of calculus of variations. While in this chapter we use PDEs to solve segmentation problems, Bayesian statistical inference formulation is used to set a common language between the discrete and continuous approaches.

Given an image $I(\mathbf{x})$ we would like to infer the delineating curve C between an object and its background. This is done via the maximization of the probability distribution function (PDF) $P(C | I)$, using Bayes law:

$$P(C | I) \propto P(I | C)P(C) . \quad (2.1)$$

The prior probability $P(C)$ will be reformulated to have “syntactic” and “semantic” components. Classical active contours methods, such as the snake algorithm of Kass et al [49] use parametric representation $C = C(p)$. In this formulation

the term $-\log P(I | C) = -\lambda \int |\nabla I(C(p))| dp$ is the “external force”. The term $-\log P(C) = \int (\alpha |C_p|^2 + \beta |C_{pp}|^2) dp$ is the “internal force”. The maximization over all possible separating curves is done by minimizing $-\log P(C | I)$. Note that here the prior is syntactic since it dictates the degree of smoothness of the curve and is not directly related to the class of images or objects to be segmented. Our main interest in this work is in a special kind of “semantic” prior that depends on another image of the object. The segmenting curve C is represented implicitly using the level-set framework [75, 25]. This approach for curve evolution has been successfully applied to extract complex object boundaries, allowing an automatic change in the topology. Moreover, the continuous, parameter-free shape representation, indicated by the positive levels of the image level-set function, enables a definition of a unique dissimilarity measure between object views.

In the level set framework for curve evolution [75], an evolving curve $C(t)$ is defined as the zero level of a level set function $\phi: \Omega \rightarrow \mathbb{R}$ at time t :

$$C(t) = \{\mathbf{x} \in \Omega | \phi(\mathbf{x}, t) = 0\}. \quad (2.2)$$

Following [11], we use the Heaviside function of ϕ

$$H(\phi(\mathbf{x})) = \begin{cases} 1 & \phi(\mathbf{x}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

as an object indicator function, assigning the positive and the negative levels of ϕ to the foreground and the background image regions, respectively. As in [11], a smooth approximation of the Heaviside function H_ϵ , rather than a step function, is used. In particular,

$$H_\epsilon(\phi) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) \quad (2.4)$$

and its derivative $\delta_\epsilon(\phi) = dH(\phi)/d\phi$ is

$$\delta_\epsilon(\phi) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + \phi^2}. \quad (2.5)$$

Figs. 2.1a-b illustrate H_ϵ and δ_ϵ near zero. Thus, the evolutionary change of ϕ ,

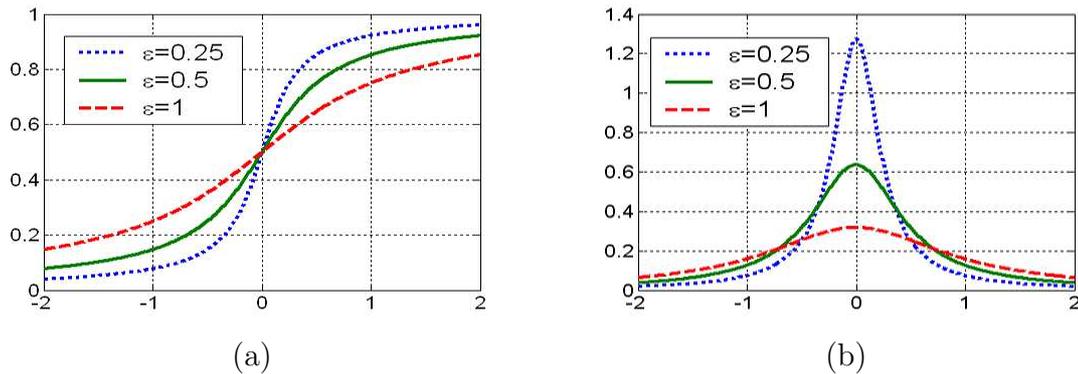


Figure 2.1: Smooth approximation of the Heaviside function H_ϵ (a) and its derivative δ_ϵ (b). Adapted from [11].

according to (3.5) is not restricted to its zero level-set. We can now rephrase our PDF as

$$P(\phi | I) \propto P(I | \phi)P(\phi) . \quad (2.6)$$

Next we elaborate on the conditional probability term $P(I | \phi)$.

2.1.2 Region-based data term

Let $I: \Omega \rightarrow \mathbb{R}^+$ denote a gray level image, where $\Omega \subset \mathbb{R}^2$ is the image domain. Let $\omega \subset \Omega$ be an open subset, not necessarily connected, of the image domain Ω . In the spirit of the Gestalt laws of similarity and proximity, we define a boundary $C = \partial\omega$ that delimits homogeneous regions in Ω . In particular, we use the two-phase formalism, in which the image is partitioned into foreground and background regions. Thus, given a feature G , e.g. the average gray level, we look for a curve C that maximizes the difference between two scalars u^+ and u^- defined as follows:

$$u^+ = A^+ \int_{\omega} G^+(I(\mathbf{x}))d\mathbf{x}, \quad u^- = A^- \int_{\Omega \setminus \omega} G^-(I(\mathbf{x}))d\mathbf{x}. \quad (2.7)$$

The superscripts $+$ and $-$ correspond to the feature values in $\omega \subset \Omega$ and in Ω/ω , respectively. Hereafter we denote $\mathbf{x} \equiv (x, y)$, $A^+ = 1/\int_{\omega} d\mathbf{x}$ and $A^- = 1/\int_{\Omega \setminus \omega} d\mathbf{x}$. The possibly different functions G^+ and G^- are defined on the object and the background domains, respectively.

In general, one can use a set of features $\{G_i\}$ corresponding to two sets of scalars $\{u_i^+\}$, $\{u_i^-\}$. The features chosen should be related to the expected image homogene-

ity. In the work of Chan and Vese [11] the image is approximated by a piecewise constant function whose values are given by $G_1^+(I(\mathbf{x})) = G_1^-(I(\mathbf{x})) = I(\mathbf{x})$. Hence $u_1^+ = \overline{I_{in}}$ and $u_1^- = \overline{I_{out}}$ are the average gray levels in the object regions and in the background regions respectively. In that formulation it is assumed that the PDF of the gray levels in each region is a Gaussian with $\sigma = 1$. More elaborate characterization of the partition can be obtained using color histograms or a Gaussian mixture model. For texture-based segmentation the Gabor filters may be used as in Sandberg et. al. [98] and in Sagiv et. al. [97]. The definition of $G(I)$ could be further extended as suggested in [119], where $u_+, u_- \in C^1(\mathbb{R}^2)$ are smooth approximations of the regions in I . In this study we use the average gray levels and the variance [119, 62]:

$$G_2^+(I) = (I(\mathbf{x}) - \overline{I_{in}})^2 \quad ; \quad G_2^-(I) = (I(\mathbf{x}) - \overline{I_{out}})^2 \quad (2.8)$$

We may now express the term $-\log P(I | \phi)$ via a region based cost functional with a well defined integration domain:

$$E_{\text{RB}}(\phi) = \sum_{i=1}^2 \int_{\Omega} [(G_i^+(I(\mathbf{x})) - u_i^+)^2 H(\phi) + (G_i^-(I(\mathbf{x})) - u_i^-)^2 (1 - H(\phi))] d\mathbf{x} \quad (2.9)$$

An elegant statistical formulation representing the region-based term was introduced in [128], followed by [78] and reviewed in [20].

The evolving boundary $C(t)$ is derived from $\phi(t)$ using (2.2). For a given $\phi(t)$ and $\{G_i\}$, the feature values u_i^+ and u_i^- are updated at each iteration according to (2.7). The level set function ϕ evolves via the gradient descent:

$$\phi_t^{\text{RB}} = \delta(\phi) \sum_{i=1}^2 [G_i^-(I(\mathbf{x})) - u_i^-]^2 - (G_i^+(I(\mathbf{x})) - u_i^+)^2 \quad , \quad (2.10)$$

where δ is the derivative of the Heaviside function H .

2.1.3 Geodesic active contour: data part

Following the pioneering work of Kass et al [49], Malladi et al [64] and independently Caselles et al [9] suggested a non-variational framework for dynamic edge integration. These works known as the Geometric Active Contours were followed by the Geodesic

Active Contour (GAC) frameworks of [51] and [10].

Let $C(s) = (x(s), y(s))$ be the parametric description of a planar contour $C: [0, L] \rightarrow \mathbb{R}^2$ where s is an arc-length parameter and L is the length of C . We denote the vector field of the image gradients by $\nabla I(x, y) = (I_x, I_y)^T = \left(\frac{\partial I(x, y)}{\partial x}, \frac{\partial I(x, y)}{\partial y} \right)^T$. The Geodesic Active Contour (GAC) term integrates an inverse edge indicator function along the contour:

$$E_{\text{GAC}}(C) = \int_0^L g_{\text{GAC}}(|\nabla I(C(s))|) ds. \quad (2.11)$$

The function g_{GAC} should be chosen such that $E_{\text{GAC}}(C)$ is minimized when the curve C is located on the maxima of the absolute values of the image gradients. As was shown in [9, 10], g_{GAC} can be any positive strictly decreasing function of the image gradients $r = |\nabla I|$ as long as $g_{\text{GAC}}(r) \rightarrow 0$ as $r \rightarrow \infty$.

The first variation of the energy term (2.11) is

$$\delta E / \delta C = (\langle \nabla g_{\text{GAC}}, \vec{n} \rangle - \kappa g_{\text{GAC}}) \vec{n} \quad (2.12)$$

where $\langle \cdot, \cdot \rangle$ denotes inner product, $\vec{n}(s) = \{-y_s(s), x_s(s)\}$ is the exterior normal to the curve C and κ is the curvature of C . Using the level set formulation, equation (2.12) takes the form $-\text{div}(g_{\text{GAC}} \frac{\nabla \phi}{|\nabla \phi|}) |\nabla \phi|$ having the relation $\vec{n} = \nabla \phi / |\nabla \phi|$.

We unorthodoxly split the GAC term, Eq. (2.11), into two terms. This splitting will be justified in subsection 2.1.6 where both terms will appear in the final energy formulation. The data term (DGAC) is given by

$$E_{\text{DGAC}}(C) = \int_0^L g_{\text{DGAC}}(|\nabla I(C(s))|) ds \quad (2.13)$$

where

$$g_{\text{DGAC}}(|\nabla I|) = -\frac{|\nabla I|^2}{1 + |\nabla I|^2}. \quad (2.14)$$

This term vanishes as the gradient magnitudes decrease to zero and attains -1 asymptotically for large gradients. Expressing this term in a level-set framework we obtain

$$E_{\text{DGAC}} = \int_{\Omega} g_{\text{DGAC}}(|\nabla I|) |\nabla H(\phi(\mathbf{x}))| d\mathbf{x}, \quad (2.15)$$

with the associated gradient descent equation:

$$\phi_t^{\text{DGAC}} = \delta(\phi) \operatorname{div} \left(g_{\text{DGAC}}(|\nabla I|) \frac{\nabla \phi}{|\nabla \phi|} \right). \quad (2.16)$$

Adapting the formulation of [11] for the edge-based terms, we use $H(\phi)$ instead of ϕ in the energy terms (see also equations (2.17, 2.19, 2.21)) and consequently $\delta(\phi)$ instead of $|\nabla \phi|$ in the gradient descent equations (also in 2.18, 2.20, 2.22). This formulation may appear different from the standard convention used in the literature, yet practically both are almost equivalent.

The GAC functional includes another geometrical term that will be described in subsection 2.1.5.

2.1.4 Edge alignment term

Segmentation can be refined by constraining the normal direction of the active contour to align with the directions of the image gradients as suggested by Kimmel and Bruckstein [53] and independently by [115]. Fig 2.2 outlines this concept. The edge alignment term (EA) defined in [52] takes the form: $E_{\text{EA}}(C) = - \int_0^L |\langle \nabla I(\mathbf{x}(s)), \vec{n}(s) \rangle| ds$. The expression for $E_{\text{EA}}(C)$ is an integration of the projection of ∇I on the normal $\vec{n}(s)$ along the curve. A minor contribution of this chapter is the level-set formulation of the alignment term:

$$E_{\text{EA}} = - \int_{\Omega} \left| \langle \nabla I, \frac{\nabla \phi}{|\nabla \phi|} \rangle \right| |\nabla H(\phi)| d\mathbf{x} \quad (2.17)$$

where $\nabla \phi(\mathbf{x})/|\nabla \phi(\mathbf{x})|$ is normal to the level-set ϕ in \mathbf{x} . The associated gradient descent equation is

$$\phi_t^{\text{EA}} = -\delta(\phi) \operatorname{sign}(\langle \nabla \phi, \nabla I \rangle) \Delta I, \quad (2.18)$$

where ΔI is the Laplacian. This equation is similar to the one derived in [52].

2.1.5 Syntactic prior: geometry

The prior probability $P(C)$ in Eq. (2.1) is determined by the minimum description length criterion (based on the Gestalt principle of simplicity). That is $P(C) \propto$

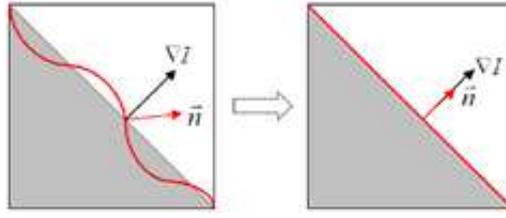


Figure 2.2: The edge *alignment* term is minimized when the scalar product between normal direction of the active contour and the image gradient direction is maximal. Image courtesy of Nir Ben Zadok.

$\exp(-\nu|C|)$, $\nu > 0$ [20], where $|C| = \int ds = \int |C_p| dp$ is the curve length. An equivalent representation, using the level set formulation, takes the form:

$$|C| = E_{\text{LEN}} = \int_{\Omega} |\nabla H(\phi(\mathbf{x}))| d\mathbf{x} \quad (2.19)$$

This functional measures the length of the curve and usually serves as an indicator for the curve smoothness [11]. Minimizing (2.19) with respect to ϕ , we obtain the associated Euler Lagrange equation for ϕ :

$$\phi_t^{\text{LEN}} = \delta(\phi) \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right). \quad (2.20)$$

Combining E_{LEN} and E_{DGAC} , (defined in Eq.(2.15)), we get the usual form of the GAC functional [10, 51],

$$E_{\text{GAC}} = \int_{\Omega} g_{\text{GAC}}(|\nabla I|) |\nabla H(\phi(\mathbf{x}))| d\mathbf{x}, \quad (2.21)$$

where $g_{\text{GAC}} = 1 + g_{\text{DGAC}} = 1/(1 + |\nabla I|^2)$. The gradient descent equation is

$$\phi_t^{\text{GAC}} = \delta(\phi) \operatorname{div} \left(g_{\text{GAC}}(|\nabla I|) \frac{\nabla \phi}{|\nabla \phi|} \right). \quad (2.22)$$

Finite difference method and the relation to the MRF pairwise term

Practically, we use the finite difference method, in which a grid of nodes spaced by a parameter h is set up over the image domain. The differential operator is approximated by finite difference equations operating on neighboring nodes. Specifically, the GAC term is discretized and by using the finite difference method for $|\nabla I|$, it

takes the form: $\sum_{\mathbf{x} \in \Omega} \tilde{g}(|I(\mathbf{x} + h) - I(\mathbf{x})|)$. The term $\mathbf{x} + h$ denotes a shift of one pixel from \mathbf{x} and obviously $\mathbf{x} + h \in \mathcal{N}_{\mathbf{x}}$ where $\mathcal{N}_{\mathbf{x}}$ is the set of nearest neighbors of \mathbf{x} . Using this formulation, it is easy to see the analogy of the GAC term to the MRF pairwise term (sum of the 2-clique potentials) [32], standing for the inverse logarithm of the conditional probability function $-\log P(I | C)$, for independent and identically distributed process.

2.1.6 Unified cost functional

A unified cost functional that is based on the data of the images to segment, on general assumptions with respect to properties of natural images, and on the segmentation curves both images takes the form:

$$E(\phi) = W^{\text{RB}} E_{\text{RB}}(\phi) + W^{\text{LEN}} E_{\text{LEN}}(\phi) + W^{\text{DGAC}} E_{\text{DGAC}}(\phi) + W^{\text{EA}} E_{\text{EA}}(\phi) \quad (2.23)$$

with equations (2.9), (2.19), (2.15) and (2.17), respectively.

Note that the GAC term (Eq. 2.21) is split into the smoothness term E_{LEN} (Eq. 2.19) and the DGAC term E_{DGAC} (Eq. 2.15) where each has its own weight. This allows more flexibility in the functional construction, in particular when the edge based term (DGAC) should be ignored while the smoothness term is most desirable. The evolution of the level-set functions ϕ in either of the images, is determined at each iteration by $\phi(t + \Delta t) = \phi(t) + \phi_t \Delta t$. The associated gradient descent equations ϕ_t are derived using the first variation of each of the terms in the functional (2.23)

$$\phi_t = W^{\text{RB}} \check{\phi}_t^{\text{RB}} + W^{\text{LEN}} \check{\phi}_t^{\text{LEN}} + W^{\text{DGAC}} \check{\phi}_t^{\text{DGAC}} + W^{\text{EA}} \check{\phi}_t^{\text{EA}}. \quad (2.24)$$

The terms $\check{\phi}_t^{\text{TERM}}$ are obtained by slight modification of the gradient descent terms ϕ_t^{TERM} determined by equations (2.10), (2.20), (2.16) and (2.18), respectively. This issue and the determination of the weights W^{TERM} for the different terms in Eq.(2.24) are discussed in section 2.2.

Refinement of the segmentation results can be obtained for images with multiple channels, $I: \Omega \rightarrow \mathbb{R}^n$, e.g. color images. The region-based term ϕ_t^{RB} ; the DGAC term ϕ_t^{DGAC} and the alignment term ϕ_t^{EA} can be presented as the sum of the contributions

of each color channel I_i . Multi-channel segmentation is particularly suitable when, for example, the object boundaries are dominant in part of the channels while the piecewise homogeneity is preserved in others. Figure 5.11 in chapter 5 demonstrates segmentation of a color image. Further exploration could address the use of Beltrami flow, introduced by Sochen, Kimmel and Malladi [101].

2.2 Setting the weights of the energy terms

When the solution to an image analysis problem is obtained by minimizing a cost functional, the issue of setting the relative weights of the energy terms is unavoidable. However, in the absence of a satisfying method for the determination of these parameters, this subject is usually marginalized. A guaranteed but time consuming approach is to set the weight parameters by an exhaustive search on the parameters grid, followed by qualitative examination of the solutions obtained. In the context of image processing, algorithms for setting parameters are considered robust if the search space can be significantly reduced or if the parameters should be only slightly tuned for a class of similar images. Being tedious and subjective this heuristic is not suitable for most real world applications.

In contrast, we propose a heuristic that adaptively determines the relative weight of the contributions of the terms of the functional to the evolution of the level set function, as expressed in Eq. (2.24). The proposed heuristic is based on the assumption that the contributions of the energy terms in each iteration should be balanced. The following observations are considered:

1. The relative weight between the length term and the area terms (region based terms) is squared as the image size is increased.
2. The weight of region based term is affected by the units of the feature chosen to characterize the regions. The simplest example is the gray level range which is usually chosen to be either $[0, 1]$ or $[0, 255]$.
3. The terms that are based on the image gradients may have high dynamic range due to noise or sharp discontinuities.

4. All terms are affected by the instantaneous form of the level set function ϕ which evolves in time. The dynamics induced by ϕ should be taken into consideration.

The suggested scheme for automatic and dynamic weight setting is as follows. Let

$$\check{\phi}_t^{\text{TERM}}(\mathbf{x}) = B(\phi_t^{\text{TERM}}(\mathbf{x})) = \begin{cases} U_B & \text{if } \phi_t^{\text{TERM}}(\mathbf{x}) > U_B \\ L_B & \text{if } \phi_t^{\text{TERM}}(\mathbf{x}) < L_B \\ \phi_t^{\text{TERM}}(\mathbf{x}) & \text{otherwise} \end{cases} \quad (2.25)$$

where

$$U_B = \text{std}(\phi_t^{\text{TERM}}(\mathbf{x})), \quad L_B = -U_B$$

Here, $\text{std}(\phi_t(\mathbf{x}))$ stands for the standard deviation of $\phi_t(\mathbf{x})$ over Ω . The functional $B(\cdot)$ operates on ϕ_t^{TERM} to bound its dynamic range. Next, the range of $|\check{\phi}_t^{\text{TERM}}|$ is normalized

$$W^{\text{TERM}} = 1 / \max_{\mathbf{x}} |\check{\phi}_t^{\text{TERM}}(\mathbf{x})|. \quad (2.26)$$

Note that the clipping (Eq. 2.25) affects only extreme values of ϕ_t^{TERM} , that is $\check{\phi}_t^{\text{TERM}}(\mathbf{x}) = \phi_t^{\text{TERM}}(\mathbf{x})$ for most $\mathbf{x} \in \Omega$. Since W is recalculated at each iteration it is time dependent. This formulation enables an automatic and adaptive determination of the weights of the energy terms.

2.3 Example: Segmentation of uterine fibroids

2.3.1 General

We present a medical application that exemplifies the principles of level set based segmentation discussed in this chapter¹. The motivation here is the detection of uterine fibroids in MR images prior to Ultrasound treatment. We found that the bottom-up segmentation approach is sufficient for accurate segmentation in most of these cases. Minimizing an energy functional composed of region-based, edge based and smoothness terms we obtained fibroid boundaries that were in good match with the delineation performed by an expert.

¹This section is based on collaboration with Nir Ben-Zadok [3].

2.3.2 Background and clinical motivation

Uterine fibroids are the most common benign gynecological pathology of women in the reproductive age. Magnetic Resonance-guided Focused Ultrasound (MRgFU) has been demonstrated as a valuable option for effectively treating uterine fibroids [27] [106]. The acquired MR images of uterine fibroids are used in order to plan the treatment and to ensure exact tumor targeting. At the end of the treatment T1 weighted contrast enhanced MR images are acquired, in order to evaluate treatment efficiency. Fig. 2.3 shows Uterine fibroids and surrounding anatomy of the uterus. Segmentation is the core task of treatment guidance, as it defines the regions

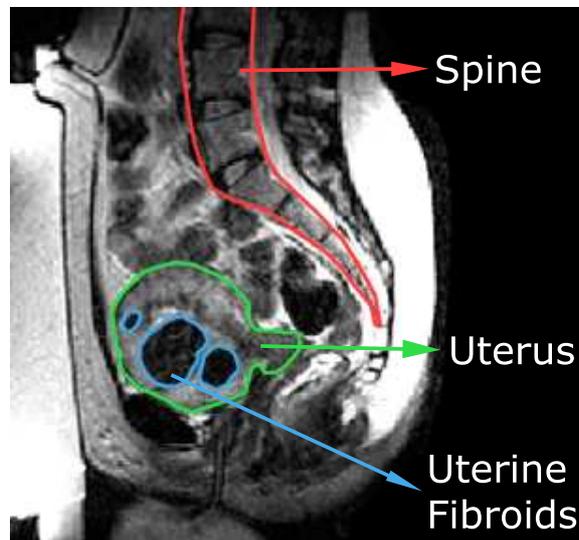


Figure 2.3: MR image of uterine fibroids and the surrounding anatomy of the uterus

to be ablated during the treatment. There are several factors that turn the task of segmenting uterine fibroids into a complex and difficult one. First, the visual characteristics of uterine fibroids (such as size, shape, location and intensity) vary between different patients and also when comparing different MR slices taken from the same patient (see Fig. 2.5). Second, the partial volume effect (referring to a situation where a pixel represents more than one kind of tissue type), blurs the intensity-difference between the fibroids and nearby tissues. We address the extraction of uterine fibroids in MR images using level-set technique. The required user interaction is minimal. The automatic processing time is few seconds using non-optimized MATLAB code. Experimental results on MR sequences, acquired from several patients, are in good agreement with expert segmentation.

2.3.3 Algorithm summary

For each image the following steps are carried out:

- Initialization - The user sets a seed point inside each fibroid. The initial level set function ϕ_0 is generated such that its zero level coincides with a default circle, encircling each of the seed points marked by the user. For fast convergence, the user may also set the circle radii (see Fig. 2.4).
- Repeat until convergence
 - Compute the values u^+ and u^- using equation 2.7. These values are the average gray levels of the object and background regions respectively, defined by $H_\epsilon(\phi(\mathbf{x}, t))$.
 - Update ϕ according to the gradient descent equation 2.24.
- The final contour is the zero level of the final level-set function ϕ .

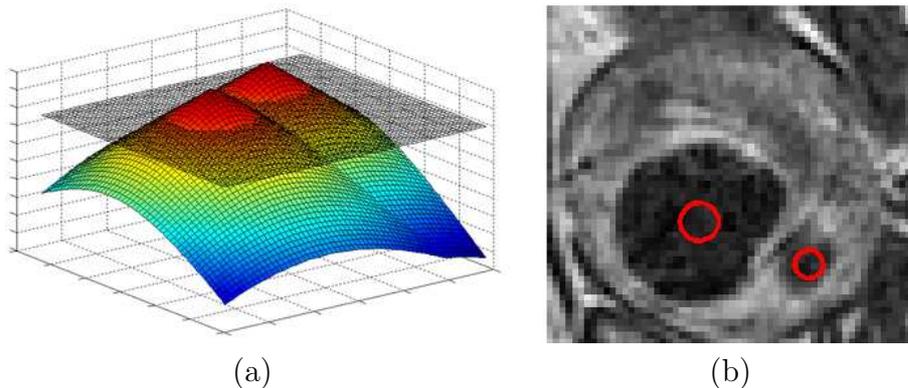


Figure 2.4: (a) Initial Level-set function (b) The corresponding initial contour

2.3.4 Experimental results

We exemplify the algorithm presented in subsection 2.3.3 on several clinical abdomen MR sequences with various forms of fibroids. The MRgFUS therapy procedures were performed by Insightec's ExAblate 2000 system, which uses a 1.5T whole-body system (Genesis Signa; GE Medical Systems, Milwaukee, Wis.) to acquire the MR image sequences. Fig. 2.5 shows a variety segmentation examples. The segmenting contours (red) delineate the fibroids in each of the images. The user only marks

the approximate center of the fibroids. Additional MR image sequences and their automatic and manual segmentation are available in [3].

2.4 Semantic prior: shape term

In the previous sections we described the edge-based, region-based and smoothness constraints that control the segmentation based on the image data and on general assumptions on natural images. We now give an outline of the shape terms, denoted by E_{SHAPE} that are derived from the semantic knowledge on the object of interest.

Denoting a prior shape representation by $\tilde{\phi}: \Omega \rightarrow \mathbb{R}$, the statistical formulation is then

$$P(\phi, T \mid I, \tilde{\phi}) \propto P(I \mid \phi, \tilde{\phi}, T)P(\phi \mid \tilde{\phi}, T)P(T, \tilde{\phi}) = P(I \mid \phi)P(\phi \mid \tilde{\phi}, T)P(T)$$

where T is the transformation that aligns ϕ and $\tilde{\phi}$ and pairwise independence of I , $\tilde{\phi}$ and T is assumed. The logarithm of the inverse conditional probability term $-\log P(\phi \mid \tilde{\phi}, T)$ can be substituted by a dissimilarity measure $D(\phi, T(\tilde{\phi}))$ between shape representations $\tilde{\phi}$ and ϕ . Some previous approaches, e.g. [57, 58], used:

$$D(\phi, T(\tilde{\phi})) = \int_{\Omega} \left(\phi(\mathbf{x}) - T(\tilde{\phi}(\mathbf{x})) \right)^2 d\mathbf{x}$$

where ϕ and $T(\tilde{\phi})$ are signed distance functions. This measure, however, is affected by the size of the background area $\Omega \setminus \omega$, see [21] and references therein. Moreover, in a cluttered image, when the weight of this measure in the segmentation functional is high, objects that do not correspond to the prior shape are ignored [23]. To avoid these drawbacks several modifications to control the integration domain have been suggested [21, 23, 96]. In all dissimilarity measures between level-set functions (represented as signed distance functions) the alignment term T is restricted to be an isometry. This restriction guarantees that the transformed level-set function $\tilde{\phi}_T = T(\tilde{\phi})$ will preserve the form of a signed distance function.

In chapter 3 we use the square difference between the Heaviside functions of ϕ and $\tilde{\phi}$ as a dissimilarity measure between the shape representations, where ϕ and $\tilde{\phi}$

are **not** necessarily distance functions².

$$D(\phi, \tilde{\phi}|T) = \int_{\Omega} \left[H(\phi(\mathbf{x})) - H(T(\tilde{\phi}(\mathbf{x}))) \right]^2 d\mathbf{x} \quad (2.27)$$

This formulation enables the introduction of the projective alignment term T between the shapes. Furthermore, D does not depend on the size of the background area $\Omega \setminus \omega$.

Denote $L = H(\phi(\mathbf{x}))$ and $\tilde{L}_T = H(T(\tilde{\phi}))$. The functions $L: \Omega \rightarrow \{0, 1\}$ and $\tilde{L}_T: \Omega \rightarrow \{0, 1\}$ are binary labeling functions. We assume that the labeling L and \tilde{L}_T are independent and identically distributed. Thus, the conditional probability based on the proposed dissimilarity measure can be formulated as:

$$P(\phi | \tilde{\phi}, T) = \frac{1}{\mathcal{N}} \exp\left(- \sum_{\mathbf{x} \in \Omega} (L(\mathbf{x}) - \tilde{L}_T(\mathbf{x}))^2\right),$$

where $\mathcal{N} > 0$ is a normalizing scalar.

The shape term (2.27) is suitable when the prior $\tilde{\phi}$ is perfect and constant in time. In chapters 4-5, a different dissimilarity measure is suggested, which is more appropriate when either $\tilde{\phi}$ is the evolving (and imperfect) segmentation of the other image or the symmetrical counterpart of ϕ .

²A significant gain from not enforcing ϕ to be a distance function is the elimination of the process of re-distancing [34, 105].

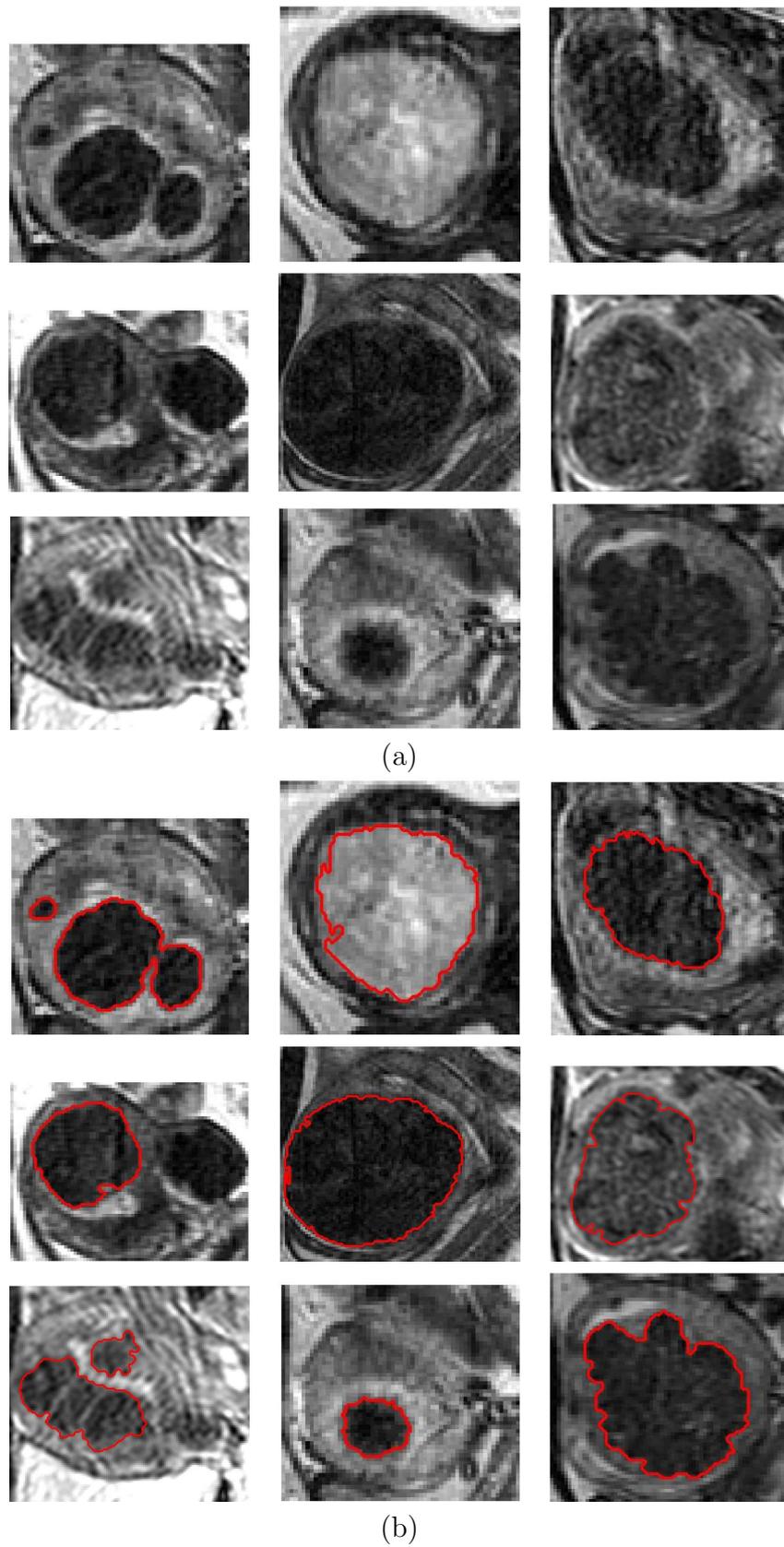


Figure 2.5: (a) MR images of various types of uterine fibroids (b) Segmentation results obtained by the level-set based segmentation algorithm.

Chapter 3

Segmentation with Shape Priors

3.1 Background and previous work

The coupled tasks of segmentation and object detection are essential for the extraction of semantic content from images. Prior knowledge on the shape of interest can significantly facilitate these processes, particularly when the object boundaries are not well defined. However the integration of shape information into the segmentation process is non-trivial. The main difficulty is the need to account for possible transformations between the prior shape and the shape being segmented.

This problem was studied extensively via template matching techniques, see [37, 45, 46, 59, 67, 117, 118] and references therein. When dealing with contours, registration via template matching is equivalent to the Hough Transform [104]. Shape is then commonly represented in a parametric way and the dimension of the representation grows with the complexity of the shape and its degrees of freedom. Template matching and Hough transform are known to yield difficult global optimization problems when the dimension is high, thus incorporation of projective transformations is difficult to accomplish. Moreover, these methods require a substantial set of corresponding points (or features), such that the matched images must have identical texture, see [47] and reference therein.

A different approach focuses on landmark based analysis, in which a training shape set, sampled and aligned, is assumed. In particular, the Active Shape Model [17] and Active Appearance Model [16], that use Principal Component Analysis (PCA)

of landmarks to model shape variability, gained popularity for their simplicity and efficiency. However, these methods are semi-automatic, require prior shape analysis and are parameterization dependent.

Variational methods solve segmentation problems by means of energy minimization, integrating top-down and bottom-up information, see [1] and references therein. Specifically, the level-set framework [75] for curve evolution has been successfully applied to extract complex object boundaries, allowing an automatic change in the topology. Prior-based segmentation methods incorporate a representation of a reference shape within the energy functional. Thus, the recovered object boundary should resemble the expected contour, in addition to being constrained by length, smoothness and compatibility with the image gray levels and gradients.

In order to introduce prior shape knowledge and a given group of transformations in the level-set formulation, a shape dissimilarity measure should be provided. Several works, for example [23, 26, 96, 110], use the distance function as the level-set and the square difference between level-sets as the shape dissimilarity measure. A symmetric and unbiased modification of this shape distance (called pseudo distance) has been recently suggested by Cremers and Soatto [21]. However, these similarity measures only account for isometric transformations and scaling, since more general transformations (such as non-isotropic scaling or perspectivity) do not preserve the characteristics of distance functions.

The statistical methodology [12, 19, 42, 58, 57, 96, 110] accounts for transformations beyond similarity and for small non-rigid deformations by using a comprehensive training set. It characterizes the probability distribution of the shapes and then measures the similarity between the evolving object boundary (or level-set function) and representatives of the training data. It is important to note that there is no distinction in this method between transformation-based and deformation-based shape variation. The modes of variation have to account for both. Moreover, the performance depends on the size and coverage of the training set.

None of the existing methods accounts for projective transformations between the prior shape and the shape of interest. The inability to deal with projective transformations is significant. In the presence of projectivity, neither similarity nor (even) the affine model provide reasonable approximation for the transformation between

the prior shape and the shape to segment. Figs. 3.1b-c show the *best* registration of a prior contour of the object shown in Fig. 3.1a to a given image, assuming similarity and affine transformations respectively. The apparent mismatch inhibits the segmentation process and prohibits accurate reconstruction of the missing parts. In contrast, the registration example in Fig. 3.1d demonstrates that planar projective transformation is a good approximation even though the elephant shape contour is roughly planar.

We suggest a novel variational approach to prior-based segmentation, that explicitly accounts for planar projective transformation, using a *single* reference object. The segmentation process is carried out concurrently with the registration of the prior shape to the shape of interest. The outcomes of the algorithm include the detection of the object of interest and correct extraction of its boundaries. The planar projective transformation between the two object views is accurately recovered as well. Neither point correspondence nor direct methods [43] are used, thus color or texture compatibility between the prior and the segmented image is needless. This is accomplished by introducing a novel shape-similarity measure, that admits a wide range of transformations, beyond similarity, and using it to generalize the Chan-Vese level-set framework [11]. The proposed region-based segmentation functional includes an explicit expression of the projective homography between the prior shape and the shape to segment.

Employing the parameterization-free shape description, enabled by the level-set formulation, we gain a significant advantage over landmark-based and template matching techniques that represent shapes by collections of points or features. The suggested distance function between the level-set representations of the matched shapes is well defined and is not depend on shapes sampling. Moreover, transformations applied on the domains of the level-set functions, transform the represented shapes correspondingly. This results in an elegant and powerful mathematical formulation to align the prior and the evolving shape, minimizing their dissimilarity measure with respect to the transformation parameters. The graceful merge of the image data with that of the projectively registered prior is the essence of the proposed contribution.

The suggested algorithm is demonstrated on a variety of images, in the presence of

perspective distortion. The successful segmentation results and the reliable estimation of the transformation parameters suggest this method as a promising tool for various segmentation and registration applications.

This chapter is organized as follows. In section 3.2 we review the segmentation framework of Chan and Vese [11], and outline our prior shape model, in particular the novel shape similarity measure. We first (section 3.3) introduce the *unlevel-sets* prior-based segmentation method. The prior shape is represented by a *generalized cone* that its base has the shape of the prior object. The intersections of the cone with the image plane in different orientations and poses correspond to possible instances of the prior shape under perspective distortion and scaling. This technique has a nice intuitive geometric interpretation and a relatively simple mathematical formulation. This method, however, only handles perspectivity (a six-parameter transformation), and is applicable to the limited class of star-shaped objects. Nevertheless, it provides the conceptual basis for the method developed in this chapter. Next, we generalize the algorithm to eight-parameter projectivity model and remove the restriction to star-shaped objects. Fundamental concepts from two-view geometry that are necessary to accomplish this generalization are presented in section 3.4. The embedding of the transformation model within the variational framework, and the minimization of the resulting functional, are considered in section 3.5. Experimental results using the eight-parameter model are provided in section 3.6.

3.2 Variational framework

3.2.1 Chan-Vese two-phase model

A general form of the region-based energy term has been derived in subsection 2.1.2 of the previous chapter. This chapter extends the two-phase segmentation functional proposed by Chan and Vese [11]. We will hereby review the Chan-Vese framework.

The Mumford and Shah [70] functional is based on the assumption that a gray level image $I: \Omega \rightarrow \mathbb{R}$ where $\Omega \subset \mathbb{R}^2$ is the image domain, can be fairly approximated by a piecewise smooth function u . The functional minimizer is the collection of the

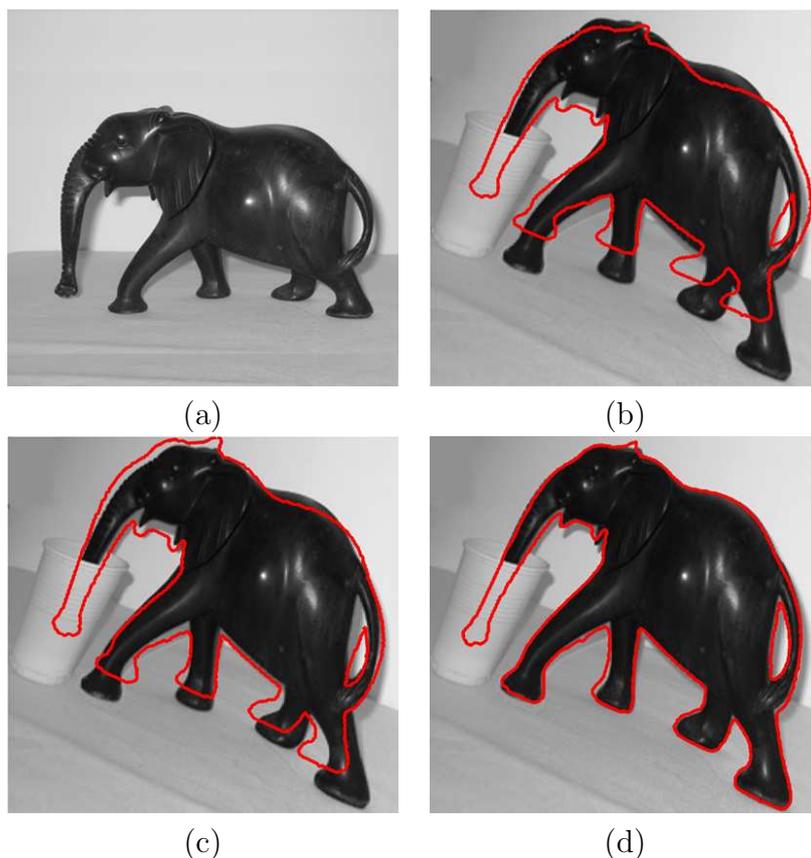


Figure 3.1: Registration, by the proposed method, of the contour of the object shown in (a) image using : (b) Similarity model (4 transformation parameters). (c) Affine model (6 transformation parameters). (d) Planar projective homography (8 transformation parameters). Note the mismatch in (b) and (c) despite the relatively small change in the camera view. Registration results of (b) and (c) are the best possible with similarity and affine transformations. They have been obtained after an extensive search of the respective parameter spaces. The corresponding transformation parameters are displayed in the last two rows of Table 3.3.

image edges C .

$$E(u, C) = \frac{1}{2} \int_{\Omega} (I - u)^2 d\mathbf{x} + \lambda \frac{1}{2} \int_{\Omega-C} |\nabla u|^2 d\mathbf{x} + \nu |C|, \quad (3.1)$$

where λ and ν are positive scalars and $\mathbf{x} = (x, y)$. Minimizing this functional is far from being trivial since the unknown function C appears in the integration domain. A reduced form of this functional is simply a restriction of u to a piecewise *constant* approximation of I , such that each connected component Ω_i , where $\cup_i \Omega_i = \Omega$ and $\Omega_i \cap \Omega_j = \emptyset$, has a constant gray level value u_i . This leads to the minimum partition

problem, where the functional takes the form:

$$E(u, C) = \frac{1}{2} \sum_i \int_{\Omega_i} (I - u_i)^2 d\mathbf{x} + \nu |C| \quad (3.2)$$

Using the level-set formulation, C is the zero level of a level-set function $\phi: \Omega \rightarrow \mathbb{R}$: $C = \{\mathbf{x} \in \Omega \mid \phi(\mathbf{x}) = 0\}$. Chan and Vese [11], inspired by the segmentation model of Mumford and Shah (3.1), suggested to segment an input image I using the Heaviside function of the evolving level-set function $H(\phi)$ as an indicator for the image object and background regions. The contour length was used as a regularization term. For the two-phase case the Chan-Vese functional takes the form:

$$\begin{aligned} E_{CV} &= \int_{\Omega} [(I - u_+)^2 H(\phi) + (I - u_-)^2 (1 - H(\phi))] \\ &+ \nu |\nabla H(\phi)| d\mathbf{x}. \end{aligned} \quad (3.3)$$

The scalars u_+ and u_- are alternately updated with the evolution of the level-set function. They take the average gray level values of the input image in the regions indicated by $\phi \geq 0$ and $\phi < 0$.

$$u_+ = \frac{\int I(\mathbf{x}) H(\phi) d\mathbf{x}}{\int H(\phi) d\mathbf{x}} \quad u_- = \frac{\int I(\mathbf{x}) (1 - H(\phi)) d\mathbf{x}}{\int (1 - H(\phi)) d\mathbf{x}} \quad (3.4)$$

Note that the integration domain of the reformulated functional (3.3) is now well-defined and known. The gradient descent equation for the evolution of ϕ is derived using the Euler-Lagrange equations for the functional (3.3):

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (I - u_+)^2 + (I - u_-)^2 \right]. \quad (3.5)$$

The energy functional (3.3) can be extended by adding a prior shape term [22]:

$$E(\phi, u_+, u_-) = E_{CV}(\phi, u_+, u_-) + \mu E_{shape}(\phi), \quad \mu \geq 0. \quad (3.6)$$

The inclusion of this shape term within the energy functional leads to three related questions:

1. What should be the representation of the prior shape?

2. What should be the similarity measure between the prior shape and the evolving segmented shape?
3. What is the group of transformations that can be accommodated by the above similarity measure?

The current work provides a comprehensive and innovative solution to these questions. Let $\tilde{\Omega} \subset \mathbb{R}^2$ be the reference image frame. The representation of the prior shape within the energy functional (3.6) is a 3D function $\tilde{\phi} : \tilde{\Omega} \rightarrow \mathbb{R}$ that embeds the contour \tilde{C} of the known shape:

$$\tilde{C} = \{\mathbf{x} \in \tilde{\Omega} \mid \tilde{\phi}(\mathbf{x}) = 0\}, \quad (3.7)$$

Positive and the negative values of $\tilde{\phi}$ correspond to object and background regions in $\tilde{\Omega}$ respectively. We suggest two alternative representations of $\tilde{\phi}$, that are discussed and compared in the subsequent sections. The formulation of $\tilde{\phi}$ leads to a definition of a dissimilarity measure with respect to the evolving segmentation. It is a weighted sum of the non-overlapping positive and negative regions of ϕ and $\tilde{\phi}$:

$$E_{shape}(\phi) = \int_{\Omega} \left(H_{\epsilon}(\phi(\mathbf{x})) - H_{\epsilon}(\tilde{\phi}(\mathbf{x})) \right)^2 d\mathbf{x} \quad (3.8)$$

This extension to the functional adds the following term to the evolution equation (3.5):

$$\delta_{\epsilon}(\phi) \left(H_{\epsilon}(\phi) - H_{\epsilon}(\tilde{\phi}) \right) \quad (3.9)$$

Thus, at each time step, ϕ will be modified in image regions where there is inconsistency between the object-background areas indicated by $H_{\epsilon}(\phi)$ and $H_{\epsilon}(\tilde{\phi})$. The change in ϕ is weighted by δ_{ϵ} , which is illustrated in Fig. 2.1b for several values of ϵ . The shape-term is further extended to incorporate possible transformations between the prior shape and the shape of interest. This is approached by applying a 3D transformation $\mathcal{T}_p : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ to $\tilde{\phi}$:

$$E_{shape}(\phi, T) = \int_{\Omega} \left(H_{\epsilon}(\phi(\mathbf{x})) - H_{\epsilon}(T(\tilde{\phi}(\mathbf{x}))) \right)^2 d\mathbf{x} \quad (3.10)$$

The reformulated energy functional becomes:

$$\begin{aligned} E(\phi) = & \int_{\Omega} \{(I - u_+)^2 H_{\epsilon}(\phi) + (I - u_-)^2 (1 - H_{\epsilon}(\phi)) \\ & + \nu |\nabla H_{\epsilon}(\phi)| + \mu [H_{\epsilon}(\phi) - H_{\epsilon}(T(\tilde{\phi}))]^2\} d\mathbf{x} \end{aligned} \quad (3.11)$$

The extended gradient descent equation for the evaluation of ϕ is:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = & \delta_{\epsilon}(\phi) \left[(I - u_-)^2 - (I - u_+)^2 + \nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right. \\ & \left. + \mu (H_{\epsilon}(\phi) - H_{\epsilon}(T(\tilde{\phi}))) \right]. \end{aligned} \quad (3.12)$$

Note that we do not enforce the evolving level-set function ϕ to resemble $T(\tilde{\phi})$. Instead, we demand similarity of the regions within the respective contours. Thus, ϕ is not necessarily a distance function. Therefore T can be used to accommodate for *planar projective transformations* between the prior contour \tilde{C} and the evolving segmenting contour C .

In section 3.3 we introduce the six-parameter model. It is based on a unique representation of the prior shape $\tilde{\phi}$, which simplifies its perspective mapping to the shape of interest. The eight-parameter model, presented in section 3.4, generalizes the framework to projectivity.

3.3 Perspectivity

3.3.1 Cone of rays

Consider a set of rays, defined by an object in 3D space and a camera center. An image is obtained by the intersection of these rays with a plane. This set is usually referred to as a *cone of rays*, although it is not a cone in the classical sense [38]. Now, suppose that this cone of rays is intersected by two planes, as shown in Fig. 3.2. Then, there exists a perspective transformation H mapping one image onto the other. This means that the images obtained by the same camera center may be mapped to one another by a plane projective transformation [28, 29, 38].

Let f and f' be the image planes of two cameras, having the same camera center,

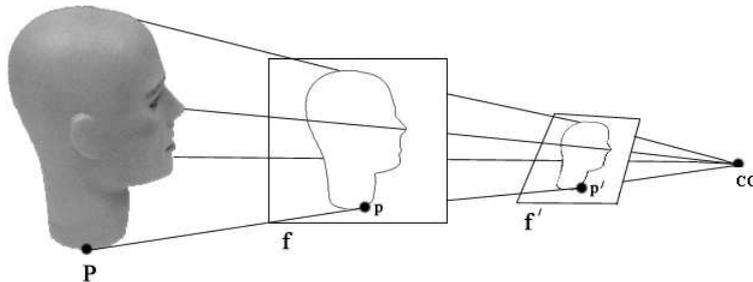


Figure 3.2: The cone of rays. An image is defined by the intersection of this set of rays with a plane. A ray between the camera center CC and a 3D world point P intersects the plane in the image points $p \in f$ and $p' \in f'$. The relation between all such image points is a perspective mapping: $p' = Hp$. Adapted from [38].

with projection matrices M and M' respectively. For simplicity, the coordinate system of the first camera is chosen to coincide with the world coordinate system, so that $M = K[\mathcal{I} \mid 0]$, where K is the 3×3 camera calibration matrix and \mathcal{I} is the identity matrix. The projection matrix of the second camera is $M' = K'[R \mid 0]$, where K' is the calibration matrix and R is the relative rotation between the coordinate systems of the cameras.

Consider two image points $p \in f$ and $p' \in f'$ of a 3D world point P . Let \mathbf{x} , \mathbf{x}' and \mathbf{X} be their corresponding homogeneous coordinates. Thus, $\mathbf{x} = K[\mathcal{I} \mid 0]\mathbf{X}$, $\mathbf{x}' = K'[R \mid 0]\mathbf{X}$ and the mapping between \mathbf{x} and \mathbf{x}' is $\mathbf{x}' = K'RK^{-1}\mathbf{x} = H\mathbf{x}$, with $H = K'RK^{-1}$. This relation can be simplified by working with a normalized coordinate system¹. When K and K' only differ by their focal length, $K'K^{-1} = \text{diag}(k, k, 1)$ where k is the ratio of the focal lengths.

The transformations of zoom and camera rotation can be expressed as a movement of the image plane while maintaining the camera center fixed. Note that these are particular cases of *perspectivity*. Handling displacements between the centers of the cameras (without parallax) requires extension of the geometric model. These steps are considered in following subsections. We proceed to introduce a prior shape representation derived from the cone of rays.

¹Internal camera calibration is beyond the scope of this work. Therefore, we will present the internal camera parameters in terms of pixels, assuming square pixels. Moreover, the origin of coordinates in the image plane is set at the camera's principal point. The term *normalized coordinate system* is adopted from [29].

3.3.2 Generalized cone

A generalized cone² or a conical surface, is a ruled surface generated by a moving line (the generator) that passes through a fixed point (the vertex) and continually intersects a fixed planar curve (the directrix) [15]. Let $P_v = (X_v, Y_v, Z_{vertex})$ denote the cone vertex, and let $p_v = (x_v, y_v)$ be the projection of the vertex on the directrix plane. We set, without loss of generality, $X_v = x_v$ and $Y_v = y_v$. Now, consider a directrix, $C = p(s) = (x(s), y(s))$ which is a closed contour, parameterized by arc-length s , of an object shape in the plane $Z = Z_{plane} = 0$. The generalized cone surface is defined by:

$$\tilde{\phi}(r, s) = \tilde{\phi}((1 - r)p(s) + rp_v) = (1 - r)Z_{plane} + rZ_{vertex} \quad (3.13)$$

where r varies smoothly from 1, that corresponds to the vertex, via 0, the directrix, to some convenient negative value.

When the vertex of the generalized cone is located at the camera center, the definition of the generalized cone coincides with that of the cone of rays, presented in subsection 3.4. It follows that by planar slicing of the generalized cone, one can generate new image views as though they had been taken with a camera under the perspective model³. Note, however that the construction of the generalized cone (3.13), implies that the first image plane is perpendicular to the principal axis. Extension of the allowable transformation between corresponding image points \mathbf{x}' and \mathbf{x} to translation is possible if the object contour is planar and parallel to the first image plane.

3.3.3 Formulation of the transformation

We will now consider the representation of the transformation $T(\tilde{\phi})$ of the prior shape in the energy functional. The recovery of the transformation parameters, given the

²The concept of generalized cone (or cylinder) in computer vision has been introduced to model 3D objects [4, 66]. Its geometrical properties have been intensively investigated, see [31, 81] and references therein.

³There is, however, one exception to this analogy. The intersection of a cone and a plane is either a closed curve, an open curve or a point. In projective geometry terminology, the latter two correspond to projection of finite points in the first image plane to infinity. We do not consider ideal points and planes at infinity. Phrasing it explicitly, our only concern is the mapping of a given closed curve to another closed curve.

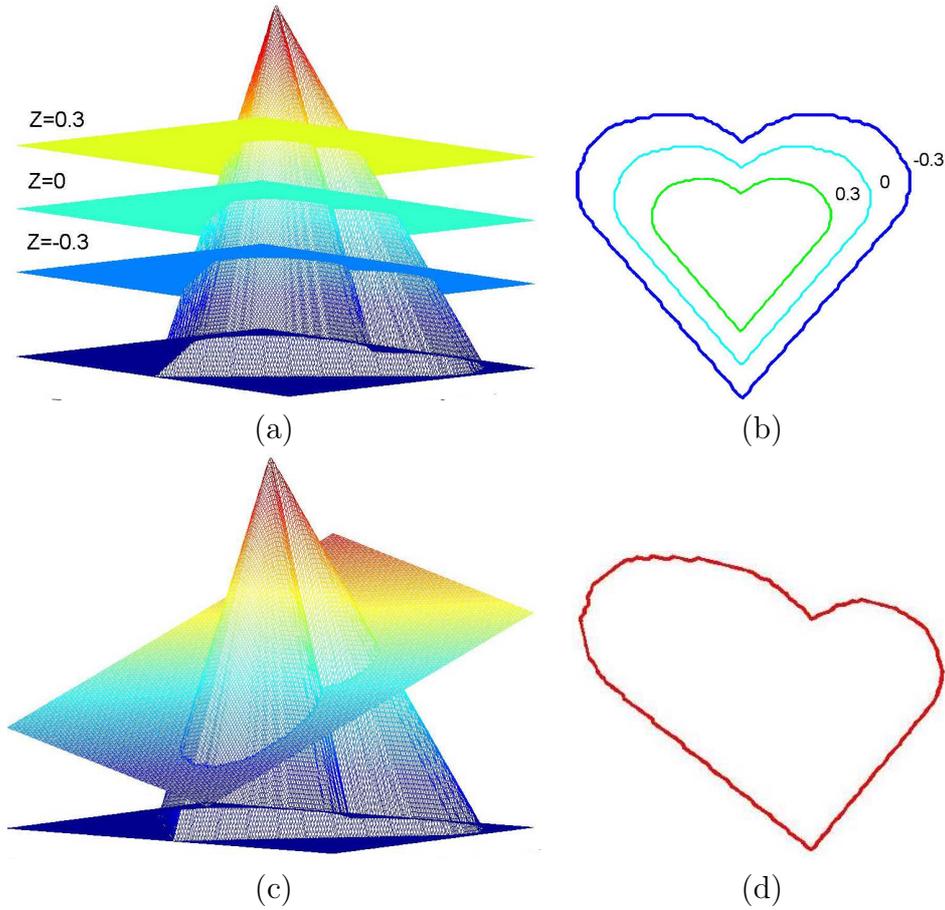


Figure 3.3: (a) A generalized cone is sliced by three planes, at $Z = 0.3$, $Z = 0$ and $Z = -0.3$. (b) The resulting intersections. (c) A generalized cone is intersected by an inclined plane: $ax + by + cz + d = 0$. (d) The resulting contour.

prior contour and the curve generated by the zero-crossing of the estimated level-set function, is described subsequently.

The reference level-set function $\tilde{\phi}$ embeds the prior contour according to (3.13). For simplicity, the vertex height Z_{vertex} , which corresponds to the focal length, will be set to 1. The prior contour \tilde{C} is the cone's directrix.

In order to minimize the energy functional (3.11), one has to apply a gradient descent process that calls for the evaluation of ϕ simultaneously with the recovery of the transformation T of the function $\tilde{\phi}$. We demonstrate this for transformations T that consist of translation and rotation of the generalized cone, and correspond to scaling, translation, rotation and perspective distortion in the image.

Scaling and general translation

Assume first that the focal length is changed. This approximately corresponds to translation of the image plane along the optical axis. In the image, the effect is scaling. Figs. 3.3a-b illustrate this notion: as the planar section of the generalized cone is closer to its vertex, the cross-section shape is smaller. Equivalently, the generalized cone can be displaced in the direction of the Z axis, while the intersecting plane remains stationary at $Z = 0$. Formally, in this case, $T(\tilde{\phi}) = \tilde{\phi} + t_z$.

To account also for translation in the image, we displace the generalized cone by $\mathbf{t} = (t_x, t_y, t_z)^T$. The corresponding shape term in the energy functional is then

$$E_{shape}(\phi) = \int_{\Omega} (H(\phi)(x, y) - H(\tilde{\phi}(x + t_x, y + t_y) + t_z))^2 dx.$$

Rotation and compound motion

Consider a tilted planar cut of the generalized cone, as shown in Figs. 3.3c-d. The resulting contour is perspectively deformed, as a function of the inclination of the intersecting plane and its proximity to the vertex of the cone. Equivalently, one may rotate the generalized cone around its vertex, and zero-cross to get the same perspective transformation.

We denote by γ , β and α the Euler rotation angles around the Z, Y and X axes respectively, in that order of rotation. They determine the relative orientation between the first and the second camera coordinate systems. The rotation matrix $R \in \mathbb{R}^{3 \times 3}$ operating on a vector $(x, y, z)^T$ takes the form:

$$R = \begin{bmatrix} c_{\beta}c_{\gamma} & c_{\beta}s_{\gamma} & -s_{\beta} \\ s_{\alpha}s_{\beta}c_{\gamma} - c_{\alpha}s_{\gamma} & s_{\alpha}s_{\beta}s_{\gamma} + c_{\alpha}c_{\gamma} & s_{\alpha}c_{\beta} \\ c_{\alpha}s_{\beta}c_{\gamma} + s_{\alpha}s_{\gamma} & c_{\alpha}s_{\beta}s_{\gamma} - s_{\alpha}s_{\gamma} & c_{\alpha}c_{\beta} \end{bmatrix} \quad (3.14)$$

where s_{α} is shorthand for $\sin(\alpha)$ and c_{α} for $\cos(\alpha)$. General rotation and translation of the generalized cone by R and \mathbf{t} is expressed as $(x', y', T(\tilde{\phi})) = R(x, y, \tilde{\phi}) + \mathbf{t}$.

Explicitly,

$$\begin{aligned}
 x' &= c_\beta c_\gamma x + c_\beta s_\gamma y - s_\beta + t_x \\
 y' &= (s_\alpha s_\beta c_\gamma - c_\alpha s_\gamma) x + (s_\alpha s_\beta s_\gamma + c_\alpha c_\gamma) y + s_\alpha c_\beta + t_y \\
 z' &= (c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma) x + (c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma) y + c_\alpha c_\beta + t_z
 \end{aligned} \tag{3.15}$$

where $z = \tilde{\phi}$ and $z' = T(\tilde{\phi})$.

3.3.4 The six-parameter algorithm

We summarize the six-parameter algorithm for image segmentation and recovery of the transformation between the current and prior object instances, assuming the following setup.

The input consists of two image frames f and f' of the same object, taken with the same camera, but under different viewing conditions. The boundary \tilde{C} of the object in f is known. The image f' has to be segmented. The world plane is assumed to be parallel to the first image plane f , and f is assumed to be at distance 1 from the camera center. The second image plane f' is tilted and shifted relative to the first one.

1. Given the contour \tilde{C} , construct a generalized cone $\tilde{\phi}$, using (3.13) with $Z_{vertex} = 1$.
2. Choose some initial level-set function ϕ , for example a right circular cone. Its zero-level set is the initial segmenting contour.
3. Set initial values (e.g. zero) for α , β , γ , t_x , t_y and t_z .
4. Compute the average gray levels of the current object and background regions, u_+ and u_- , using (3.4). Recall that the gray level function of the image frame f' is I .
5. Translate (by \mathbf{t}) and then rotate (by R) the prior shape representation $\tilde{\phi}$, using the coordinate transformation (3.15) with the current estimate of the transformation parameters.

6. Update ϕ using the gradient descent equation (3.12).
7. Update the transformation parameters $\alpha, \beta, \gamma, t_x, t_y$ and t_z using the derivatives of the cost functional (3.11) with respect to each parameter.
8. Repeat steps 4-7 until convergence.

The algorithm is exemplified in Fig. 3.4. The prior image is shown in Fig. 3.4a; its segmentation is known. Fig. 3.4b shows a new instance of the mannequin head, rotated and translated with respect to the reference pose. The hat creates significant occlusion. Note that the head is placed on a plastic base of a similar grey-level. The initial segmenting contour, obtained by zero-crossing the initial level-set function (right circular cone) is shown. Successful segmentation using the six-parameter algorithm is shown in Fig. 3.4c. Note the precise tracing of the profile and the recovery of the occluded crown, despite the perspective distortion. Fig. 3.4d is the generalized cone that represents the prior shape (a). The final state of the level-set function ϕ is presented in Fig. 3.4e. Note that ϕ and $\mathcal{T}_p(\tilde{\phi})$ resemble in terms of their Heaviside functions - that is by their zero-crossings (the final contour), but not in their entire shapes. Since the actual transformation was not measured, the recovered transformation parameters are confirmed in Fig. 3.4f by comparing the final segmenting contour with the prior shape, transformed according to the recovered parameters.

3.4 From six to eight parameters

3.4.1 Planar projective homography

To generalize the admissible geometric relation between two corresponding shape contours we review the concept of *planar projective homography*. The equivalence of geometric projectivity and algebraic homography is supported by the theorems of Desargues [102]. Planar projective homography (projectivity) is a mapping $H : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ such that the points p_i are collinear if and only if $H(p_i)$ are collinear (projectivity preserves lines) [38, 102].

The relation between corresponding views of points on a plane Π (world plane) in a 3D space, as is illustrated in Fig. 3.5, can be modeled by a planar homography

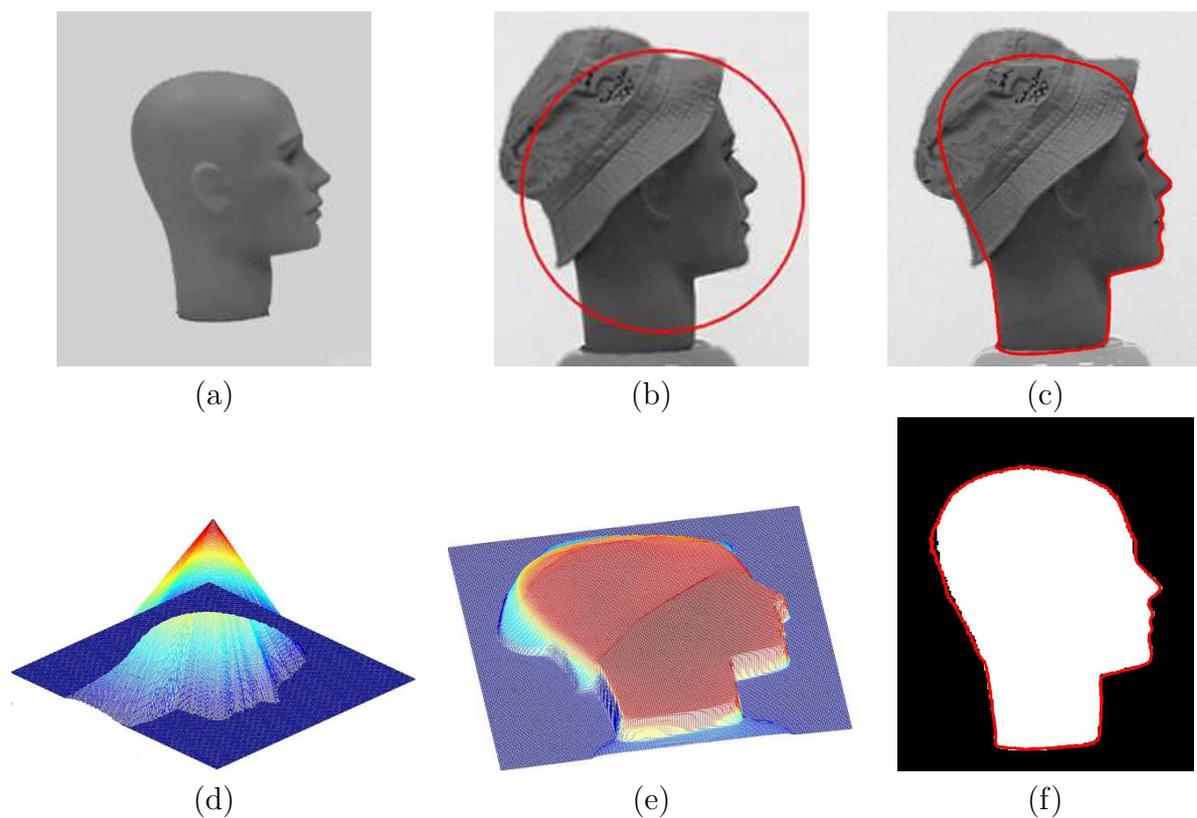


Figure 3.4: Segmentation results of a mannequin head image. (a) Prior image (mannequin head). (b) The image to be segmented. The mannequin head is rotated, translated and partly occluded. The initial contour is shown. (c) Successful segmentation using the six-parameter algorithm. (d) The generalized cone based on the prior contour of (a). Only the positive part of the cone is shown. (e) The final state of the level-set function ϕ . (f) Validation: The dark shape is the prior silhouette, transformed according to the recovered transformation parameters. Shown red is the final contour, as in (c).

induced by the plane. An explicit expression for the induced homography can be derived as follows: Let p and p' be two views of a world point $P \in \Pi$, in two camera frames f and f' respectively. Let \mathbf{x} , \mathbf{x}' and \mathbf{X} be their corresponding homogeneous coordinates. As in subsection 3.3.1, M and M' are the projection matrices of the respective cameras, where $M = K[\mathcal{I} \mid 0]$. However, since the cameras generally do not have a common center, $M' = K'[R \mid \mathbf{t}]$, where \mathbf{t} is the relative translation between the cameras. Thus, $\mathbf{x} = K[\mathcal{I} \mid 0]\mathbf{X}$ and $\mathbf{x}' = K'[R \mid \mathbf{t}]\mathbf{X}$.

Let \mathbf{n} be the unit normal vector to the plane Π , and let $d > 0$ denote the distance of Π from the optical center of the first camera. The linear transformation from \mathbf{x} to

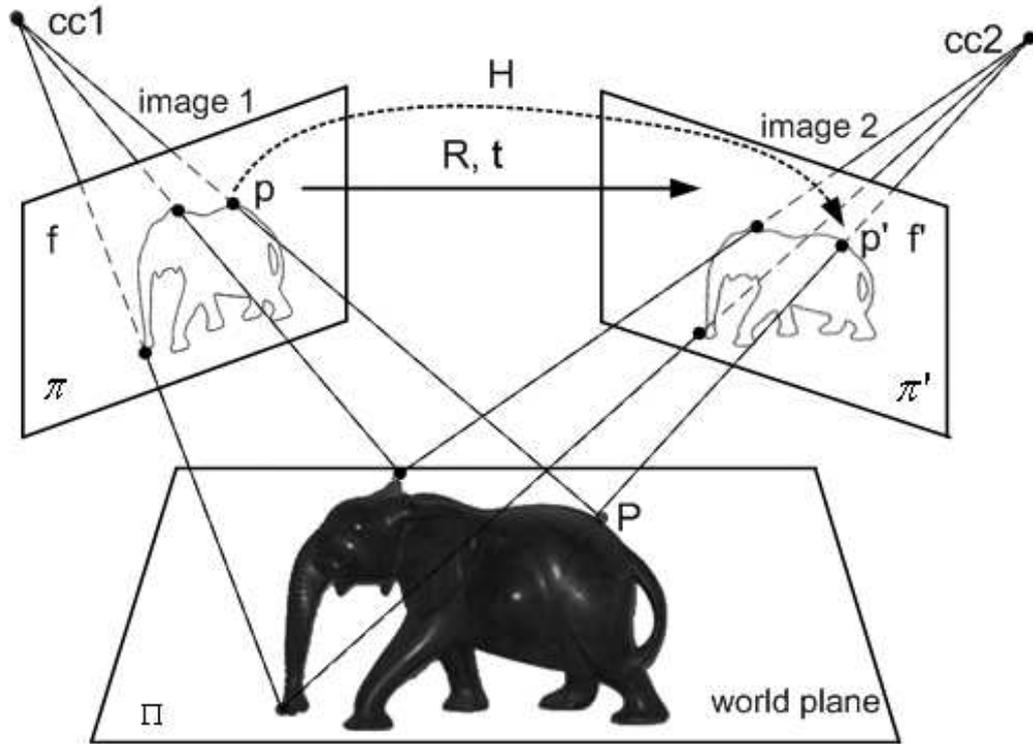


Figure 3.5: The homography induced by a plane. The ray that intersects the first image plane at a point p meets the world plane Π at a point P . The point P is projected to a point p' in the second image plane. The mapping from p to p' is the homography, denoted by H , induced by the plane Π . Illustrated after [38]. In this research, the corresponding points p and p' are on corresponding planar contours.

\mathbf{x}' can be expressed as

$$\mathbf{x}' = K' \left(R + \frac{1}{d} \mathbf{t} \mathbf{n}^T \right) K^{-1} \mathbf{x} = H \mathbf{x} . \quad (3.16)$$

Using the same assumptions as in subsection 3.3.1, we set K and K' to the identity matrix \mathcal{I} . Thus, the planar homography matrix takes the form

$$H = R + \frac{1}{d} \mathbf{t} \mathbf{n}^T . \quad (3.17)$$

A detailed derivation can be found in [29, 38, 63]. The matrix H is determined by the translation and rotation between the two cameras $\{R, \mathbf{t}\}$, and by the structure parameters $\{\mathbf{n}, d\}$ of the world plane Π . Note that only the ratio \mathbf{t}/d can be recovered from H . We proceed to show how the structure of the homography can be used to recover the relation between the camera frames and thus between the respective shape

contours.

3.4.2 Implicit recovery of the homography

A nonsingular homography between two image planes f and f' is determined (up to a scale factor) by four pairs of corresponding points $p \in f$ and $p' \in f'$, with no triplet of points in either plane being collinear. Consider the homography matrix H ,

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (3.18)$$

and let $\mathbf{x} = (x, y, 1)$ and $\mathbf{x}' = (x', y', 1)$ be the homogeneous representations of the points p and p' , such that $\mathbf{x}' = H\mathbf{x}$. The eight unknowns of H (the ratios of its nine entries) can be recovered by solving at least four pairs of equations of the form:

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \quad (3.19)$$

Classic approaches recover H by solving an over-determined set of equations like (3.19). The translation and rotation (R, \mathbf{t}) between the image planes, and the scene structure (\mathbf{n}, d) , are recovered by decomposition of the known homography matrix (see [29, 38, 63] and references therein). In contrast, our novel approach calculates the homography directly in its *explicit* form (3.17). Rather than relying on point correspondence, we match two corresponding contours of the shape of interest using calculus of variations. Note that since the recovery of the homography and the segmentation process are simultaneous, only the reference shape is known in advance. The prior shape is registered to the shape being segmented as part of its detection and extraction.

3.4.3 Explicit recovery of the homography

We now use the explicit formulation of the homography (3.17) to reformulate equation (3.19). Consider first the special case in which the world plane and the first image plane coincide, and d is set to 1. In this case, the normal to the world plane, \mathbf{n}

is $(0, 0, 1)$. Substituting d and \mathbf{n} in (3.17), we obtain the entries of the homography matrix H : $h_{13} = R_{13} + t_x$, $h_{23} = R_{23} + t_y$, $h_{33} = R_{33} + t_z$ and $h_{ij} = R_{ij}$ otherwise.

Generally, the world plane is not perpendicular to the optical axis of the first camera, thus $\mathbf{n} \neq (0, 0, 1)$. As before, we represent \mathbf{n} using the coordinate system of the first camera. The unit vector \mathbf{n} is obtained by first rotating the vector $(0, 0, 1)$ by an angle ξ around the y -axis and then by an angle ψ around the x -axis. Hence, $\mathbf{n} = (-\sin \xi, \sin \psi \cos \xi, \cos \psi \cos \xi)$. Substituting R , \mathbf{t} and \mathbf{n} in (3.17), we obtain the components of the homography matrix H :

$$\begin{aligned}
h_{11} &= \cos \beta \cos \gamma - \frac{t_x}{d} \sin \xi \\
h_{12} &= \cos \beta \sin \gamma + \frac{t_x}{d} \sin \psi \cos \xi \\
h_{13} &= -\sin \beta + \frac{t_x}{d} \cos \psi \cos \xi \\
h_{21} &= \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma - \frac{t_y}{d} \sin \xi \\
h_{22} &= \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma + \frac{t_y}{d} \sin \psi \cos \xi \\
h_{23} &= \sin \alpha \cos \beta + \frac{t_y}{d} \cos \psi \cos \xi \\
h_{31} &= \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma - \frac{t_z}{d} \sin \xi \\
h_{32} &= \cos \alpha \sin \beta \sin \gamma - \sin \alpha \sin \gamma + \frac{t_z}{d} \sin \psi \cos \xi \\
h_{33} &= \cos \alpha \cos \beta + \frac{t_z}{d} \cos \psi \cos \xi
\end{aligned} \tag{3.20}$$

Substituting the entries of H in (3.19), one can explicitly relate shape points between the two views. We will use the calculus of variations to relate the shape contours and recover the translation, rotation and structure parameters.

3.5 Objective functional

3.5.1 Representation of the prior shape

The generalized cone representation of the prior shape, discussed in section 3.3, is inadequate when the shape contour is seen from two different view points and the world plane is in general position. Furthermore, zero-crossing the transformed generalized cone $H(T(\tilde{\phi}))$, as described in 3.3.3, requires $T(\tilde{\phi})$ to be single-valued. This restricts the group of admissible shape contours and the range of allowable transformations. We therefore present an alternative representation of the prior shape that can be

applied to general planar shapes.

Consider a prior image $f : \Omega \rightarrow \mathbb{R}^+$ with labeled object and background regions Ω_+ and Ω_- respectively. The prior shape is represented by a function $\tilde{\phi} \in \mathbb{R}^3$:

$$\tilde{\phi}(x, y) = \begin{cases} 1 & f(x, y) \in \Omega_+ \\ 0 & f(x, y) \in \Omega_- \end{cases} \quad (3.21)$$

Let $T : (\tilde{\phi}(x, y), x, y) \rightarrow (\tilde{\phi}'(x', y'), x', y')$ denote the projective transformation, that maps a point (x, y) and its $\tilde{\phi}$ value to a (projectively) equivalent point (x', y') , with $\tilde{\phi}'(x', y') = \tilde{\phi}(x, y)$. Thus, the binary function $\tilde{\phi}'(x', y')$ is obtained from $\tilde{\phi}(x, y)$ by coordinate transformation.

Expressions for x' and y' are obtained by substituting the explicit entries for the homography matrix (3.20) in (3.19). For example, when $\xi = \psi = 0$ and $d = 1$ the expressions are:

$$x' = \frac{c_\beta c_\gamma x + c_\beta s_\gamma y - s_\beta + t_x}{(c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma) x + (c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma) y + c_\alpha c_\beta + t_z} \quad (3.22)$$

$$y' = \frac{(s_\alpha s_\beta c_\gamma - c_\alpha s_\gamma) x + (s_\alpha s_\beta s_\gamma + c_\alpha c_\gamma) y + s_\alpha c_\beta + t_y}{(c_\alpha s_\beta c_\gamma + s_\alpha s_\gamma) x + (c_\alpha s_\beta s_\gamma - s_\alpha c_\gamma) y + c_\alpha c_\beta + t_z} \quad (3.23)$$

The representation (3.21) of the shape prior is simpler than that using the generalized cone (section 3.3), but leads to complex expressions for the transformed coordinates x' and y' . For example, the coordinates given in Eqs. (3.22 - 3.23) are the quotients of the corresponding expressions in (3.15). In the proposed algorithm, at each time step one re-evaluates the homography matrix entries (3.20), based on the estimated transformation parameters. The coordinate transformation T is applied to the representation $\tilde{\phi}$ of the prior shape. The transformed representation $T(\tilde{\phi}(x, y), x, y)$ is substituted in the gradient descent equation for ϕ (3.12).

3.5.2 Recovery of the transformation

In order to evolve the level set function ϕ according to (3.12), one has to simultaneously recover the transformation $T(\tilde{\phi}, x, y)$ of the prior level-set function. The transformation parameters $(\alpha, \beta, \gamma, t_x/d, t_y/d, t_z/d, \psi$ and $\xi)$ are evaluated via

the gradient descent equations obtained by minimizing the energy functional (3.11) with respect to each of them. The general gradient descent equation for each of the transformation parameters (denoted here by η) is of the form:

$$\frac{\partial \eta}{\partial t} = 2\mu \int_{\Omega} \delta_{\epsilon}(T(\tilde{\phi})) \left(H_{\epsilon}(\phi) - H_{\epsilon}(\mathcal{T}_P(\tilde{\phi})) \right) \frac{\partial \mathcal{I}_P(\tilde{\phi}, \eta)}{\partial \eta} d\mathbf{x} \quad (3.24)$$

where

$$\begin{aligned} \frac{\partial \mathcal{I}_P(\tilde{\phi}, \eta)}{\partial \eta} &= \frac{\partial \mathcal{I}_P(\tilde{\phi})}{\partial x} \left(\frac{\partial x}{\partial x'} \frac{\partial x'}{\partial \eta} + \frac{\partial x}{\partial y'} \frac{\partial y'}{\partial \eta} \right) \\ &+ \frac{\partial \mathcal{I}_P(\tilde{\phi})}{\partial y} \left(\frac{\partial y}{\partial x'} \frac{\partial x'}{\partial \eta} + \frac{\partial y}{\partial y'} \frac{\partial y'}{\partial \eta} \right) \end{aligned} \quad (3.25)$$

The partial derivatives $\partial T(\tilde{\phi})/\partial x$ and $\partial T(\tilde{\phi})/\partial y$ are computed numerically using the finite difference method. The derivatives $\partial x'/\partial \eta$ and $\partial y'/\partial \eta$ with respect to each transformation parameter η have been derived analytically. From the implicit function theorem we obtain

$$\begin{aligned} \frac{\partial x}{\partial x'} &= \frac{\frac{\partial y'}{\partial y}}{\frac{\partial y'}{\partial y} \frac{\partial x'}{\partial x} - \frac{\partial y'}{\partial x} \frac{\partial x'}{\partial y}} \\ \frac{\partial x}{\partial y'} &= \frac{\frac{\partial x'}{\partial y}}{\frac{\partial y'}{\partial x} \frac{\partial x'}{\partial y} - \frac{\partial y'}{\partial y} \frac{\partial x'}{\partial x}} \\ \frac{\partial y}{\partial x'} &= \frac{\frac{\partial y'}{\partial x}}{\frac{\partial y'}{\partial x} \frac{\partial x'}{\partial y} - \frac{\partial y'}{\partial x} \frac{\partial x'}{\partial y}} \\ \frac{\partial y}{\partial y'} &= \frac{\frac{\partial x'}{\partial x}}{\frac{\partial y'}{\partial y} \frac{\partial x'}{\partial x} - \frac{\partial y'}{\partial x} \frac{\partial x'}{\partial y}} \end{aligned} \quad (3.26)$$

Finally, the equations for $\frac{\partial x'}{\partial x}$, $\frac{\partial x'}{\partial y}$, $\frac{\partial y'}{\partial x}$, $\frac{\partial y'}{\partial y}$ are evaluated by differentiating the expressions (3.22 - 3.23) for x' and y' with respect to x and y .

3.5.3 The eight-parameter algorithm

Algorithm

We summarize the proposed algorithm assuming the following setup. The inputs are two image frames f and f' of the same object, taken with identical cameras, but from

different viewing positions. The segmentation of the reference object in f is known and its contour is assumed to be approximately planar. The segmented reference object is used to construct the prior shape representation $\tilde{\phi}$ according to (3.21).

1. Choose some initial level-set function ϕ , for example a right circular cone. Its zero level-set is the initial segmenting contour.
2. Set initial values (e.g. zero) for the transformation parameters $\alpha, \beta, \gamma, t_x, t_y, t_z, \xi$ and ψ . If d is unknown, set $d = 1$.
3. Compute the average gray level values of the current object and background pixels, u_+ and u_- , using (3.4). Recall that the gray level function of the image frame f' is I .
4. Apply a coordinate transformation to the prior shape representation $\tilde{\phi}$ (substitute (3.20) in (3.19)) with the current estimate of the transformation parameters.
5. Update ϕ using the gradient descent equation (3.12).
6. Update the transformation parameters $\alpha, \beta, \gamma, t_x, t_y, t_z, \psi$ and ξ as explained in 3.5.4.
7. Repeat steps 3-6 until convergence.

The segmentation and registration steps are interleaved, thus realizing the power of the prior-based segmentation approach. Note that the algorithm directly recovers the transformation parameters, thus the cumbersome task of decomposing the homography matrix is avoided [29, 38, 63].

3.5.4 Optimization

Global optimization problems are common in computer vision. Here, the cost functional (3.11) may have several local minima with respect to the parameters and the evolving level set function ϕ . Direct update of the parameters via their the derivatives (3.24) may lead to an undesired local minimum, as exemplified in Fig. 3.6.

It is known [109] that the global minimum of a general function (or functional) above a continuous domain cannot be found in finite time. Nevertheless, global optimization strategies can be successfully applied to well-behaved functions. In step 6 of the algorithm, the transformation parameters are determined by multidimensional minimization, using the Nelder–Mead downhill simplex algorithm [71] followed by the gradient based Quasi-Newton method [14]. The direct search method of Nelder-Mead is based on evaluating a function at the vertices of a simplex, then iteratively shrinking the simplex as better points are found until some desired bound is obtained. The nine eight-dimensional vertices are randomly initialized within the limits of the parameters. The outcome of the Nelder-Mead algorithm is used as the starting point for Quasi-Newton minimization employing the partial derivatives (3.24) for fine tuning of the search results. The simplex algorithm works, in our examples, moderately well. In most cases, the globally minimal cost (corresponding to shape overlap) was identical to the result of quasi-Newton search from the zero starting point. The Quasi-Newton algorithm has a quadratic rate of convergence, but when applied by itself finds the local minimum in the vicinity of its initial point.

The stability of the resulting shape with respect to small errors in the minimizing transformation parameters was examined numerically on several images and transformation examples. The value of the shape term in the functional is taken as a measure for the shape variation. It was found that the minimizing shape varies smoothly with the change in the transformation parameters. The results differ slightly from image to image. In Fig. 3.7 we depict the results of varying one parameter while keeping the seven others fixed in the ground truth. The degree of stability is defined by the Hessian of the shape term at that point. Few parameters are more sensitive to such a variation. For a better understanding of the stability the correlations (or higher order derivatives) should be calculated. This is under current study.

3.6 Experimental results

The algorithm developed in this chapter is capable of prior-based segmentation and shape registration in the presence of perspective distortion. We demonstrate its operation on a variety of examples. In each example, we show the input image (the image

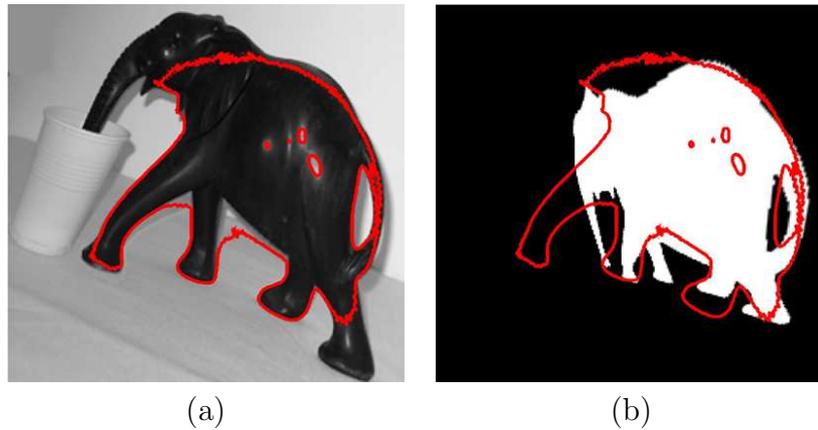


Figure 3.6: Failed segmentation and registration due to convergence to a local minimum. (a) A wrong segmenting contour (red), drawn on the input image. (b) The same contour, drawn on the incorrectly transformed prior shape. To avoid convergence to an incorrect local minimum, we use the Nelder–Mead downhill simplex algorithm followed by the gradient based Quasi-Newton minimization.

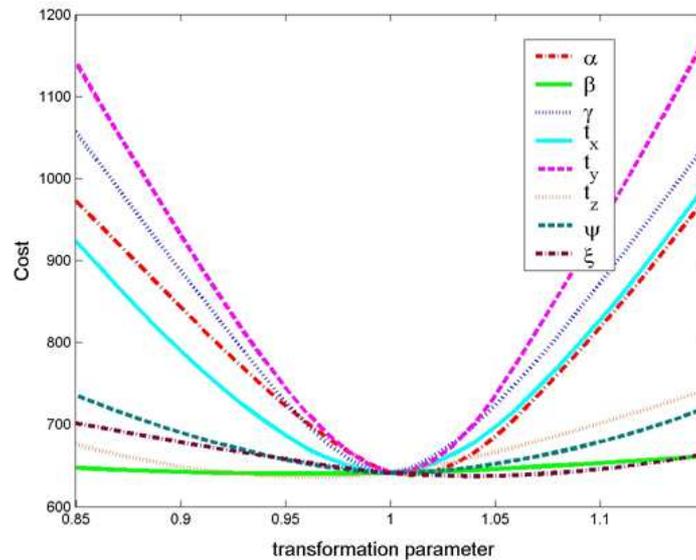


Figure 3.7: The shape dissimilarity measure (Eq. 3.10) as a function of each of the transformation parameters while keeping the seven others fixed in the ground truth.

to segment), and the reference image, from which the prior-shape was obtained. In addition to the prior-based segmentation results, we illustrate the mismatch between the input image and the reference image, and verify the transformation parameters estimated by the algorithm. In all the experiments we set $dt = 0.1$, $\epsilon = 1$ and $d = 1$. The contributions of each term in the gradient descent equation of ϕ (3.12) are normalized to $[-1, 1]$, thus ν and μ are set to 1. Modifying the weight of the prior term is done by normalizing the term to $[-\mu, \mu]$ with $\mu > 0$. Examples for such modifications are shown in Fig. 3.10f-i. In all the examined examples the number of iterations⁴ needed for convergence was between 50 and 100. The reader is also referred to the web page [82] which contains video clips that visually demonstrate the convergence of the segmentation processes for the examples shown in Fig. 3.9, Fig. 3.10 and Fig. 3.13.

We start by presenting two synthetic examples (Fig. 3.8 and Fig. 3.9) that allow comparison with the ground truth, and proceed to show segmentation of various real images. The first example (Fig. 3.8 and Table 3.1) demonstrates successful segmentation and registration for a wide range of projective transformations between a prior shape, shown in Fig. 3.8a and its different appearances, presented in Figs. 3.8b-f. The recovered transformation parameters are compared with the true ones in Table 3.1. Without the prior shape term, the final contour (red) may either incorrectly segment the white objects (as in Fig 3.8g) or the black unconnected object (as in Fig 3.8h), depending on the initial contour. When the prior shape term is over-stressed, the final segmenting contour ignores the actual image data (Fig 3.8i). This figure also displays the significant mismatch between the prior shape and the shape to segment.

Consider next the synthetic reference image shown in Fig. 3.9a, that contains several components of different sizes and gray levels. The prior shape was obtained by thresholding. The image to segment is shown in Fig. 3.9b, together with the initial contour. The input image is a noisy⁵, transformed and corrupted version of the image in Fig. 3.9a. Successful segmentation is demonstrated in Fig. 3.9c. The recovered transformation parameters are validated by comparison to the true transformation parameters in Table 3.2. The misalignment between the image to segment and the

⁴Repetitions of steps 3-6 in the eight parameter algorithm.

⁵Zero-mean Gaussian noise with STD equal to 14% of the grey-level dynamic range.

Transformation parameter	α	β	γ	t_x	t_y	t_z	ψ	ξ
Fig 3.8c True	0.200 ⁰	-0.200 ⁰	-15.00 ⁰	-5.00	5.00	0.100	0.100 ⁰	-0.200 ⁰
Recovered	0.212 ⁰	-0.214 ⁰	-14.45 ⁰	-5.09	4.66	0.098	0.123 ⁰	-0.269 ⁰
Fig 3.8d True	0.200 ⁰	0.100 ⁰	20.00 ⁰	-10.00	10.00	-0.050	-0.500 ⁰	-0.500 ⁰
Recovered	0.197 ⁰	0.104 ⁰	20.68 ⁰	-9.82	9.91	-0.040	-0.472 ⁰	-0.495 ⁰
Fig 3.8e True	-0.300 ⁰	-0.300 ⁰	-50.00 ⁰	-10.00	10.00	0.100	0.700 ⁰	0.700 ⁰
Recovered	-0.302 ⁰	-0.300 ⁰	-49.74 ⁰	-10.19	10.00	0.111	0.718 ⁰	0.678 ⁰
Fig 3.8f True	-0.300 ⁰	-0.300 ⁰	60.00 ⁰	-10.00	15.00	0.300	-0.300 ⁰	0.700 ⁰
Recovered	-0.299 ⁰	-0.309 ⁰	60.21 ⁰	-10.16	14.96	0.309	-0.281 ⁰	0.701 ⁰

Table 3.1: A comparison of the recovered and true transformation parameters, for the examples shown in Fig. 3.8

reference image is shown in Fig. 3.9d. Without using the prior shape, the input image is difficult to segment, see Fig. 3.9e.

Real images are used in all other examples, Figs. 3.10-3.13. Consider the two views of the toy elephant shown in Figs. 3.10a-b. As seen in Fig. 3.10c, the segmenting contour precisely tracks the outline of the elephant, and recovers the missing part of its trunk. This segmenting contour is the zero level-set of the final evolving function ϕ shown in Fig. 3.10d. The accuracy of the recovered transformation is visually verified in Fig. 3.10e. To demonstrate the influence of the prior shape weight on the final segmentation, we have conducted several experiments with various values of μ , shown in Figs. 3.10f-i. In Fig. 3.10f, the prior term is over-stressed ($\mu \gg 1$) thus the data is ignored and the transformation between the prior and the image to segment is not recovered. Note the significant misalignment between the prior shape (red) and the input image. Fig. 3.10g shows segmentation without prior ($\mu = 0$). When the weight of the shape term is low, as demonstrated in Fig. 3.10h ($\mu = 0.5$) and in Fig. 3.10i ($\mu = 0.8$), the final segmenting contour occupies the gap between the object boundaries and the transformed prior. Further stressing the contour length term ($\nu > 1$) a result similar to Fig. 3.10g is obtained. We compared the homography matrix generated from the estimated transformation parameters with that obtained using 13 manually selected corresponding point pairs, see Table 3.3. The comparison has been done for the implicit homography parameters that were recovered directly (row 2) and those that were computed from the explicit homography matrix entries

according to Eq. 3.20 (row 3). The values of the parameters recovered assuming similarity transformation (Fig. 3.1a) and affine transformation (Fig. 3.1b) are reported in rows 4-5.

Transformation parameter	α	β	γ	t_x	t_y	t_z	ψ	ξ
True values	0.075 ⁰	0.075 ⁰	7.5 ⁰	-10.0	10.0	-0.1	0.075 ⁰	0.15 ⁰
Recovered Values	0.069 ⁰	0.075 ⁰	7.4 ⁰	-10.0	9.6	-0.1	0.086 ⁰	0.14 ⁰

Table 3.2: A comparison of the recovered and true transformation parameters, for the example shown in Fig. 3.9.

Homography matrix entry	h_{11}	h_{12}	h_{13}	h_{21}	h_{22}	h_{23}	h_{31}	h_{32}
Point correspondence	0.91	0.28	3.2	-0.31	0.71	34.05	-0.0003	0.0019
Proposed algorithm, implicit homography	0.92	0.27	2.88	-0.29	0.70	32.7	-0.0003	0.0017
Proposed algorithm, explicit homography	0.92	0.27	2.80	-0.30	0.70	33.35	-0.0003	0.0018
Similarity approximation	0.88	0.37	1.81	-0.37	0.88	35.5	0	0
Affine approximation	0.92	0.30	0.09	-0.37	0.78	31.01	0	0

Table 3.3: A comparison of the homography matrix entries obtained using the proposed algorithm and via manually selected corresponding point pairs, for Fig. 3.10. The last two rows of the table display the transformation parameters obtained assuming similarity and affine transformations respectively. These values correspond to the registration results shown in Fig. 3.1b-c. Note, that not all the similarity parameters are independent.

Fig. 3.11 demonstrates successful prior-based segmentation and registration of the scissors despite the holes and the specular reflection. This is a challenging example, since the overlap between the reference image and the image to segment is small, and localized in the over-exposed part of the image. In Figs. 3.12c,g, the bottle-opener is accurately segmented regardless of the significant projective distortion and the over-exposure at the bottom-right part of the opener. Note the difference in the grey-level distribution between the two views. Figs. 3.12h-i show registration of the prior contour to the given image, assuming similarity and affine transformations respectively. The similarity and affine approximations recover the rotation in the image plane, but fail to capture the perspective deformation. The apparent mismatch inhibits the

segmentation process and prohibits accurate reconstruction of the missing parts.

Finally, Figs. 3.13a-b are a day and night satellite views of Europe respectively. The images were adapted from NASA:*earthobservatory.nasa.gov*. Successful segmentation is demonstrated in Fig. 3.13c. The coast line of the continent and the islands are precisely extracted, despite the spatial brightness variations. The accuracy of the recovered transformation is visually verified in Fig. 3.13d. Note the significant mismatch, Fig. 3.13e, between the two views. Without using the prior shape, segmentation of Fig. 3.13b is difficult. As seen in Fig. 3.13f, parts of Scandinavia and the eastern Adriatic shore are incorrectly segmented⁶.

3.7 Concluding remarks

Given an image and a prior shape, segmentation and registration can be viewed as two aspects of a single underlying problem. The essence of the method proposed in this chapter is concurrent segmentation and registration, as mutually supporting processes within a unified variational framework. The keys to the success is the introduction of a shape similarity measure that enables the incorporation of the homography formulation within the objective functional. Explicit encoding of a geometric transformation with such a high descriptive power allowed to break away from the common statistical representation of transformations in prior-based segmentation.

The ability to precisely extract the actual object contour based on the image data and its expected shape is critical in many applications. A robot manipulator guided by the segmentation of Fig. 3.10g, for example, might break the elephant trunk or attempt to pierce through its specular reflections, mistakenly interpreted as holes. In this sense, segmentation that is based on Fig. 3.10c is preferable.

The Chan-Vese two-phase segmentation model, currently employed in our scheme, requires the average grey-levels of the foreground and background to be distinct. Nevertheless, as seen in the bottle opener, scissors and Europe examples (Figs. 3.11-3.13), the use of the shape prior allows to substantially relax the homogeneity requirements. The suggested method is modular in the sense that the underlying level-set segmen-

⁶A similar segmentation result (no prior), showing similar difficulties, appeared in [11], Fig. 11.

tation model can be upgraded. For example, it will be possible to accommodate cluttered background by using the Vese-Chan multi-phase segmentation model [119].

Expressing the planar-projective transformation within the energy functional (3.11) in terms of relative camera-object motion (R, \mathbf{t}) and plane structure (\mathbf{n}, d) rather than via the homography matrix is a considerable shortcut toward the recovery of these parameters. The cumbersome task of decomposing the homography matrix [29, 38, 63] is avoided. Also, note that reference images are often taken in controlled conditions. If the prior image plane is perpendicular to the optical axis, the search for the eight homography parameters can be reduced to six by setting the structure parameters (ψ, ξ) to zero, thus simplifying the registration procedure. For non-calibrated cameras the homography can be fully recovered in its implicit form (recovery of $h_{11} \dots h_{32}$ instead of α, β etc.). Then only eight parameters are recovered and the need for decoupling between the internal and external camera parameters does not arise. The mathematical formulation for the recovery of the (implicit) homography parameters is obtained by substituting the expressions (3.19) directly in (3.25) for $\eta = h_{11} \dots h_{32}$. The homography parameters recovered using this formulation are shown in Table 3.3, second row.

In principle, the planar projective transformation model requires the observed contour points of the 3D object to be coplanar and to take part in the prior shape outline as well. In practice, this assumption can be relaxed for moderate transformations. As seen in the satellite image example (Fig. 3.13), an excellent result is obtained even though the coastline of Europe is clearly non-planar. The elephant example (Fig. 3.10) demonstrates successful application of the suggested method despite the evident violation of the assumption. The geometric model can be extended beyond planar projective homography to handle non-planar transformations. In a work by Brox, Rosenhahn and Weickert [6], that followed this research, three dimensional object projections were used as priors for segmentation. A comprehensive solution for general 3D objects is a subject for further study. Another interesting future direction would be to consider framework where all the components are perspective invariant, including arc-length, etc, as done for example in the affine invariant active contours (without shape prior) by Olver et al [73, 74].

In active contour methods that do not employ a prior shape, the selection of the

initial contour might be crucial for correct segmentation and designation of segments as object or background. Figs. 3.8f-g demonstrate possible segmentation results obtained without a shape prior, that are initialization dependent. Prior shape knowledge facilitates the detection of the desired object, reducing the influence of the initialization on the final result, as demonstrated in the successful segmentation results in Figs. 3.8b-e.

The suggested approach can be extended to deal with non-rigid objects in two ways. If the non-rigidity can be modeled in parametric form, as in the case of articulated objects, the additional parameters could be incorporated within the transformation model. Otherwise, one may integrate the proposed method with a statistical scheme based on a training set, e.g., Rousson and Paragios [96]. This would allow to capture the non-rigidity via the statistical component, and the projective transformation using the approach that we suggest.

The successful segmentation results and the reliable estimation of the transformation parameters encourage future research. Potential applications include perspective-invariant search in image databases, registration and structure recovery in stereo imaging, and object tracking in video sequences.

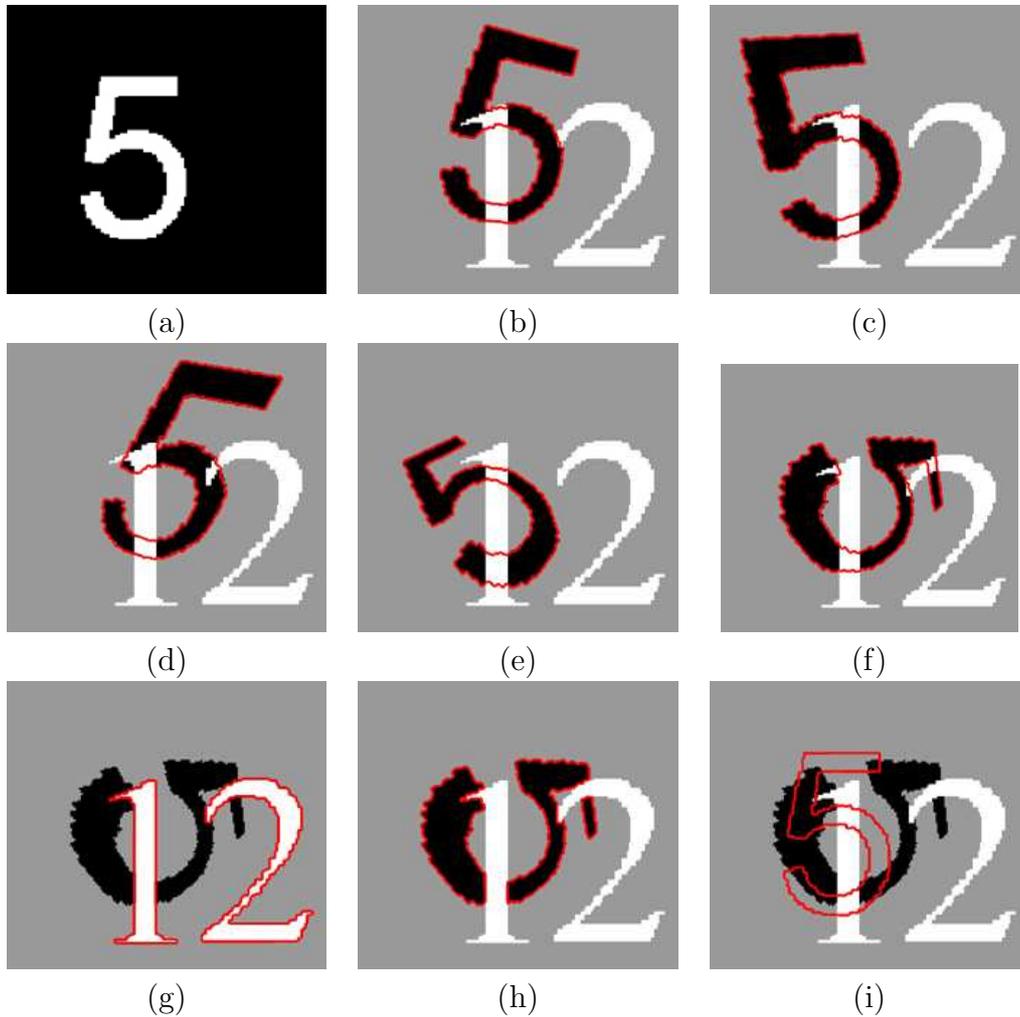


Figure 3.8: Segmentation results of the digit five. (a) Prior shape. (b)-(f) Successful segmentation results using the prior shape. Note the significant perspective distortions of the shape that simulate quite wide camera view point changes with respect to (a). A comparison of the recovered parameters with the true ones is shown in Table 3.1. (g)-(h) Segmentation without prior shape term. The final contour may either incorrectly segment the white objects (g) or the black disconnected object (h), depending on the initial contour. (i) The prior shape term is over-stressed. The final segmenting contour ignores the actual image data. This image also demonstrates the misalignment between the prior shape (red) and the input image.

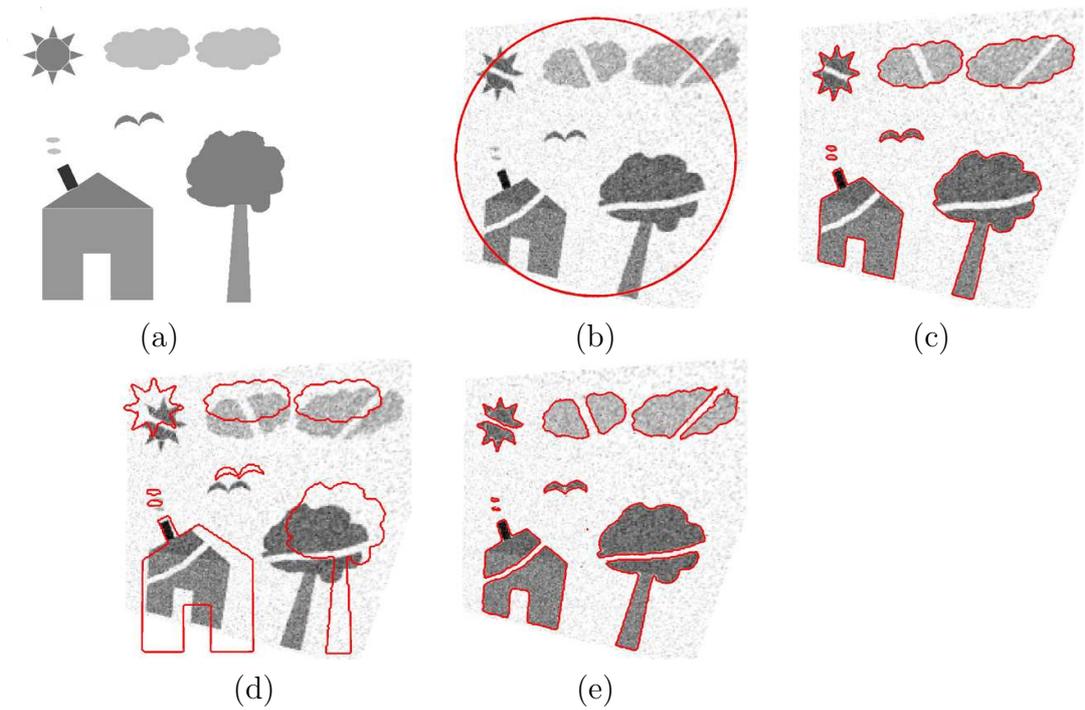


Figure 3.9: Segmentation results of a synthetic image. (a) A reference image, containing several components of different sizes and gray levels. The prior shape was obtained by thresholding. (b) The input image: a noisy, transformed and corrupted version of (a), with the initial contour (red). (c) Correct segmentation (red). (d) The unregistered prior shape (red) drawn on the input image. The reference (prior) image and the image to segment are clearly misaligned. (e) Segmentation (red) without using the prior shape.

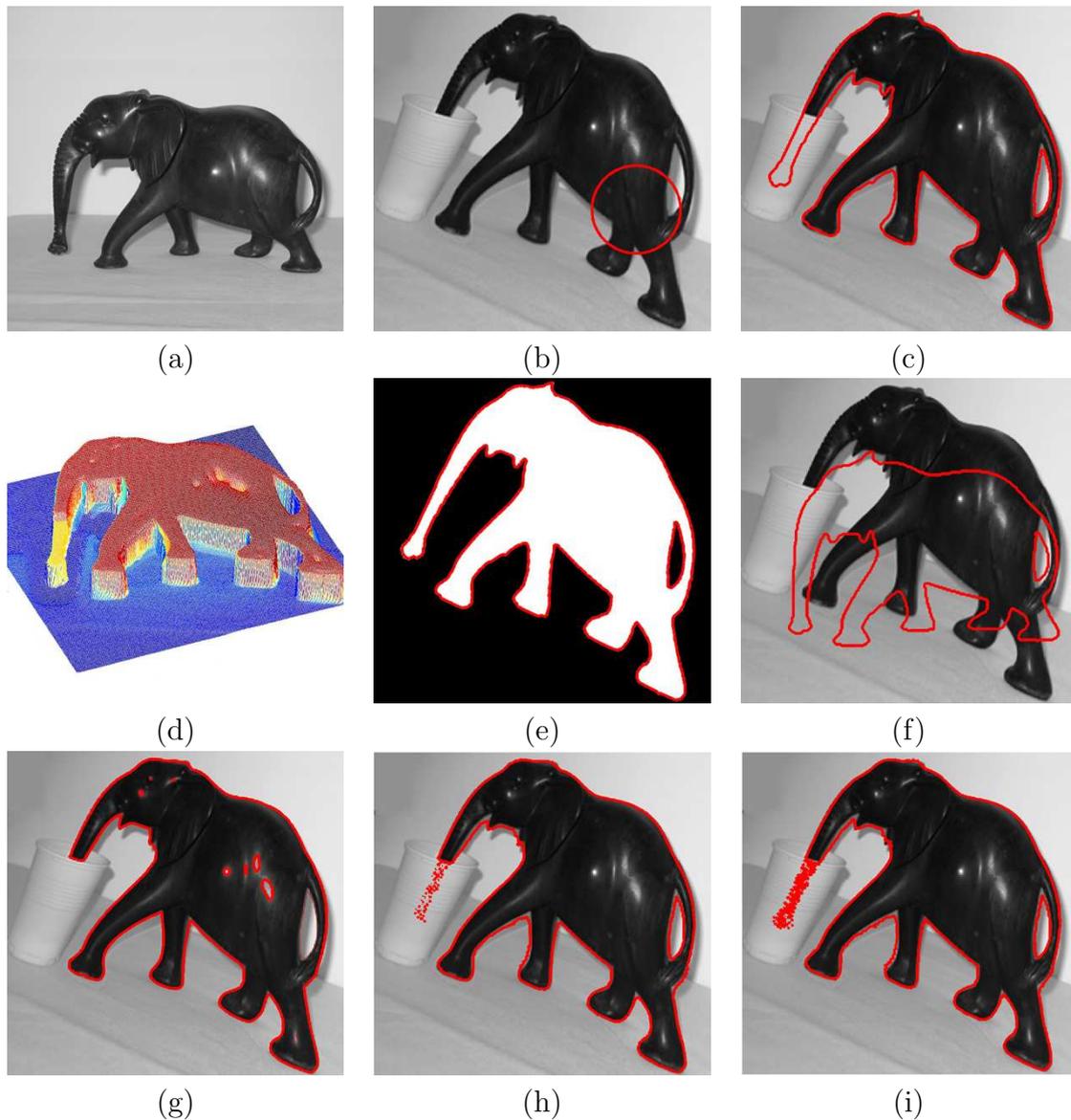


Figure 3.10: Segmentation results of a toy elephant image. (a) A reference image (elephant), from which the prior shape was obtained. (b) The input image. The initial contour used for the segmentation process is shown (red). (c) Successful segmentation: the final contour is shown (red). (d) The final form of the evolving level-set function ϕ . (e) Verification of the estimated transformation parameters: The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. (f)-(i) The suggested method is demonstrated for several values of prior shape weight (μ). (f) The prior shape is over-stressed ($\mu \gg 1$) thus the data is ignored and the transformation between the prior and the image to segment is not recovered. Note the significant misalignment between the prior contour (red) and the input image. (g) $\mu = 0$. The prior shape does not effect the segmentation result. (h)-(i) Low shape term weight ($\mu < 1$). The final segmenting contour occupies the gap between the object boundaries and the transformed prior contour. Further stressing the contour length will yield result similar to (g). (h) $\mu = 0.5$ (i) $\mu = 0.8$.

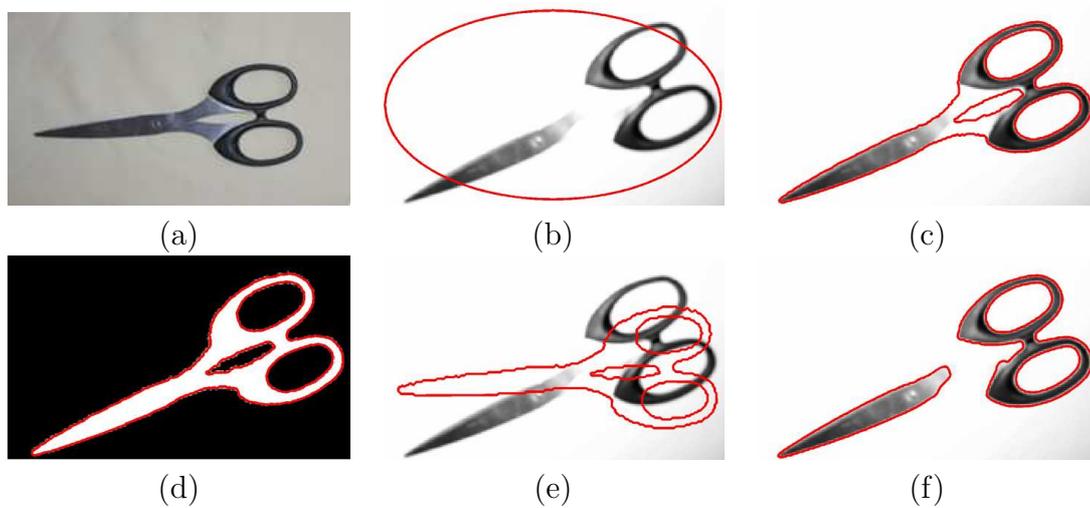


Figure 3.11: Segmentation results of the scissors image. (a) A reference image (scissors), from which the prior shape was obtained. (b) The input image, taken from a different viewpoint and with missing parts due to over-exposure. (c) Successful segmentation. The missing contour parts are accurately completed. (d) Verification of the estimated transformation parameters: The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. (e) The misalignment between the prior shape (red) and the input image. (f) Segmentation (red) without using the prior shape.

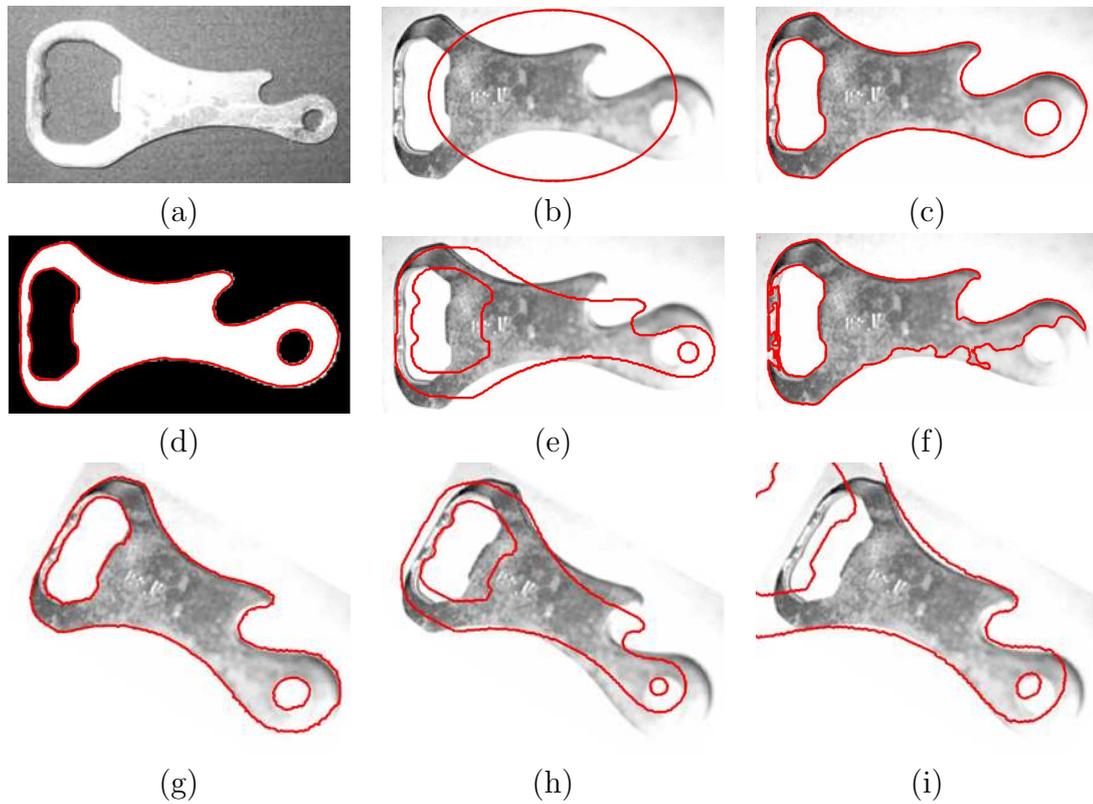


Figure 3.12: Segmentation results of a bottle opener: (a) A reference image (bottle opener), from which the prior shape was obtained. (b) The input image, taken from a different viewpoint and with over-exposure. (c) Successful segmentation. The missing contour parts are accurately completed. (d) Verification of the estimated transformation parameters: The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. (e) The misalignment between the prior shape (red) and the input image. (f) Segmentation (red) without using the prior shape. (g) Successful segmentation with wider camera view, using the reference image (a). (h)-(i) Registration of the prior contour to the given image using (h) similarity approximation and (i) affine approximation. Note the apparent mismatch. The similarity and the affine approximations cannot capture the perspective distortion.

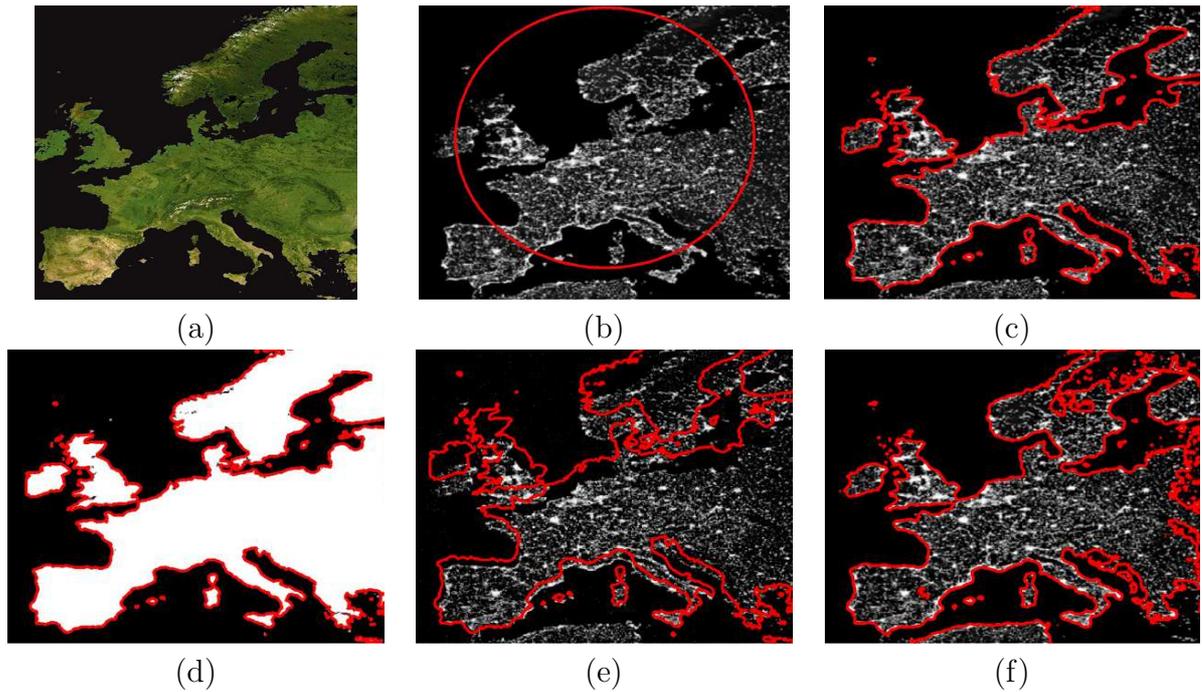


Figure 3.13: Segmentation results of Europe image taken from a satellite. (a) A reference image of Europe, from which the prior shape was obtained. (b) Input image: A satellite image of Europe at night. The reference and prior image were adapted from *earthobservatory.nasa.gov*. (c) Successful segmentation: the final contour (red) accurately traces the coast lines. (d) Verification of the estimated transformation parameters: The final contour as in (c), drawn on the projectively transformed prior shape according to the recovered homography. (e) The misalignment between the prior shape (red) and the input image. (f) Segmentation (red) without using the prior shape. The dark areas in Scandinavia and in the eastern Adriatic shore are incorrectly segmented.

Chapter 4

Mutual Segmentation

4.1 Background and related works

This chapter addresses the segmentation of an image pair of the same object in different appearances and poses. We can then redefine segmentation as the task of extracting the common object in both images. Uncertainty regarding the object boundaries can be resolved having two instances instead of one. Nevertheless, this setting provides only an implicit prior. The “double” segmentation problem is thus far from being trivial.

The proposed framework is useful when explicit prior knowledge is limited or not available. Instead, two images of the same object are given. Had good segmentation been possible in either image, it could have supported the segmentation of the other. However, in the absence of additional information on the object of interest, segmentation of each image by itself is prone to errors. Specifically, background regions can be mistakenly labeled as foreground (excess) or vice versa (deficiency). The resulting (possibly erroneous) segmentation of one of the images cannot provide a reliable prior for the other.

We therefore suggest a coupled segmentation process for both images, in which the information gained in the evolving segmentation of one image is a *dynamic* prior for the other. The segmentation and thus the prior information are refined along the process. We call this mutually supporting evolution process *mutual segmentation*.

Recent works, similar in their spirit include, [95], [116] and [124]. In the *cosegmen-*

tation method of [95], both images are simultaneously partitioned assuming that the common property of the foreground regions is their color probability density function (modeled by histograms) which should be also significantly different from the color distribution of the background. The *unsupervised* segmentation algorithm presented in [116], uses the best SIFT matches of Harris-Affine features to extract the common objects in image pairs. Similarly to [124], we presume that the object instances resemble in their *shapes*, thus having the benefit of being specific to the particular object of interest and insensitive to color and (in most cases) illumination variation. Yet, while [124] handles images corrupted only by noise, where at least one of them nearly contains sufficient information to be segmented by itself, we deal with cluttered images corrupted by noise and occlusions.

The mutual segmentation approach goes beyond the concepts of shape-based segmentation, because a well-defined shape prior is not available and the matching is between two possibly corrupted and noisy images. The main difficulty resides in labeling regions where the aligned images do not overlap. Obviously, erroneous foreground-background classifications spoil the segmentation of both images. Fig. 4.1 exemplifies labeling ambiguity. The regions pointed by the red arrows could be attributed either to the boot (according to one of the images) or to the background (according to the other).

The conflict between two possible interpretations of jointly segmented images has never been addressed before. Note that when more than two object instances are available, this ambiguity can be resolved by applying a majority rule [26]. Having only two images, we favor the image partitioning that minimizes a biased shape dissimilarity measure between the images. The definition of this biased shape term is one of the important contributions of the proposed study.

We suggest a framework for *mutual segmentation* of two images of the same object, related by projective transformation. Segmentation is carried out concurrently with registration of the evolving contours. The foundation of the proposed method is the construction of two level set functions, one for each image. Their evolution is controlled by the data contained in the associated image together with the shape information gained in the segmentation process of the other image. The shape information is embedded in a biased shape dissimilarity measure that accommodates

either deficient or excess parts in the extracted object shape. This measure is also invariant to planar projective transformations. The outcomes of the proposed algorithm include segmentation of the object appearances in both images and the recovery of the homography that aligns them.

The chapter is organized as follows. In the following, we introduce the biased shape dissimilarity measure which plays a key role in the proposed mutual segmentation algorithm. The embedding of the transformation model within the variational framework and the minimization of the resulting functional are also considered. The mutual segmentation algorithm together with the implementation details are presented in section 4.3. Experimental results are provided in section 4.4. Section 4.5 concludes this chapter.



Figure 4.1: The inherent labeling ambiguity demonstrated on two boot images. (a-b) Two instances of a boot. (c) Superposition of the aligned boot images. There is an inherent labeling ambiguity in aligned image regions that correspond to different objects. The regions pointed by the red arrows could be attributed either to the boot (according to one of the images) or to the background (according to the other).

4.2 Mutual segmentation with projectivity

We consider the segmentation of two images, I_1 and I_2 that depict two imperfect (occluded, noisy etc.) instances of an object. The segmentation is done by performing a joint maximum a priori probability (MAP) approximation via an alternating maximization. Indeed since a reliable prior $\tilde{\phi}$ is not available each evolving contour in its turn is employed as a prior for the other. The two level set functions ϕ_1 and ϕ_2 are alternately evolved. At even iterations the segmenting level-set is $\phi = \phi_1$ and

the prior is given by $\tilde{\phi} = \phi_2$. At odd iterations $\phi = \phi_2$ and $\tilde{\phi} = \phi_1$. The statistical inference problem takes the form:

$$P(\phi, \tilde{\phi}, T \mid I_1, I_2) \propto P(I_1, I_2 \mid \phi, \tilde{\phi}, T)P(\phi \mid \tilde{\phi}, T)P(\tilde{\phi})P(T) .$$

We assume that $P(T)$ is flat such that no projective transformation is a-priori preferred. The assumption that the two views are not unrelated is incorporated by initializing T by the identity transformation. The crucial observation is that $P(C \mid \tilde{C}, T)$ (or equivalently $P(\phi \mid \tilde{\phi}, T)$) which is defined via a dissimilarity measure between curves is *not* symmetric with respect to the two curves. This is so since $P(\phi_1 \mid \phi_2, T)$ is not necessarily equal to $P(\phi_2 \mid \phi_1, T)$. The symmetric shape dissimilarity measure defined in Eq. (2.27) is thus not suitable for mutually evolving level-set functions. In the next subsection we present a biased shape dissimilarity term which is more appropriate.

4.2.1 Biased shape dissimilarity measure

Consider the image pair in Figure 4.2a-b. Both have deficiencies. In Fig. 4.2a the hoop is absent. In Fig. 4.2b a portion of the creature's leg was erased. When the final segmentation of Fig. 4.2a is the prior for the segmentation of Fig. 4.2b and vice versa, the imperfections of each segmentation spoil the other, as shown in Fig. 4.2e-f. Note that the left leg is incorrectly segmented in Fig. 4.2e, while the hoop in Fig. 4.2f is not segmented at all.

The images in Fig. 4.2c-d contain superfluous hoops located in different places. When each segmentation is the prior for the other, using the unbiased dissimilarity measure in Eq. (2.27), the contours of the superfluous hoops in Fig. 4.2c-d undesirably appear in the segmentation shown in Fig. 4.2h,g respectively.

The discrimination between integral object parts (leg, hoop) and other surrounding objects (superfluous hoops) raises a fundamental question which extends beyond the scope of the current work and relates to perceptual organization of images in general. Given more than two images, this difficulty can be tackled by employing a majority decision rule to determine the object-background partition. However, for mutual segmentation of two images, another decision tool or source of information is

needed. For simplicity, we assume either of the following “world states”:

1. **The images to segment have (mostly) deficiencies.** Application of this rule is demonstrated in Fig. 4.2i-j to resolve labeling ambiguities between Fig. 4.2a and Fig. 4.2b. Thus, for example, the missing part in the leg of the creature, shown in Fig. 4.2b has been labeled as part of the object.
2. **The images to segment have (mostly) excess parts.** When this rule is applied to Fig. 4.2c-d, the superfluous hoops are labeled as background, as shown in Fig. 4.2k-l.
3. **The prior shape is perfect.** Examples for the application of this rule are shown in Fig. 4.2e-h.

Refer again to the dissimilarity measure in Eq. (2.27). The cost functional integrates the non-overlapping object-background regions in both images indicated by $H(\phi)$ and $H(T(\tilde{\phi}))$. This is equivalent to a pointwise exclusive-or (xor) operation integrated over the image domain. We may thus rewrite the functional as follows:

$$D(\phi, \tilde{\phi} | T) = \int_{\Omega} \left[H(\phi) \left(1 - H(\tilde{\phi}_T) \right) + (1 - H(\phi)) H(\tilde{\phi}_T) \right] d\mathbf{x} \quad (4.1)$$

To simplify the expression we denote $T(\tilde{\phi}) \equiv \tilde{\phi}_T$. Note that the expressions (2.27) and (4.1) are identical, since $H(\phi)$ is equal to $(H(\phi))^2$, when $H(\phi)$ is a strictly binary function. There are two types of disagreement between the labeling of $H(\phi)$ and $H(\tilde{\phi}_T)$. The left term in (4.1) does not vanish if there exist image regions labeled as *object* by the image data (ϕ) and labeled as *background* by the shape prior ($\tilde{\phi}_T$). The right term in (4.1) does not vanish if there exist image regions labeled as *background* by the image data and labeled as *object* by $\tilde{\phi}_T$. Inserting a weight parameter $\mu \geq 0$, the relative contributions of the terms are changed.

$$E_{\text{MUTUAL}}(\phi, \tilde{\phi}, T) = \int_{\Omega} \left[\mu H(\phi) \left(1 - H(\tilde{\phi}_T) \right) + (1 - H(\phi)) H(\tilde{\phi}_T) \right] d\mathbf{x} \quad (4.2)$$

The associated gradient descent equation for ϕ is then:

$$\phi_t^{\text{MUTUAL}} = \delta(\phi) [H(\tilde{\phi}_T) - \mu(1 - H(\tilde{\phi}_T))] \quad (4.3)$$

Now, if excess parts are assumed, the left penalty term should be dominant, thus $\mu > 1$. Otherwise, if deficiencies are assumed, the right penalty term should be dominant and $\mu < 1$.

4.2.2 Projective invariance

The mutual shape term E_{MUTUAL} in (4.2) depends on the transformation T between the currently evolved level set function ϕ and the level-set function associated with the other image $\tilde{\phi}$. We assume that T can be modeled by planar projective transformation. Let $C(t)$ and $\tilde{C}(t)$ be the two *planar* active contours associated with the images I and \tilde{I} , respectively. Recall that $C(t)$ and $\tilde{C}(t)$ are the zero levels of the level set functions $\phi(t)$ and $\tilde{\phi}(t)$ (respectively) at time t . Let $p \in C$ and $p' \in \tilde{C}$ denote corresponding points on C and \tilde{C} . Their homogeneous coordinates \mathbf{x} and \mathbf{x}' are related by *planar projective homography*, i.e. $\mathbf{x}' = \mathcal{H}\mathbf{x}$. Recall, from the previous chapter (Eqs. 3.18 and 3.19) that

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}},$$

where h_{ij} are the entries of the homography matrix \mathcal{H} . Equivalently we can define $T(\phi(\mathbf{x})) \equiv \phi(\mathbf{x}') = \phi(\mathcal{H}\mathbf{x})$, where \mathcal{H} is a linear transformation applied on the homogeneous coordinate of ϕ . Thus, given ϕ and $\tilde{\phi}$, we would like to infer the homography \mathcal{H} that minimizes the “distance” (4.2) between ϕ and $\tilde{\phi}_T$. The eight unknown ratios of the homography matrix entries, $\hat{h}_k = h_{ij}/h_{33}$ where $k = 1 \dots 8$, are re-calculated at each iteration for the currently updated ϕ and $\tilde{\phi}$. The PDEs for \hat{h}_k are obtained by differentiating (4.2) with respect to each.

$$\frac{\partial \hat{h}_k}{\partial t} = \int_{\Omega} \delta(T(\tilde{\phi})) [(1 - H(\phi)) - \mu H(\phi)] \frac{\partial T(\tilde{\phi})}{\partial \hat{h}_k} d\mathbf{x} \quad (4.4)$$

Derivation of $\frac{\partial T(\tilde{\phi})}{\partial \hat{h}_k}$ can be done as in the previous chapter, subsection 3.5.2.

At each iteration, the level set functions of the image pair are evolved alternately with the recovery of the transformation that aligns them. In subsection 4.3.1 we discuss further aspects of this process.

4.2.3 Unified cost functional

A unified cost functional based on the data of the images to segment, on general assumptions with respect to properties of natural images, and on the segmentation curves both images takes the following form:

$$\begin{aligned} E(\phi) &= W^{\text{RB}} E_{\text{RB}}(\phi) + W^{\text{LEN}} E_{\text{LEN}}(\phi) + W^{\text{GAC}} E_{\text{GAC}}(\phi) + W^{\text{EA}} E_{\text{EA}}(\phi) \\ &+ W^{\text{MUTUAL}} E_{\text{MUTUAL}}(\phi), \end{aligned} \quad (4.5)$$

with the equations (2.9, 2.19, 2.21, 2.17, 4.2).

Note that the smoothness term E_{LEN} (eq. 2.19) is contained in the GAC term (eq.2.21) as discussed in subsection 2.1.5. It however appears in the complete energy functional to enable flexibility in the functional construction, in particular when the GAC term should be ignored while the smoothness term is most desirable. Refer for example for Fig. 4.3. Practically either of the weights W^{GAC} or W^{LEN} is set to zero. The evolution of the level-set functions ϕ in either of the images, is determined at each iteration by $\phi(t + \Delta t) = \phi(t) + \phi_t \Delta t$. The associated gradient descent equations ϕ_t are derived using the first variation of the functional (4.5)

$$\phi_t = W^{\text{RB}} \phi_t^{\text{RB}} + W^{\text{LEN}} \phi_t^{\text{LEN}} + W^{\text{GAC}} \phi_t^{\text{GAC}} + W^{\text{EA}} \phi_t^{\text{EA}} + W^{\text{MUTUAL}} \phi_t^{\text{MUTUAL}}, \quad (4.6)$$

The determination of W^{TERM} is discussed in section 2.2.

4.3 Implementation

The essence of the proposed method is the simultaneous evolution of two level-set functions. Each evolves on its corresponding image and is controlled by the data of the associated image and by the level-set function associated with the other image. The planar projective transformation between these two level-set function is updated at each iteration. The algorithm is quite robust to the selection of initial level-set function $\phi_0(\mathbf{x})$. The only limitation is that image regions labeled as foreground in the first iteration, i.e $\omega_0 = \{\mathbf{x} \mid \phi_0(\mathbf{x}) \geq 0\}$, will contain a significant portion of the object to be segmented, such that the calculated image features will approximately

characterize the object region. Formally, we assume that $G^+(I(\omega_0)) \approx G^+(I(\hat{\omega}))$, where $\hat{\omega}$ is the actual object region in the image. When there exists an estimate of the average gray levels of either the foreground or the background image regions, this restriction can be eliminated.

We run the algorithm until the following stopping condition is met:

$$\max(d(\phi_1), d(\phi_2)) < s$$

where s is a predefined threshold and

$$d(\phi_i) = \sum_{\mathbf{x} \in \Omega} |H(\phi_i^{t+\Delta t}(\mathbf{x})) - H(\phi_i^t(\mathbf{x}))| \quad i = 1, 2.$$

Here, $\phi_i^{t+\Delta t}(\mathbf{x})$ is the i -est level set function, at time $t + \Delta t$.

4.3.1 Recovery of the transformation parameters

Minimizing the cost functional (4.2) with respect to the eight unknown ratios of the homography matrix entries is a complicated computational task. Direct update of the parameters via their the derivatives (4.4) may lead to an undesired local minimum as discussed in subsection 3.5.4. Here we suggest a slight modification to the optimization process used in section 3.5 We thus suggest to perform a rough search in the 8 dimensional parameter space working on a coarse to fine set of grids before applying the gradient based Quasi-Newton method [14]. The former search, done only *once*, significantly reduces the search space, constraining it to the region of attraction of the global minimum. The gradient based algorithm, applied in every iteration, tunes of the search result based on the updated level-set functions. Further discussion on the application of the Quasi-Newton optimization and its significance for the recovery of the transformation parameters can be found in subsection 3.5.4.

Since the registration of the evolved level-set functions toward each other is done independently, the mean error (or max error) between the entries of the matrices $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}^{-1}$ (or $\mathcal{H}_{2 \rightarrow 1}$ and $\mathcal{H}_{1 \rightarrow 2}^{-1}$) could be used as a criterion for obtaining globally optimal of the transformation parameters. Here we assume that the probability of achieving a minimum, which is not a global one, that satisfies $\mathcal{H}_{1 \rightarrow 2} \approx \mathcal{H}_{2 \rightarrow 1}^{-1}$ is small.

Table 4.1 and Table 4.2 present the estimated errors in the recovered homographies based on this comparison. The exact formalism of the error estimate is discussed in Section 4.4.

4.3.2 Algorithm

We summarize the proposed algorithm assuming the following setup. The input is two images I_1 and I_2 of the same object, taken from different viewpoints. The object contours are approximately coplanar. Two level-set functions ϕ_i , $i = 1, 2$ that correspond to images I_i are alternately evolved, based on the data of the corresponding image and the other level-set function.

1. Choose initial level-set functions ϕ_i , $i = 1, 2$, for example, standard circular (or elliptic) cones. The intersections of the initial level-sets with the corresponding image domains form the initial contours.
2. The homography matrices $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}$ are initialized to the identity matrices.
3. For each image I_i , compute the values u_+ and u_- using Eq. (2.7), based on the current object-background partition, defined by the corresponding level-set function.
4. At even iterations evolve the level-set function $\phi = \phi_1$ using the other level-set function $\tilde{\phi} = \phi_2$ as a prior. At odd iterations evolve the level-set function $\phi = \phi_2$ using the other level-set function $\tilde{\phi} = \phi_1$ as a prior.
5. Apply the corresponding projective transformation on the instantaneous prior $\tilde{\phi}$ using Eq. (3.19) with the parameters estimated in the preceding time step.
6. Update ϕ using the gradient descent equation (4.6).
7. Update the transformation parameters h_k using the derivatives (4.4). The relation $\mathcal{H}_{1 \rightarrow 2} = \mathcal{H}_{2 \rightarrow 1}^{-1}$ can be used to speed up (or enhance) the recovery of the transformation parameters.

8. Repeat steps 3-7 until convergence. A convergence criterion is stated in Section 4.3.

4.4 Experiments

Fig.		\hat{h}_1	\hat{h}_2	\hat{h}_3	\hat{h}_4	\hat{h}_5	\hat{h}_6	\hat{h}_7	\hat{h}_8
4.2	$\mathcal{H}_{1 \rightarrow 2}$	0.858	0.406	-12.12	-0.298	0.964	-27.93	-0.0002	-0.0009
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	0.868	0.408	-12.11	-0.302	0.969	-27.90	-0.0001	-0.0009
	$\text{Err}(\hat{h}_i)$	0.0058	0.0025	0.0004	0.0067	0.0026	0.0005	0.3333	0
4.3	$\mathcal{H}_{1 \rightarrow 2}$	1.0502	0.1496	0.67	-0.1271	0.9857	11.41	-0.0001	0.0008
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	1.0573	0.1504	1.2125	-0.1283	1.0027	11.79	-0.0001	0.0007
	$\text{Err}(\hat{h}_i)$	0.0034	0.0027	0.2882	0.0047	0.0085	0.0164	0	0.0667
4.4	$\mathcal{H}_{1 \rightarrow 2}$	1.0898	-0.1315	11.4	0.1276	1.0277	-11.55	-0.0018	0.0003
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	0.9921	-0.1289	10.88	0.2303	1.0272	-11.25	-0.0010	0.0018
	$\text{Err}(\hat{h}_i)$	0.0469	0.0100	0.0233	0.2870	0.0002	0.0132	0.2857	0.7143
4.5	$\mathcal{H}_{1 \rightarrow 2}$	1.003	-0.1158	8.58	0.0359	0.9188	-0.16	0.0005	-0.0001
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	1.002	-0.1158	8.61	0.0364	0.9189	-0.12	0.0005	-0.0001
	$\text{Err}(\hat{h}_i)$	0.0005	0	0.0017	0.0069	0.0001	0.1429	0	0
4.6	$\mathcal{H}_{1 \rightarrow 2}$	0.9469	0.1461	17.92	-0.2128	1.031	-29.9	-0.0019	-0.0009
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	0.9503	0.1564	17.97	-0.2240	1.025	-29.6	-0.0018	-0.0007
	$\text{Err}(\hat{h}_i)$	0.0018	0.0340	0.0014	0.0256	0.0029	0.0050	0.0270	0.1250
4.7	$\mathcal{H}_{1 \rightarrow 2}$	0.9162	0.3374	-12.75	-0.2865	0.9016	8.79	-0.0005	-0.0003
	$\mathcal{H}_{2 \rightarrow 1}^{-1}$	0.9183	0.3396	-12.53	-0.2882	0.9034	8.81	-0.0005	-0.0003
	$\text{Err}(\hat{h}_i)$	0.0011	0.0032	0.0087	0.0030	0.0010	0.0011	0	0

Table 4.1: Comparison of the entries of the matrices $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}^{-1}$ obtained through the registration phase in the mutual segmentation algorithm applied to image pairs presented in Figs. 4.2-4.7. The error estimate is calculated according to $\text{Err}(\hat{h}_i) = |\hat{h}_i - \hat{p}_i| / |\hat{h}_i + \hat{p}_i|$, where \hat{h}_i and \hat{p}_i are the i -est entries of the normalized matrices $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}^{-1}$, respectively.

We exemplify the mutual segmentation algorithm on image pairs related by projective transformations. The input images are shown with the initial and final segmenting contours. The mismatch between the respective object views is demonstrated by superposition of the images. The accuracy of recovered homographies is tested by a comparison between the $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}^{-1}$. Table 4.1 exemplifies such comparisons done on the homographies recovered for the image pairs shown in Figs. 4.2-4.7. The term $\text{Err}(\hat{h}_i) = |\hat{h}_i - \hat{p}_i| / |\hat{h}_i + \hat{p}_i|$ is our suggested error measure for the recovered

Fig. No	Mean Error
Fig. 4.2	0.0440
Fig. 4.3	0.0488
Fig. 4.4	0.1726
Fig. 4.5	0.019
Fig. 4.6	0.0278
Fig. 4.7	0.0023

Table 4.2: The mean of the errors over the entries of the homography matrices $\mathcal{H}_{1 \rightarrow 2}$ that were estimated in each of the experiments. The estimated errors for each entry of the respective matrices are presented in Table 4.1.

transformation parameter \hat{h}_i , where \hat{h}_i and \hat{p}_i are the i -est entries of the normalized matrices $\mathcal{H}_{1 \rightarrow 2}$ and $\mathcal{H}_{2 \rightarrow 1}^{-1}$, respectively. Table 4.2 presents the mean error for each of the estimated homography matrices.

In all the experiments we set $dt = 0.1$ and $\epsilon = 1$. The weights of the gradient descent terms (4.6) are adaptively determined as described in section 2.2. Figure 4.3 shows two images of a hand taken from two different view points. The misalignment between the hand instances is shown in Fig. 4.3e. Successful segmentation of both images, using the mutual segmentation algorithm with $\mu < 1$, is demonstrated in Fig. 4.3c-d. Fig. 4.3f-h demonstrate unsuccessful segmentation of each image by itself. Fig. 4.3g-h display two possible segmentations of the noisy instance of the hand. In the segmentation shown in Fig. 4.3h the smoothness term, Eq. (2.19), has been stressed by multiplying its weight W^{LEN} by 2. The segmenting contour is thus smoother but does not extract precisely the narrow regions between the fingers.

The boot images in Fig. 4.4a-b were mutually segmented using the proposed algorithm, with $\mu < 1$. The delineating contour (shown in Figure 4.4d-e) traces precisely the boot boundaries while correctly completing the occluded parts. The misalignment between the boot instances is shown in Fig. 4.4c. The necessity of the biased shape dissimilarity measure is demonstrated in Fig. 4.4f-g. In these figures we used the unbiased dissimilarity measure (Eq. (2.27)) and the evolving segmentation of each image spoiled the segmentation of the other.

Fig. 4.5 demonstrates mutual segmentation of two images of a license plate with corrupted digits. Fig. 4.5a-b present the license plate images with the initial contour

(red). The misalignment between the images is shown in Fig. 4.5c. Assuming excess parts we set $\mu > 1$. Successful mutual segmentation results are shown in Fig. 4.5d-e. For a comparison, fig. 4.5f-g display undesired segmentation results obtained when each image is segmented by itself.

Fig. 4.6 shows the ability of the proposed method to detect the common object, stop sign, in an image pair despite the clutter. The images have been downloaded from different web sites (see caption). Note that the poses and surroundings of the stop signs are different. As in all other examples, the algorithm was tested on gray level images, thus the color cue is not used. The contours (red) in Fig. 4.6c-d precisely extract the common object based on its shape. Fig. 4.6e-f present the final level-set functions.

Mutual segmentation of the chess pieces shown in fig. 4.7 demonstrates the ability of the algorithm to deal with both clutter and partial occlusions. Note that though the objects segmented are definitely not planar, the comparisons of the homographies between the object contours in Table 4.1, show that homography is a reasonable approximation for the transformation.

4.5 Concluding remarks

We presented a method for concurrent, mutually-supporting segmentation of two images of the same object, taken from different view points. Having *two* images instead of one provides redundancy that is employed by using each instance to guide the segmentation of the other. Unlike previous methods, the concept of a perfect shape prior is replaced by information gathered from incomplete instances.

Segmentation is metaphorically similar to cliff climbing. Prior-based segmentation is analogous to situations where someone climbs first and secures a rope to the cliff. If this is not possible, the combined effort of at least a duo is needed. The two climb in turns: at each stage one person holds the cliff and helps the other climb. The main contribution of this chapter is the formulation of this duo shape term, that enables solution of the mutual segmentation problem.

Having two object instances is helpful in regions where the aligned images agree, but there is an inherent ambiguity where they don't. In this chapter, we address this

ambiguity via the biased shape dissimilarity measure. Note that if more than two images are available, the ambiguity can be resolved by majority rule. We consider this issue in the next chapter that deals with symmetry based segmentation. The prior shape term of equation 5.26 in sub-subsection 5.5.2 handles similar scenario.

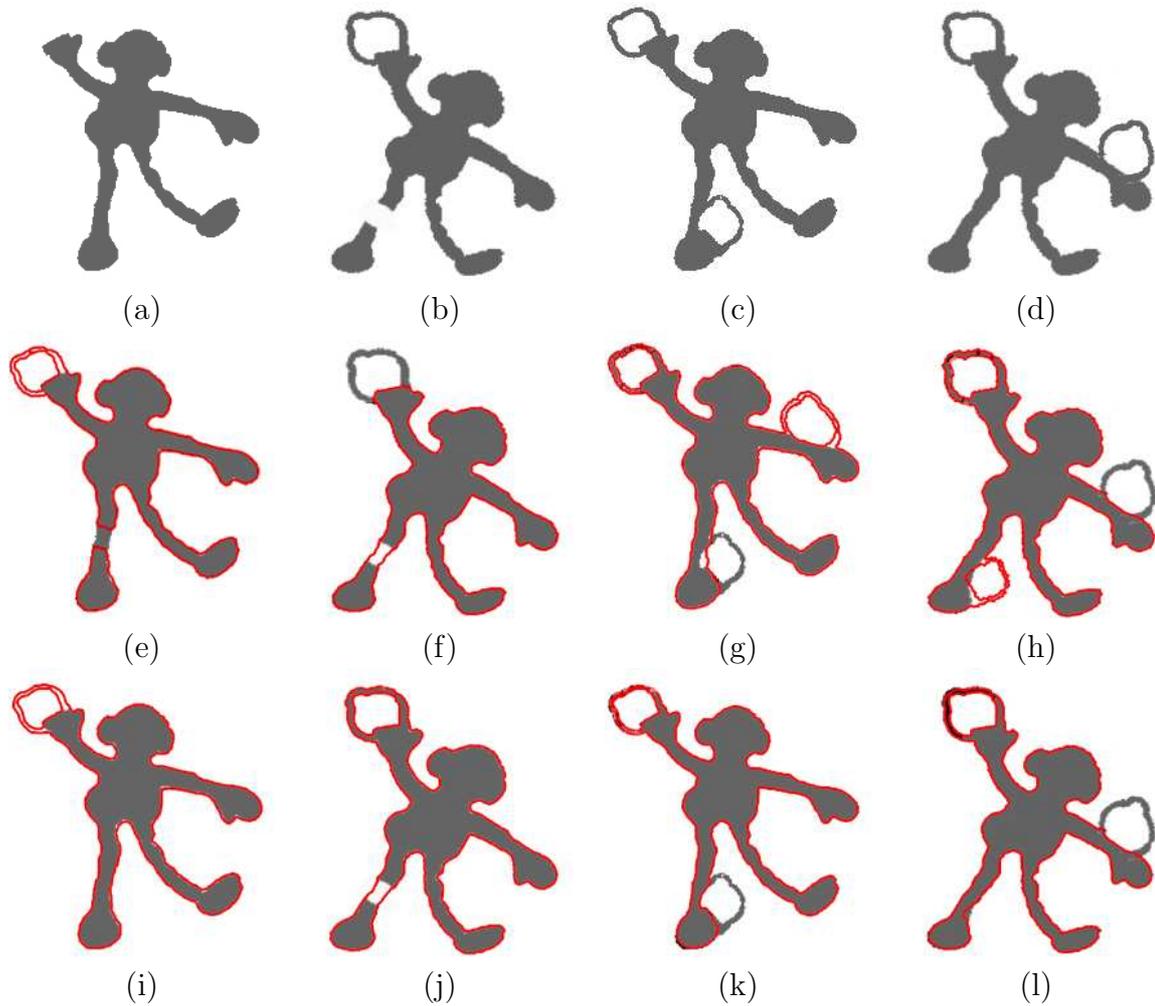


Figure 4.2: A comparison of the mutual segmentation results using biased and unbiased dissimilarity measures. (a,b) Input images with deficiencies (leg, hoop). (c,d) Input images with excess parts (extra hoops). (e) Segmentation (red) of the image in (a) using (b) as a prior. (f) Segmentation of the image in (b) using (a) as a prior. (g) Segmentation of the image in (c) using (d) as a prior. (h) Segmentation of the image in (d) using (c) as a prior. (i,j) Mutual segmentation results for images (a) and (b) respectively. (k,l) Mutual segmentation results for images (c) and (d) respectively. The images are related by projective transformation. The recovered parameters are shown and compared in Table 4.1.

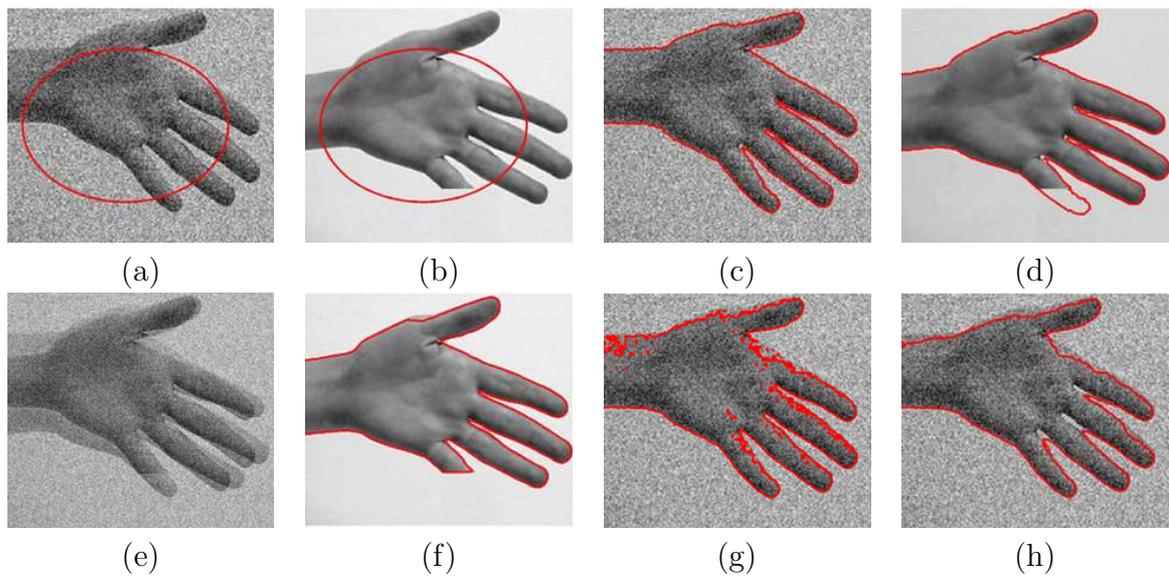


Figure 4.3: Segmentation results of the hand images. (a,b) Noisy (a) and corrupted (b) images of the same object taken from different view points. The initial contours are drawn in red. (c,d) Successful mutual segmentation results (red). (e) Superposition of the two images to demonstrate the misalignment. (f- h) Segmentation of each image by itself. The noisy image (g,h) was segmented twice with different weights of smoothness term: (g) The contribution of the smoothness term $W_{\text{LEN}}(t)(\phi_t^{\text{LEN}})$ was restricted to $[-1, 1]$ (refer to section 2.2 for details). The contour “mistakenly” follows image gradients that are due to noise. (h) The smoothness term $W_{\text{LEN}}(t)(\phi_t^{\text{LEN}})$ was further stressed, i.e. its contributions were multiplied by two. The segmenting contour (red) is smoother but the gaps between the fingers are not well extracted.

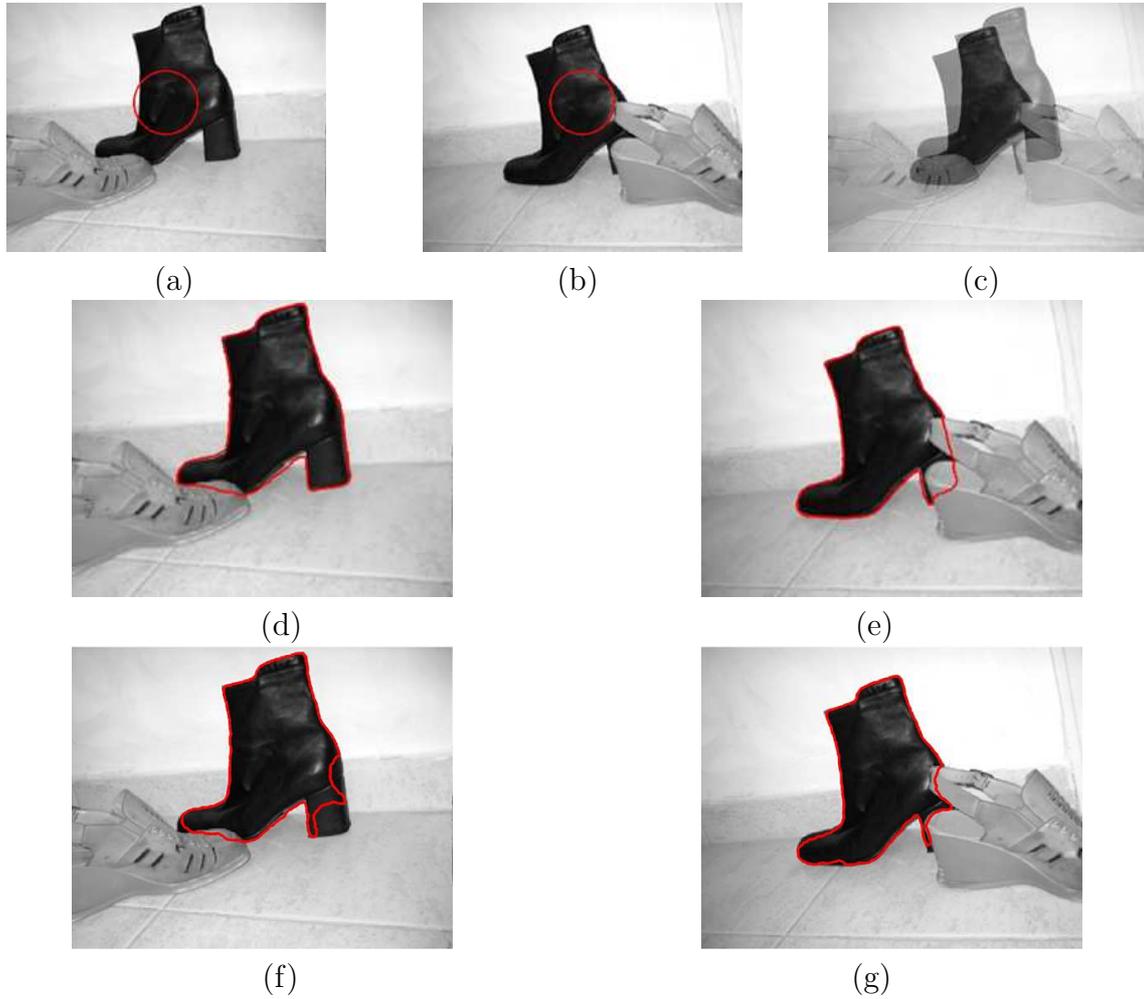


Figure 4.4: Segmentation results of the boot images.(a,b) Input images with their initial contours (red). The images are of the same object (boot) taken from different viewpoints. (c) Superposition of the two images to demonstrate the misalignment. (f,g) Successful mutual segmentation results (red). (d,e) Segmentation using an unbiased shape dissimilarity measure, Eq.(2.27). The evolving segmentation of each image spoiled the other.

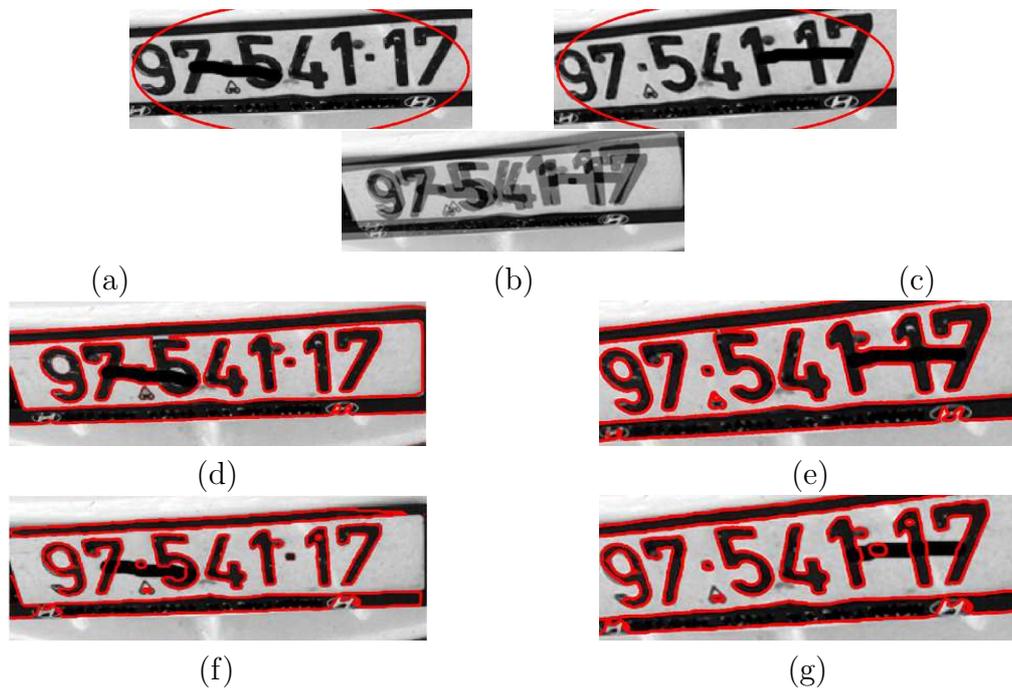


Figure 4.5: Segmentation results of the license plate images. (a)-(b) Input images with their initial contours (red). (c) Superposition of the two images to demonstrate the misalignment. (d)-(e) Segmentation of each license plate image by itself using the Chan-Vese level-set method for segmentation. (f)-(g) Successful *mutual* segmentation of license plate images with corrupted digits taken from two different viewpoints.

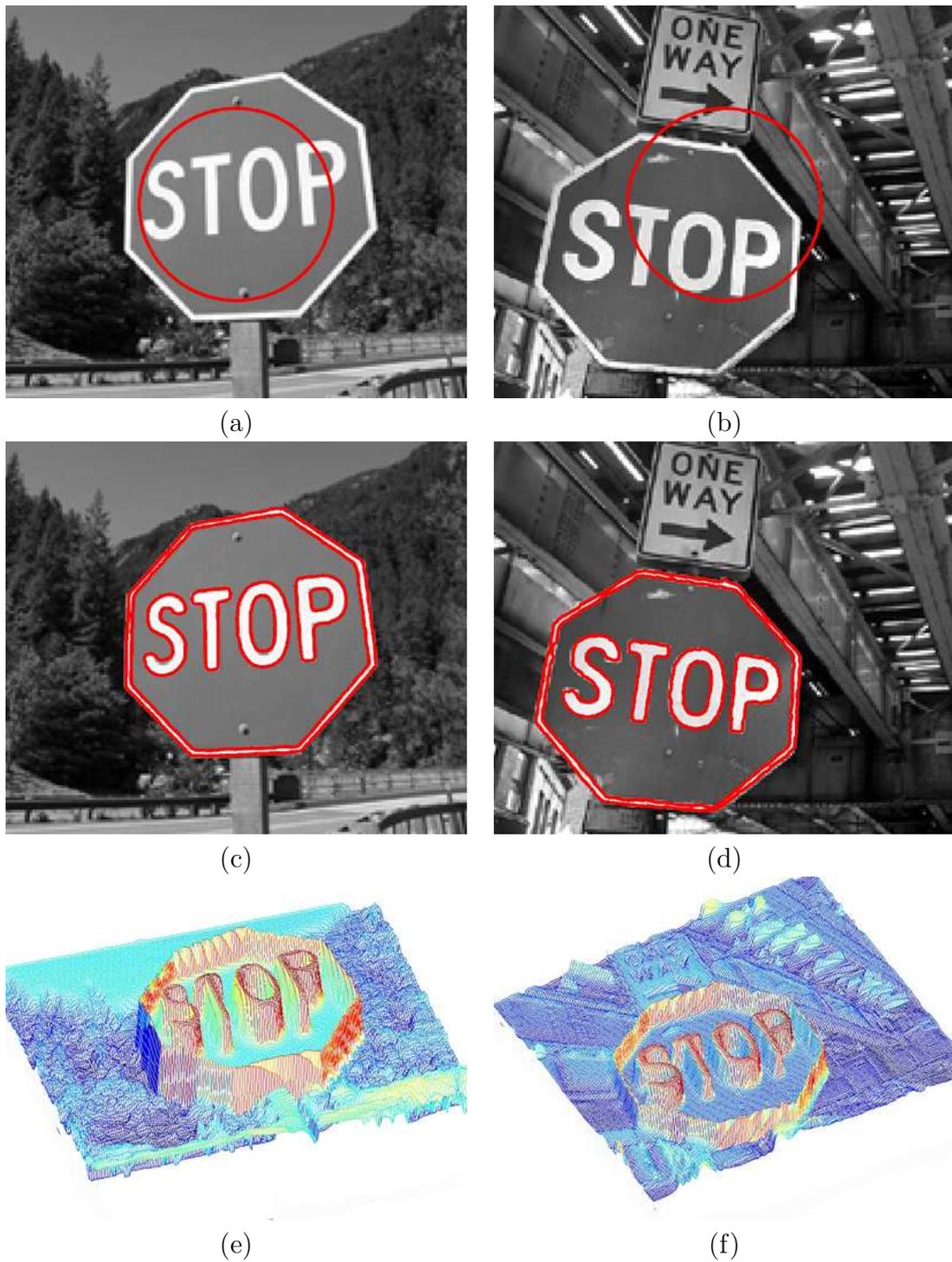


Figure 4.6: Segmentation results of the stop sign images. (a, b) Different images of stop signs taken from different viewing positions with their initial contours (red). (a) Original image courtesy of Erik Henne, URL: www.erikhenne.indexxxx.com. (b) Original image courtesy of the Friedman archives, URL: www.friedmanarchives.com/Chicago. (c,d) Successful detection of stop signs using the proposed *mutual* segmentation algorithm. (e,f) Final level set functions of the stop-sign images obtained by using the *mutual* segmentation algorithm.

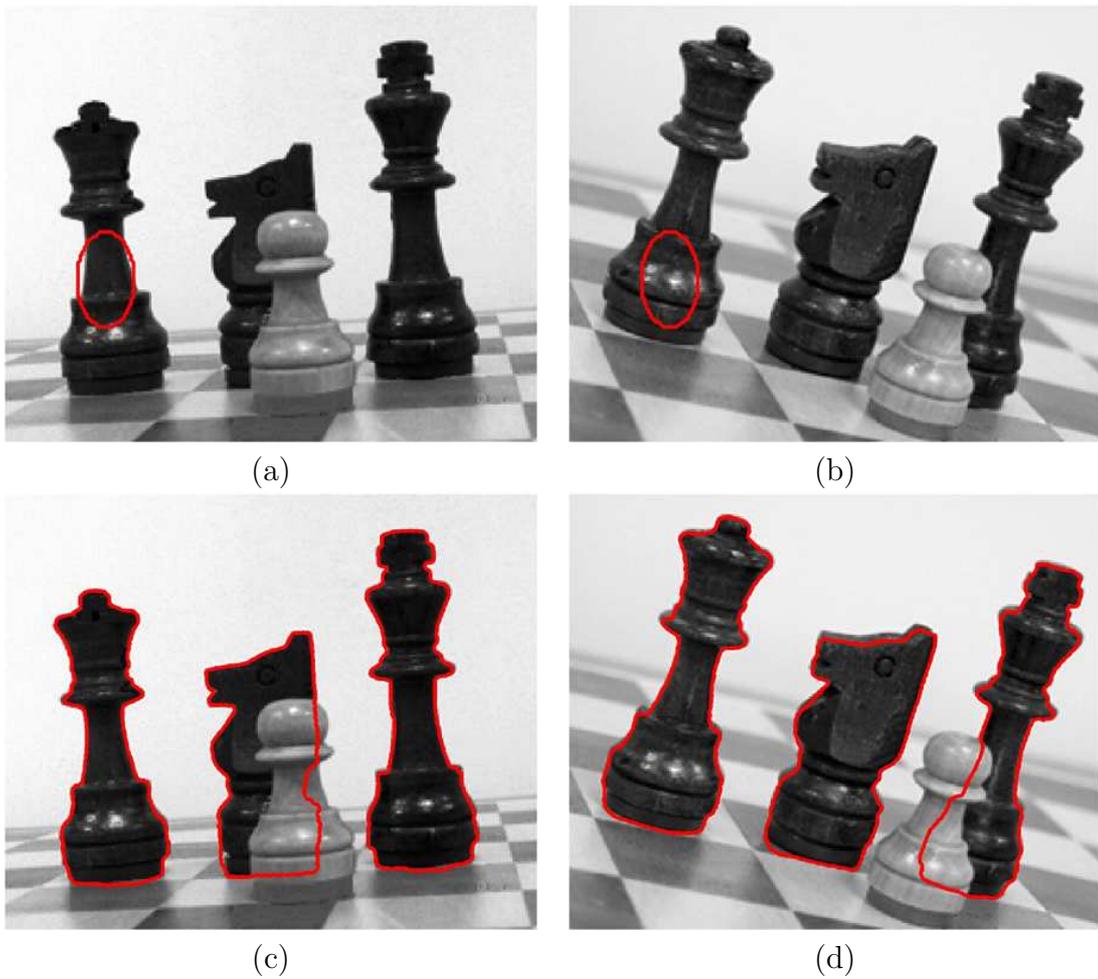


Figure 4.7: Segmentation results of the chess images. (a,b) Input images with their initial contours (red). (c, d) Successful segmentation of chess pieces using the proposed *mutual* segmentation algorithm. The final delineating contour overcomes clutter and occlusions.

Chapter 5

Segmentation with Symmetry Constraint

5.1 Background and related works

Shape symmetry is a useful visual feature for image understanding [111]. This research employs symmetry for object detection and segmentation. In the presence of noise, clutter, distortion, shadows, occlusions or assimilation with the background, segmentation becomes challenging. In these cases, object boundaries do not fully correspond to edges in the image and may not delimit homogeneous image regions. Hence, classic region-based and edge based segmentation techniques are not sufficient. We therefore suggest a novel approach to facilitate segmentation of objects that are *known* to be symmetrical, by using their symmetry property as a shape constraint. The model presented is applicable to objects with either rotational or bilateral (reflection) symmetry distorted by projectivity.

The proposed segmentation method partitions the image into foreground and background domains, where the foreground region is known to be approximately symmetrical up to planar projective transformation. The boundary of the symmetrical region (or object) is inferred by minimizing a cost functional. This functional imposes the smoothness of the segmenting contour as well as its alignment with the local maxima of image edges and the homogeneity of the regions it delimits. The assumed symmetry property of the object to extract provides an essential additional

constraint.

There has been intensive research on symmetry related to human and computer vision. We mention a few of the classical and of the most recent papers. Most natural structures are only approximately symmetrical, therefore there is a great interest, pioneered by the work of [126] and followed by [48], in defining a measure of symmetry. There is a mass of work on symmetry and symmetry-axis detection [7, 18, 50, 54, 65, 69, 80, 113]. Recovery of 3D structure from symmetry is explored in [30, 100, 107, 108, 114, 122, 127]. There are a few works that use symmetry for grouping [103] and segmentation [35, 56, 61, 123]. The majority of the symmetry-related chapters consider either bilateral symmetry [30, 100, 127] or rotational symmetry [18, 41, 80, 125]. Some studies are even more specific, for example [68] suggests a symmetry measure for bifurcating structures, [44] handles tree structures and [41, 122] demonstrate the relation between symmetry and perspectivity on simple geometrical shapes such as rectangles.

The suggested method uncommonly treats the symmetrical object shape as a single entity and not as a collection of landmarks or feature points. This is accomplished by modeling the image by a level-set function, assigning the positive levels to the object domain. Taking the role of object indicator functions (or labeling functions) level-sets are most adequate for dynamic shape representation. A dissimilarity measure between objects is defined as a function of the weighted sum of pixels with contradicting labeling. Moreover, any transformation applied to the image domain changes, via the coordinate system, the defined level-set function (chapter 3). The shape of the region bounded by its zero level is transformed correspondingly, as illustrated in Fig. 5.1.

We define the concept of *symmetrical counterpart* in the context of image analysis. When the imaged object has a bilateral symmetry, the symmetrical counterpart image is obtained by a vertical (or horizontal) flip of the image domain. When the imaged object has a rotational symmetry, the symmetrical counterpart image is provided by a rotation of the image domain. In the same manner we define the symmetrical counterpart of a level-set (or a labeling) function. The symmetry constraint is imposed by minimizing the dissimilarity measure between the evolving level-set (labeling) function and its symmetrical counterpart. The proposed segmentation approach is thus

fundamentally different from other methods that support segmentation by symmetry [35, 56, 61, 123].

When symmetry is distorted by perspectivity, the detection of the underlying symmetry becomes non-trivial, thus complicating symmetry aided segmentation. In this case, even a perfectly symmetrical image is not identical to its symmetrical counterpart. We approach this difficulty by performing registration between the symmetrical counterpart level-set functions. The registration process is justified by showing that an image of a symmetrical object, distorted by a projective transformation, relates to its symmetrical counterpart by a planar projective homography. A key result presented in this manuscript is the structure of this homography, which is determined, up to well defined limits, by the distorting projective transformation. A significant gain from this result is the waiving of the potentially obstructive phase of symmetry axis detection - an essential step in all previous symmetry oriented frameworks.

Figs. 5.2-5.3 illustrate the main idea of the proposed framework. Fig. 5.2a shows an approximately symmetrical object (its upper-left part is erased) that underwent a perspective distortion. Fig. 5.2b is a reflection of Fig. 5.2a with respect to the vertical symmetry axis of the image domain. Note however that this is not the symmetry axis of the object view, which is unknown. We call Fig. 5.2b the *symmetrical counterpart* of Fig. 5.2a. Fig. 5.2a-b can be considered as two views of the same object. Fig. 5.2b can be aligned to Fig. 5.2a by applying a perspective transformation different from the counter reflection, as shown in Fig. 5.2c. Superposition of Fig. 5.2a and Fig. 5.2c yields the complete non-corrupted object view as shown in Fig. 5.2d. In the course of the iterative segmentation process the symmetrical counterpart of the object delineated provides a dynamic shape prior and thus facilitates the recovery of the hidden or vague object boundaries.

Fig. 5.3 demonstrates the detection of symmetrical objects. In Fig. 5.3a only one of the flowers imaged has rotational symmetry (up to an affine transformation). The symmetrical counterpart image (Fig. 5.3b) is generated by rotation of the image domain. Fig. 5.3c shows the superposition of the images displayed in Fig. 5.3a,b. Fig. 5.3d presents the superposition of the original image (Fig. 5.3a) and its symmetrical counterpart aligned to it. Note that the alignment between the two images was not obtained by the counter rotation but by an affine transformation.

This chapter contains two fundamental, related contributions. The first is the use of an intrinsic shape property - symmetry - as a prior for image segmentation. This is made possible by a group of theoretical results related to symmetry and projectivity which are the essence of the second contribution. Specifically we present the structure of the homography that relates an image (or a level-set function) to its symmetrical counterpart. The unknown projective transformation that distorts the object symmetry can be recovered from this homography under certain conditions. These conditions are specified, defining the concept of *symmetry preserving transformation*. We show that the transformation applied to a symmetrical image (or a labeling function) cannot be recovered from the homography that aligns symmetrical counterparts if it does not distort the image (or labeling function) symmetry. We propose a measure for the ‘distance’ between a labeling function and its aligned symmetrical counterpart. We call it the *symmetry imperfection measure*. This measure is the basis of the symmetry constraint that is incorporated within a unified functional for segmentation, integrating region-based, edge-based and smoothness constraints. The suggested segmentation method is demonstrated on various images of approximately symmetrical objects distorted by planar projective transformation.

This chapter is organized as follows. In the next section (section 5.2) we review the main concepts of shape representation via level-sets. The mathematical formulation has been already introduced in chapters 2-3. It is represented here from a different viewpoint to emphasize the importance of this unique shape representation in the proposed symmetry-based segmentation algorithm. In section 5.3 the concepts of symmetry and projectivity are reviewed. The main theoretical contribution resides in subsection 5.3.3 which establishes the key elements of the suggested study. In particular, the structure of the homography that relates between symmetrical counterpart images is analyzed. In section 5.4, a measure of *symmetry imperfection of approximately symmetrical* images, based on that homography, is defined. We use this measure to construct a symmetry-shape term. Implementation details and further implications are presented in section 5.5. Experimental results are provided in section 5.6. We conclude this chapter in section 5.7.

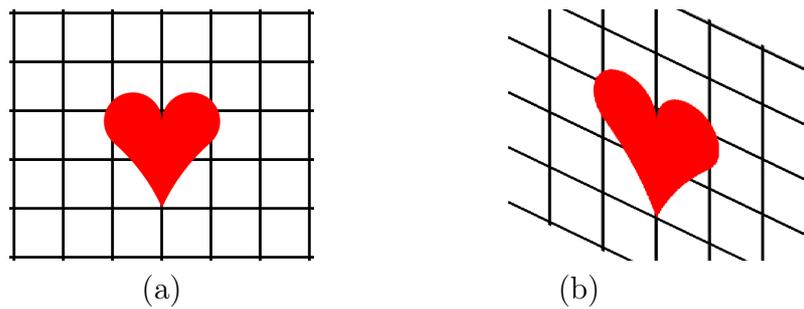


Figure 5.1: Shape transformation. The coordinate transformation applied to the domain of the object indicator function transforms the represented shape accordingly.

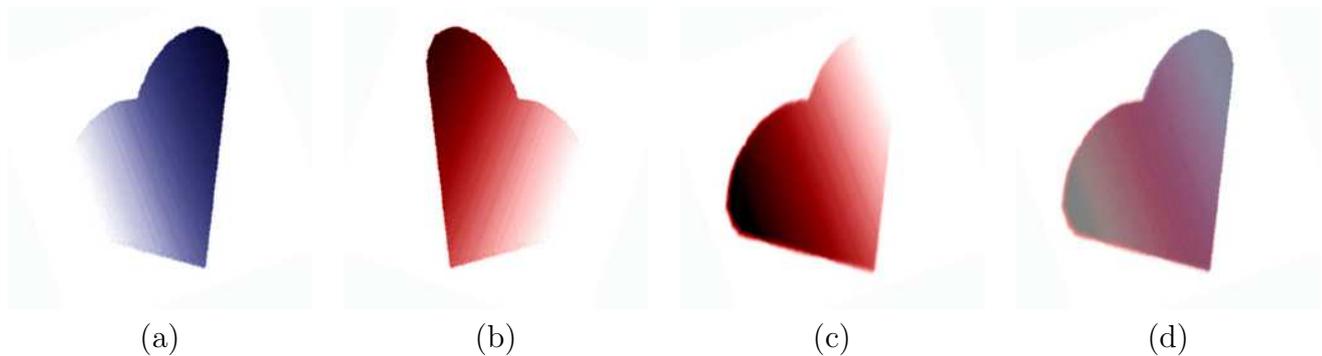


Figure 5.2: (a) Image of approximately symmetrical object distorted by perspective. (b) The *symmetrical counterpart* of (a) is the reflection of (a) along the vertical symmetry axis of the image domain. (c) The image (b) aligned to (a). (d) Superposition of the image (a) with the image (c). The corrupted shape is recovered.

5.2 Level sets framework and shape representation

In chapters 1- 2 we have discussed the advantages of the level-set representation. The implicit contour representation is non-parametric and allows an automatic changes in the contour topology. These characteristics are also most desirable for the dynamic representation of the *shape* of the image region bounded by the zero level - an essential element of proposed segmentation framework - which is based on shape symmetry.

5.2.1 Notation

Let I denote an image defined on the domain $\Omega \subset \mathbb{R}^2$. We define a closed contour $C \in \Omega$, that partitions the image domain Ω into two disjoint open subsets: ω and $\Omega \setminus \omega$. Without loss of generality, the image pixels $\{\mathbf{x} \mid \mathbf{x} \in \omega\}$ will be attributed to

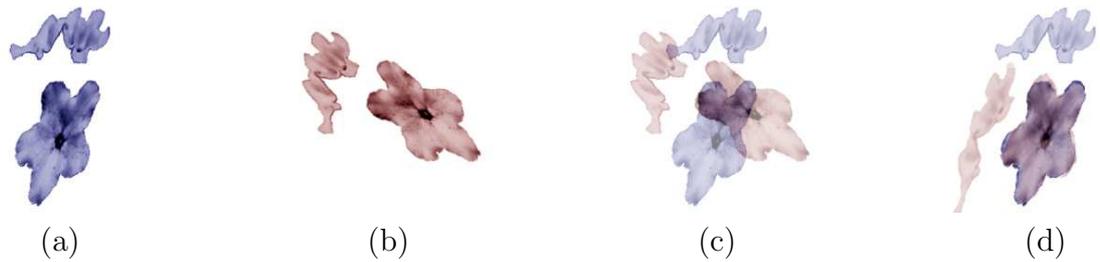


Figure 5.3: (a) The image contains an object with rotational symmetry distorted by affine transformation. (b) The *symmetrical counterpart* of (a) – Rotation of the image plane (a) by 90° . (c) Superposition of (a) and (b). (d) Superposition of the image (a) with the image (b) aligned to (a). The object with rotational symmetry can be extracted.

the object domain. We use the level-set formulation for an implicit representation of the evolving boundary $C(t)$ - which is defined as the zero level of a level set function $\phi: \Omega \rightarrow \mathbb{R}$ at time t . The heaviside function of ϕ (defined in 2.3) assigns the positive levels of the level-set function to the object domain and the negative levels to the background. We can equivalently define an object indicator function or a labeling function $L: \mathbb{R}^2 \rightarrow \{0, 1\}$ where,

$$L(\mathbf{x}, t) = \begin{cases} 1 & \mathbf{x} \in \omega \\ 0 & \mathbf{x} \in \Omega \setminus \omega \end{cases} \quad (5.1)$$

The binary function $L(\mathbf{x}, t)$ is determined by the evolving level-set function at time t . This function thus embeds the dynamics of the shape bounded by the propagating contour. It also plays an important role in defining shape symmetry, for example, if $L(\mathbf{x}, t) = L(x, y, t)$ embeds a shape with left-right bilateral symmetry then $L(x, y, t) = L(-x, y, t)$. In general, transformation of the embedded shape is obtained by a coordinate transformation \mathcal{H} of the image domain i.e. $L_{\mathcal{H}}(t) = L(\mathbf{x}', t)$, where $\mathbf{x}' = \mathcal{H}\mathbf{x}$, as illustrated in Fig. 5.1. This issue will be hereby discussed.

5.2.2 Shape term

The shape of the region bounded by the zero level of the evolving level-set function is of much importance when a prior shape constraint is incorporated within the cost

functional. In chapters 3 and 4 the shape term was a dissimilarity measure between the evolving level-set function and a representation of the prior shape. We showed there that the challenging part is the alignment process between the prior shape and the shape evolved in the course of the segmentation. Shape transformations must be then taken into consideration.

Shape transformations

Recall, (chapter 3) that shape transformation is obtained by a coordinate transformation of the image domain Ω . Let \mathcal{H} denote a 3×3 matrix that represents a planar transformation¹. The matrix \mathcal{H} operates on homogeneous vectors $\mathbf{X} = (\mathbf{x}, 1)^T = (x, y, 1)^T$. We define an operator $h: \Omega \rightarrow \tilde{\Omega}$, where $\tilde{\Omega} \in \mathbb{P}^2$ is the projective domain, such that $h(\mathbf{x}) = \mathbf{X}$. Let $\mathbf{X}' = (X', Y', b) = \mathcal{H}\mathbf{X}$ define the coordinate transformation of the projective domain $\tilde{\Omega}$. The entries of the 2-vector $\mathbf{x}' \in \Omega$ are the ratios of the first two coordinates of \mathbf{X}' and the third one, i.e. $\mathbf{x}' = (x', y') = (X'/b, Y'/b)$. Equivalently one can use the ‘inverse’ operator $h^{-1}: \tilde{\Omega} \rightarrow \Omega$ and write $h^{-1}(\mathbf{X}') = \mathbf{x}'$. Transformation of the object indicator function $L(\mathbf{x})$ by a planar transformation \mathcal{H} will be denoted by $L \circ \mathcal{H} \equiv L(\mathcal{H}\mathbf{x})$, where $\mathcal{H}\mathbf{x}$ is a shorthand to $h^{-1}(\mathcal{H}(h(\mathbf{x})))$. We can think on L as a function defined on \mathbb{R}^2 where Ω is the support of the image. This way the operation \mathcal{H} maps vectors in \mathbb{R}^2 to vectors in \mathbb{R}^2 and is well defined. Note that the shape of the support may be changed under the action of the operator/matrix \mathcal{H} .

Shape dissimilarity measure

Let \tilde{L} be a binary representation of the prior shape. The segmentation process is defined by the evolution of the object indicator function L . A shape constraint takes the form $D(L(\mathbf{x}), \tilde{L}(\mathcal{H}\mathbf{x})) < \epsilon$ where D is a dissimilarity measure between $L(\mathbf{x})$ and the aligned prior shape representation $\tilde{L}(\mathcal{H}\mathbf{x})$. The matrix \mathcal{H} represents that alignment and is recovered concurrently with the segmentation process.

When only a single image is given, such prior is not available. Nevertheless, if an object is known to be symmetrical, its replicative form, induced by the symmetry,

¹Note that \mathcal{H} denotes the transformation while H denotes the Heaviside function.

can be used. Specifically, we treat the image and its symmetrical counterpart (e.g. its reflection or rotation) as if they are two different views of the same object. The instantaneous symmetrical counterpart of the evolving shape provides a dynamic shape prior. The symmetry dissimilarity measure is based on a theoretical framework established in section 5.3. Section 5.4 considers the incorporation of the symmetry constraint within a level-set framework for segmentation.

5.3 Symmetry and projectivity

5.3.1 Symmetry

Symmetry is an intrinsic property of an object. An object is symmetrical with respect to a given operation if it remains invariant under that operation. In $2D$ geometry these operations relate to the basic planar Euclidean isometries: reflection, rotation and translation. We denote the symmetry operator by S . The operator S is an isometry that operates on homogeneous vectors $\mathbf{X} = (\mathbf{x}, 1)^T = (x, y, 1)^T$ and is represented as

$$S = \begin{vmatrix} \mathbf{s}R(\theta) & \mathbf{t} \\ \mathbf{0}^T & 1 \end{vmatrix} \quad (5.2)$$

where \mathbf{t} is a $2D$ translation vector, $\mathbf{0}$ is the null $2D$ vector, R is a 2×2 rotation matrix and \mathbf{s} is the diagonal matrix $\text{diag}(\pm 1, \pm 1)$.

Specifically, we consider *either* of the following transformations:

1. S is translation if $\mathbf{t} \neq \mathbf{0}$ and $\mathbf{s} = R(\theta) = \text{diag}(1, 1)$.
2. S is rotation if $\mathbf{t} = \mathbf{0}$, $\mathbf{s} = \text{diag}(1, 1)$ and $\theta \neq 0$.
3. S is reflection if $\mathbf{t} = \mathbf{0}$, $\theta = 0$ and \mathbf{s} is either $\text{diag}(-1, 1)$ for left-right reflection or $\text{diag}(1, -1)$ for up-down reflection.

In the case of reflection, the symmetry operation *reverses orientation*, otherwise (translation, rotation and inversion) it is *orientation preserving*.

The particular case of translational symmetry requires an infinite image domain. Hence, it is not specifically considered in this manuscript.

Definition 5.3.1 (Symmetrical image). *Let S denote a symmetry operator as defined in (5.2). The operator S is either reflection or rotation. The image $I: \Omega \rightarrow \mathbb{R}^+$ is symmetrical with respect to S if*

$$I(\mathbf{x}) = I(S\mathbf{x}) \equiv I \circ S \quad (5.3)$$

The concept of symmetry is intimately related to the notion of invariance. We say that a vector or a function L is invariant with respect to the transformation (or operator) S if L is in the kernel of $(S - 1)$ i.e. $SL = L$. It means in particular that the respective eigen value of S is 1. Inversely we can call the transformation (or operator) that has an eigenvalue 1 a symmetry operator and say that L is symmetric with respect to the operation S if L is invariant to S .

Since we are interested in symmetry in terms of shape and not in terms of gray levels, we will further consider the *object indicator function* $L: \Omega \rightarrow \{0, 1\}$ defined in subsection 5.2.1. Hereafter, we use the shorthand notation for the coordinate transformation of the image domain as defined in subsection 5.2.2.

Definition 5.3.2 (Symmetrical counterpart). *Let S denote a symmetry operator. The object indicator function $\hat{L}(\mathbf{x}) = L \circ S(\mathbf{x}) = L(S\mathbf{x})$ is the symmetrical counterpart of $L(\mathbf{x})$.*

L is symmetrical iff $L = \hat{L}$.

We claim that the object indicator function of a symmetrical object distorted by a projective transformation is related to its symmetrical counterpart by projective transformation different from the defining symmetry. Before we proceed proving this claim we recall the definition of projective transformation.

5.3.2 Projectivity

This subsection follows the definitions in [38].

Definition 5.3.3. *A planar projective transformation (projectivity) is a linear transformation represented by a non-singular 3×3 matrix \mathcal{H} operating on homogeneous vectors, $\mathbf{x}' = \mathcal{H}\mathbf{x}$, as in equation 3.18.*

Important specializations of the group formed by projective transformation are the affine group and the similarity group which is a subgroup of the affine group. These groups form a hierarchy of transformations. A similarity transformation is represented by

$$\mathcal{H}_{\text{SIM}} = \begin{bmatrix} \kappa R(\theta) & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (5.4)$$

where R is a 2×2 rotation (by θ) matrix and κ is an isotropic scaling. When $\kappa = 1$, \mathcal{H}_{SIM} is the Euclidean transformation denoted by \mathcal{H}_E . An affine transformation is obtained by multiplying the matrix \mathcal{H}_{SIM} with

$$\mathcal{H}_A = \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}. \quad (5.5)$$

K is an upper-triangular matrix normalized as $\text{Det}K = 1$. The matrix \mathcal{H}_P defines the “essence” of the projective transformation and takes the form:

$$\mathcal{H}_P = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{v}^T & v \end{bmatrix}, \quad (5.6)$$

where $\mathbf{1}$ is the 2-identity matrix. A projective transformation can be decomposed into a chain of transformations of a descending (or ascending) hierarchy order,

$$\mathcal{H} = \mathcal{H}_{\text{SIM}}\mathcal{H}_A\mathcal{H}_P = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix} \quad (5.7)$$

where $v \neq 0$ and $\mathbf{A} = \kappa RK + \mathbf{t}\mathbf{v}^T$ is a non-singular matrix.

5.3.3 Theoretical results

In this subsection we consider the relation between an image (or object indicator function) of a symmetrical object distorted by planar projective transformation \mathcal{H} and its symmetrical counterpart. The object indicator function L will be treated as a binary image and will be called for simplicity - *image*.

Recall from the previous subsection that a symmetrical image (or labeling func-

tion) L with respect to a symmetry operator S is invariant to S . We next consider the invariant of symmetrical image that underwent a planar projective transformation \mathcal{H} .

Theorem 5.3.4. *Let $L_{\mathcal{H}} = L(\mathcal{H}\mathbf{x})$ denote the image obtained from the symmetrical image L by applying a planar projective transformation \mathcal{H} . If L is invariant to S , i.e. $L(\mathbf{x}) = L(S\mathbf{x}) = L(S^{-1}\mathbf{x})$ then $L_{\mathcal{H}}$ is invariant to $S_{\mathcal{H}}$ where $S_{\mathcal{H}} = \mathcal{H}^{-1}S\mathcal{H} \equiv \mathcal{H}^{-1}S^{-1}\mathcal{H}$.*

Proof. We need to prove that $L_{\mathcal{H}}(\mathbf{x}) = L_{\mathcal{H}}(S_{\mathcal{H}}\mathbf{x})$.

We define $\mathbf{y} = \mathcal{H}\mathbf{x}$.

$$\begin{aligned} L_{\mathcal{H}}(S_{\mathcal{H}}\mathbf{x}) &= L_{\mathcal{H}}(\mathcal{H}^{-1}S\mathcal{H}\mathbf{x}) = L(S\mathcal{H}\mathbf{x}) \\ &= L(S\mathbf{y}) = L(\mathbf{y}) \\ &= L(\mathcal{H}\mathbf{x}) = L_{\mathcal{H}}(\mathbf{x}) \end{aligned} \tag{5.8}$$

□

Alternatively we can rewrite Eq.(5.8) as follows:

$$\begin{aligned} L_{\mathcal{H}}(\mathbf{x}) &= L_{\mathcal{H}}(\mathcal{H}^{-1}\mathcal{H}\mathbf{x}) = L(\mathcal{H}\mathbf{x}) \\ &= L(\mathbf{y}) = L(S\mathbf{y}) \\ &= L(\mathcal{H}\mathcal{H}^{-1}S\mathbf{y}) = L_{\mathcal{H}}(\mathcal{H}^{-1}S\mathbf{y}) \\ &= L_{\mathcal{H}}(\mathcal{H}^{-1}S\mathcal{H}\mathbf{x}) = L_{\mathcal{H}}(S_{\mathcal{H}}\mathbf{x}) \end{aligned} \tag{5.9}$$

he chain of equalities in (5.9) is equivalent to the following sequence of operations:

1. Apply the inverse of the projective (distorting) transformation, \mathcal{H}^{-1} on $L_{\mathcal{H}}$ to generate a symmetrical image L .
2. Apply the symmetry operation S on L , under which it remains invariant.
3. Apply the projective transformation matrix \mathcal{H} on L to obtain back the image $L_{\mathcal{H}}$.

We now use the result obtained in *theorem 1* to define the structure of the homography that aligns symmetrical counterpart images.

Theorem 5.3.5. Let $L_{\mathcal{H}}$ denote the image obtained from the symmetrical image L by applying a planar projective transformation \mathcal{H} . Let $\hat{L}_{\mathcal{H}} = L_{\mathcal{H}}(S\mathbf{x}) \equiv L_{\mathcal{H}} \circ S$ denote the symmetrical counterpart of $L_{\mathcal{H}}$ with respect to a symmetry operation S . The image $L_{\mathcal{H}}$ can be obtained from its symmetrical counterpart $\hat{L}_{\mathcal{H}}$ by applying transformation represented by a 3×3 matrix of the form:

$$M = S^{-1}\mathcal{H}^{-1}S\mathcal{H}. \quad (5.10)$$

Proof. We need to prove that $L_{\mathcal{H}}(\mathbf{x}) = \hat{L}_{\mathcal{H}}(M\mathbf{x})$, where $M = S^{-1}\mathcal{H}^{-1}S\mathcal{H}$.

The image $L_{\mathcal{H}}$ is invariant to $S_{\mathcal{H}}$ thus $L_{\mathcal{H}}(\mathbf{x}) = L_{\mathcal{H}}(S_{\mathcal{H}}\mathbf{x})$. By definition, $\hat{L}_{\mathcal{H}}(\mathbf{x}) = L_{\mathcal{H}}(S\mathbf{x})$.

From the above equations and *theorem 1*, defining $\mathbf{y} = S^{-1}S_{\mathcal{H}}\mathbf{x}$, we get:

$$\begin{aligned} L_{\mathcal{H}}(\mathbf{x}) &= L_{\mathcal{H}}(S_{\mathcal{H}}\mathbf{x}) = L_{\mathcal{H}}(SS^{-1}S_{\mathcal{H}}\mathbf{x}) \\ &= L_{\mathcal{H}}(S\mathbf{y}) = \hat{L}_{\mathcal{H}}(\mathbf{y}) \\ &= \hat{L}_{\mathcal{H}}(S^{-1}S_{\mathcal{H}}\mathbf{x}) = \hat{L}_{\mathcal{H}}(S^{-1}\mathcal{H}^{-1}S\mathcal{H}\mathbf{x}) \\ &= \hat{L}_{\mathcal{H}}(M\mathbf{x}). \end{aligned} \quad (5.11)$$

The image $L_{\mathcal{H}}$ can be generated from its symmetrical counterpart $\hat{L}_{\mathcal{H}}$ either by applying the inverse of the symmetry operation S or by a projective transformation M which is different from S^{-1} .

Let M_{INV} denote a 3×3 non-singular matrix such that $M_{\text{INV}} = \mathcal{H}^{-1}S^{-1}\mathcal{H}S$. M_{INV} is a projective transformation since $\mathcal{H}M_{\text{INV}} = S^{-1}\mathcal{H}S$ is a projective transformation according to:

$$S^{-1}\mathcal{H}S = S^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & 1 \end{bmatrix} S = \begin{bmatrix} \mathbf{A}' & \mathbf{t}' \\ \mathbf{v}'^T & v' \end{bmatrix} = \mathcal{H}' \quad (5.12)$$

where \mathcal{H} is scaled such that $v = 1$.

Thus, $M = M_{\text{INV}}^{-1}$ represents a projective transformation. It is easy to prove that $M \neq S^{-1}$, when S is not the identity matrix. Assume to the contrary that there exists a non-singular \mathcal{H} and a symmetry operation S such that $M = S^{-1}$. Then, from (5.10), $S^{-1} = S^{-1}\mathcal{H}^{-1}S\mathcal{H}$. Thus, $\mathcal{H} = S\mathcal{H}$, which implies that either S is the

identity matrix or \mathcal{H} is singular, in contradiction to the assumptions. \square

The next theorem gives tighter characterization of M .

Theorem 5.3.6. *Let $L_{\mathcal{H}}$ denote the image obtained from the symmetrical image L by applying transformation \mathcal{H} . Let M denote the matrix that relates $L_{\mathcal{H}}$ to its symmetrical counterpart $\hat{L}_{\mathcal{H}}$. The matrices M and \mathcal{H} belong to the same subgroup of transformations.*

Proof. The proof is divided into several sub-cases. When S is either rotation or translation, M is the product of Euclidean transformations (S, S^{-1}) and either Euclidean, similarity or affine transformations ($\mathcal{H}, \mathcal{H}^{-1}$). Thus M belongs to the respective subgroup of transformations. When $S = \text{diag}(\mathbf{s}, 1)$ is reflection, i.e. \mathbf{s} is either $(1, -1)$ or $(-1, 1)$ the claim can be simply proved by matrix multiplications. Specifically, when $\mathcal{H} = \mathcal{H}_{\text{SIM}}$ represents a similarity transformation, as defined in (5.4):

$$M = \begin{bmatrix} R(2\theta) & \mathbf{t}' \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \mathbf{t}' = \kappa^{-1} \mathbf{t}^T [\mathbf{1} - \text{diag}(\mathbf{s})] R(\theta)^T \quad (5.13)$$

When \mathcal{H} represents an affine transformation:

$$\mathcal{H} = \mathcal{H}_A = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

$$M = \begin{bmatrix} A' & \mathbf{t}' \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (5.14)$$

where,

$$A' = \frac{1}{\text{Det}A} \begin{bmatrix} a_1 a_4 + a_3 a_2 & 2a_2 a_4 \\ 2a_1 a_3 & a_1 a_4 + a_3 a_2 \end{bmatrix}, \quad \mathbf{t}' = \frac{1}{\text{Det}A} \mathbf{t}^T [\mathbf{1} - \text{diag}(\mathbf{s})] A.$$

\square

Next, we argue that \mathcal{H} cannot be explicitly and fully recovered from M and S , if the operation of \mathcal{H} (or one of its factors) on a symmetrical image does not distort its symmetry.

Definition 5.3.7 (symmetry preserving transformation). *Let L be a symmetrical image with respect to S . The projective transformation matrix \mathcal{H} is symmetry preserving if*

$$L_{\mathcal{H}} = L(\mathcal{H}\mathbf{x}) = L(S\mathcal{H}\mathbf{x}) \quad (5.15)$$

Lemma 5.3.8. *Let \mathcal{H} denote a symmetry preserving transformation with respect to the symmetry operation S . Then \mathcal{H} and S commute, i.e. $S\mathcal{H} = \mathcal{H}S$.*

Proof. L is symmetrical, i.e. $L(\mathbf{x}) = L(S\mathbf{x})$. Applying the transformation \mathcal{H} on L we obtain: $L(\mathcal{H}\mathbf{x}) = L(\mathcal{H}S\mathbf{x})$. Since \mathcal{H} is symmetry preserving, $L(\mathcal{H}\mathbf{x})$ is symmetrical, thus $L(\mathcal{H}\mathbf{x}) = L(S\mathcal{H}\mathbf{x})$. The latter two equations prove the claim. \square

Theorem 5.3.9. *Consider the image $L_{\mathcal{H}} = L(\mathcal{H}\mathbf{x})$, where L is a symmetrical image. Let $\hat{L}_{\mathcal{H}} = L(S\mathcal{H}\mathbf{x})$ denote its symmetrical counterpart. Let the matrix M satisfy Eq. (5.10), where $\hat{L}_{\mathcal{H}}(M\mathbf{x}) = L_{\mathcal{H}}(\mathbf{x})$. If \mathcal{H} can be factorized such that $\mathcal{H} = \mathcal{H}_S\tilde{\mathcal{H}}$ and \mathcal{H}_S denotes a symmetry preserving transformation with respect to S , then \mathcal{H} cannot be recovered from M .*

Proof.

$$\begin{aligned} M &= S^{-1}\mathcal{H}^{-1}S\mathcal{H} \\ &= S^{-1}(\mathcal{H}_S\tilde{\mathcal{H}})^{-1}S\mathcal{H}_S\tilde{\mathcal{H}} = S^{-1}\tilde{\mathcal{H}}^{-1}\mathcal{H}_S^{-1}S\mathcal{H}_S\tilde{\mathcal{H}} \\ &= S^{-1}\tilde{\mathcal{H}}^{-1}\mathcal{H}_S^{-1}\mathcal{H}_S S\tilde{\mathcal{H}} \quad \text{since } S \text{ and } \mathcal{H}_S \text{ commute} \\ &= S^{-1}\tilde{\mathcal{H}}^{-1}S\tilde{\mathcal{H}} \end{aligned} \quad (5.16)$$

\square

5.3.4 Examples

The claims above are exemplified for four particular cases. Consider, first, the image of the symmetrical object and its left-right reflection shown in Fig. 5.4a-b. Suppose that the image symmetry has been distorted by an Euclidean transformation of the form:

$$\mathcal{H}_E = \begin{bmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the Euclidean transformation is an isometry and thus preserves the object symmetry. However, it draws the symmetry axis of the object away from the symmetry axis of the image, rotating it by angle θ and translating it by t_x . Fig. 5.4a relates to its symmetrical counterpart by:

$$M = S^{-1}\mathcal{H}_E^{-1}S\mathcal{H}_E = \begin{bmatrix} \cos 2\theta & -\sin 2\theta & 2t_x \cos \theta \\ \sin 2\theta & \cos 2\theta & 2t_x \sin \theta \\ 0 & 0 & 1 \end{bmatrix},$$

where $S = (-1, 1, 1)$. Fig. 5.4a can thus be obtained from Fig. 5.4b by a rotation with angle 2θ and translation by $2R(\theta)[t_x, 0]^T$. Note that any translation parallel to the symmetry axis (in this case t_y) cannot be recovered from M . This expression for M implies that in the case of pure Euclidean transformation, the deviation of the object's symmetry axis from that of the image can be recovered from M . Obviously M is an Euclidean transformation.

Consider, next the images shown in Fig. 5.4c-d. The object is distorted by a projective transformation \mathcal{H}_P :

$$\mathcal{H}_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v_1 & v_2 & 1 \end{bmatrix}$$

The relation between the two images can be described by:

$$M = S^{-1}\mathcal{H}_P^{-1}S\mathcal{H}_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2v_1 & 0 & 1 \end{bmatrix}$$

When $v_2 \neq 0$ the object shape is distorted but its symmetry is preserved, thus v_2 cannot be recovered from M .

The last example, Fig. 5.6 demonstrates the particular case of translational symmetry. The object (line of elephants) is invariant to translation along the x direction, as long as the image domain is infinite. Fig. 5.6b is obtained by applying projec-

tive transformation $\mathcal{H} = \mathcal{H}_P$ on Fig. 5.6a. Corresponding pairs of sub-images, for example the sub-images bordered by the red and the blue dashed lines are symmetrical counterparts. Each sub-image can be obtained from the other by a projective transformation other than a pure translation, t_x . Let

$$S = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then,

$$M = \begin{bmatrix} 1 + t_x v_1 + t_x^2 v_1^2 & t_x v_2 + t_x^2 v_1 v_2 & t_x^2 v_1 \\ 0 & 1 & 0 \\ -t_x v_1^2 & -t_x v_1 v_2 & 1 - t_x v_1 \end{bmatrix}.$$

Handling images with translational symmetry requires infinite image domain. Hence the discussion on translational invariance is only theoretical.

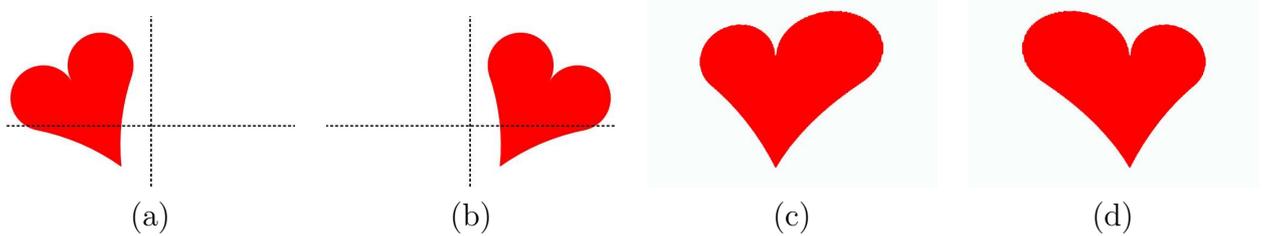


Figure 5.4: (a) An image of a symmetrical object transformed by an Euclidean transformation: planar rotation by θ and then translation by \mathbf{t} . (b) The symmetrical counterpart of the image in (a). (c) An image of a symmetrical object distorted by projective transformation. (d) The symmetrical counterpart of the image in (c).

Finally, consider the object shown in Fig. 5.5. The object's appearance is invariant to rotation by $n\pi/2$, where n is an integer. Let S denote rotation by $\pi/2$ i.e.

$$S = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Fig. 5.5b is obtained by applying the affine transformation $\mathcal{H} = \mathcal{H}_{\text{SIM}}\mathcal{H}_A$ on Fig. 5.5a.

Specifically, in this example

$$\mathcal{H} = \begin{bmatrix} R(\theta) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \text{where} \quad K = \begin{bmatrix} k_1 & k_2 \\ 0 & 1/k_1 \end{bmatrix}.$$

The scalars k_1 and k_2 determine the non-isotropic scaling: ratio and orientation. Fig. 5.5c is the symmetrical counterpart of Fig. 5.5b. It can be also obtained by applying M^{-1} on Fig. 5.5b, where M is of the form:

$$M = \begin{bmatrix} k_1^2 & k_1 k_2 & 0 \\ k_1 k_2 & 1/k_1^2 + k_2^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note, that the rotation θ has no effect on M .

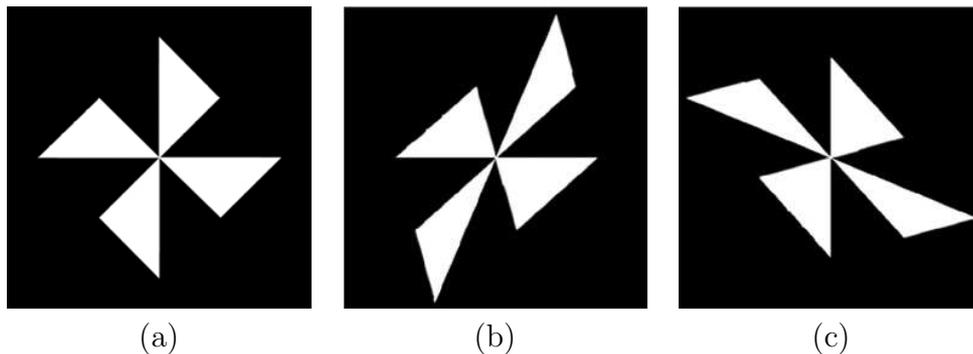


Figure 5.5: (a) Schematic draw of a vane. The vane is symmetrical under rotation by $n\pi/4$, where n is an integer. (b) The vane (as in (a)) distorted by affine transformation. (c) The symmetrical counterpart of (b). This image can be generated from (b) either by clockwise rotation by 90° or by applying a projective transformation which captures the affine distortion of the vane.

5.4 Symmetry based segmentation

5.4.1 Symmetry imperfection

In the current subsection we formalize the concept of *approximately symmetrical* shape. Specifically, two alternative definitions for ϵ -*symmetrical* images and *symmetry imperfection* measures are suggested.

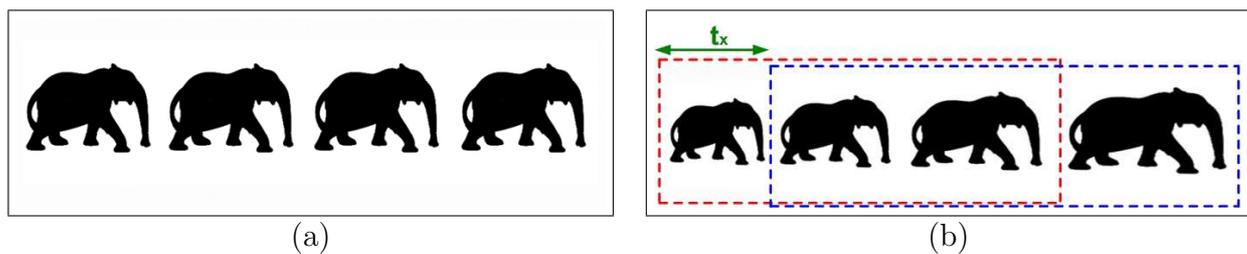


Figure 5.6: (a) A line of successive elephants. The image is symmetrical under translation along the x direction if the image domain is extended to infinity. (b) A perspective distortion of the image in (a). Corresponding pairs of sub-images, for example the sub-images bordered by the red and the blue dashed lines are symmetrical counterparts. Each sub-image can be obtained from the other by a projective transformation other than a pure translation, t_x .

The symmetry imperfection measure is of much importance when segmentation is constrained by symmetry. Finding the object contour that minimizes this measure implies that the eventual object shape is imposed to be symmetrical. When symmetry is distorted by projectivity, defining a measure of symmetry imperfection is much less trivial. We approach this difficulty relying on the results presented by *theorems* 1,2 in the previous section. The following measures for symmetry imperfection are based on a dissimilarity measure between two labeling functions. We denote this measure by $\mathcal{D} = D(L_1, L_2)$, where \mathcal{D} is a function of the contradicting labeling in L_1 and L_2 . The exact form of \mathcal{D} will be discussed in subsection 5.4.3.

Recall that a symmetrical binary image distorted by projective transformation is invariant to the operator $S_{\mathcal{H}}$ as defined in equation (5.8). This is the basis of the first definition.

Definition 5.4.1 (ϵ -symmetrical image; Symmetry imperfection (1)). *The image L^ϵ is ϵ -symmetrical with respect to the symmetry operator S if $\mathcal{D} = D(L^\epsilon, L^\epsilon \circ S) \leq \epsilon$, where \mathcal{D} is a dissimilarity measure (distance function) and ϵ is a small and positive scalar. The measure $D(L^\epsilon, L^\epsilon \circ S)$ quantifies the symmetry imperfection of L^ϵ . Let $L_{\mathcal{H}}^\epsilon$ denote the image obtained from L^ϵ by applying a planar projective transformation \mathcal{H} . The measure for the symmetry imperfection of the perspectively distorted image $L_{\mathcal{H}}^\epsilon$ is defined by $D(L_{\mathcal{H}}^\epsilon, L_{\mathcal{H}}^\epsilon \circ S_{\mathcal{H}})$, where $S_{\mathcal{H}} = \mathcal{H}^{-1}S\mathcal{H}$.*

Although this definition is natural, it not always applicable, since usually the

projective transformation \mathcal{H} is unknown. We therefore use an alternative equivalent definition of symmetry imperfection. This alternative measure involves the concept of the symmetrical counterpart image (see definition 5.3.2) and the homography M that aligns symmetrical counterpart images.

Definition 5.4.2 (Symmetry imperfection (2)). *Let $L_{\mathcal{H}}^{\epsilon}$ be ϵ -symmetrical image (with respect to S) distorted by planar projective homography \mathcal{H} . We assume that S is known, but \mathcal{H} is unknown. Let $\hat{L}_{\mathcal{H}}^{\epsilon} = L_{\mathcal{H}}^{\epsilon} \circ S$ denote the symmetrical counterpart of $L_{\mathcal{H}}^{\epsilon}$. The distance function $D(L_{\mathcal{H}}^{\epsilon}, \hat{L}_{\mathcal{H}}^{\epsilon} \circ M)$ measures to the symmetry imperfection of $L_{\mathcal{H}}^{\epsilon}$.*

Lemma 5.4.3. *The two definitions are equivalent*

Proof. Direct consequence of the identity $\hat{L}_{\mathcal{H}}^{\epsilon} \circ M = L_{\mathcal{H}}^{\epsilon} \circ S_{\mathcal{H}}$ □

We assume that M is known or precisely recovered via registration between $\hat{L}_{\mathcal{H}}$ and $L_{\mathcal{H}}$.

In the following subsections we show how M can be recovered during the segmentation process by minimizing the term $\mathcal{D}_{\mathcal{H}} = D(L_{\mathcal{H}}^{\epsilon}, \hat{L}_{\mathcal{H}}^{\epsilon} \circ M)$ with respect to M . The recovery of M can be enhanced when \mathcal{H} is either Euclidean or affine transformation, following the result of *theorem 3*. In that case only 4 or 6 entries of M should be recovered. The matrix \mathcal{H} can be recovered from M up to a symmetry preserving transformation (*theorem 4*).

5.4.2 Symmetry constraint

In the previous section the symmetrical counterparts were either images or labeling or object indicator functions. Using level set formulation, we will now refer to level-sets and their Heaviside functions, $H(\phi)$. Recall that $H(\phi(t))$ is an indicator function of the estimated object regions in the image at time t .

Let $\hat{\phi}: \Omega \rightarrow \mathbb{R}$ denote the symmetrical counterpart of ϕ with respect to a symmetry operation S . Specifically, $\hat{\phi}(\mathbf{x}) = \phi(S\mathbf{x})$ where S is either reflection or rotation. We assume that S is known. Let M denote the planar projective transformation that aligns $H(\hat{\phi})$ to $H(\phi)$ i.e. $H(\hat{\phi}) \circ M = H(\hat{\phi}(M\mathbf{x})) = H(\hat{\phi}_M)$. M captures the deviation of the object symmetry axis from that of the image and the projective

transformation that distorts its symmetry. Note, however, that this information is not known in advance. The matrix M is recovered by a registration process held concurrently with the segmentation, detailed in subsection 5.4.4.

Let $\mathcal{D} = D(H(\phi), H(\hat{\phi}_M))$ denote a dissimilarity measure between the evolving shape representation and its symmetrical counterpart. Note that if M is correctly recovered and ϕ captures a perfectly symmetrical object (up to projectivity) then \mathcal{D} should be zero. \mathcal{D} thus quantifies the symmetry imperfection of ϕ which is not caused by the projectivity. Whenever this distortion is due to false detection of the object boundaries (caused by noise, occlusions, clutter, etc.) and not a feature of the object shape, \mathcal{D} defines an adequate symmetry constraint. A possible definition of \mathcal{D} is:

$$D(\phi, \hat{\phi} \mid M) = \int_{\Omega} \left[H(\phi(\mathbf{x})) - H(\hat{\phi}_M) \right]^2 d\mathbf{x}. \quad (5.17)$$

This measure is identical to the measure suggested in chapter 3 for prior-based segmentation. Nevertheless, since $H(\hat{\phi})$ is identical to $H(\phi)$ up to an isometry, it is subject to the same distortions in its symmetry and thus cannot replace a well defined prior. Hence, a different formulation is needed, to be described in the following subsection.

5.4.3 Biased shape dissimilarity measure

Consider, for example, the approximately bilateral symmetrical (up to projectivity) images shown in Fig. 5.7a,d. The objects symmetry is distorted by either deficiencies or excess parts. We would like to use the symmetry to overcome these shape distortions. Nevertheless, incorporating the unbiased shape constraint (according to Eq. 5.17) in the cost functional for segmentation, results in the undesired segmentation shown in Fig. 5.7b,e. The symmetrical counterpart of a level-set function ϕ is as imperfect as ϕ . To support a correct evolution of ϕ by $\hat{\phi}$, we have to account for the specific type of corruption.

The problem is identical to the one presented in chapter 4. We used then the observation that an integral over the pointwise exclusive-or (xor) operations between the respective pixels can replace the sum-square-differences in equation. (5.17). Similarly to equation 4.1 in chapter 4, we will rewrite the dissimilarity measure between

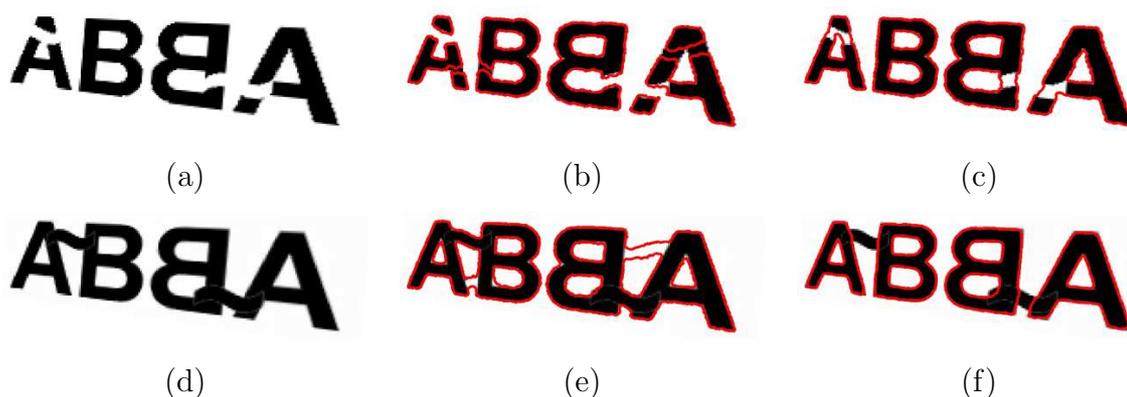


Figure 5.7: Segmentation results using biased and unbiased dissimilarity measures. (a, d) Images of symmetrical objects up to a projective transformation. The objects are distorted either by deficiencies (a) or by excess parts (d). (b, e) Segmentation (red) of the images in (a) and (d) respectively, using *unbiased* dissimilarity measure between ϕ and its transformed reflection as in Eq. (5.17). Object segmentation is further spoiled due to the imperfection in its reflection. (c, f) Successful segmentation (red) using the *biased* dissimilarity measure as in Eq. (5.19).

the evolving shape and its symmetrical counterpart, as follows:

$$D(\phi, \hat{\phi} \mid M) = \int_{\Omega} \left[H(\phi) \left(1 - H(\hat{\phi}_M) \right) + (1 - H(\phi)) H(\hat{\phi}_M) \right] d\mathbf{x} \quad (5.18)$$

Note that the expressions (5.17) and (5.18) are identical, since $H(\phi) = (H(\phi))^2$. There are two types of ‘disagreement’ between the labeling of $H(\phi)$ and $H(\hat{\phi}_M)$. The first additive term in the right hand side of (5.18) is not zero for image regions labeled as *object* by ϕ and labeled as *background* by its symmetrical counterpart $\hat{\phi}$. The second additive term of (5.18) is not zero for image regions labeled as *background* by ϕ and labeled as *object* by $\hat{\phi}$. We can change the relative contribution of each term by a relative weight parameter $\mu \neq 0$:

$$E_{\text{SYM}} = \int_{\Omega} \left[\mu H(\phi) \left(1 - H(\hat{\phi}_M) \right) + (1 - H(\phi)) H(\hat{\phi}_M) \right] d\mathbf{x} \quad (5.19)$$

The associated gradient equation for ϕ is then:

$$\phi_t^{\text{SYM}} = \delta(\phi) [H(\hat{\phi}_M) - \mu(1 - H(\hat{\phi}_M))] \quad (5.20)$$

Now, if excess parts are assumed, the left penalty term should be dominant, setting $\mu > 1$. Otherwise, if deficiencies are assumed, the right penalty term should be dominant, setting $\mu < 1$. Fig. 5.7c,f show segmentation of symmetrical objects with either deficiencies or excess parts, incorporating the shape term (5.19) within the segmentation functional. We used $\mu = 0.5$ and $\mu = 2$ for the segmentation of Fig. 5.7c and Fig. 5.7f, respectively. In chapter 4 similar dissimilarity measure was used to quantify the ‘distance’ between the alternately evolving level-set functions of two different object views.

5.4.4 Recovery of the transformation

In the previous subsection we suggested a biased dissimilarity measure E_{SYM} (Eq. 5.19) that is based on the shape symmetry. This measure indicates the misalignment between $H(\phi)$ and its symmetrical counterpart $H(\hat{\phi})$. We now look for the optimal alignment matrix M that minimizes E_{SYM} .

Using the notation defined in subsection 5.2.2, we denote the coordinate transformation \mathcal{M} of $\hat{\phi}(\mathbf{x})$ as follows:

$$\mathcal{M}H(\hat{\phi}(\mathbf{x})) = H(\hat{\phi}(\mathbf{x})) \circ M = H(\hat{\phi}(M\mathbf{x})) \equiv H(\hat{\phi}_M) \quad (5.21)$$

We assume that the matrix M is a planar projective homography, as defined in equation (3.18). The eight unknown ratios of its entries $\hat{m}_k = m_{ij}/h_{33}$, $\{i, j\} = \{1, 1\}, \{1, 2\}, \dots, \{3, 2\}$ are recovered through the segmentation process, alternately with the evolution of the level set function ϕ . The PDEs for \hat{m}_k are obtained by minimizing (5.19) with respect to each.

$$\frac{\partial \hat{m}_k}{\partial t} = \int_{\Omega} \delta(\hat{\phi}_M) [(1 - H(\phi)) - \mu H(\phi)] \frac{\partial \mathcal{M}(\hat{\phi})}{\partial \hat{m}_k} d\mathbf{x} \quad (5.22)$$

where,

$$\frac{\partial \mathcal{M}(\hat{\phi})}{\partial \hat{m}_k} = \frac{\partial \mathcal{M}(\hat{\phi})}{\partial x} \left(\frac{\partial x}{\partial x'} \frac{\partial x'}{\partial \hat{m}_k} + \frac{\partial x}{\partial y'} \frac{\partial y'}{\partial \hat{m}_k} \right) + \frac{\partial \mathcal{M}(\hat{\phi})}{\partial y} \left(\frac{\partial y}{\partial x'} \frac{\partial x'}{\partial \hat{m}_k} + \frac{\partial y}{\partial y'} \frac{\partial y'}{\partial \hat{m}_k} \right) \quad (5.23)$$

Refer to chapter 3, subsection 3.5.2 for detailed derivation of (5.22).

5.4.5 Unified segmentation functional

Symmetry-based, edge-based, region-based and smoothness constraints can be integrated to establish a comprehensive cost functional for segmentation:

$$E(\phi) = E_{\text{BU}} + W^{\text{SYM}} E_{\text{SYM}}(\phi), \quad (5.24)$$

The cost functional E_{BU} is composed of the low-level image constraints defined in equation (2.23), where BU stands for bottom-up. The top-down flow of information is incorporated via the symmetry term, E_{SYM} , defined in equation (5.19).

The evolution of ϕ at each time step $\phi(t + \Delta t) = \phi(t) + \phi_t$ is determined by

$$\phi_t = \phi_t^{\text{BU}} + W^{\text{SYM}} \phi_t^{\text{SYM}}, \quad (5.25)$$

where ϕ_t^{BU} is the gradient descent equation of E_{BU} according to equation (2.24) and ϕ_t^{SYM} is the gradient descent equation (5.20). The determination of the weights W^{TERM} is discussed in section 2.2.

5.5 Implementation and further implications

Segmentation is obtained by minimizing a cost functional that incorporates symmetry as well as region-based, edge-based and smoothness constraints. The minimizing level-set function is evolved concurrently with the registration of its instantaneous symmetrical counterpart to it. The symmetrical counterpart is generated by a flip or rotation of the coordinate system of the propagating level-set function. The evolution of the level-set function is controlled by the constraints imposed by the data of the associated image and by the its aligned symmetrical counterpart. The planar projective transformation between the evolving level-set function and its symmetrical counterpart is updated at each iteration.

5.5.1 Algorithm

We summarize the proposed algorithm for segmentation of a symmetric object distorted by projectivity.

1. Choose an initial level-set function $\phi(t = 0)$ that determines the initial contour within the image.
2. Set initial values for the transformation parameters \hat{m}_k . For example, set M to the identity matrix.
3. Compute u_+ and u_- according to (2.7), based on the current contour interior and exterior, defined by $\phi(t)$.
4. Generate the symmetrical counterpart of $\phi(\mathbf{x})$, $\hat{\phi}(\mathbf{x}) = \phi(S\mathbf{x})$. The symmetry operator S is assumed to be known.
5. Update the matrix M by recovering the transformation parameters m_k according to (5.22)
6. Update ϕ using the gradient descent equation (5.25).
7. Repeat steps 3-6 until convergence.

5.5.2 Segmentation with multiple symmetrical counterparts

Consider the segmentation of an object with rotational symmetry. Let $\phi(t)$ denote its corresponding level-set function at time t and $L_{\mathcal{H}} = H(\phi(t))$ denote the respective object indicator function. If $L_{\mathcal{H}}$ is invariant to rotation by α degrees, where $\alpha \leq 2\pi/3$ then more than one symmetrical counterpart level-set functions can be used to support the segmentation. Specifically, the number of supportive symmetrical-counterpart level-set functions is $N = \lfloor 2\pi/\alpha \rfloor - 1$, since the object is invariant to rotations by $n\alpha^0$, where $n = 1 \dots N$. We denote this rotation by $R_{n\alpha}$. In that case, the symmetry constraint takes the following form:

$$E_{\text{SYM}} = \sum_{n=1}^N \int_{\Omega} \left[\mu H(\phi) \left(1 - H(\hat{\phi}_n \circ M_n) \right) + (1 - H(\phi)) H(\hat{\phi}_n \circ M_n) \right] d\mathbf{x}, \quad (5.26)$$

where, $\hat{\phi}_n = \phi(R_{n\alpha}\mathbf{x}) = \phi \circ R_{n\alpha}$ and M_n is the homography that aligns $\hat{\phi}_n$ to ϕ . The computational cost using this symmetry constraint is higher since N homographies M_n should be recovered. However, the eventual segmentation results are better, as demonstrates Fig. 5.12.

5.5.3 Segmentation with partial information

The proposed segmentation algorithm assumes that S is known. This information is essential for generating the symmetrical counterpart. We will now consider the cases in which only partial information on S is available. For example, the object view is known to have a bilateral symmetry however it is not known if the bilateral symmetry is horizontal or vertical. Alternatively, the image has a rotational symmetry but the precise rotation degree is not known.

Let $S = R_\alpha$ be a symmetry operator which is equivalent to a planar clockwise rotation by α degrees. Consider the case where the symmetrical counterpart is mistakenly generated by a planar rotation of β degrees, thus $\hat{L}_{\mathcal{H}}^\beta = L_{\mathcal{H}}(R_\beta \mathbf{x}) \equiv L_{\mathcal{H}} \circ R_\beta$. The symmetry imperfection measure that will be eventually minimized, takes the following form: $D(L_{\mathcal{H}}, \hat{L}_{\mathcal{H}}^\beta \circ \tilde{M})$. The matrix \tilde{M} aligns $\hat{L}_{\mathcal{H}}^\beta$ to $L_{\mathcal{H}}$. The next theorem relates M and \tilde{M} .

Theorem 5.5.1. *Let L be an image with rotational symmetry with respect to $S = R_\alpha$. The operator R_α is a planar clockwise rotation by α degrees. Let $L_{\mathcal{H}}$ be a perspective distortion of L by \mathcal{H} . Let $\hat{L}_{\mathcal{H}} = L_{\mathcal{H}}(R_\alpha \mathbf{x})$ be the symmetrical counterpart of $L_{\mathcal{H}}$ and let M be a planar projective homography such that $L_{\mathcal{H}}(\mathbf{x}) = \hat{L}_{\mathcal{H}}(M\mathbf{x})$. Suppose that $\hat{L}_{\mathcal{H}}^\beta = L_{\mathcal{H}}(R_\beta \mathbf{x})$ is the mistaken symmetrical counterpart of $L_{\mathcal{H}}$, where R_β is a clockwise rotation of β° , $\beta \neq \alpha$. We claim that $L_{\mathcal{H}} = \hat{L}_{\mathcal{H}}^\beta(\tilde{M}\mathbf{x})$, where $\tilde{M} = MR_{\alpha-\beta}$.*

Proof. By definition, $\hat{L}_{\mathcal{H}}(\mathbf{x}) = L_{\mathcal{H}}(R_\alpha \mathbf{x})$ and $\hat{L}_{\mathcal{H}}^\beta(\mathbf{x}) = L_{\mathcal{H}}(R_\beta \mathbf{x})$. This implies that,

$$\hat{L}_{\mathcal{H}}(\mathbf{x}) = \hat{L}_{\mathcal{H}}^\beta(R_{\alpha-\beta} \mathbf{x}) \quad (5.27)$$

Using *theorem 2* for the special case of $S = R_\alpha$ and the above definition of $\hat{L}_{\mathcal{H}}$, we get:

$$\hat{L}_{\mathcal{H}}(\mathbf{x}) = L_{\mathcal{H}}(M^{-1} \mathbf{x}). \quad (5.28)$$

From equations (5.27) and (5.28), we get: $L_{\mathcal{H}}(M^{-1} \mathbf{x}) = \hat{L}_{\mathcal{H}}^\beta(R_{\alpha-\beta} \mathbf{x})$. The latter equation implies that

$$L_{\mathcal{H}}(\mathbf{x}) = \hat{L}_{\mathcal{H}}^\beta(MR_{\alpha-\beta} \mathbf{x}) = \hat{L}_{\mathcal{H}}^\beta(\tilde{M}\mathbf{x}),$$

defining $\tilde{M} = MR_{\alpha-\beta}$.

The homography \tilde{M} is equivalent to M up to a rotation by $\alpha - \beta$ degrees. \square

Theorem 5 implies that when the symmetry operator is an unknown rotation, the symmetrical counterpart level-set function can be generated by an arbitrary rotation (e.g. $\pi/2$). In this case the recovered homography is equivalent to the homography of the *true* symmetrical counterpart up to a planar rotation. The distorting projective transformation \mathcal{H} can be recovered up to a similarity transformation. The result presented in *theorem 5* is also applicable to object with bilateral symmetry. An image (or any function defined on \mathbb{R}^2) reflected along its vertical axis can be aligned to its reflection along its horizontal axis by a rotation of $R = \pi$. Thus, $S_{LR} = R(\pi)S_{UD}$ where S_{LR} is an horizontal reflection and S_{UD} is a vertical reflection. Similarly, $S_{UD} = R(\pi)S_{LR}$. Using a mistaken symmetrical counterpart, the homography of the true symmetrical counterpart can be obtained from the recovered homography by a rotation of π degrees.

5.6 Experiments

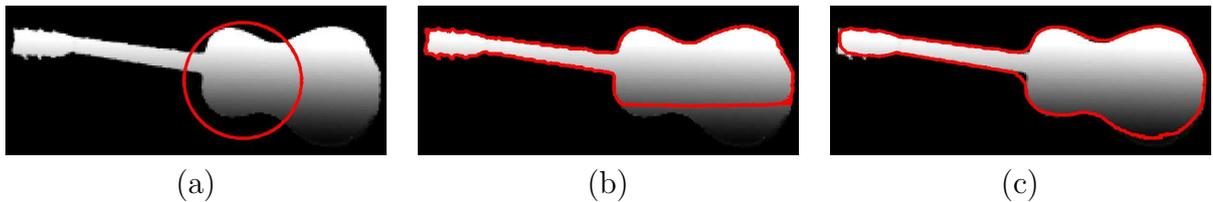


Figure 5.8: Segmentation results of a guitar. (a) Input image of a symmetrical object with the initial segmentation contour (red). (b) Segmentation (red) without the symmetry constraint. (c) Successful segmentation (red) with the proposed algorithm.

We demonstrate the proposed algorithm for the segmentation of approximately symmetrical objects in the presence of projective distortion. The images are displayed with the initial and final segmenting contours. Segmentation results are compared to those obtained using the functional (2.23) presented in chapter 2 which does not include the symmetry shape-term term. The contribution of each term in the gradient descent equation (5.25) is bounded to $[-1, 1]$ as explained in section 2.2. In Fig. 5.8

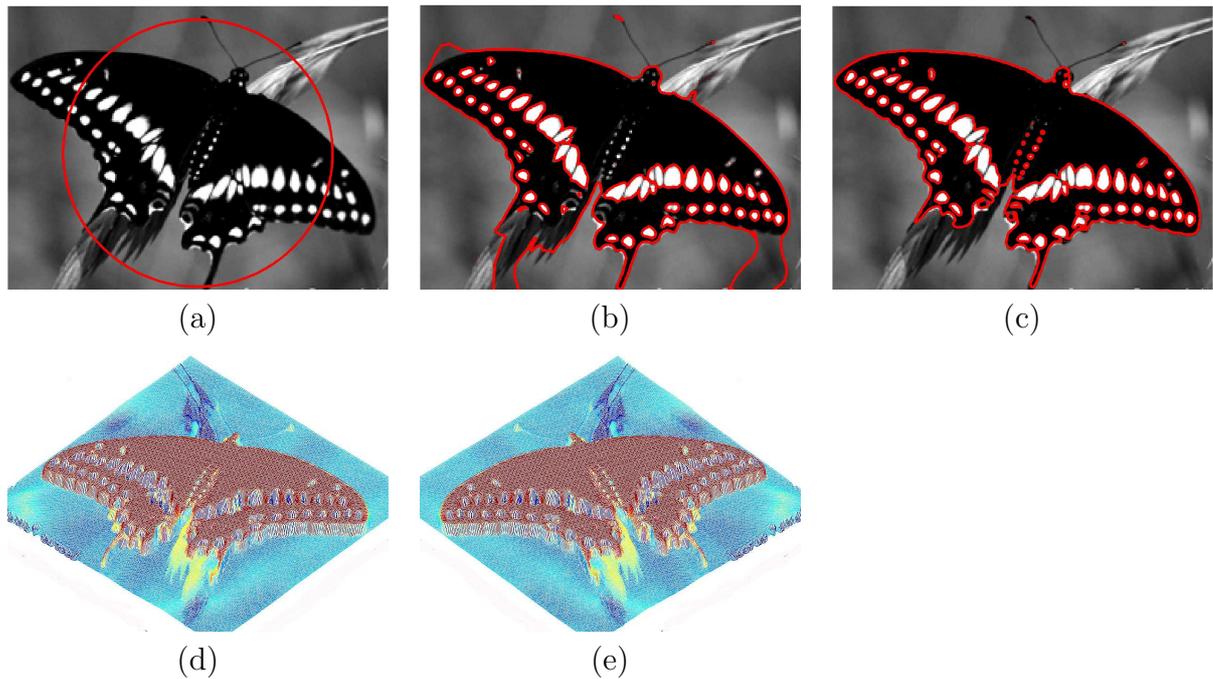


Figure 5.9: Segmentation results of a butterfly image. (a) Input image of a symmetrical object with the initial segmentation contour (red). (b) Segmentation (red) without the symmetry constraint. (c) Successful segmentation (red) with the proposed algorithm. (d) Final form of the level set function ϕ . (e) The symmetrical counterpart of the level-set function in (d). Original image courtesy of George Payne. URL: <http://cajunimages.com/images/butterfly%20wallpaper.jpg>

the upper part of the guitar is used to extract its lower part correctly. In the butterfly image, Fig. 5.9, a left-right reflection of the evolving level set function is used to support accurate segmentation of the left wing of the butterfly. The deviation between the symmetry axis of the butterfly and that of the image has been recovered concurrently with the segmentation process. In Fig. 5.10 the approximately symmetrical object (man's shadow) is extracted based on the bilateral symmetry of the object. In the swan example shown in Fig. 5.11 we used both color and symmetry cues for correct extraction of the swan and its reflection. In the last example, Fig. 5.12, segmentation of an object with approximate rotational symmetry is demonstrated. In this example we could theoretically use four symmetrical counterparts to support the segmentation. Note, however, that using only two symmetrical counterparts (instead of one) significantly improves the results. One of the symmetrical counterparts of the final level-set function is shown in 5.12f. We used the symmetry constraint for

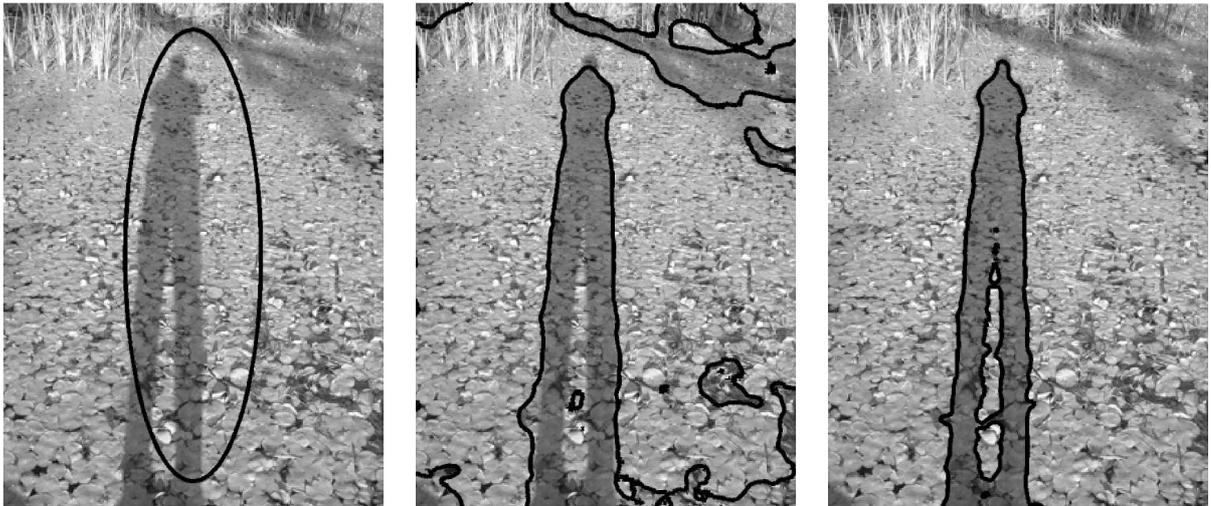


Figure 5.10: Segmentation results of a man-shadow image. (a) Input image of a symmetrical object with the initial segmentation contour (black). (b) Segmentation (black) without the symmetry constraint. (c) Successful segmentation (black) with the proposed algorithm. Original image courtesy of Amit Jayant Deshpande. URL: <http://web.mit.edu/amitd/www/pics/chicago/shadow.jpg>

multiple symmetrical counterpart level-set functions according to equation (5.26).

5.7 Concluding remarks

This chapter contains two fundamental, related contributions. The first is a variational framework for the segmentation of symmetrical objects distorted by perspective. The proposed segmentation method relies on theoretical results related to symmetry and perspective which are the essence of the second contribution.

As opposed to previous approaches to symmetry, the symmetrical object is considered as a single entity and not as a collection of landmarks or feature points. This is accomplished by using the level-set formulation and assigning (w.l.o.g.) the positive levels of the level-set function to the object domain and the negative levels to the background. An *object*, in the proposed framework, is represented by the support of its respective labeling function. As the level-set function evolves, the corresponding labeling function and thus the represented object change accordingly.

A key concept in the suggested study is the symmetrical counterpart of the evolving level-set function, obtained by either rotation or reflection of the level-set function

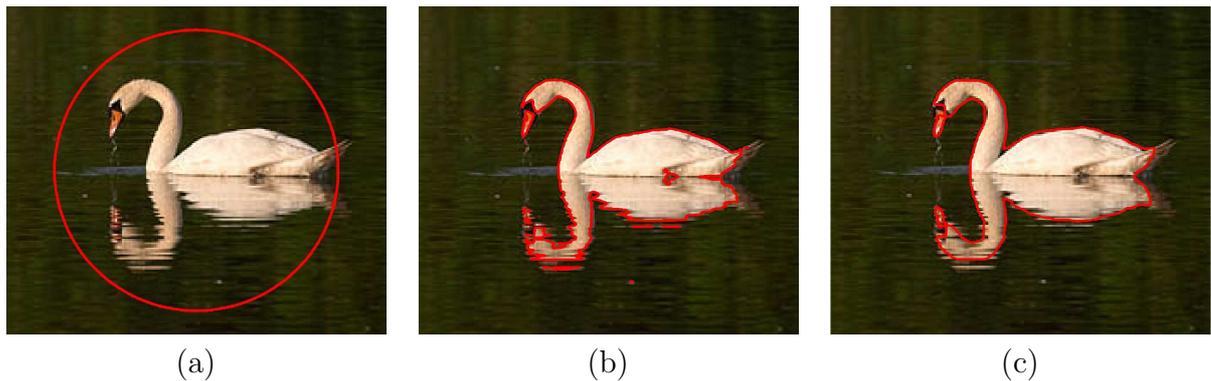


Figure 5.11: Segmentation results of a swan reflection image. (a) Input image of a symmetrical object with the initial segmentation contour (red). (b) Segmentation (red) without the symmetry constraint. (c) Successful segmentation (red) with the proposed algorithm. Original image courtesy of Richard Lindley. URL: <http://www.richardlindley.co.uk/links.htm>

domain. We assume that the object to segment underwent a planar projective transformation, thus the source level-set function and its symmetrical counterpart are not identical. We show that these two level-set functions are related by planar projective homography.

Due to noise, occlusion, shadowing or assimilation with the background, the propagating object contour is only approximately symmetrical. We define a symmetry imperfection measure which is actually the ‘distance’ between the evolving level-set function and its symmetrical counterpart aligned to it. A symmetry constraint based on the symmetry imperfection measure is incorporated within the level-set based functional for segmentation. The homography matrix that aligns the symmetrical counterpart level-set functions is recovered concurrently with the segmentation.

We show in *theorem 5.3.5* that the recovered homography is determined by the projective transformation that distorts the image symmetry. This implies that the homography obtained by the alignment process of the symmetrical counterpart functions (or images) can be used for partial recovery of the 3D structure of the imaged symmetrical object. As implied from *theorem 5.3.9*, full recovery is not possible if the projection of the 3D object on the image plane preserves its symmetry.

The algorithm suggested is demonstrated on various images of objects with either bilateral or rotational symmetry, distorted by projectivity. Promising segmentation

results are shown. In this manuscript only shape symmetry is considered. The proposed framework can be extended considering symmetry in terms of gray levels or color. The notion of symmetrical counterpart can be thus applied to the image itself and not to binary (or level-set) functions. Possible applications are de-noising, super-resolution from a single symmetrical image and inpainting. All are subjects for future research.

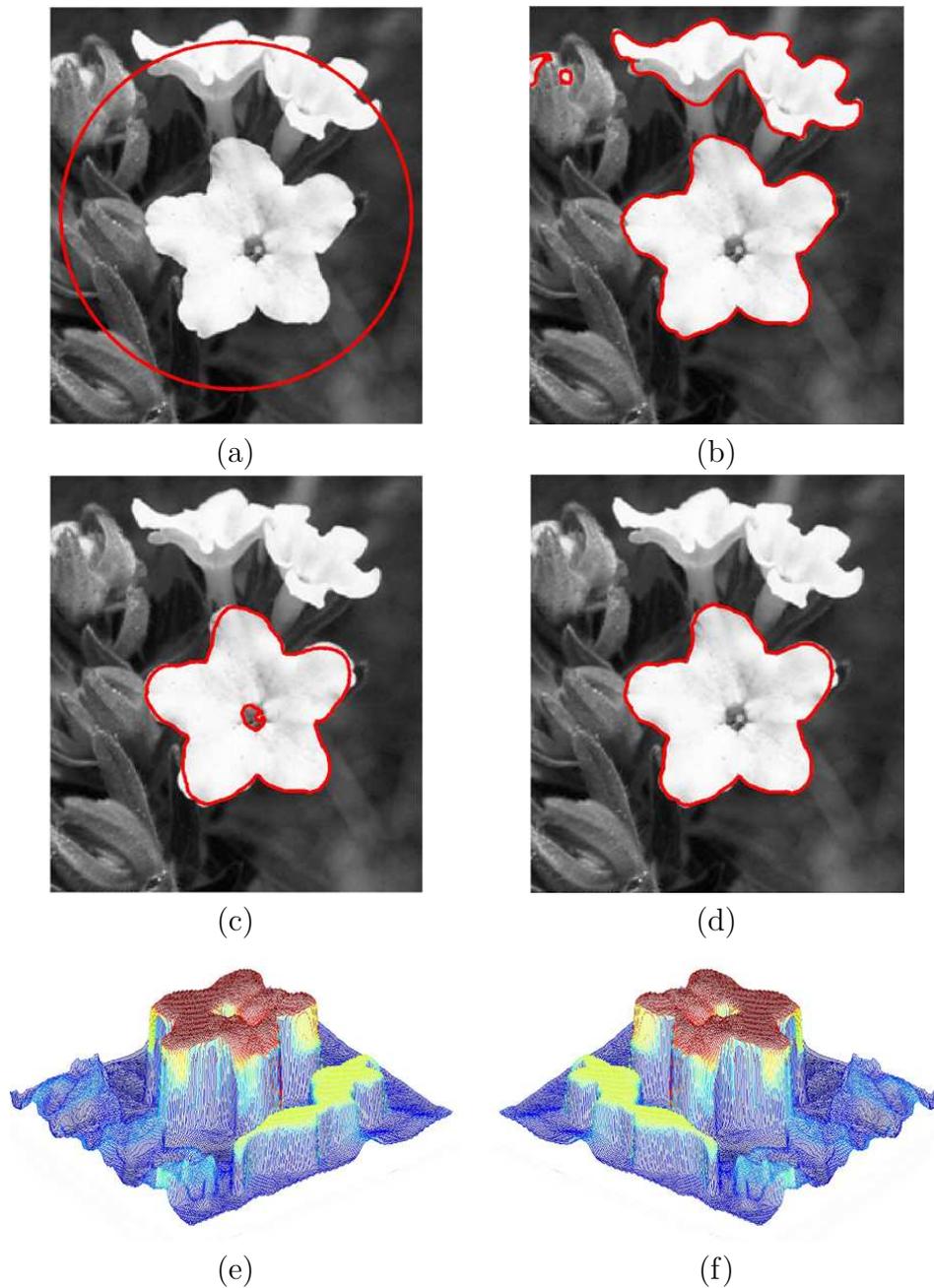


Figure 5.12: (a) Input image of an object (flower) depicted with rotational symmetry. The initial segmentation contour (red) is drawn. (b) Segmentation (red) without the symmetry constraint. (c) Segmentation (red) with the proposed algorithm using a single symmetrical counterpart. (d) Segmentation (red) with the proposed algorithm using two symmetrical counterparts. (e) Final level-set function. (f) One of the symmetrical counterparts of the level-set function (e), obtained by rotation of 90° . Original image courtesy of Kenneth R. Robertson. URL: <http://inhs.uiuc.edu>

Chapter 6

Segmentation of Mouse Brain Atlas

6.1 Propagating distributions for the segmentation of mouse brain atlas

Automatic extraction of volumes of interest (VOIs) within experimental dataset of brain images can be facilitated by the availability of an atlas. The atlas consists of delineated and tagged anatomical structures of reference brain images. Usually, due to the high accuracy required from an atlas, segmentation is carried out manually by an expert. This manual segmentation is time consuming and laborious, therefore, automatic method for atlas segmentation is sought.

The Mouse Brain Library (MBL) is a neuroinformatic resource of brain images of inbred strain of mice [93]. The database contains a large collection of sparse coronal or horizontal brain sections. It is used for quantitative traits analysis - a genetic mapping approach, which is based on a comparison of the morphometric and genetic variations between strains. An automatic segmentation of the region of interest (ROI) within each slice can significantly facilitate the process of 3-D segmentation which plays a key role in traits analysis.

In this chapter, we address the problem of volumetric atlas segmentation based on a high resolution dataset in the presence of tight memory constraints. Segmentation is therefore carried out slice-by-slice where the successful segmentation of one section provides the initial conditions for the next one. Slices are sequentially partitioned into two inhomogeneous regions - the ROI (sliced anatomical structure of interest) and the

background. The intensities and spatial locations of the regions are modeled by two distinct three dimensional Gaussian mixtures. The Gaussian Mixture Model (GMM) parameters adaptively propagate across slices, possibly in a bidirectional manner.

The foundation of the proposed method is the construction of a unified edge-based region-based segmentation functional. A key contribution is a unique region-based energy term that considers the Kullback-Leibler (K-L) divergence [55] between the probability density functions (PDFs) of the partitions. Since histological data is less noisy than the images obtained by other modalities, the magnitudes and directions of the image gradients can be used to enhance the segmentation.

We use the level set framework [75] for segmentation, where the segmenting curve is the zero level of a level set function that evolves subject to some predefined constraints. Being parameterization-free, the level-set model allows automatic changes in the active contour topology. This feature has particular importance when the volume traced along the slices splits or merges.

Methods that use GMM for segmentation have been presented before, e.g. [99]. The Kullback-Leibler (K-L) divergence has been used to as a measure of pseudo-distance between GMM distributions in [33]. The works of [128, 78] introduce a variational framework for image segmentation using the PDFs of the partitions. Slice-by-slice approaches for volumetric segmentation that use level-sets include [76, 77, 2]. Aykac et al [2] suggest an active contour algorithm for the segmentation of mouse spleen in micro-CT. However, their method, based on the piecewise constant image model of Chan and Vese is not suitable for segmentation of regions with heterogeneous intensities. Paragios [76, 77] proposed a user guided model based on prior shape-knowledge for cardiac image analysis. Such shape-based approaches depend on the availability of reference shapes and credible alignment. Yet, since transformations between corresponding or consecutive slices are non-parametric they cannot be modeled reliably. Therefore, shape constraints, based on prior knowledge, are likely to deflect the contour evolution from the actual image data.

The proposed segmentation framework is capable of handling challenging segmentation tasks. This is accomplished by modeling the distribution functions of the image data and by imposing compatibility of the partitions to the model. The alignment of the ROI boundary with image gradients is constrained as well. The algorithm is

demonstrated on images of histological atlas of mouse brain [36]. Results compare well with manual segmentation.

6.2 Algorithm description

We first present the segmentation of a single slice by minimizing a unified region-based and edge-based energy terms and associated gradient descent equations are presented. The slice-by-slice segmentation framework is then described.

6.2.1 Level-set representation of the region of interest

In this section we describe a level-set framework for image partitioning into two disjoint regions. Let $I: \Omega \rightarrow \mathbb{R}^+$ denote a gray level image, where $\Omega \subset \mathbb{R}^2$ is the image domain. Let ω^+ and ω^- be open subsets of Ω , denoting the region of interest (ROI) and the background, respectively. The optimal boundary $C \in \Omega$, $C = \partial\omega^+$ is obtained by minimizing a cost functional $E(C|I)$. The curve C is implicitly represented as the zero level of a level set function. Recall that for any arbitrary function $f: \Omega \rightarrow \mathbb{R}$ defined on domain Ω the following holds:

$$\int_{\mathbf{x} \in \omega^\pm} f(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{x} \in \Omega} f(\mathbf{x}) H(\pm\phi(\mathbf{x})) d\mathbf{x}.$$

Again, H denotes the Heaviside function defined in equation 2.3.

6.2.2 The K-L divergence term

Region based segmentation schemes are established on piecewise smoothness approximations of images [70]. Usually the homogeneity of semantic regions within the image is in terms of gray levels [11]. Here, we propose a segmentation framework that extends the concept of region homogeneity to include probability density functions of the pixels feature vectors. This allows to extract ROIs with heterogeneous intensities. The fundamentals of our region-based energy term partially rely on the works of [128, 11, 78]. Let $G(\mathbf{x}) = (x, y, I(x, y))$ be a vector representation of the pixel $(x, y) \in \Omega$. We assume that $G(\mathbf{x})$ are drawn from two distinct probability density

functions, p^+ if $\mathbf{x} \in \omega^+$ and p^- otherwise. Specifically, p^+ and p^- are Gaussian mixture models of $(\omega^+, I(\omega^+))$ and $(\omega^-, I(\omega^-))$, respectively. Assuming an independent and identically distributed conditional PDF, we get:

$$P(G(\mathbf{x})|\phi) = \prod_{\mathbf{x} \in \omega^+} p^+(\mathbf{x}) \prod_{\mathbf{x} \in \omega^-} p^-(\mathbf{x}), \quad (6.1)$$

hence,

$$\log P(G(\mathbf{x})|\phi) = \int_{\mathbf{x} \in \omega^+} \log p^+(\mathbf{x}) d\mathbf{x} + \int_{\mathbf{x} \in \omega^-} \log p^-(\mathbf{x}) d\mathbf{x}. \quad (6.2)$$

In Paragios and Rousson [78] Eq. 6.2 is maximized. In contrast, we maximize the *K-L divergences* between p^+ and p^- within the partitions to obtain homogeneity in regions in terms of their PDFs. The resulting region-based energy term, to be minimized, takes the form:

$$E_{KL} = -[D_{KL}(p^+ \| p^-) + D_{KL}(p^- \| p^+)] d\mathbf{x}, \quad (6.3)$$

where,

$$D_{KL}(p^+ \| p^-) = \int_{\mathbf{x} \in \omega^+} p^+(\mathbf{x}) \log \frac{p^+(\mathbf{x})}{p^-(\mathbf{x})} d\mathbf{x} \quad (6.4)$$

$$D_{KL}(p^- \| p^+) = \int_{\mathbf{x} \in \omega^-} p^-(\mathbf{x}) \log \frac{p^-(\mathbf{x})}{p^+(\mathbf{x})} d\mathbf{x}. \quad (6.5)$$

The optimal partitioning is inferred by maximizing the asymmetric divergences $D_{KL}(p^+ \| p^-)$ and $D_{KL}(p^- \| p^+)$ within ω^+ and ω^- , respectively. In the proposed framework, the PDFs p^+ and p^- are multivariate Gaussian densities

$$p^+(\mathbf{x}|\theta^+) = \sum_{i=1}^{N^+} w_i^+ \frac{\exp\{-\frac{1}{2}(\mathbf{x} - \mu_i^+)^T \Sigma_i^{+^{-1}} (\mathbf{x} - \mu_i^+)\}}{\sqrt{(2\pi)^d |\Sigma_i^+|}} \quad (6.6)$$

$$p^-(\mathbf{x}|\theta^-) = \sum_{i=1}^{N^-} w_i^- \frac{\exp\{-\frac{1}{2}(\mathbf{x} - \mu_i^-)^T \Sigma_i^{-^{-1}} (\mathbf{x} - \mu_i^-)\}}{\sqrt{(2\pi)^d |\Sigma_i^-|}} \quad (6.7)$$

where, $\theta^\pm = \{w_i^\pm, \mu_i^\pm, \Sigma_i^\pm\}_{j=1}^{N^\pm}$ are the parameters of the Gaussian Mixture Model and $d = 3$.

The gradient descent equation associated with E_{KL} is:

$$\phi_t^{KL} = \delta(\phi)[D_{KL}(p^+ \| p^-) - D_{KL}(p^- \| p^+)], \quad (6.8)$$

where δ is the derivative of the Heaviside function H . The parameters θ^+ and θ^- are first estimated given the segmentation of the predecessor slice and then updated at every iteration along the evolution of the delineating contour. The GMMs p^+ and p^- are alternately updated with the evolution of ϕ . We use the EM technique [24] based on a maximum likelihood principle (ML) to estimate GMM parameters.

6.2.3 Unified cost functional

Histological images are usually less noisy than those obtained by other modalities. We can therefore construct a cost functional which is based on the image gradients, the smoothness constraint and the estimated parameters based on the adjacent slices:

$$E(\phi) = W^{\text{KL}} E_{\text{KL}}(\phi) + W^{\text{GAC}} E_{\text{GAC}}(\phi) + W^{\text{LEN}} E_{\text{LEN}}(\phi) + W^{\text{EA}} E_{\text{EA}}(\phi), \quad (6.9)$$

according to equations (6.3), (2.21), (2.19), and (2.17) for E_{RB} , E_{GAC} and E_{EA} respectively ¹. The evolution of the level-set functions ϕ at each iteration is determined by $\phi(t + \Delta t) = \phi(t) + \phi_t$. The associated gradient descent equations ϕ_t are derived using the first variation of the functional (6.9) above,

$$\phi_t(\phi) = W^{\text{KL}}(t) \phi_t^{\text{KL}} + W^{\text{GAC}}(t) \phi_t^{\text{GAC}} + W^{\text{LEN}}(t) \phi_t^{\text{LEN}} + W^{\text{EA}}(t) \phi_t^{\text{EA}}, \quad (6.10)$$

The weight terms $W^{\text{TERM}}(t, \mathbf{x})$ in (6.9) adaptively changing scalars. Their determination is explained in section 2.2.

6.2.4 Slice by slice segmentation

The segmentation algorithm proposed in the previous section is applied to images of consecutive brain slices. Given a sequence of images $\{\dots, I_{s-1}, I_s, I_{s+1}, \dots\}$, the final estimates of either $(\phi_{s-1}, \theta_{s-1})$ or $(\phi_{s+1}, \theta_{s+1})$ are used to initialize (ϕ_s, θ_s) . The level-set function ϕ_s and the GMM parameters θ_s are updated alternately. We assume that the segmentation of few sparsely sampled slices is done manually. Once the segmentation is known the PDF parameters can be computed accordingly. This

¹One may set the weights of the edge-based terms W^{GAC} and W^{EA} to zero when the data is noisy and the image gradients are not reliable.

information then propagates in both directions. This mostly automatic scheme facilitates significantly the process of segmentation, while keeping it away from undesired drifts. In the post-processing phase, the delineating contours of the image sequence are stacked together to generate a model of the VOI.

6.3 Experiments

The images utilized in this study are sampled coronal views of a 3D brain atlas of a C57BL/6J mouse. To create the atlas, consecutive 17.9 μm thick horizontal sections were collected from a freshly frozen adult male brain using low distortion cyrosectioning system [72] while simultaneously imaging the block face. The sections were stained for acetylcholinesterase and imaged to yield a pixel pitch of $8\mu\text{m}$. The volume reconstruction was guided by the block face images. The alignment was refined by section-to-section registration with rigid-body AIR [120, 121, 93]. Finally, the volume was downsampled to yield an isotropic $17.9\mu\text{m}$ /voxel atlas.

Figure 6.1 exemplifies segmentation of the last slice in a sequence of 10 consecutive slices. Figure 6.1a-b shows the reference slice without and with the manual segmentation, respectively. Figure 6.1c is the slice to segment.demonstrates the gradual change of the ROI by displaying the image of slice #509 together with the ROI contour (red) of slice#500. The slice-by-slice segmentation algorithm and the expert segmentation is shown are Fig. 6.1e-f, respectively. To allow qualitative comparison between the manual (blue) and the automatic segmentation using the slice-by-slice method (red) we present a superposition of the respective contours, shown in Fig. 6.1g. Figure 6.1h, obtained using a region-based term that is established on a piecewise constant image model [11] as used in [2]. A quantitative comparison between the manual segmentation and the segmentation obtained by the proposed algorithm is shown in Table 6.1.

Figure 6.2 demonstrates successful segmentation when two disjoint ROIs that belong to the same structure merge. Figure 6.2a-b shows the reference slice without and with the manual segmentation, respectively. Figure 6.2c is the slice to segment. Figure 6.2d shows the first segmentation step where the reference slice boundary initiates the segmentation process. A comparison between the slice-by-slice segmentation algorithm and the expert segmentation is shown in Figs. 6.2e-f, respectively

Figure 6.3 demonstrates successful segmentation when the ROI splits into two regions. Figure 6.3a-b shows the reference slice without and with the manual segmentation, respectively. Figure 6.3c is the slice to segment. Figure 6.3d shows the first segmentation step where the reference slice boundary initiates the segmentation process. A comparison between the slice-by-slice segmentation algorithm and the expert segmentation is shown in Figs. 6.3e-f, respectively.

Slice#	501	502	503	504	505	506	507	508	509
Error percentage	.87	.93	1.03	1.31	1.41	1.48	1.73	1.93	2.14

Table 6.1: Correspondence between the slice-by-slice automatic segmentation and the expert segmentation. The table presents the percentage of pixels with dissimilar labeling with respect to the total number of pixels for slices 501 – 509

6.4 Concluding remarks

This chapter addressed the segmentation of structures in the mouse brain from histological data, using different, yet natural source of priors. Namely, the statistical information on the shape and the intensity distributions of consecutive slices was employed. The key contribution is the incorporation of the K-L divergence in a fidelity term that compares the PDFs of adjacent slices. This modification suggest a novel way to handle meaningful image regions with heterogeneous appearance. The region of interest (ROI) in each slice was extracted by minimizing a cost functional that imposed the alignment of the ROI's boundary with the image gradients and consistency with nearby slices. It is a step towards the use of a comprehensive prior that consists of shape, color and textural information.

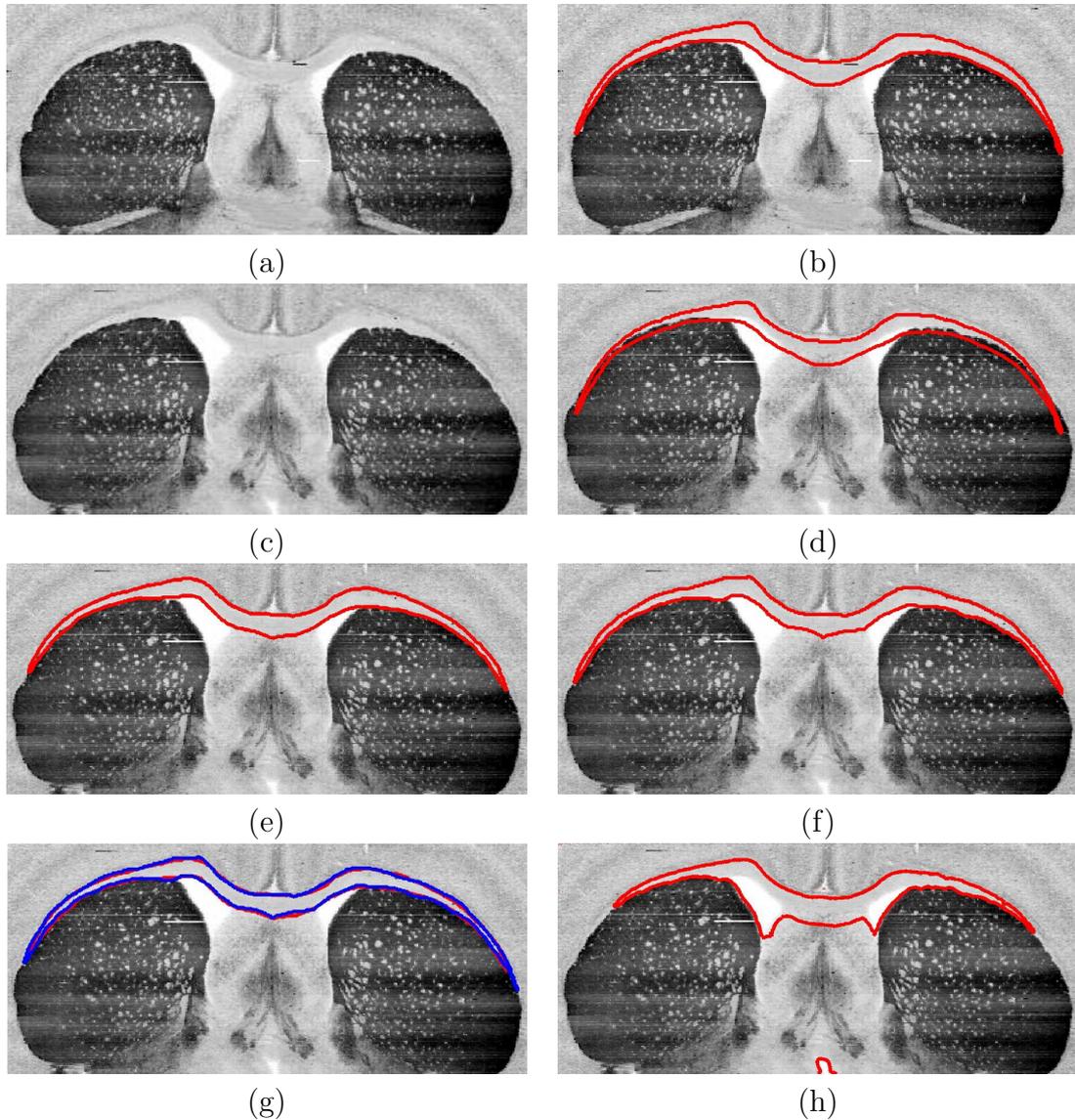


Figure 6.1: Segmentation of the last slice in a 10 slice sequence. (a) Slice #500- the first slice in the sequence;(b) Manual segmentation of slice #500;(c) Slice #509- the last slice in the sequence (d) The contour (red) of slice #500 is shown on slice #509 - to demonstrate the mismatch between the ROIs boundaries. (e) Segmentation of slice #509 using the proposed algorithm; (f) Manual segmentation of slice #509 (g) A comparison between the contour obtained by the slice-by-slice algorithm (red) to the contour delineated by an expert (blue) (h) Segmentation (undesired) of slice #509 with the Chan-Vese algorithm.

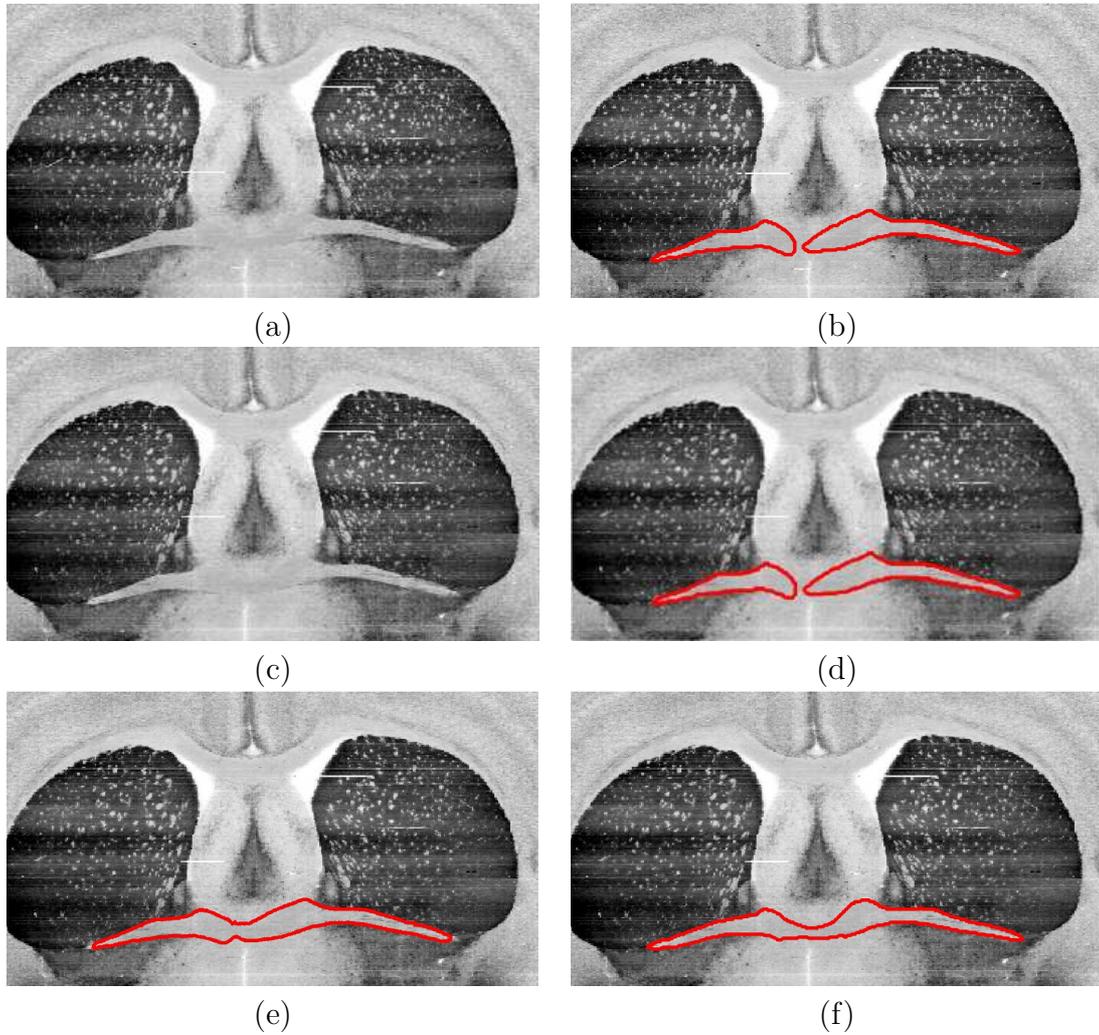


Figure 6.2: Slice-by-slice segmentation captures the merge of two regions of interest. (a) Slice #320- the first slice in the sequence; (b) Manual segmentation of slice #320; (c) Slice #321- the next slice in the sequence (d) The contour (red) of slice #320 is shown on slice #321 - to demonstrate the topological changes between the respective ROIs boundaries. (e) Segmentation of slice #321 using the proposed algorithm; (f) Manual segmentation of slice #321

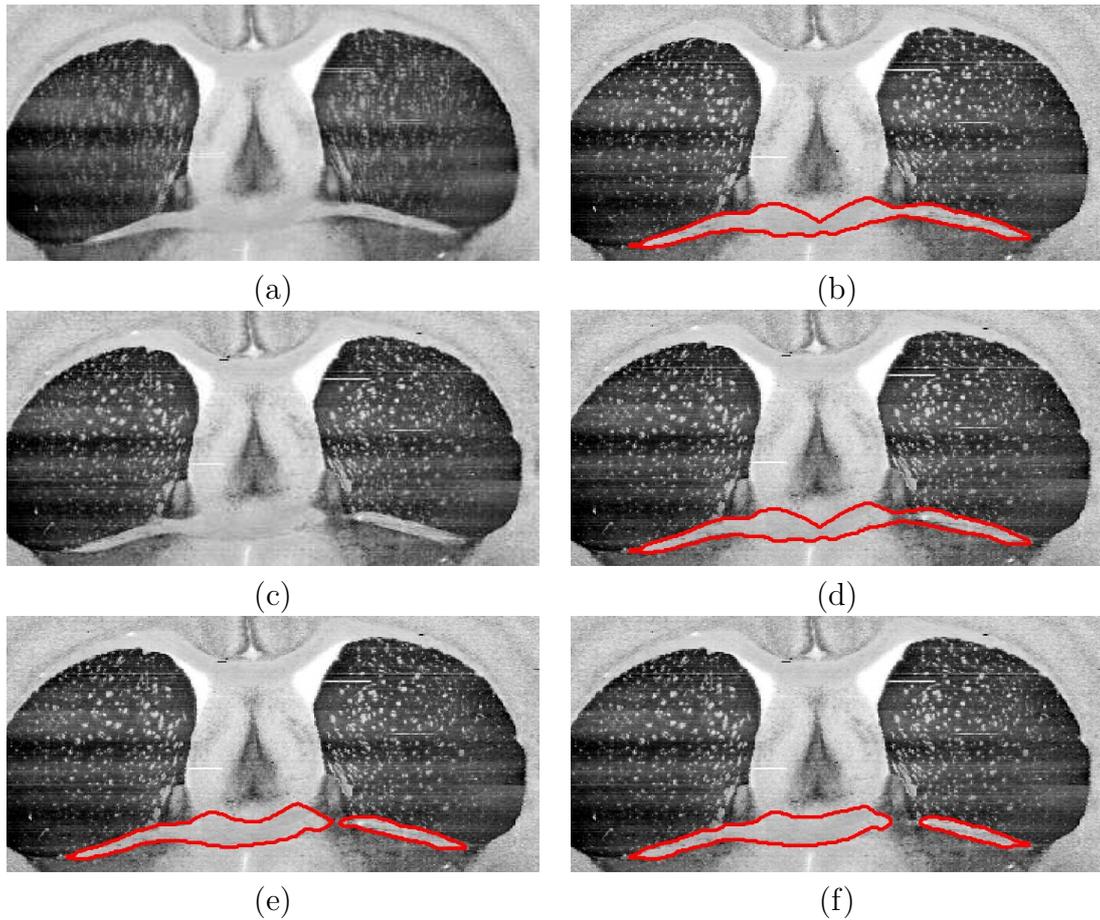


Figure 6.3: Slice-by-slice segmentation captures the split into two regions. (a) Slice #322; (b) Manual segmentation of slice #322; (c) Slice #324 (d) The contour (red) of slice #322 is shown on slice #324 - to demonstrate the topological changes between the respective ROIs boundaries. (e) Segmentation of slice #324 using the proposed algorithm; (f) Manual segmentation of slice #324

Chapter 7

Discussion and Future Directions

Segmentation has a key role in the fascinating field of image understanding. This thesis addresses the image partitioning problem using a unified framework, incorporating low-level image features (such as intensities and edges) together with high-level visual knowledge. The dominant player in this top-down flow of information is the shape of the regions of interest. We also present an algorithm that uses expected intensity distribution.

We use a variational formulation which allows a convenient way to integrate information from several sources. This is accomplished by the construction of a cost functional that is composed of bottom-up image data terms and top-down flow of external information. The functional is minimized to obtain the object contour represented by the zero level of a level-set function.

Shape priors are obtained from other instances of the same or similar object. When the object is taken from different, unknown viewpoints - the alignment of the object views should be addressed. The main difficulty resides in a “chicken and egg” problem, where the boundary of one of the object views is registered towards the boundary of the other while being evolved. Previous works resolved this difficulty assuming a similarity transformation between the views. Similarity transformations, however, rarely approximate satisfactorily the actual transformation between objects taken close to the camera. Usually a bank of shapes and its statistics were used to account for transformations beyond similarity.

One of the main novelties of this thesis is the inclusion of the two view geom-

etry model within a variational framework for segmentation to account for planar projective transformation. This was enabled using a unique shape representation together with a novel dissimilarity measure between the object to segment and the prior shape. The variational registration process interleaved with the segmentation, is an effective alternative to the classical computer vision problem of point correspondence. The entire object contours are matched rather than a collection of feature points or landmarks. The process of shape alignment is carried out under the assumption that the object boundaries are approximately coplanar. This restriction can be removed, extending the geometric model by increasing the number of the transformation parameters.

Consider now the more complicated case where neither of the two object instances can be segmented by itself. In a process we call *mutual segmentation*, the evolving object contour in each image provides a dynamic prior for the segmentation of the other object view. Again we assume that the two views are related by planar projective transformation. The homography between the object views is recovered concurrently with the mutual segmentation. The mutual segmentation scheme provides a non-standard tool to distinguish between the foreground image regions and the background. The objects are regions with similar appearance in both images. In this dissertation, similarity is defined in terms of shape. Future study may address mutual segmentation of objects with similar intensity distributions or texture.

Shape symmetry is an important cue for image understanding. When the object taken is known to be symmetrical, the symmetry property forms a significant shape constraint and thus can be used to facilitate segmentation. We address the segmentation of a single instance of a symmetrical object. We define the concept of a *symmetrical counterpart* image of an object with known type of symmetry, as a flip or rotation of the source image. The symmetrical counterpart image is considered as another view of the same object, providing an implicit shape prior. The suggested formulation is based on a theoretical result which shows that symmetrical counterpart images are related by planar projective homography. The limits on the ability to recover this homography are stated as well. These theorems are part of a set of theoretical results that link between symmetry, two view geometry and level-set representation. The implications of the theoretical foundation of the symmetry-based

method go far beyond segmentation. One of the interesting problems, related to symmetry, is 3D reconstruction of symmetrical objects. Exploring this subject with the proposed theoretical approach is a promising research direction.

The main body of the thesis is devoted to shape based object segmentation, using either an explicit and well-defined shape prior, a dynamic prior (mutual segmentation) or a shape symmetry constraint. The need to account for non-parametric shape deformations is crucial in bio-medical applications. Using shape statistics from a comprehensive shape database as in [19, 58] is one of the possible solutions. Nevertheless, in the presence of significant morphological changes, other sources of prior information such as texture or intensity distribution should be used.

We present a slice-by-slice framework for the volumetric segmentation of mouse brain structures, obtained from histological data. The region of interest deforms slowly across the slices but the deformations may be arbitrary. Moreover, topological changes of the region boundaries, such as splits and merges, frequently occur. We found that the spatial-intensity distribution of the neighboring slices is the most suitable prior. Based on this observation we incorporate the Kullback-Leibler divergence in a fidelity term that compares the PDFs of adjacent slices. This modification suggest a novel way to handle meaningful image regions with heterogeneous appearance.

A promising extension of this thesis involves feedback provided by the user to guide the segmentation process. This direction is mostly suitable for bio-medical applications. It will allow the physician to interactively transfer his/her knowledge to the segmentation algorithm via partial region labeling or control points. User interactive frameworks became very common recently [94]. User-interactive variational method was suggested by Liang, McInerney and Terzopoulos [60]. Yet, a coherent level-set framework that mediates online between the evolving segmentation based on the inherent image properties and an external information flow is yet to be explored.

This research suggests promising methods to solve hard segmentation problems. It also paves the way to advanced prior-based segmentation algorithms that will make use of banks of shapes, statistics on a variety of characteristic features and more elaborate geometrical models.

Bibliography

- [1] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, 2002.
- [2] D. Aykac, J. R. Price, and J.S. Wall. 3d segmentation of mouse spleen in micro-ct via active contours. *IEEE Nuclear Science Symposium Conference Record*, 3:1542–1545, Oct. 2005.
- [3] N. Ben-Zadok, T. Riklin-Raviv, and N. Kiryati. Automatic segmentation of uterine fibroids for MR-guided focused ultrasound treatment planning. Technical report, Tel Aviv university, 2007.
- [4] T.O. Binford. Visual perception by computer. In *Proc. IEEE Conf. Systems and Control*, December 1971.
- [5] E. Borenstein and S. Ullman. Class-specific, top-down segmentation. In *Proceedings of the European Conference on Computer Vision*, page II: 109 ff., 2002.
- [6] T. Brox, B. Rosenhahn, and J. Weickert. Three dimensional shape knowledge for joint image segmentation and pose estimation. In *DAGM-Symposium*, pages 109–116, 2005.
- [7] A.M. Bruckstein and D. Shaked. Skew-symmetry detection via invariant signatures. *Pattern Recognition*, 31(2):181–192, 1998.
- [8] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.
- [9] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik*, 66(1):1–31, 1993.

- [10] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, Feb. 1997.
- [11] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, Feb. 2001.
- [12] Y. Chen, H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggs, and E.A. Geiser. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3):315–328, Dec. 2002.
- [13] L.D. Cohen and I. Cohen. Finite-element methods for active contour models and balloon for 2-D and 3-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131–1147, 1993.
- [14] T.F. Coleman. Optimization toolbox: Quasi newton methods. The Mathworks, Inc. <http://www.mathworks.com/access/helpdesk/help/toolbox/optim/ug>.
- [15] Cone, in mathematics. The Columbia Encyclopedia, 6th Edition. Columbia University Press, 2001.
- [16] G.J. Cootes, T.F. and Edwards and C.J. Taylor. Active appearance models. In *Proceedings of the European Conference on Computer Vision*, volume 2, pages 484–498, 1998.
- [17] 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, January 1995.
- [18] H. Cornelius and G. Loy. Detecting rotational symmetry under affine projection. In *Proceedings of the International Conference on Pattern Recognition*, pages 292–295, 2006.
- [19] D. Cremers, T. Kohlberger, and C. Schnorr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, Sep. 2003.

- [20] D. Cremers, M. Rousson, and R. Deriche. Review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, April 2007.
- [21] D. Cremers and S. Soatto. A pseudo-distance for shape priors in level set segmentation. In *Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, pages 169–176, 2003.
- [22] D. Cremers, N. Sochen, and C. Schnorr. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In *Intl. Conf. on Scale-Space Theories in Computer Vision*, pages 388–400, June 2003.
- [23] D. Cremers, N. Sochen, and C. Schnörr. Multiphase dynamic labeling for variational recognition-driven image segmentation. *International Journal of Computer Vision*, 66(1):67–81, January 2006.
- [24] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [25] A. Dervieux and F. Thomasset. A finite element method for the simulation of rayleigh-taylor instability. *Lecture Notes in Mathematics*, 771:145–158, 1979.
- [26] A. Duci, A. Yezzi, S. Mitter, and S. Soatto. Region matching with missing parts. In *Proceedings of the European Conference on Computer Vision*, volume 3, pages 48–64, May 2002.
- [27] R.D. Edwards, J.G. Moss, M.A. Lumsden, O. Wu, L.S. Murray, S. Twaddle, and G.D. Murray. Uterine-artery embolization versus surgery for symptomatic uterine fibroids. *The New England Journal of Medicine*, 356(4):360–370, January 2007.
- [28] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.
- [29] O. Faugeras, Q.T. Luong, and T. Papadopoulos. *The Geometry of Multiple Images*. MIT Press, 2001.

- [30] R. Fawcett, A. Zisserman, and J.M. Brady. Extracting structure from an affine view of a 3d point set with one or 2 bilateral symmetries. *Image and Vision Computing*, 12(9):615–622, November 1994.
- [31] D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
- [32] S. Geman and D. Geman. Stochastic relaxation, gibbs distribution and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.
- [33] J. Goldberger, S. Gordon, and H. Greenspan. An efficient image similarity measure based on approximation of kl-divergence between two gaussian mixtures. In *Proceedings of the International Conference on Computer Vision*, pages 487–493, 2003.
- [34] J. Gomes and O.D. Faugeras. Level sets and distance functions. In *Proceedings of the European Conference on Computer Vision*, pages 588–602, 2000.
- [35] A. Gupta, V.S.N. Prasad, and L.S. Davis. Extracting regions of symmetry. In *ICIP*, pages III: 133–136, 2005.
- [36] C. Gustafson, O. Tretiak, L. Bertrand, and Y. Nissanov. Design and implementation of software for assembly and browsing of 3d brain atlases. *Computer Methods and Programs in Biomedicine*, 74:53–61, 2004.
- [37] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision*, volume II. Addison-Wesley, 1993. pp. 593-594.
- [38] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2003.
- [39] S. Hochstein and M. Ahissar. View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36:791–804, December 2002.
- [40] M. Holtzman-Gazit, R. Kimmel, N. Peled, and D. Goldsher. Segmentation of thin structures on volumetric medical images. *IEEE Transactions on Image Processing*, 15(2):354–363, 2006.

- [41] W. Hong, Y. Yang, and Y. Ma. On symmetry and multiple view geometry: Structure, pose and calibration from single image. *International Journal of Computer Vision*, 60:241–265, 2004.
- [42] X. Huang, Z. Li, and D. Metaxas. Learning coupled prior-shape and appearance models for segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI04*, volume I, pages 60–69, 2004.
- [43] M. Irani and P. Anandan. All about direct methods. In W. Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory and Practice*. Springer-Verlag, 1999.
- [44] H. Ishikawa, D. Geiger, and R. Cole. Finding tree structures by grouping symmetries. In *Proceedings of the International Conference on Computer Vision*, pages 1132–1139, 2005.
- [45] A.K. Jain, Y. Zhong, and M.P. Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, 71(2):109–129, December 1998.
- [46] R. Jain, R. Kasturi, and B.G. Schunck. *Machine Vision*. McGraw-Hill, 1995. pp. 482-483.
- [47] F. Jurie and M. Dhome. Real time 3d template matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 791–797, 2001.
- [48] K. Kanatani. Symmetry as a continuous feature: Comment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(3):246–247, March 1997.
- [49] M. Kass, A.P. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, Jan. 1988.
- [50] Y. Keller and Y. Shkolnisky. An algebraic approach to symmetry detection. In *Proceedings of the International Conference on Pattern Recognition*, pages III: 186–189, 2004.

- [51] S. Kichenassamy, A. Kumar, P.J. Olver, A. Tannenbaum, and A.J. Yezzi. Gradient flows and geometric active contour models. In *Proceedings of the International Conference on Computer Vision*, pages 810–815, 1995.
- [52] R. Kimmel. Fast edge integration. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging Vision and Graphics*. Springer-Verlag, 2003.
- [53] R. Kimmel and A.M. Bruckstein. Regularized laplacian zero crossings as optimal edge integrators. *International Journal of Computer Vision*, 53(3):225–243, 2003.
- [54] N. Kiryati and Y. Gofman. Detecting symmetry in grey level images: The global optimization approach. *International Journal of Computer Vision*, 29(1):29–45, 1998.
- [55] S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.
- [56] A. Laird and J. Miller. Hierarchical symmetry segmentation. In *Proc. SPIE, Intelligent Robots and Computer Vision IX: Algorithms and Techniques*, volume 1381, pages 536–544, Feb. 1991.
- [57] M. Leventon, O. Faugeraus, W.E.L. Grimson, and W. Wells. Level set based segmentation with intensity and curvature priors. In *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 4–11, June 2000.
- [58] M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume I, pages 316–323, 2000.
- [59] M.D. Levine. *Vision in Man and Machine*. McGraw-Hill, 1985. pp. 46-53.
- [60] J. Liang, T. McInerney, and D. Terzopoulos. United snakes. *Medical Image Analysis*, 10(2):215–233, April 2006.

- [61] T.L. Liu, D. Geiger, and A.L. Yuille. Segmenting by seeking the symmetry axis. In *Proceedings of the International Conference on Pattern Recognition*, pages 994–998, 1998.
- [62] L.M. Lorigo, O. Faugeras, Grimson W.E.L., R. Keriven, R. Kikinis, A. Nabavi, and C.F. Westin. Codimension two-geodesic active contours for the segmentation of tabular structures. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 444–451, 2000.
- [63] Y. Ma, S. Soatto, J. Košecká, and S. S. Sastry. *An Invitation to 3-D Vision*. Springer-Verlag, 2003.
- [64] 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):158–175, 1995.
- [65] G. Marola. On the detection of the axes of symmetry of symmetric and almost symmetric planar images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(1):104–108, 1989.
- [66] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, 1982.
- [67] M.I. Miller and L. Younes. Group actions, homeomorphisms and matching: a general framework. *International Journal of Computer Vision*, 41:61–84, July 2001. Special issue on statistical and computational theories of vision: Part 2.
- [68] D. Milner, S. Raz, H. Hel-Or, D. Keren, and E. Nevo. A new measure of symmetry and its application to classification of bifurcating structures. *Pattern Recognition*, 40:2237–2250, 2007.
- [69] D.P. Mukherjee, A. Zisserman, and J.M. Brady. Shape from symmetry—detecting and exploiting symmetry in affine images. *Phil. Trans. of the Royal Society of London*, 351:77–106, 1995. Series A.

- [70] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–684, 1989.
- [71] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.
- [72] J. Nissanov and O. Bertrand, L. Tretiak. Cryosectioning distortion reduction using tape support. *Microsc Res Tech*, 53:239–240, 2001.
- [73] P.J. Olver, G. Sapiro, and A. Tannenbaum. Differential invariant signatures and flows in computer vision: a symmetry group approach. In B. Ter Haar Romeny, editor, *Geometry Driven Diffusion in Computer Vision*, page 255. Kluwer Acad. Publ.
- [74] P.J. Olver, G. Sapiro, and A. Tannenbaum. Invariant geometric evolutions of surfaces and volumetric smoothing. *SIAM J. Applied Math*, 57:176–194, 1997.
- [75] S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
- [76] N. Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *International Journal of Computer Vision*, 50:345–362, December 2002.
- [77] N. Paragios. Variational methods and partial differential equations in cardiac image analysis. *IEEE International Symposium on Biomedical Imaging*, 1(15-18):17–20, April 2004.
- [78] N. Paragios and R. Deriche. Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, 13:249–268, 2002.
- [79] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

- [80] V.S.N Prasad and L. Davis. Detecting rotational symmetries. In *Proceedings of the International Conference on Computer Vision*, pages 954–961, 2005.
- [81] K.G. Rao and G. Medioni. Generalized cones: Useful geometric properties. In *Computer Vision, Graphics, and Image Processing*, pages 185–208, 1992.
- [82] T. Riklin-Raviv. Segmentation examples. <http://www.eng.tau.ac.il/~tammy>, 2007.
- [83] T. Riklin-Raviv, N. Ben-Zadok, N. Kiryati, N. Sochen, Gefen S., L. Bertand, and J. Nissanov. Segmentation of mouse brain atlas via propagating distributions. Technical report, Tel Aviv university, 2007.
- [84] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Unlevel-sets: Geometry and prior-based segmentation. In *Proceedings of the European Conference on Computer Vision*, volume 4, pages 50–61, 2004.
- [85] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Prior-based segmentation by projective registration and level sets. In *Proceedings of the International Conference on Computer Vision*, volume I, pages 204–211, 2005.
- [86] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Segmentation by level-sets and symmetry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1015–1022, June 2006.
- [87] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Prior-based segmentation and shape registration in the presence of projective distortion. *International Journal of Computer Vision*, 72(3):309–328, May 2007.
- [88] T. Riklin-Raviv, N. Sochen, and N. Kiryati. On symmetry, perspectivity and level-set based segmentation. *submitted*, 2007.
- [89] T. Riklin-Raviv, N. Sochen, and N. Kiryati. Shape-based mutual segmentation. *submitted*, 2007.
- [90] T. Riklin-Raviv, N. Sochen, N. Kiryati, N. Ben-Zadok, Gefen S., L. Bertand, and J. Nissanov. Propagating distributions for segmentation of brain atlas. In

- IEEE International Symposium on Biomedical Imaging*, pages 1304–1307, April 2007.
- [91] T. Riklin-Raviv and N. Sochen, N. and Kiryati. Mutual segmentation with level-sets. In *The fifth IEEE computer society workshop on perceptual organization in computer vision*, June 2006.
- [92] M. Rochery, I.H. Jermyn, and J. Zerubia. Higher-order active contours. *International Journal of Computer Vision*, 69(1):27–42, August 2006.
- [93] G.D. Rosen et al. Informatics center for mouse genomics: The dissection of complex traits of the nervous system. *Neuroinformatics*, 1:327–342, 2003.
- [94] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *Acm Transaction on Graphics*, August 2004.
- [95] C. Rother, T. Minka, A. Blake, and V. Kolmogorov. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into MRFs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 993–1000, 2006.
- [96] M. Rousson and N. Paragios. Shape priors for level set representation. In *Proceedings of the European Conference on Computer Vision*, pages 78–92, 2002.
- [97] C. Sagiv, N. Sochen, and Y.Y. Zeevi. Integrated active contours for texture segmentation. *IEEE Transactions on Image Processing*, 15(6):1633–1646, June 2006.
- [98] B. Sandberg, T. Chan, and L. Vese. A level-set and gabor based active contour for segmenting textured images. Technical report, UCLA CAM report, July 2002.
- [99] A. Shahar and H. Greenspan. Probabilistic spatial-temporal segmentation of multiple sclerosis lesions. *IEEE International Symposium on Biomedical Imaging*, 1(15-18):440–443, April 2004.

- [100] I. Shimshoni, Y. Moses, and M. Lindenbaum. Shape reconstruction of 3D bilaterally symmetric surfaces. *International Journal of Computer Vision*, 39(2):97–110, 2000.
- [101] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7:310–318, 1998. Special Issue on Geometric Active Diffusion.
- [102] C. E. Springer. *Geometry and Analysis of Projective Spaces*. Freeman, 1964.
- [103] J.S. Stahl and S. Wang. Globally optimal grouping for symmetric boundaries. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1030–1037, June 2006.
- [104] G.C. Stockman and A.K. Agrawala. Equivalence of hough curve detection to template matching. *Communications of the ACM*, 20(11):820–822, 1977.
- [105] J. Strain. Fast tree-based redistancing for level set computations. *Journal of Computational Physics*, 152(2):664–686, 1999.
- [106] C.M. Tempany, E.A. Stewart, N. McDannold, B.J. Quade, F.A. Jolesz, and K. Hynynen. MR imaging-guided focused ultrasound surgery of uterine leiomyomas: A feasibility study. *Radiology*, 226:897–905, 2003.
- [107] D. Terzopoulos, A.P. Witkin, and M. Kass. Symmetry-seeking models and 3d object reconstruction. *International Journal of Computer Vision*, 1:211–221, 1987.
- [108] S. Thrun and B. Wegbreit. Shape from symmetry. In *Proceedings of the International Conference on Computer Vision*, pages 1824–1831, 2005.
- [109] A. Törn and A. Žilinskas. *Global Optimization*, volume 350 of *Lecture Notes in Computer Science*. Springer, 1989.
- [110] A. Tsai, A. Yezzi, Jr., W.M. Wells, III, C. Tempany, D. Tucker, A. Fan, W.E.L. Grimson, and A.S. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137–154, Feb. 2003.

- [111] C.W. Tyler. *Human Symmetry Perception and Its Computational Analysis*. Lawrence Erlbaum Associates, 2002.
- [112] S. Ullman. *High-Level Vision: Object Recognition and Visual Cognition*. MIT Press, 1996.
- [113] L.J. Van Gool, T. Moons, D. Ungureanu, and A. Oosterlinck. The characterization and detection of skewed symmetry. *Computer Vision and Image Understanding*, 61(1):138–150, January 1995.
- [114] L.J. Van Gool, T. Moons, D. Ungureanu, and E. Pauwels. Symmetry from shape and shape from symmetry. *International Journal of Robotics Research*, 14(5):407–424, October 1995.
- [115] A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. In *Proceedings of the International Conference on Computer Vision*, pages 149–154, 2001.
- [116] A. Vedaldi and S. Soatto. Local features, all grown up. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1753–1760, 2006.
- [117] R.C. Veltkamp and M. Hagedoorn. Shape similarity measure, properties and constructions. In *Principles of Visual Information Retrieval*, pages 467–476, London, UK, 2000. Springer-Verlag.
- [118] R.C. Veltkamp and M. Hagedoorn. State of the art in shape matching. In *Principles of Visual Information Retrieval*, pages 87–119. Springer-Verlag, London, UK, 2001.
- [119] L.A. Vese and T.F. Chan. A multiphase level set framework for image segmentation using mumford and shah model. *International Journal of Computer Vision*, 50(3):271–293, 2002.
- [120] R.P. Woods, S.T. Grafton, C.J. Holmes, S.R. Cherry, and J.C. Mazziotta. Automated image registration: 1. general methods and intrasubject, intramodality validation. *J Comput Assist Tomogr*, 22:139–152, 1998.

- [121] R.P. Woods, S.T. Grafton, J.D. Watson, N.L. Sicotte, and J.C. Mazziotta. Automated image registration: 2. intersubject validation of linear and nonlinear models. *J Comput Assist Tomogr*, 22:153–165, 1998.
- [122] Y.A. Yang, K. Huang, S. Rao, W. Hong, and Y. Ma. Symmetry-based 3-d reconstruction from perspective images. *Computer Vision and Image Understanding*, 99:210–240, August 2005.
- [123] Y.A. Yang, S. Rao, K. Huang, W. Hong, and Y. Ma. Geometric segmentation of perspective images based on symmetry groups. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 1251–1258, 2003.
- [124] A. Yezzi, L. Zöllei, and T. Kapur. A variational framework to integrate segmentation and registration through active contours. *Medical Image Analysis*, 7:171–185, 2003.
- [125] H. Zabrodsky and D. Peleg, S. and Avnir. Completion of occluded shapes using symmetry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 678–679, June 1993.
- [126] H. Zabrodsky, S. Peleg, and D. Avnir. Symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1154–1166, December 1995.
- [127] H. Zabrodsky and D. Weinshall. Using bilateral symmetry to improve 3d reconstruction from image sequences. *Computer Vision and Image Understanding*, 67:48–57, 1997.
- [128] S.C. Zhu and A.L. Yuille. Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.