Amar Mitiche
Ismail Ben Ayed

SPRINGER TOPICS IN SIGNAL PROCESSING

Variational and Level Set Methods in Image Segmentation

Springer

Springer Topics in Signal Processing

Volume 5

For further volumes:
http://www.springer.com/series/8109

Amar Mitiche · Ismail Ben Ayed

Variational and Level Set Methods in Image Segmentation

Springer

Prof. Amar Mitiche
Institut National de la Recherche
Scientifique (INRS)
800, de la Gauchetière Ouest
Montreal, Quebec, Canada
mitiche@emt.inrs.ca

Dr. Ismail Ben Ayed
Institut National de la Recherche
Scientifique (INRS)
800, de la Gauchetière Ouest
Montreal, Quebec, Canada
benayedi@emt.inrs.ca

ISSN 1866-2609 e-ISSN 1866-2617
ISBN 978-3-642-15351-8 e-ISBN 978-3-642-15352-5
DOI 10.1007/978-3-642-15352-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010935288

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

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

Contents

1

INTRODUCTION

The subject of this book is *image segmentation* by variational methods. The focus is on formulations which use closed curves to define the segmentation regions and on a level set implementation of the corresponding active curve evolution algorithms.

In broad terms, image segmentation consists of dividing an image domain into disjoint regions according to a characterization of the image within or in-between the regions. Therefore, segmenting an image is to divide its domain into relevant components.

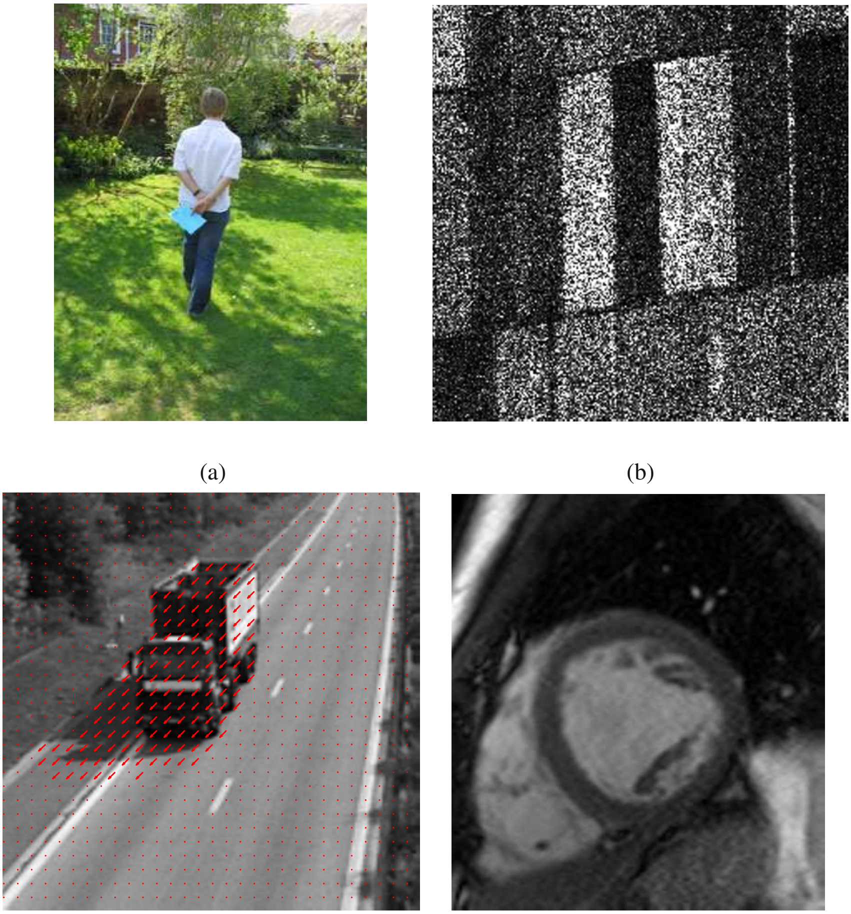
Many useful applications require image segmentation. They generally involve images too numerous or too complex to treat manually, even partly. Images can be of various kinds. Images sensed in some segment of the electromagnetic spectrum, such as digital pictures acquired by ordinary cameras, and satellite and medical images, such as radar and polarimetric, come to mind first, but there are other important ones, such as sonar and thermal [1], depth maps, and motion fields. Here following are a few examples.

Figure 1.1 (a) shows a color image acquired by an ordinary camera. The goal of segmentation would be, for instance, to detect the person. Figure 1.1 (b) depicts a synthetic aperture radar (SAR) image of the Landes forest, France. A radar uses wave scattering to record a complex signal. The SAR image is the squared norm of this complex signal [2]. Figure 1.1 (c) shows the image of a truck moving on a highway taken by a fixed ordinary camera. An estimate of the field of optical velocities (optical flow), depicted by arrows, is superimposed on this image at the points of a grid over the image domain. Here, segmentation is on the basis of optical flow. The segmentation boundary separates the figure (the truck image) from ground (the complement of the truck image), each of which is described by a piecewise constant image motion field. As a final example, Figure 1.1 (d) shows an MRI (magnetic resonance imaging) scan of the heart. The goal of segmentation would be, for instance, to detect the ventricular cavity.

Figure 1.2 displays, for the images of Figure 1.1, the type of image division one may look for.

To partition an image into regions, segmentation uses a *characterization* of the image in each region, i.e., descriptions which distinguish each region from the others.

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_1,

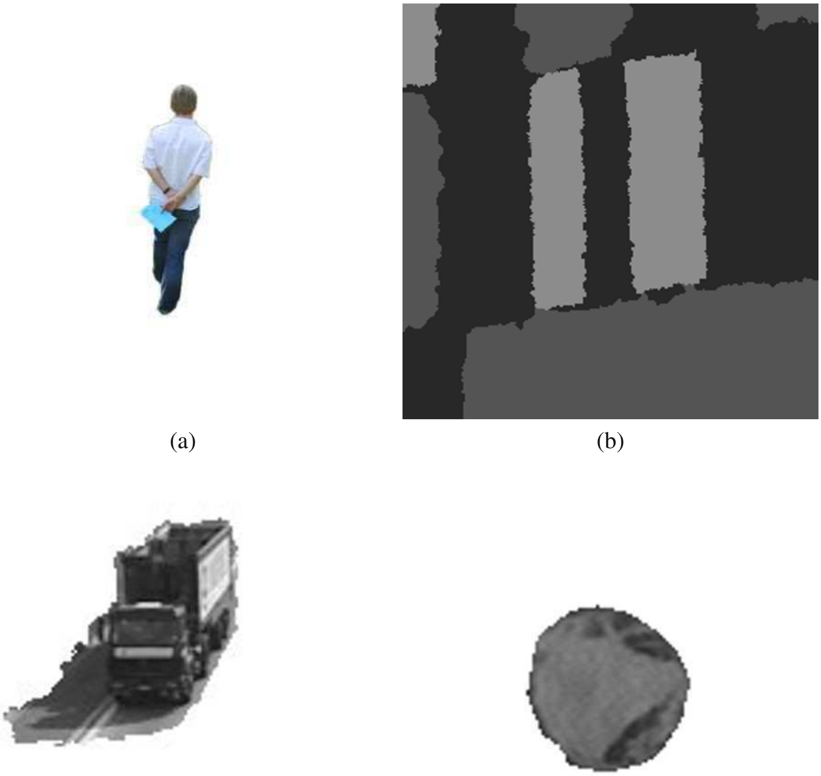


(a) (b)

(c) (d)

Fig. 1.1. (a) Color image acquired by an ordinary camera, (b) SAR image of the Landes region, (c) Traffic highway scene, for a motion based segmentation (an estimate of the optical flow field is superimposed on the image) and, (d) MRI scan of a heart for the detection of the ventricular cavity.

The most common of the many and diverse ways to characterize the segmentation regions is to use a statistical parametric image model. In this case, each segmentation region is characterized by its image distribution parameters, i.e., the regions are assumed to differ by their parameters. For luminance images, scanned by ordinary cameras, the constant distribution model and, more generally, the Gaussian model, are often assumed to be applicable. However, although prevalent, they are often not applicable. For instance, natural images and textures are better represented by a Weibull distribution [3, 4]. Images of the type SAR (synthetic aperture radar) in



(a) (b)

(c) (d)

Fig. 1.2. Segmentations for the (a) color image (detection of the person based on a non parametric description of the image), (b) SAR image of the Landes region (based on a Gamma model description of the image), (c) Traffic highway scene based on image motion (the segmentation includes the moving truck shadow) and, (d) the heart image (detection of the ventricular cavity).

remote sensing, and of the type polarimetric, in medical imaging and remote sensing, require more descriptive models. The applicable models are the Gamma distribution (SAR) and the complex Wishart (polarimetric) [2, 5, 6, 7].

Non parametric distributions can also be used to characterize a region image for segmentation [8, 9, 10]. In this case, there is no need to learn or assume a specific analytical model. Instead, a region image is described by histograms of filter responses on the image. Common filters include the identity, edge strength measurements, and

Gabor filters [11]. These histograms of filter responses, often called *features*, can be regarded as empirical marginal densities of the underlying density of the image.

Variational formulations of image segmentation minimize an objective functional which evaluates the merit of an image domain partition with respect to a set of constraints on both the image and the partition [12]. In general, variational formulations easily allow statistical models. Image constraints give rise to functional terms called *data* terms, or *observation* terms, which measure how close an image fits a model. As such, these terms measure the *error* of fit to the model [13]. Partition constraints yield partition *priors* which usually describe a geometric aspect of the partition boundaries, such as their length. The Mumford and Shah functional [14] is fundamental:

$$\mathscr{E}(g,\Gamma) = \int_{\Omega\backslash K} (g_0 - g)^2 dxdy + \rho \int_{\Omega\backslash\Gamma} \|\nabla g\|^2 dxdy + \lambda \rho l(\Gamma) \tag{1.1}$$

The functional, $\mathscr{E}$, is often referred to as an *energy*. It is associated with the partition determined by image g and boundaries K. The observed image is g_0; Ω is the image domain; l designates length, and ρ (scale), λ (weight) are positive constants. The first term is the data term. It measures how closely g approximates the observed image data g_0. The last term is a length prior. As we will see in subsequent chapters, its effect is to bias the minimization of $\mathscr{E}$ toward partitions with smooth boundaries and to avoid small partition fragments. The second term is a typical *smoothness* term to bias the minimization toward a smoothly varying approximation g everywhere except across the boundaries Γ. The partition regions are not explicit in this functional. However, they can be made explicit, as in the following two-region Chan-Vese functional [15]:

$$\mathscr{E}(\gamma,a,b) = \int_{\mathbf{R}_\gamma} (g_0 - a)^2 dxdy + \int_{\mathbf{R}_\gamma^c} (g_0 - b)^2 dxdy + \lambda \int_\gamma ds \tag{1.2}$$

where $\gamma : [0,1] \to \mathbb{R}^2$ is a C^1 closed parametric plane curve, $\mathbf{R}_\gamma$ is the interior of γ, $\mathbf{R}_\gamma^c$ its complement, and a, b are the averages of g_0 within $\mathbf{R}_\gamma$ and $\mathbf{R}_\gamma^c$, respectively. The last integral is the length of γ. The functional is called a *curve evolution functional* because the minimization equations with respect to γ are curve evolution equations. As a result, γ is called an *active curve*. Variables a and b are, more generally, called *region parameters*. They are the averages of g_0 within $\mathbf{R}_\gamma$ and $\mathbf{R}_\gamma^c$, respectively, because the corresponding minimization equations defines them as such. Basically, minimization of (1.2) seeks the best piecewise constant approximation of the original image, g_0, over a two-region partition of the image domain.

There are several key issues in segmentation by active curve functionals. One issue is *multiregion* segmentation, in which several active curves are used simultaneously [16]. The problem is to ensure that the image domain division defined by the solution curves is a partition, i.e., a set of non intersecting regions which cover the image domain, and do so in a computationally efficient manner.

Another issue is image *modeling*. When the segmentation regions differ by their image distribution, the problem is to determine appropriate image models to use in the objective functional. We mentioned three specific models, namely the Gamma distribution for SAR images, the complex Wishart to describe polarimetric images, and the Weibull distribution for natural and texture images. In general, image modeling, which, to a large extent, is independent of segmentation, is a laborious and intricate task.

There is also the question of determining the proper *number* of segmentation regions [17]. Currently, most of the formulations assume that this number is given beforehand, which simplifies the problem significantly. However, the number of regions is not known in many applications. The problem, then, is to allow it to be a variable in the functional minimization process.

Segmentation by *motion* is a topic of major importance because motion is a fundamental dimension in visual perception, both human and machine [18, 19]. Image domain partitioning by means of motion can be done using either optical flow, the apparent motion field of the environmental surfaces projection, or the actual movement of the environmental objects. In either case, segmentation corresponds to the division of the image domain into distinctly moving objects.

Finally, there is the subject of *shape priors*, where the segmentation focuses on geometric patterns about which some knowledge is available beforehand, such as its shape modulo some plane transformations [20, 21]. The problem is to describe the shape so as to assist in driving the active curve to coincide with the pattern boundary outline.

This book is structured based on these issues, except shape priors for which, however, there are sources which offer a good view of the subject and literature, for instance [20, 22, 23, 24, 25, 26, 27]. Following a presentation of the fundamental active curve and level set segmentation methods, each issue will be the subject of a chapter.

The efficient solution of the key problems in image segmentation, of the sort we just mentioned, promises to enable a rich array of useful applications. Some of the current major application areas of segmentation are:

Remote Sensing [28, 29]. Agricultural remote sensing, which originated in the 1950s, is a long-standing application. Today, most agricultural producers use remote sensing for a variety of purposes. For instance, remote sensing can inform them on crop disease, insect infestation, weed proliferation, and weather damage, as well provide them with crop inventory, water resources mapping, grazing land repartition and status, and soil composition. This information, most of which is attainable through image segmentation, affords producers precise farming management and monitoring capabilities.

Remote sensing is also a vital tool to study ecological systems, of interest to scientists and practitioners from various disciplines such as geology, forestry, agronomy, hydrology, and environmental management. The purpose is to evaluate and monitor ecological resources. Image segmentation assists in determining the geographical repartition of the ecological units to study.

Medical image interpretation [30]. Medical imaging is becoming a major component of patients care and medical research. It is presently used to diagnose various clinical conditions such as diseases. It is also the focus of extensive research for the critical impact it can have on health care in general. Segmentation is essential in medical image interpretation. For instance, segmentation serves functional neuroimaging to map brain functions in PET (positron emission tomography) or MRI (magnetic resonance imaging) scans. Neuroimaging is an important research tool in fields such as cognitive neuroscience and cognitive psychology, to investigate the relationship between cognition and neural structure and activity. Segmentation of CT and MRI images, or of other scans, such as ultrasound and thermal, are also used to diagnose diseases, as in mammographies, images of the lungs, or cardiac images.

Robotics. Robotics applications of computer vision are long standing and numerous. Image segmentation is of fundamental importance in the field because it can assist a visually guided mobile robot to navigate autonomously and react to the presence of objects in its visual field [31]. A variety of images can be used for this purpose, such as color [15], range [32], optical flow [33, 34], and images from multiple viewpoints, such as stereo [35, 36].

Visual field monitoring. Security is of great current concern in a wide range of practical domains. Applications include monitoring traffic, securing sensitive sites of human activity such as airports, and protecting private property. Functions of a visual surveillance system include change detection, event recognition, and tracking targets such as airborne devices and people.

This book on image segmentation describes several variational methods from an algorithmic viewpoint. Each method is developed from an objective functional which embeds constraints on both the segmentation image domain partition and the data within each of the partition segments. The necessary conditions to optimize the objective functional are then derived and solved numerically. Most of the methods described in the subsequent chapters use closed regular curves to define the segmentation boundaries. These curves become the segmentation variables and are computed using a level set representation, i.e., each curve is represented by the zero level set of a function defined on the image domain. The other variables are parameters which describe the observed data in each segmentation region. For methods which use optical flow to segment the image, the flow can be computed concurrently with the segmentation. In this case, the flow enters the objective functional as a variable vector field subject to constraints which allow its well posed definition and stable computation.

Although this book covers exclusively variational and level set methods, there are other schemes in use. Image segmentation was first studied along two traditional veins: *edge detection*, to find relevant segmentation boundaries, and *region growing*, to reach whole regions from seed regions [37, 38]. Edge detection is generally based on the image gradient, combined with some form of smoothing and thresholding, to compute an edge map. The most serious problem with edge detection as a segmentation tool is that the computed edges do not, in general, form closed contours, necessitating external and often *ad hoc* processes to close them [39, 40].

Region growing, which can be combined with its counterpart, region splitting, uses region statistics to determine how and when to grow or split intermediate regions. In general, region growing and splitting produce undesirably fragmented segmentations with irregular boundaries [38, 41]. Moreover, the final outcome is sensitive to initialization. Both region growing and edge detection for image segmentation are still in use in applications where these disadvantages can be overcome by experience.

Traditional methods were followed by Markov random field (MRF) modeling methods [1, 42, 43, 44] which view an image as a realization of an MRF over the image positional array. This description can be quite effective but minimization of the objective function is often computationally voracious.

Graph cut combinatorial methods, which view segmentation as a discrete label assignment problem, have been of intense interest recently [45], although they have been used for binary images much earlier [46]. Several studies have shown that graph cuts can be useful in image analysis [47]. For instance, very fast methods have been implemented for image segmentation [45, 48, 49], motion and stereo segmentation [50, 51], tracking [52], and restoration [53]. Current graph cut methods require improvements to include data terms of a more general definition that those they currently accommodate.

The remainder of this book is organized as follows. A focused literature review and the corresponding references are conveniently given in each chapter.

Chapter 2 contains mathematical formulas which are used repeatedly in the subsequent chapters. This includes Euler-Lagrange equations, descent formulas of optimization, level set evolution equations and their discretization. It also contains equations used in motion induced segmentation, namely optical flow estimation functionals, and the optical flow equations of three-dimensional (3D) rigid body motion.

The basic methods of variational and level set image segmentation are described in *Chapter* 3. The Mumford and Shah functional [14] is presented first. This is the fundamental formulation. All the methods we study or mention in this book, including motion segmentation methods, use the Mumford and Shah functional as a template, modulo terms to fit specific applications and an implementation via active curves and level sets. The active curves are variable closed regular plane curves used to define the segmentation boundaries. Application specific terms include geometric priors on the segmentation boundaries, statistical models of the image distribution in the segmentation regions, and priors which affect the number of segmentation regions. The presentation of the Mumford and Shah functional is followed by the Leclerc discrete implementation [54], the Zhu and Yuille active curve formulation [55], and the Chan and Vese level set transposition [15]. Both the piecewise constant image model and the Gaussian generalization are treated in this chapter. This chapter also reviews edge based segmentation methods, namely, the Snakes scheme [56] and the geodesic active contour [57, 58], which move a curve according to boundary information rather than region information.

The level set methods treated in Chapter 3 are two-region partitioning methods. These require a single active curve. Level set image partitioning into more regions than two (multiregion segmentation) requires several curves. This raises the problem of ascertaining that the curves converge so as to define a partition of the image domain, and do so efficiently. *Chapter* 4 describes in some detail several schemes, and makes a statement on their computational efficiency.

Chapter 5 reviews image models. As mentioned earlier, image segmentation uses a characterization of the image data within each region, i.e., image descriptions which distinguish each region from the others. The descriptions validity largely affects the segmentation accuracy, i.e., the descriptions must fit the class of images to segment. The chapter reviews segmentation functionals of the general form

$$\mathscr{D} + \lambda \int_{\gamma} ds, \tag{1.3}$$

where $\mathscr{D}$ is a data term which measures how well the image fits a specified statistical description and λ a positive constant to weigh the contribution of the curve length term against the data term. This chapter investigates various data terms, each corresponding to a particular statistical image description. Some of these descriptions are derived from Bayesian estimation, whereas others are built on information theoretic concepts or measures of discrepancy (or affinity) between distributions. Each statistical description will result in a particular active curve segmentation algorithm. Specifically, we will study the following parametric distributions to model the image within the segmentation regions: The Gaussian, univariate and multivariate, followed by the more general Weibull distribution, the Gamma and generalization to the Exponential family (which includes the Gamma, Poisson, Bernoulli, and others), and the complex Wishart. We will also describe image segmentation by non-parametric distributions which, in the digital case, uses empirical image distributions (histograms) to describe the image within the segmentation regions. In general, the data term for such distributions evaluates the discrepancy or the similarity between two distributions. To this effect, we will study segmentation using a reference distribution (of a region of interest), segmentation by the image domain labeling of maximum mutual information and, finally, by maximum or *a priori* known separation between the segmentation regions. For all the methods, examples are given to illustrate their working.

Although the number of segmentation regions is unknown in most applications, most current active curve methods assume that it is given beforehand. In the methods examined in the preceding chapters, the number of regions is taken as a constant in the segmentation functional and consequently in its optimization. How to allow it to vary is an important question which has been generally avoided and which is taken up in *Chapter* 6. We will look at methods which use a *region merging prior* in the segmentation functional to alter the number of regions. Under the effect of this term, the effective number of regions, equal to some maximum number of regions initially, decreases automatically during the objective functional optimization to be, ideally, the desired number of regions. We will examine two region merging priors. One is related to the regions logarithmic area and has an entropic interpretation [17]; the other is proportional to the number of regions and has a minimum description length (MDL) interpretation [55, 59, 60].

Chapter 7 covers optical flow segmentation. First, it presents a minimum description length formulation. This is an application to optical flow of the Leclerc description of the piecewise constant image segmentation [54], a discrete version of the Mumford and Shah formulation. This is followed by a general linear parametric level set method. The functional data term of this active curve formulation uses the Horn and Shunck gradient constraint equation (Chapter 2) where each optical flow component function is written as a linear combination of basis functions. It thus measures the conformity of a linear parametric representation of optical flow in each image segment to the image brightness pattern spatio-temporal variations. The coefficients of the motion components as linear combinations of basis functions characterize each image segment.

The subject of *Chapter* 8 is image segmentation in terms of the movement of real objects: the segmentation regions correspond to differently moving objects in space. Three methods are described. All assume that the objects are rigid, in which case the motion is a screw, i.e., the composition of a three-dimensional (3D) rotation and a three-dimensional translation. Motion is relative to the viewing system so that simultaneous movement of objects and viewing system is allowed. The three methods differ essentially in the way 3D motion is described in the functional data term. One method [61] uses the optical flow 3D rigid body constraint (given in Chapter 2). Therefore, it eliminates optical flow from the formulation and introduces the 3D kinematic screw, i.e., the parameters of the instantaneous rotation and translation of each moving object, as well as the imaged scene depth. Another method [62] uses the essential parameters constraint (given in Chapter 2). Therefore, it describes the rigid motion by the essential parameters, which are parameters related analytically to the rigid body motion screw parameters of translation and rotation. Optical flow remains in the formulation: the 3D essential parameters of each moving object and optical flow are the unknowns to estimate, along with the segmentation boundaries described by regular closed plane curves. The last method studied [63] is a hybrid method: it segments the image according to optical flow, as in the methods of Chapter 7, but the objective functional contains additional terms to constrain optical flow to be the result of rigid 3D motion, via the essential parameters.

The book concludes with the Appendix which contains three items, the optical flow estimation algorithms of Horn and Schunck and of Aubert, Kornprobst and Deriche, and a rudimentary stereoscopic image construction from an image sequence 3D interpretation. This information can be useful when dealing with motion based segmentation, or simply with implementations of optical flow estimation.

References

1. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical mrf model," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216–1231, 2000.
2. J. W. Goodman, "Laser speckle and related phenomena," in *Statistical Properties of Laser Speckle Patterns*, 1974, pp. 9–75.
3. I. Ben Ayed, N. Hennane, and A. Mitiche, "Unsupervised variational image segmentation/classification using a weibull observation model," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3431–3439, 2006.
4. J. M. Geusebroek and A. W. M. Smeulders, "A six-stimulus theory for stochastic texture," *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 7–16, 2005.
5. F. Goudail and P. Réfrégier, "Contrast definition for optical coherent polarimetric images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 947–951, 2004.
6. I. Ben Ayed, A. Mitiche, and Z. Belhadj, "Multiregion level set partitioning on synthetic aperture radar images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 793–800, 2005.
7. ——, "Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.
8. S. Jehan-Besson, M. Barlaud, G. Aubert, and O. Faugeras, "Shape gradients for histogram segmentation using active contours," in *International Conference on Computer Vision (ICCV)*, 2003, pp. 408–415.
9. D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 518–526, 2004.
10. O. V. Michailovich, Y. Rathi, and A. Tannenbaum, "Image segmentation using active contours driven by the bhattacharyya gradient flow," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2787–2801, 2007.
11. S. C. Zhu, Y. Wu, and D. Mumford, "Filters, random fields, and minimax entropy (frame): Toward a unified theory for texture modeling," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 1996, pp. 686–693.
12. G. Aubert and P. Kornpbrost, *Mathematical problems in image processing: Partial differential equations and the calculus of variations*. Springer Verlag, 2006.
13. R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.
14. D. Mumford and J. Shah, "Boundary detection by using functionals," *Computer Vision and Image Understanding*, vol. 90, pp. 19–43, 1989.
15. T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
16. A. Mansouri, A. Mitiche, and C. Vázquez, "Multiregion competition: A level set extension of region competition to multiple region partioning," *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
17. I. Ben Ayed and A. Mitiche, "A region merging prior for variational level set image segmentation," *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2301–2313, 2008.
18. A. Mitiche, *Computational Analysis of Visual Motion*. Kluwer Academic, 1994.
19. A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: A synopsis of current problems and methods," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.

20. Y. Chen, H. D.Tagare, S. R. 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*, vol. 50, no. 3, pp. 315–328, 2002.
21. D. Cremers, "Nonlinear dynamical shape priors for level set segmentation," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1–7.
22. M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2000, pp. 316–323.
23. M. Rousson and N. Paragios, "Shape priors for level set representations," in *European Conference on Computer Vision (ECCV)*, vol. 2, 2002, pp. 78–92.
24. ——, "Prior knowledge, level set representations & visual grouping," *International Journal of Computer Vision*, vol. 76, no. 3, pp. 231–243, 2008.
25. D. Cremers, S. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *International Journal of Computer Vision*, vol. 69, no. 3, pp. 335–351, 2006.
26. T. F. Chan and W. Zhu, "Level set based shape prior segmentation," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 1164–1170.
27. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 137–154, 2003.
28. V. Adamchuk, R. L. Perk, and J. S. Schepers, "Applications of remote sensing in site-specific management," University of Nebraska Cooperative, Tech. Rep. EC 03-702, 2003.
29. S. Lhermitte, J.Verbesselt, J.Jonckheere, J. Aardt, K. Nackaerts, W.W.Verstraeten, and P.Coppin, "Hierarchical image segmentation based on similarity of ndvi time series," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 506–521, 2008.
30. I. N. Bankman, *Handbook of Medical Image Processing and Analysis, 2nd Edition*. Elsevier, 2007.
31. N. E. Mortensen, *Progress in Autonomous Robot Research*. Nova Science Publishers, 2008.
32. E. Natonek, "Fast range image segmentation for servicing robots," in *International Conference on Robotics and Automation (ICRA)*, 1998, pp. 406–411.
33. D. Cremers, "A multiphase level set framework for motion segmentation," in *Scale Space Theories in Computer Vision*, 2003, pp. 599–614.
34. C. Vázquez, A. Mitiche, and R. Laganière, "Joint segmentation and parametric estimation of image motion by curve evolution and level sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 782–793, 2006.
35. O. Faugeras, *Three Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MIT Press, 1993.
36. R. Laganière, H. Hajjdiab, and A. Mitiche, "Visual reconstruction of ground plane obstacles in a sparse view robot environment," *Graphical Models*, vol. 68, no. 3, pp. 282–293, 2006.
37. A. Rosenfeld and A. Kak, *Digital Picture Processing, 2nd Edition*. Academic Press, 1982.
38. T. Pavlidis, *Algorithms for Graphics and Image Processing*. Computer Science Press, 1981.
39. J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
40. R. Deriche, "Optimal edge detection using recursive filtering," in *International Conference on Computer Vision (ICCV)*, 1987, pp. 501–505.

41. R. F. White, "Change detection in sar imagery," *International Journal of Remote Sensing*, vol. 12, no. 2, pp. 339–360, 1991.
42. S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
43. Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classificiation using a hierarchical markovian model," in *International Conference on Computer Vision (ICCV)*, 1995, pp. 169–174.
44. Y. Boykov, O. Veksler, and R. Zabih, "Markov random fields with efficient approximations," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 1998, pp. 648–655.
45. Y. Boykov and G. Funka Lea, "Graph cuts and efficient n-d image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
46. D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posterori estimation for binary images," *Journal of the Royal Statistical Society*, vol. B: 51, no. 2, pp. 271–279, 1989.
47. Y. Boykov and O. Veksler, "Graph cuts in vision and graphics: Theories and applications," in *Workshop on Mathematical Methods in Computer Vision*, 2005, pp. 79–96.
48. O. Veksler, "Efficient graph-based energy minimization methods in computer vision, thesis," Cornell University, Tech. Rep., Jul. 1999.
49. D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2005, pp. 755–762.
50. T. Schoenemann and D. Cremers, "High resolution motion layer decomposition using dual-space graph cuts," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–7.
51. S. Roy, "Stereo without epipolar lines: A maximum-flow formulation," *International Journal of Computer Vision*, vol. 34, no. 2-3, pp. 147–161, 1999.
52. J. Xiao and M. Shah, "Motion layer extraction in the presence of occlusion using graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1644–1659, 2005.
53. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
54. Y. G. Leclerc, "Constructing simple stable descriptions for image partitioning," *International Journal of Computer Vision*, vol. 3, no. 1, pp. 73–102, 1989.
55. S. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 118, no. 9, pp. 884–900, 1996.
56. M. Kass, A. P. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
57. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
58. S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi, "Gradient flows and geometric active contour models," in *International Conference on Computer Vision (ICCV)*, 1995, pp. 810–815.
59. T. Kadir and M. Brady, "Unsupervised non-parametric region segmentation using level sets," in *International Conference on Computer Vision (ICCV)*, 2003, pp. 1267–1274.
60. T. Brox and J. Weickert, "Level set segmentation with multiple regions," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213–3218, 2006.

61. H. Sekkati and A. Mitiche, “Concurrent 3-d motion segmentation and 3-d interpretation of temporal sequences of monocular images,” *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 641–653, 2006.
62. A. Mitiche and H. Sekkati, “Optical flow 3d segmentation and interpretation: A variational method with active curve evolution and level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1818–1829, 2006.
63. H. Sekkati and A. Mitiche, “Joint optical flow estimation, segmentation, and 3d interpretation with level sets,” *Computer Vision and Image Understanding*, vol. 103, no. 2, pp. 89–100, 2006.

2

INTRODUCTORY BACKGROUND

The variational methods of image segmentation discussed in this book minimize functionals. In this chapter, we review some formulas we use repeatedly in the definition and minimization of these functionals: Euler-Lagrange equations, gradient descent minimization, level set representation. We also review optical flow basic expressions used in motion based segmentation.

2.1 Euler-Lagrange equations

Most of the image segmentation methods we study in this book are variational methods which use closed regular plane curves to define an image domain partition. The objective functionals they minimize have these curves as arguments. They are, typically, the sum of integrals of one of two types, namely, integrals along a regular closed plane curve and integrals over the region enclosed by such a curve. The Euler-Lagrange equations corresponding to these variable domain integrals can be derived using standard calculus of variations and, in particular, the Euler-Lagrange equations corresponding to definite integrals [1].

2.1.1 Definite integrals

Let x_1 and x_2 be fixed real numbers and consider an integral of the form

$$\mathscr{E}(y) = \int_{x_1}^{x_2} g(x,y,y')dx \tag{2.1}$$

where $y = y(x)$ is a twice differentiable real function, $y' = \frac{dy}{dx}$, and g is a function twice differentiable with respect to any of its arguments, x, y, and y'. Assuming there is a function y satisfying the end point conditions $y(x_1) = y_1$ and $y(x_2) = y_2$, which minimizes (2.1), then y satisfies the Euler-Lagrange differential equation

$$\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right) = 0 \tag{2.2}$$

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_2,

A free end point endpoint at $x = x_1$ is expressed by the condition

$$\left.\frac{\partial g}{\partial y'}\right|_{x_1} = 0 \tag{2.3}$$

A similar expression holds for a free end point at $x = x_2$:

$$\left.\frac{\partial g}{\partial y'}\right|_{x_2} = 0 \tag{2.4}$$

For an integral involving several dependent variables $y(x), ..., z(x)$, of the form

$$E(y, ..., z) = \int_{x_1}^{x_2} g(x, y, ..., z, y', ..., z')dx \tag{2.5}$$

where $y' = \frac{dy}{dx}, ..., z' = \frac{dz}{dx}$, there is an Euler-Lagrange equation similar to (2.2) for each dependent variable:

$$\begin{aligned} \tfrac{\partial g}{\partial y} - \tfrac{d}{dx}\left(\tfrac{\partial g}{\partial y'}\right) &= 0 \\ &\cdots \\ \tfrac{\partial g}{\partial z} - \tfrac{d}{dx}\left(\tfrac{\partial g}{\partial z'}\right) &= 0 \end{aligned} \tag{2.6}$$

In subsequent chapters, we will encounter integrals involving scalar functions of two independent variables. For an integral of the form

$$\mathscr{E}(w) = \int_{\mathbf{R}} g(x, y, w, w_x, w_y)dxdy \tag{2.7}$$

where $\mathbf{R}$ is some fixed region of $\mathbb{R}^2$, $w = w(x, y)$ assumes some prescribed values at all points on the boundary $\partial\mathbf{R}$ of $\mathbf{R}$, w_x and w_y are the partial derivatives of w, and g is twice differentiable with respect to each of its arguments, the Euler-Lagrange equation is

$$\frac{\partial g}{\partial w} - \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial w_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial g}{\partial w_y}\right) = 0 \tag{2.8}$$

When w is not specified on $\partial\mathbf{R}$, we have

$$\frac{\partial g}{\partial w_x}\frac{dy}{ds} - \frac{\partial g}{\partial w_y}\frac{dx}{ds} = 0 \quad \text{on} \ \partial\mathbf{R} \tag{2.9}$$

where s is arc length and $(x(s), y(s))$ the corresponding parametric representation of $\partial\mathbf{R}$.

2.1.2 Variable domain of integration

Let $\gamma : s \in [0,1] \to (x(s),y(s)) \in \mathbb{R}^2$ be a simple closed plane curve parametrized by arc length, and $\mathbf{R}_\gamma$ its interior (the region it encloses). The segmentation functionals we will encounter in the subsequent chapters typically contain a term of the form

$$\int_{\mathbf{R}_\gamma} f(x,y)dxdy \tag{2.10}$$

where f is a scalar function, i.e., independent of γ. They also typically contain the term

$$\int_\gamma ds \tag{2.11}$$

which is the length of γ. Consider the following functional:

$$\mathscr{E}(\gamma) = \int_{\mathbf{R}_\gamma} f(x,y)dxdy + \lambda \int_\gamma ds \tag{2.12}$$

where f is a scalar function and λ is a positive constant. To determine the Euler-Lagrange equation corresponding to the minimization of (2.12) with respect to γ (we assume that the problem is to minimize $\mathscr{E}$, but the following discussions apply to maximization as well), the first integral in the functional is first transformed into a simple integral as follows. Let

$$P(x,y) = -\frac{1}{2}\int_0^y f(x,z)dz \tag{2.13}$$

and

$$Q(x,y) = \frac{1}{2}\int_0^x f(z,y)dz \tag{2.14}$$

Then, using the Green's theorem [2], we have

$$\int_{\mathbf{R}_\gamma} f(x,y)dxdy = \int_{\mathbf{R}_\gamma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \int_\gamma Pdx + Qdy = \int_0^l \left(Px' + Qy'\right) ds \tag{2.15}$$

where $x' = \frac{dx}{ds}$ and $y' = \frac{dy}{ds}$. Applying (2.6) to the last integral in (2.15), i.e., using $g(s,x,y,x',y') = P(x(s),y(s))x'(s) + Q(x(s),y(s))y'(s)$, yields

$$\begin{aligned} \frac{\partial g}{\partial x} - \frac{d}{ds}\left(\frac{\partial g}{\partial x'}\right) &= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) y' = fy' \\ \frac{\partial g}{\partial y} - \frac{d}{ds}\left(\frac{\partial g}{\partial y'}\right) &= \left(-\frac{\partial Q}{\partial x} + \frac{\partial P}{\partial y}\right) x' = -fx' \end{aligned} \tag{2.16}$$

Therefore, orienting γ so that the outward normal is

$$\mathbf{n} = (y', -x'), \tag{2.17}$$

the component of the Euler-Lagrange equation corresponding to (2.12) due to its first term is

$$f\mathbf{n} \tag{2.18}$$

The second integral of (2.12) is rewritten as

$$\int_0^l \left(x'^2 + y'^2\right)^{\frac{1}{2}} ds \tag{2.19}$$

Application of (2.6), i.e., using $g(s,x,y,x',y') = \left((x'(s))^2 + (y'(s))^2\right)^{\frac{1}{2}}$, gives

$$\begin{aligned}
\frac{\partial g}{\partial x} - \frac{d}{ds}\left(\frac{\partial g}{\partial x'}\right) &= -\frac{d}{ds}\left(\frac{x'}{\left((x')^2 + (y')^2\right)^{\frac{1}{2}}}\right) = \kappa y' \\
\frac{\partial g}{\partial y} - \frac{d}{ds}\left(\frac{\partial g}{\partial y'}\right) &= -\frac{d}{ds}\left(\frac{y'}{\left((x')^2 + (y')^2\right)^{\frac{1}{2}}}\right) = -\kappa x'
\end{aligned} \tag{2.20}$$

where κ is the curvature of γ given by, independently of the parametrization [3],

$$\kappa = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{\frac{3}{2}}} = \operatorname{div}\left(\frac{y'}{\left((x')^2 + (y')^2\right)^{\frac{1}{2}}}, \frac{-x'}{\left((x')^2 + (y')^2\right)^{\frac{1}{2}}}\right) = \operatorname{div}\left(\frac{\mathbf{n}}{\|\mathbf{n}\|}\right) \tag{2.21}$$

with $x'' = \frac{d^2x}{ds^2}$, $y'' = \frac{d^2y}{ds^2}$, and div is the divergence operator. Therefore, the contribution of the second integral of (2.12) to the Euler-Lagrange equation is

$$\kappa\mathbf{n} \tag{2.22}$$

This gives the Euler-Lagrange equation corresponding to $\mathscr{E}(\gamma)$:

$$(f + \lambda\kappa)\mathbf{n} = \mathbf{0} \tag{2.23}$$

In the computer vision literature, the left-hand side of (2.23) is referred to as the *functional derivative* of $\mathscr{E}$ with respect to γ, the sum of the functional derivatives of its two component terms. It is noted $\frac{d\mathscr{E}}{d\gamma}$, or $\frac{\partial\mathscr{E}}{\partial\gamma}$ when $\mathscr{E}$ is also dependent on other variables, as it is the case in the methods discussed in this book.

In a subsequent chapter, we will encounter the following integral, called a *geodesic* functional :

$$\int_\gamma h ds \tag{2.24}$$

where $h = h(x(s), y(s))$ is a positive scalar function (the length integral in (2.11) is a special case, with $h = 1$). Using previous arguments, its functional derivative, with respect to γ, is

$$\frac{d}{d\gamma}\left(\int_{\gamma} h ds\right) = (\langle \nabla h, \mathbf{n}\rangle + h\kappa)\mathbf{n} \tag{2.25}$$

and the corresponding Euler-Lagrange equation is

$$(\langle \nabla h, \mathbf{n}\rangle + h\kappa)\mathbf{n} = \mathbf{0} \tag{2.26}$$

where $\langle . \rangle$ denotes the scalar product, and $\nabla h = \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}\right) = (h_x, h_y)$ is the spatial gradient of h.

The Euler-Lagrange equations corresponding to the region integral (2.10) and to the geodesic functional (2.24) can be generalized to volume and surface integrals, respectively. Although not used in this book, such integrals appear, for instance, in spatio-temporal image segmentation [4]. Let S be a closed regular surface [3] in $\mathbb{R}^3$. The Euler-Lagrange equation corresponding to $\int_S h dS$ is

$$(\langle \nabla h, \mathbf{n}\rangle + 2h\kappa)\mathbf{n} = \mathbf{0}, \tag{2.27}$$

where $\mathbf{n}$ is the external unit normal to S and κ its mean curvature function [3]. The factor 2 appearing in this equation is due to the definition of the mean curvature. The equation corresponding to the volume integral $\int_{\mathbf{R}_S} f dV$, where $\mathbf{R}_S$ is the interior of S, is

$$f\mathbf{n} = \mathbf{0} \tag{2.28}$$

A detailed proof of these formulas can be found in [4].

Image segmentation functionals can involve integrals with $\mathbf{R}_\gamma$-dependent parameters. A typical such integral (called a data term) is

$$\int_{\mathbf{R}_\gamma} (I - \mu)^2 dxdy \tag{2.29}$$

where I is the image and μ is its mean over $\mathbf{R}_\gamma$. The minimization of these functionals can use the Euler-Lagrange equations obtained by assuming that the region parameters are fixed (i.e., independent of γ), leading to iterations of a greedy two-step algorithm, one step to minimize with respect to the parameters, the other to minimize with respect to the curve assuming that the parameters are fixed. However, other frameworks, such as the shape calculus, which uses shape gradients [5, 6, 7], can take the integrand dependence on γ into account to derive the necessary conditions for a minimum, leading to iterations of gradient descent (fastest descent). We will see examples of both schemes in subsequent chapters.

2.2 Descent methods for unconstrained optimization

2.2.1 Real functions

Consider the problem of finding an unconstrained local minimum of a C^1 function $f : \mathbf{x} \in \mathbb{R}^N \rightarrow f(\mathbf{x}) \in \mathbb{R}$ (assuming such a minimum exists). Let $\mathbf{x} : t \geq 0 \rightarrow \mathbf{x}(t) \in \mathbb{R}^N$ be C^1, and $g(t) = f(\mathbf{x}(t))$. We have

$$\frac{dg}{dt} = \langle \nabla f, \frac{d\mathbf{x}}{dt} \rangle \tag{2.30}$$

Therefore, if $\mathbf{x}$ varies according to

$$\begin{aligned} \frac{d\mathbf{x}}{dt}(t) &= -\alpha(t)d(\mathbf{x}(t)) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \tag{2.31}$$

where $\alpha(t) \in \mathbb{R}^+$ and $\langle \nabla f, d \rangle > 0$, function g will vary according to

$$\frac{dg}{dt} = -\alpha \langle \nabla f, d \rangle \tag{2.32}$$

Because $\nabla f = 0$ is a necessary condition for a local minimum of f and $-\alpha \langle \nabla f, d \rangle$ is negative, $\mathbf{x}$ varying by (2.31) will converge to a local minimum of f, assuming such a minimum exists on the trajectory of $\mathbf{x}$. Methods of unconstrained minimization based on (2.31) are called *descent* methods. Most often $d = \nabla f$ is used (*gradient*, or *fastest*, descent). The scaling function α is often predetermined. For instance, $\alpha(t) =$ constant, or $\alpha(t) = 1/t$ [8]. In general, descent methods are discretized and implemented as follows.

```
1. $k = 0$ ; $\mathbf{x}^0 = \mathbf{x}_0$
2. Repeat until a test of convergence is verified
       $d^k = d(x^k)$
       $\alpha_k = \arg\min_{\alpha \geq 0} f(\mathbf{x}^k - \alpha d^k)$
       $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_k d^k$
       $k \leftarrow k + 1$
```

Similar vectorial formulas apply to vectorial functions $F = (f_1, ..., f_n)^t$ by treating each component function f_i as described.

2.2.2 Integral functionals

Consider the problem of minimizing (2.1), i.e., $\mathscr{E}(y) = \int_{x_1}^{x_2} g(x, y, y')dx$. Let y vary in time, i.e., y is embedded in a one-parameter family of functions indexed by (algorithmic) time t, $y = y(x,t)$. Consider the functional

$$\mathscr{E}(y,t) = \int_{x_1}^{x_2} g(x, y(x,t), y'(x,t))dx \tag{2.33}$$

where $y' = \partial y/\partial x$. Let us calculate the derivative of $\mathscr{E}$ with respect to time:

$$\begin{aligned}
\frac{\partial \mathscr{E}}{\partial t} &= \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial y'}\frac{\partial y'}{\partial t} \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial y'}\frac{\partial}{\partial t}\left(\frac{\partial y}{\partial x}\right) \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial y'}\frac{\partial}{\partial x}\left(\frac{\partial y}{\partial t}\right) \right) dx
\end{aligned}$$

Integration by parts of the second term of the integrand gives

$$\frac{\partial \mathscr{E}}{\partial t} = \frac{\partial g}{\partial y'}\frac{\partial y}{\partial t}\bigg]_{x_1}^{x_2} + \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y} - \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y'}\right) \right) \frac{\partial y}{\partial t} dx \tag{2.34}$$

Assuming that $\frac{\partial y}{\partial t}(x_1,t) = \frac{\partial y}{\partial t}(x_2,t)\ \forall t$, which is the case for each of the two component functions of the closed curves we consider in this book, we finally have

$$\frac{\partial \mathscr{E}}{\partial t} = \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y} - \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y'}\right) \right) \frac{\partial y}{\partial t} dx \tag{2.35}$$

Therefore, varying y according to

$$\frac{\partial y}{\partial t} = -\left(\frac{\partial g}{\partial y} - \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y'}\right) \right), \tag{2.36}$$

i.e.,

$$\frac{\partial y}{\partial t} = -\frac{\partial \mathscr{E}}{\partial y}, \tag{2.37}$$

implies

$$\frac{\partial \mathscr{E}}{\partial t} = -\int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y} - \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y'}\right) \right)^2 \leq 0 \tag{2.38}$$

As a result, $\mathscr{E}$ decreases in time. As with real functions, starting from $y(0) = y_0$, y will converge to a local minimum of $\mathscr{E}$, assuming such a minimum exists. Equation (2.36) is the fastest descent equation to minimize functional (2.33). Functionals of several dependent variables are treated in a similar way.

As an example, to minimize (2.12), we adopt the descent equation

$$\frac{\partial \gamma}{\partial t} = -(f + \lambda\kappa)\mathbf{n} \tag{2.39}$$

In the computer vision literature, the partial differential equation (2.39) is seen as the evolution equation of curve γ along its normal at speed $-(f + \lambda\kappa)$, and γ is called an *active curve*. Direct implementation of this equation, as with the algorithm for real functions above, will iteratively displace each of the points of the curve.

However, a curve can split and join during evolution, and such changes in the curve topology are quite difficult, if at all possible, to effect by an explicit representation of the curve as a set of mobile marker points. An efficient and numerically stable implementation of such descent equations is via *level sets* which we outline next.

2.3 Level sets

Let γ be a simple closed curve in the image domain Ω. In the problems we address, such curves are made to move so as to converge to delimit the desired segmentation regions. Consider a curve which evolves according to a velocity vector in the direction of its normal at each point (Figure 2.1 a),

$$\mathbf{V} = V\mathbf{n} \tag{2.40}$$

The velocities we will encounter in this book have speeds of one of three distinct types.

Type 1. V is a function of the curvature of the evolving curve.

Type 2. V is of the form $\langle \mathbf{F}, \mathbf{n} \rangle$ where $\mathbf{F}$ is a vector field dependent on position and possibly time (via the underlying image function, for instance) but not on the curve. Such terms are called advection speeds in [9].

Type 3. V is a scalar function which depends on position and time but is not of the other two types.

The closed curves delineating a segmentation region can split or merge during evolution. These changes in a curve topology are at best extensively cumbersome to effect using an explicit representation of a curve as a set of marker points and an implementation of the evolution by a descent algorithm such as given in Section 2.2 for the minimization of real functions. Level sets offer an efficient and numerically stable alternative implementation.

Let Γ be the set of smooth (C^2) plane curves $\gamma : s \in [0,1] \to \gamma(s) \in \Omega$ which are closed, simple, and regular [3]. Let an active curve be represented by a one-parameter (algorithmic time) family of curves in Γ, i.e, a function $\gamma : s,t \in [0,1] \times \mathbb{R}^+ \to \gamma(s,t) = (x(s,t), y(s,t), t)) \in \Omega \times \mathbb{R}^+$ such that $\forall t$ curve $\gamma_t : s \to (x(s,t), y(s,t))$ is in Γ. With the level set implementation, an active curve γ is represented implicitly as the zero level set of a function $\phi : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$:

$$\forall s,t \quad \phi(\gamma(s,t)) = \phi(x(s,t), y(s,t), t) = 0 \tag{2.41}$$

The total derivative of (2.41) with respect to time gives, assuming ϕ is sufficiently smooth,

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial \phi}{\partial t} = \langle \nabla\phi, \frac{\partial \gamma}{\partial t} \rangle + \frac{\partial \phi}{\partial t} = 0 \tag{2.42}$$

Because $\frac{\partial \gamma}{\partial t} = V\mathbf{n}$, we have

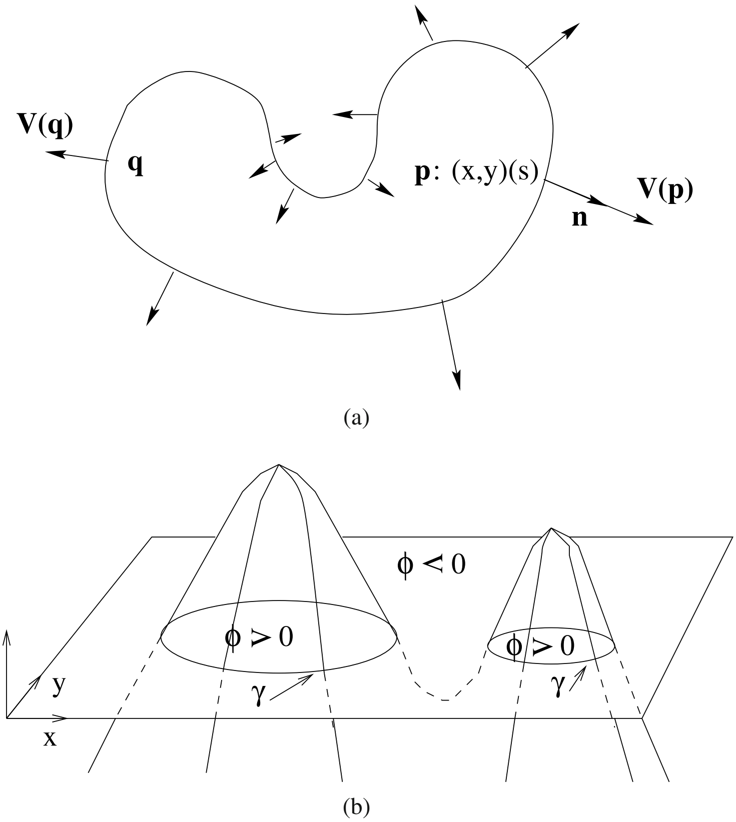


Fig. 2.1. (a) Active curve γ is moving at all times and at each point according to a velocity vector which is along its normal; (b) active curve γ is represented implicitly by the zero level of the graph of function ϕ. In this figure, γ is split into two regular curves while ϕ remains a function. In general, such a possibility cannot be implemented by an explicit representation of γ as a set points.

$$\frac{\partial \phi}{\partial t} = -V \langle \nabla \phi, \mathbf{n} \rangle \tag{2.43}$$

We also have

$$\forall s \in [0,1] \quad \frac{\partial \phi}{\partial s} = \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial s} = \langle \nabla \phi, \frac{\partial \gamma}{\partial s} \rangle = 0 \tag{2.44}$$

Therefore, because $\frac{\partial \gamma}{\partial s}$ is the tangent to the curve at s, $\nabla \phi$ is normal to the curve. With the convention that $\mathbf{n}$ is oriented outward and ϕ is positive inside its zero level set, we have

$$\mathbf{n} = -\frac{\nabla\phi}{\|\nabla\phi\|} \tag{2.45}$$

Substitution of (2.45) in (2.43) gives the temporal evolution of ϕ:

$$\frac{\partial\phi}{\partial t} = V\|\nabla\phi\| \tag{2.46}$$

For type 1 speeds, the curvature is given in terms of the level set function by:

$$\kappa = \operatorname{div}\left(\frac{\mathbf{n}}{\|\mathbf{n}\|}\right) = -\operatorname{div}\left(\frac{\nabla\phi}{\|\nabla\phi\|}\right) = -\frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2+\phi_y^2)^{\frac{3}{2}}} \tag{2.47}$$

A priori, the level set evolution is specified for points on the level set function zero level. Therefore, one must define extension velocities [9] to evolve the level set function elsewhere. For instance, the extension velocity at a point is the velocity at the point closest to it on the evolving curve. Extension velocities can also be defined so that the level set function is at all times the distance function from the evolving curve. Both of these definitions, often implemented via narrow banding [9], require the initial curves intersect the regions they segment. This is important when a region has unconnected components. An alternative robust to initialization, which we use in all the methods described in this book, extends the expression of the velocity on the evolving curve to the image domain when it can be evaluated at each point, as it is often the case [4, 10].

At all times t, active curve γ subject to velocity (2.40) can be recovered as the zero level set of level set function ϕ evolving according to (2.46). Regardless of variations in the topology of the active curve, ϕ remains a function (Figure 2.1 b). Another advantage of the level set implementation is that region membership is explicitly maintained and readily available in the level set representation because the sign of ϕ determines which points are inside curve γ, and which are outside. This information is computationally very expensive to determine with an explicit representation of active curves.

The book of Sethian [9] is about efficient and numerically stable discretization of evolution equations such as those corresponding to the level set segmentation functionals in this book. It also contains several examples of applications in various domains. Velocities of the types 1, 2, and 3 are discretized differently as summarized below [9]. Because the velocity of an active curve is, in general, a compound of velocities of the three types, the discretization of (2.46) can be written as follows:

$$\phi_{ij}^{k+1} = \phi_{ij}^k + \triangle t \begin{cases} +V_{ij}^k\left((D_{ij}^{0x})^2 + (D_{ij}^{0y})^2\right)^{\frac{1}{2}} & \text{for type 1 terms} \\ \left[\begin{array}{l}\Big(\max(F_{1ij}^k,0)D_{ij}^{-x} + \min(F_{1ij}^k,0)D_{ij}^{+x} \\ +\max(F_{2ij}^k,0)D_{ij}^{-y} + \min(F_{2ij}^k,0)D_{ij}^{+y}\Big)\end{array}\right] & \text{for type 2 terms} \\ \left(\max(V_{ij}^k,0)\nabla^+ + \min(V_{ij}^k,0)\nabla^-\right) & \text{for type 3 terms} \end{cases} \tag{2.48}$$

where i,j are indices on the discretization grid of Ω, k is the iteration index, F_1, F_2 are the coordinates of **F** appearing in the general expression of terms of type 2. Finite difference $x-$derivative operators D^{+x} (forward scheme), D^{-x} (backward scheme), and D^{0x} (central scheme), are applied to ϕ at i,j and iteration k, i.e., $D_{ij}^{+x}, D_{ij}^{-x}, D_{ij}^{0x}$ in (2.48) stand for $D^{+x}(\phi^k)_{ij}, D^{-x}(\phi^k)_{ij}, D^{0x}(\phi^k)_{ij}$ and are given by

$$\begin{aligned}
D_{ij}^{+x} &= \phi_{i+1,j}^k - \phi_{ij}^k \\
D_{ij}^{-x} &= \phi_{ij}^k - \phi_{i-1,j}^k \\
D_{ij}^{0x} &= \frac{1}{2}(\phi_{i+1,j}^k - \phi_{i-1,j}^k)
\end{aligned}$$

Similar comments and expressions apply to the $y-$derivative operators D^{+y}, D^{-y}, and D^{0y}. Finally, ∇^+ and ∇^- are defined by

$$\begin{aligned}
\nabla^+ = & \Big(\max(D_{ij}^{-x},0)^2 + \min(D_{ij}^{+x},0)^2 \\
& + \max(D_{ij}^{-y},0)^2 + \min(D_{ij}^{+y},0)^2\Big)^{\frac{1}{2}} \\
\nabla^- = & \Big(\max(D_{ij}^{+x},0)^2 + \min(D_{ij}^{-x},0)^2 \\
& + \max(D_{ij}^{+y},0)^2 + \min(D_{ij}^{-y},0)^2\Big)^{\frac{1}{2}} \qquad (2.49)
\end{aligned}$$

2.4 Optical flow

Let I be an image sequence, considered a differentiable function,

$$I : (x,y,t) \in \Omega \times]0,T[\mapsto I(x,y,t) \in \mathbb{R}^+, \qquad (2.50)$$

where x,y are image coordinates, Ω is an open subset of the real plane to represent the image domain, and T is the time of duration of the image sequence.

2.4.1 The gradient equation

Let **P** be a point on an imaged environmental surface moving relative to the viewing system. The viewing system is symbolized as in Figure 2.3. The image **p** of **P** on the viewing system projection plane can be viewed as moving along a trajectory in the space-time (actual time) domain $x-y-t$. Let $c(t) = (x(t), y(t), t)$ be the Cartesian parametric representation of this trajectory. Finally, let the restriction of I to c be $h = Ioc$, where o designates composition, i.e., $h(t) = I(x(t), y(t), t)$. The hypothesis that h is constant in time, i.e., that I does not vary along the motion trajectory c of **p**,[1] leads to the *gradient equation* of Horn and Schunck [11]:

---

[1] This is strictly true for points on a Lambertian surface in translation relative to the viewing system, under constant, uniform lighting.

$$\frac{dh}{dt} = \frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t}\frac{dt}{dt} = 0 \tag{2.51}$$

If we designate the partial derivatives of I by I_x, I_y, I_t, and $\frac{dx}{dt}, \frac{dy}{dt}$ by u, v, we have

$$I_x u + I_y v + I_t = 0 \tag{2.52}$$

Vector (u,v) is the *optical velocity* of **p**, the velocity of the projection of **P** onto the image plane. The field of optical velocities over the image domain is the *optical flow* [12]. The gradient equation is also referred to as the *optical flow constraint*. It can be written in vector form:

$$\langle \nabla I, W \rangle + I_t = 0 \tag{2.53}$$

where $\nabla I = (I_x, I_y)$ is the spatial gradient of I and $W = (u,v)$. The projection $W^{\perp}$ of W on ∇I is given by

$$W^{\perp} = \langle \frac{\nabla I}{\|\nabla I\|}, W \rangle \frac{\nabla I}{\|\nabla I\|} \tag{2.54}$$

or, using (2.53):

$$W^{\perp} = \left(-\frac{I_t}{\|\nabla I\|} \right) \frac{\nabla I}{\|\nabla I\|} \tag{2.55}$$

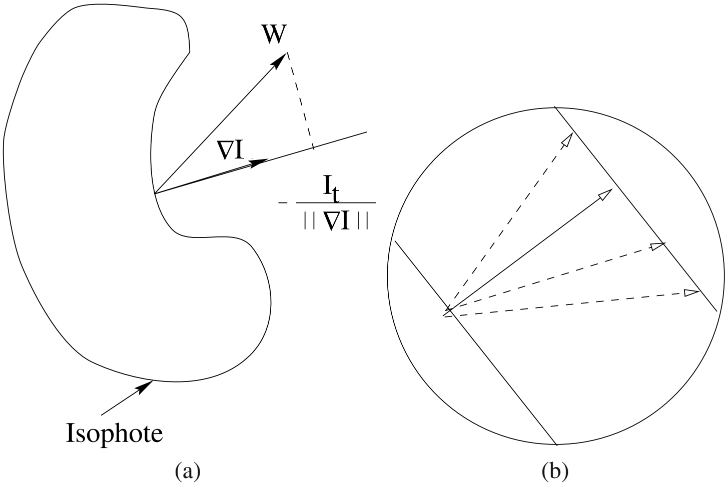
This projection can be estimated from the first order spatio-temporal variations of I (Figure 2.2). Therefore, the gradient equation gives the component of the optical velocity vector in the direction of the image gradient, i.e., the direction normal to the isophote, and only this component. This is a manifestation of the *aperture problem*: the movement of a straight edge seen through an aperture is ambiguous because only the component of motion in the direction perpendicular to the edge is determined (Figure 2.2).

2.4.2 The Horn and Schunck formulation

Horn and Schunck minimize the following functional to estimate optical flow:

$$E(u,v) = \int_{\Omega} (I_x u + I_y v + I_t)^2 dxdy + \lambda \int_{\Omega} (\|\nabla u\|^2 + \|\nabla v\|^2) dxdy \tag{2.56}$$

where λ is a positive constant to weigh the relative contribution of the two terms of the functional. The first integral, the data term, measures the conformity of the motion field to the image sequence first-order spatial and temporal variations. The second integral is a regularization term which measures the smoothness of the motion field. The Euler-Lagrange equations corresponding to (2.56) are two coupled partial differential equations:



(a) (b)

Fig. 2.2. (a) The projection of the optical flow on the image gradient can be estimated from the image first-order spatio-temporal data; whenever $\nabla I \neq 0$ it is equal to $-\frac{I_t}{\|\nabla I\|}$; (b) the aperture problem: the movement of the straight edge seen though an aperture (the circular window in this figure) is ambiguous because only the component of motion in the direction perpendicular to the edge is determined (the solid arrow).

$$\begin{aligned} I_x(I_x u + I_y v + I_t) - \lambda \nabla^2 u &= 0 \\ I_y(I_x u + I_y v + I_t) - \lambda \nabla^2 v &= 0, \end{aligned} \tag{2.57}$$

to which the (Neumann) boundary conditions are added:

$$\begin{aligned} \frac{\partial u}{\partial \mathbf{n}} &= 0 \\ \frac{\partial v}{\partial \mathbf{n}} &= 0 \end{aligned} \tag{2.58}$$

where ∇^2 designates the Laplacian and $\frac{\partial}{\partial \mathbf{n}}$ indicates differentiation in the direction of the normal $\mathbf{n}$ to the image domain boundary $\partial \Omega$.

There are very efficient numerical implementations to solve (2.57) [13]. Refer to the appendix for more details.

The main problem with the Horn and Shunck method comes from the occurrence of the Laplacian in (2.57), which causes isotropic smoothing and, therefore, blurred motion boundaries. The method of [14] formulates the problem so that smoothing of motion is inhibited across motion boundaries. This formulation is described next.

2.4.3 The Aubert, Kornprobst, and Deriche formulation

The study in [14, 15] investigates the following generalization of the Horn and Schunck functional:

$$E(u,v) = \int_{\Omega} (I_x u + I_y v + I_t)^2 dxdy + \lambda \int_{\Omega} (\rho(\|\nabla u\|) + \rho(\|\nabla v\|))dxdy \tag{2.59}$$

where ρ is a function of class C^2. With $\rho(z) = z^2$, (2.59) reduces to the Horn and Schunck functional (2.56).

The Euler-Lagrange equations corresponding to (2.59) are

$$\begin{aligned} I_x(I_x u + I_y v + I_t) &= \frac{\lambda}{2} \operatorname{div} \left(\rho'(\|\nabla u\|) \frac{\nabla u}{\|\nabla u\|} \right) \\ I_y(I_x u + I_y v + I_t) &= \frac{\lambda}{2} \operatorname{div} \left(\rho'(\|\nabla v\|) \frac{\nabla v}{\|\nabla v\|} \right), \end{aligned} \tag{2.60}$$

with the boundary conditions

$$\begin{aligned} \frac{\rho'(\|\nabla u\|)}{\|\nabla u\|} \frac{\partial u}{\partial \mathbf{n}} &= 0 \\ \frac{\rho'(\|\nabla v\|)}{\|\nabla v\|} \frac{\partial v}{\partial \mathbf{n}} &= 0 \end{aligned} \tag{2.61}$$

The Aubert-Deriche-Kornprobst function

$$\rho(s) = 2\sqrt{1+s^2} - 2 \tag{2.62}$$

follows conditions which require that it allows smoothing of motion along motion boundaries and inhibits it across [14, 15].

An efficient implementation of the formulation is described in [14, 15, 16]. Refer to the appendix for more details.

Next, we will examine the relationship between optical flow and the three-dimensional structure and motion of rigid objects. We will need this relationship in Chapter 8 on image segmentation based on the movement of real objects.

2.4.4 Optical flow of rigid body motion

We symbolize physical space by the Euclidean space $\mathbb{R}^3$ and the viewing system by an orthonormal direct coordinate system $\mathscr{S} = (\mathbf{O}; \mathbf{I}, \mathbf{J}, \mathbf{K})$ and central projection

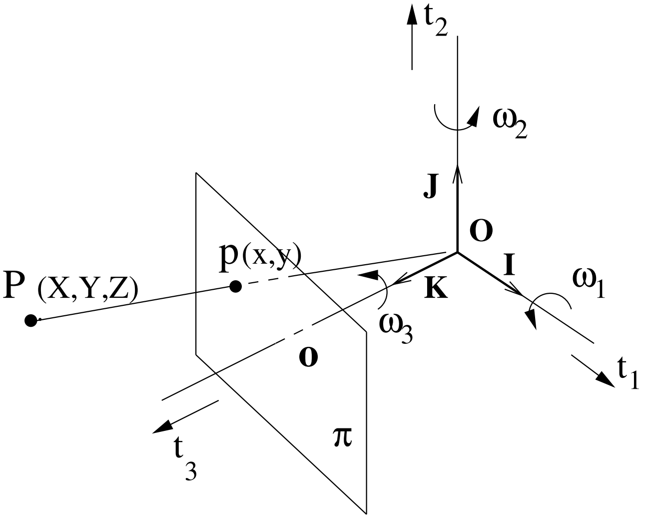


Fig. 2.3. The viewing system is symbolized by a direct orthonormal coordinate system $\mathscr{S} = (\mathbf{O};\mathbf{I},\mathbf{J},\mathbf{K})$ and central projection through **O** on plane π (the image plane) parallel to plane $P_{\mathbf{IJ}}$ and at focal distance f from **O**.

through **O** on plane π (the image plane) parallel to plane $P_{\mathbf{IJ}}$ and at distance f from **O** (the focal length). This is illustrated in Figure 2.3.

Let $t \rightarrow \mathscr{B}(t)$ be a rigid body moving in space and **P** a point of $\mathscr{B}$. Let **p** be the image of **P** on π. If $\mathbf{P} = X\mathbf{I}+Y\mathbf{J}+Z\mathbf{K}$ and $\mathbf{p} = x\mathbf{I}+y\mathbf{J}$, we have the following projection relations:

$$x = f\frac{X}{Z} \qquad y = f\frac{Y}{Z} \tag{2.63}$$

If the kinematic screw of the (rigid) motion of $\mathscr{B}$ relative to $\mathscr{S}$ is $(\omega,\mathbf{T})$ then the velocity of **P** is $\mathbf{P}' = \mathbf{T}+\omega\times\mathbf{OP}$. Substitution of this expression in the derivative with respect to time of each equation in (2.63) gives the following expression of optical velocity [17]:

$$\begin{aligned} u &= \tfrac{1}{Z}(ft_1 - xt_3) - \tfrac{xy}{f}\omega_1 + \tfrac{f^2+x^2}{f}\omega_2 - y\omega_3 \\ v &= \tfrac{1}{Z}(ft_2 - yt_3) - \tfrac{f^2+y^2}{f}\omega_1 + \tfrac{xy}{f}\omega_2 + x\omega_3 \end{aligned} \tag{2.64}$$

where $\mathbf{T} = (t_1,t_2,t_3)$ and $\omega = (\omega_1,\omega_2,\omega_3)$. Substitution of (2.64) in the gradient equation (2.52) gives the *optical flow 3D rigid body constraint*:

$$\langle \mathbf{s},\tau\rangle + \langle \mathbf{q},\omega\rangle + I_t = 0 \tag{2.65}$$

where vectors τ, **s**, and **q** are given by

$$\tau = \frac{\mathbf{T}}{Z}, \qquad \mathbf{s} = \begin{pmatrix} fI_x \\ fI_y \\ -xI_x - yI_y \end{pmatrix}, \qquad \mathbf{q} = \begin{pmatrix} -fI_y - \frac{y}{f}(xI_x + yI_y) \\ fI_x + \frac{x}{f}(xI_x + yI_y) \\ -yI_x + xI_y \end{pmatrix} \tag{2.66}$$

Pulling $\frac{1}{z}$ out of each equation in (2.64), equating the resulting expressions, and making a change of variables, leads to the following depth-free homogeneous linear equation [18, 19]:

$$\langle \mathbf{d}, \mathbf{e} \rangle = 0 \tag{2.67}$$

where $\mathbf{d} = (x^2, y^2, f^2, 2xy, 2xf, 2yf, -fv, fu, -uy + vx)$, and $\mathbf{e}$ is the vector of *essential parameters*:

$$\begin{aligned} &e_1 = -\omega_3 t_3 - \omega_2 t_2, && e_2 = -\omega_3 t_3 - \omega_1 t_1, && e_3 = -\omega_2 t_2 - \omega_1 t_1 \\ &e_4 = \tfrac{\omega_2 t_1 + \omega_1 t_2}{2}, && e_5 = \tfrac{\omega_1 t_3 + \omega_3 t_1}{2}, && e_6 = \tfrac{\omega_2 t_3 + \omega_3 t_2}{2} \\ &e_7 = t_1, && e_8 = t_2, && e_9 = t_3 \end{aligned} \tag{2.68}$$

References

1. R. Weinstock, *Calculus of variations*. Dover, 1974.
2. J. E. Marsden and A. J. Tromba, *Vector calculus*. W. H. Freeman and Company, 1976.
3. M. P. Do Carmo, *Differential geometry of curves and surfaces*. Prentice Hall, 1976.
4. A. Mitiche, R. Feghali, and A. Mansouri, "Motion tracking as spatio-temporal motion boundary detection," *Journal of Robotics and Autonomous Systems*, vol. 43, pp. 39–50, 2003.
5. M. C. Delfour and J. P. Zolesio, *Shapes and Geometries: Analysis, Differential Calculus and Optimization*. SIAM series on Advances in Design and Control, 2001.
6. S. Jehan-Besson, M. Barlaud, G. Aubert, and O. Faugeras, "Shape gradients for histogram segmentation using active contours," in *International Conference on Computer Vision (ICCV)*, 2003, pp. 408–415.
7. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients?" *SIAM Journal of Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.
8. M. Minoux, *Programmation mathématique*. Dunod, Vol. 1, 1983.
9. J. A. Sethian, *Level set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
10. A. Mansouri and J. Konrad, "Multiple motion segmentation with level sets," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 201–220, 2003.
11. B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 17, pp. 185–203, 1981.
12. J. J. Gibson, *The perception of the visual world*. Houghton Mifflin, 1950.
13. A. Mitiche and A. Mansouri, "On convergence of the Horn and Schunck optical flow estimation method," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 848–852, 2004.
14. G. Aubert, R. Deriche, and P. Kornprobst, "Computing optical flow via variational techniques," *SIAM Journal of Applied Mathematics*, vol. 60, no. 1, pp. 156–182, 1999.
15. R. Deriche, P. Kornprobst, and G. Aubert, "Optical-flow estimation while preserving its discontinuities: A variational approach," in *Asian Conference on Computer Vision (ACCV)*, 1995, pp. 71–80.
16. G. Aubert and P. Kornpbrost, *Mathematical problems in image processing: Partial differential equations and the calculus of variations*. Springer Verlag, 2006.
17. H. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proceedings of the Royal Society of London, B*, vol. 208, pp. 385–397, 1981.
18. X. Zhuang and R. Haralick, "Rigid body motion and the optical flow image," in *International Conference on Artificial Intelligence Applications*, 1984, pp. 366–375.
19. A. Mitiche, *Computational Analysis of Visual Motion*. Plenum Press, New York, 1994.

3

BASIC METHODS

The purpose of this chapter is to give a synopsis of fundamental variational image segmentation methods. Reviewed are the Mumford and Shah formulation and its basic discrete implementations, the Zhu and Yuille region competition version by curve evolution, the Chan and Vese level set form and, finally, the active curve edge detection interpretations by Snakes and geodesic active contours.

3.1 The Mumford and Shah model

The Mumford and Shah variational formulation is the core of variational image segmentation. It has been the subject of an impressive number of theoretical studies, for instance [1, 2, 3, 4, 5]. It bestirred intense research and influenced numerous methodological and algorithmic studies. Most of the other formulations can be explained as implementations, applications, or variants. The Mumford-Shah formulation [1] expresses segmentation as joint image smoothing and boundary detection. It seeks a smooth approximation $M : \mathbf{x} = (x,y) \in \Omega \subset \mathbb{R}^2 \to M(\mathbf{x}) \in \mathbb{R}$ of the original image $I : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ and a set K of discontinuities, to represent the image boundaries, by minimizing the functional

$$\mathscr{F}_{MS}(M,K) = \int_{\Omega} (I - M)^2 d\mathbf{x} + \rho \int_{\Omega \setminus K} \|\nabla M\|^2 d\mathbf{x} + \lambda \rho l(K) \tag{3.1}$$

where ρ and λ are positive constants, l is length, and $\nabla M = (M_x, M_y) = \left(\frac{\partial M}{\partial x}, \frac{\partial M}{\partial y}\right)$ is the spatial gradient of M.

The first term, referred to as the *data term*, biases the solution towards an image close to the observed image. The second term is a *smoothness prior* for a solution which varies smoothly everywhere but across the edges K. The last term is a *length prior* for regular boundaries.

Variational segmentation algorithms have commonly followed a simplified version of the Mumford-Shah functional where the search for an optimal approximation is limited to the set of piecewise constant images:

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_3,

$$\mathscr{F}_{MS}(M,K) = \sum_k \int_{\mathbf{R}_k} (I - M_k)^2 d\mathbf{x} + \lambda l(K) \tag{3.2}$$

where $\{\mathbf{R}_k\}$ is the partition determined by K and

$$M(\mathbf{x}) = M_k \quad \forall \mathbf{x} \in \mathbf{R}_k \tag{3.3}$$

In this case, the problem is to determine a partition and a set of values M_k, each corresponding to a constant approximation of I within a region. This limiting case corresponds to $\rho \to \infty$, and is known as the cartoon limit [1].

The mathematical framework of existence, uniqueness, and other mathematical properties of the functional minimizers is beyond the scope of this book [2, 3]. Instead, we will simply examine some algorithmic aspects of interpretations of the Mumford and Shah formulation.

3.1.1 Bayesian interpretation

Image segmentation can be stated as a Bayesian maximum *a posteriori* (MAP) estimation [6, 7, 8, 9, 10, 11, 12], which determines, over all possible partitions of the image domain, a partition of a MAP probability given an image I

$$\hat{\mathscr{P}}_\Omega = \arg\max_{\mathscr{P}_\Omega} P(\mathscr{P}_\Omega | I) = \arg\max_{\mathscr{P}_\Omega} P(I|\mathscr{P}_\Omega)P(\mathscr{P}_\Omega) \tag{3.4}$$

Assuming conditional independence of $I(\mathbf{x})$ and $I(\mathbf{y})$ for $\mathbf{x} \neq \mathbf{y}$ and taking the negative of the natural logarithm in (3.4), this Bayesian estimation is converted to the following minimization problem:

$$\hat{\mathscr{P}}_\Omega = \arg\min_{\mathscr{P}_\Omega} \sum_k \int_{\mathbf{x}\in\mathbf{R}_k} -\log P(I|\mathbf{R}_k) d\mathbf{x} - \log P(\mathscr{P}_\Omega) \tag{3.5}$$

Most of the variational segmentation algorithms optimize a variant of this functional [9]. The first term is a data term which measures the likelihood of the image data given image models within the partition regions. Generally, the region models are specified by the choice of a parametric distribution. In this case, each region $\mathbf{R}_k$ is characterized by a set of parameters which has to be estimated concurrently with the optimal partition. A typical choice of the region model is the Gaussian distribution:

$$P(I|\mathbf{R}_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(I-\mu_k)^2}{2\sigma_k^2}} \tag{3.6}$$

With this particular choice, the set of parameters of region $\mathbf{R}_k$ is $\{\mu_k, \sigma_k\}$, where μ_k and σ_k denote, respectively, the mean and standard deviation of the image data within $\mathbf{R}_k$. The prior term $P(\mathscr{P}_\Omega)$ imposes prior information on the desired partition. Commonly, the prior introduces geometric constraints on the solution. The length prior is typical:

$$P(\mathscr{P}_\Omega) \propto e^{-\alpha l(K)} \tag{3.7}$$

where α is a positive constant. If we further assume σ_k to be the same for all regions, i.e., $\sigma_k = \sigma \ \forall k$, and if we set α equal to $\frac{\lambda}{\sigma}$, the Bayesian estimation in (3.4) reduces to the minimization of the piecewise constant Mumford-Shah functional. The functional in (3.5) can be viewed as a statistical generalization of the cartoon model, where piecewise constant approximations of the image data within the segmentation regions are replaced by arbitrary probability distributions.

3.1.2 Graduated non convexity implementation

Blake and Zisserman [13] used the *graduated non convexity* continuation method to minimize the Mumford and Shah functional. For simplicity, the method is often described for one-dimensional images. Consider the one-dimensional version of the Mumford-Shah formulation, where the problem is to determine a smooth approximation h of a function h_0 defined on a real interval ω and a set of discontinuities K by minimizing

$$F_{MS}(h,K) = \int_\omega (h-h_0)^2 dx + \rho \int_{\omega \setminus K} h'^2 dx + \lambda \rho \, card(K) \tag{3.8}$$

where *card* denotes the cardinality. We can write a discrete version of F_{MS} as follows:

$$F^0_{MS}(h,K) = \sum_i^n (h^i - h^i_0)^2 + f^0(h^i - h^{i-1}) \tag{3.9}$$

where h^i is the value at point i of h using a discretization grid of points $i = 1,2,\ldots,n$, on Ω, f^0 is defined by

$$f^0(t) = \begin{cases} t^2 & \text{if} \quad t \le 1 \\ 1 & \text{if} \quad t > 1 \end{cases}$$

and K is the set of points verifying

$$K = \{i | h^i - h^{i-1}| > 1\} \tag{3.10}$$

This functional is not convex and can have several local minima. Graduated non convexity is a continuation method which first solves an easy problem using a convex version of F_{MS}. The unique global minimum serves as an initial approximation to solve a nearby non convex problem. The process is iterated to solve a sequence of problems which gradually tends to the original problem. Consider the following version of F_{MS}:

$$F^p_{MS}(h,k) = \sum_i^n (h^i - h^i_0)^2 + f^p(h^i - h^{i-1}) \tag{3.11}$$

where f^p is defined by

$$f^p(t) = \begin{cases} t^2, \quad \text{if} \quad |t| < \frac{1}{r} \\ 1 - \frac{(|t|-r)^2}{4p}, \quad \text{if} \quad \frac{1}{r} < |t| \leq r \\ 1, \quad \text{if} \quad r \leq |t| \\ r = \sqrt{4p+1} \end{cases}$$

One can show that f^p is convex when $p \geq 1$ and, therefore, a global minimum of f^1 can be computed. The graduated non convexity scheme seeks a minimum by decreasing p from 1 to 0. In summary, the graduated non convexity is a continuation scheme which iterates two steps, one minimizes a *coarse* version of the functional and the other uses the result as an initialization to minimize a finer version.

3.2 The minimum description length method of Leclerc

Leclerc stated image segmentation in the discrete case, i.e., for discrete images on a discrete domain, as the problem of determining the shortest description of an image in terms of an a priori specified language [7]. He applied this *minimum description length* (MDL) principle [14] using entropy coding to code approximations of the image to segment. MDL is related to MAP.

3.2.1 MDL and MAP

The MAP strategy seeks a model image M^{opt} which maximizes the model conditional probability given the observed image I:

$$M^{opt} = \arg\max_{M \in \mathscr{M}} P(M|I) = \arg\max_{M \in \mathscr{M}} P(I|M)P(M) \tag{3.12}$$

where $\mathscr{M}$ is the set of model images in which we seek a solution. The *MDL* strategy minimizes

$$|\mathscr{L}_I(I|M)|_b + |\mathscr{L}_M(M)|_b \tag{3.13}$$

where $\mathscr{L}_M(M)$ is a language describing the model images, $\mathscr{L}_I(I/|M)$ the language describing the observed image given a model image, and $|.|_b$ is the number of bits of description. With the piecewise constant model, $\mathscr{M}$ is the set of piecewise constant images underlying an observed image.

In information theory, when the probabilities of the discrete set of observations (data) to describe are known, one can build an optimal descriptive language so that the expected number of bits per description is minimal [15]. In this case, the number of bits describing a given observation is the negative base 2 logarithm of the probability of the observation. For instance, if $\mathscr{L}_M^{opt}$ is the optimal language describing model images M ($M \in \mathscr{M}$) with known probabilities, we have

$$|\mathscr{L}_M^{opt}(M)|_b = -\log_2 P(M) \tag{3.14}$$

Similarly, we have

$$|\mathscr{L}_I^{opt}(I|M)|_b = -\log_2 P(I|M) \tag{3.15}$$

where $\mathscr{L}_I^{opt}$ is the optimal language describing the image given a model. Consequently, the MDL strategy seeks a model image which minimizes

$$-\log_2 P(I|M) - \log_2 P(M) \tag{3.16}$$

and, therefore, is equivalent to the MAP strategy. The choice of either strategy depends on whether it is more convenient to specify a descriptive language or probabilities.

3.2.2 The piecewise constant image model

Assuming that an observed digital image I is the sum of a piecewise constant image M and noise N, both considered discrete, segmentation can be stated as the following optimization problem:

$$\left(M^{opt}, N^{opt}\right) = \arg \min_{M,N: I=M+N} |\mathscr{L}_M(M)|_b + |\mathscr{L}_N(N)|_b \tag{3.17}$$

where $\mathscr{L}_M$ and $\mathscr{L}_N$ are languages describing M and N, respectively. The problem amounts to finding

$$M^{opt} = \arg\min_M |\mathscr{L}_M(M)|_b + |\mathscr{L}_N(I-M)|_b \tag{3.18}$$

To solve this problem, one needs to specify the languages $\mathscr{L}_M$ and $\mathscr{L}_N$, as well as a computationally feasible procedure for finding M^{opt}. Consider piecewise constant images. A region in such images can be described by its boundary and its constant intensity. Region boundaries can be represented by a chain code of edge elements between square pixels. With this description, the number of bits to code a region is proportional to the number of elements in the chain code representation of the region boundary plus a constant to specify the constant intensity in the region and the first element in the chain. As a result, the code length to describe a piecewise constant image M is proportional to the number of regions plus the length of the segmentation boundaries.

In a piecewise constant image, edges are locations which border two regions of different values. This instructs us to approximate the length of the segmentation boundaries as follows:

$$\frac{1}{2} \sum_{i \in D} \sum_{j \in \mathscr{N}_i} (1 - \delta(M_i - M_j)) \tag{3.19}$$

where D is the discrete image domain, indices indicate pixels, $\mathscr{N}_i$ is some fixed neighborhood of pixel i, and δ is given by

$$\delta(z) = \begin{cases} 1 & \text{for} \quad z = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

The division by 2 accounts for each edge being counted twice in the sum of (3.19). The code length of specification of a piecewise constant image M can be approximated as follows:

$$|\mathscr{L}_M(M)|_b = \frac{\beta}{2} \sum_{i \in D} \sum_{j \in N_i} (1 - \delta(M_i - M_j)) \tag{3.21}$$

β is the sum of two terms: (a) the code length of specification of an element in the chain and (b) the code length of specification of the constant intensity and the first element in the chain divided by the average region boundary length.

Now recall that information theory instructs us to set the minimum number of bits required to describe the noise, i.e., the difference between the image and the piecewise constant model, equal to the negative base 2 logarithm of the probability of the noise [15]:

$$|\mathscr{L}_N(I - M)|_b = -\log_2 P(N) \tag{3.22}$$

Assuming the values of the noise at pixels are independent and identically distributed realizations of the same discrete random process, we have

$$-\log_2 P(N) = -\log_2 \prod_{i \in D} P(N_i) \tag{3.23}$$

For a noise that is discretized Gaussian,

$$P(N_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{N_i^2}{2\sigma^2}}, \tag{3.24}$$

the minimum number of bits required to specify N over the image domain grid is, after some algebraic manipulations,

$$|\mathscr{L}_N(I - M)| = c + a \sum_{i \in D} \left(\frac{N_i}{\sigma}\right)^2 = c + a \sum_{i \in D} \left(\frac{I_i - M_i}{\sigma}\right)^2 \tag{3.25}$$

where c and a are constants which depend on the number of pixels in the image and σ. Adding this expression to (3.21) and dropping the additive constants, the problem in (3.18) becomes

$$M^{opt} = \arg\min_M \ L(M)$$

with

$$L(M) = a \sum_{i \in D} \left(\frac{I_i - M_i}{\sigma}\right)^2 + \frac{\beta}{2} \sum_{i \in D} \sum_{j \in N_i} (1 - \delta(M_i - M_j)) \tag{3.26}$$

3.2.3 Numerical implementation

The objective functional (3.26) is not differentiable because of the presence of the δ function. A stochastic method such as simulated annealing can carry out its minimization. However, such methods are notoriously expensive computationally. An advantageous alternative is to embed the minimization (3.26) in a family of minimizations indexed by the parameter of a differentiable approximation of the δ function, and use a *continuation* method [7, 16]. We can index the continuation on the parameter s of the following Gaussian approximation of δ:

$$\delta(M_i - M_j) \quad \leftarrow \quad e_{ij}(M,s) = e^{\frac{-(M_i-M_j)^2}{(s\sigma)^2}} \tag{3.27}$$

Without loss of generality, we use $s\sigma$ in (3.27) rather that just s to simplify some subsequent expressions. The substitution (3.27) inscribes the original objective functional in a family of functions indexed by s

$$L(M,s) = a \sum_{i \in D} \frac{(I_i - M_i)^2}{\sigma^2} + \frac{\beta}{2} \sum_{i \in D} \sum_{j \in \mathcal{N}_i} (1 - e_{ij}(M,s)) \tag{3.28}$$

As s approaches zero, we to tend to the original functional

$$\begin{aligned} \lim_{s \to 0} \; e_{ij}(M,s) &= 0 \\ \lim_{s \to 0} \; L(M,s) &= L(M) \end{aligned} \tag{3.29}$$

For s tending to infinity, the second term in $L(M)$ tends to zero and I is the unique minimum of $L(M,s)$. Consequently, continuation solves a first easy problem at large s; it then gradually decreases s, at each value of which it solves the corresponding problem using the solution to the previous problem for initial approximation. The process continues up to a sufficiently small s. If $s_n, n = 1, 2, \ldots$, is the decreasing sequence of s, the algorithm summary is:

1. For $n = 1$, find the minimum M^1 of $L(M, s_1)$ for some sufficiently large s_1
2. Repeat until convergence
   a) Set $n = n + 1$ and $s_{n+1} = rs_n$ for some $0 < r < 1$
   b) Find a minimum M^{n+1} of $L(M, s_{n+1})$ by a descent algorithm

Step 2.b iterates updates of M starting from the initial value M^n. This algorithm is conceptually similar to the graduated non convexity algorithm in [13]; it tracks a local minimum from a coarse to a fine version of the objective function.

The necessary conditions for a local minimum of $L(M,s)$ result in an equation for each pixel:

$$\frac{\partial L(M,s)}{\partial M_i} = \frac{2a}{\sigma^2}(M_i - I_i) + \frac{2\beta}{(s\sigma)^2}\sum_{i\in\mathscr{N}_i} e_{ij}(M,s)(M_i - M_j) \tag{3.30}$$

This system of equations can be written in vector form:

$$\mathbf{B} + \mathbf{A}(M,s)M = 0 \tag{3.31}$$

where

$$\begin{aligned} \mathbf{A}_{i,i}(M,s) &= \frac{2a}{\sigma^2} + \frac{2\beta}{(s\sigma)^2}\sum_{j\in\mathscr{N}_i} e_{ij}(M,s)(M_i - M_j) \\ \mathbf{A}_{i,j}(M,s) &= \begin{cases} -\frac{2\beta}{(s\sigma)^2} e_{ij}(M,s) \text{ if } j \in \mathscr{N}_i \\ 0 \quad \text{otherwise} \end{cases} \\ b_i &= \frac{-2aI_i}{\sigma^2} \end{aligned} \tag{3.32}$$

System (3.31) is nonlinear. However, it can be solved efficiently by the following Gauss-Seidel iterations applied at each iteration to the diagonally dominant linear system of equations obtained by evaluating the exponential terms at the previous iteration:

$$M_i^{k+1} = \frac{-1}{\mathbf{A}_{i,i}^k}\left(b_i + \sum_{j\neq i}\mathbf{A}_{i,j}^k u_j^k\right) = \frac{I_i + \frac{\beta}{as^2}\sum_{j\in N_i} e_{i,j}^k M_j^k}{1 + \frac{\beta}{as^2}\sum_{j\in N_i} e_{i,j}^k} \tag{3.33}$$

where $\mathbf{A}_{i,j}^k$ are the coefficients of $\mathbf{A}(M,s)$ computed at the previous iteration and $e_{i,j}^k$ is given by $e_{i,j}^k = e_{i,j}(M^k, s)$.

3.3 The region competition algorithm

In this section, we examine the *region competition* algorithm [6] which combines curve evolution and region merging for optimization of the Bayesian estimation functional in (3.5). Assume that the number of regions is fixed and equal to N, and consider the problem of finding an optimal partition into N regions according to

$$\hat{\mathscr{P}}_\Omega = \arg\min_{\mathscr{P}_\Omega} \mathscr{F}_{Zhu-Yuille} \tag{3.34}$$

with

$$\mathscr{F}_{Zhu-Yuille} = \sum_{k=1}^{N}\left(\int_{\mathbf{R}_k} -\log P(I|\mathbf{R}_k)d\mathbf{x} + \frac{\lambda}{2}\oint_{\partial\mathbf{R}_k} ds + \nu\right) \tag{3.35}$$

where $\partial\mathbf{R}_k$ is the boundary of region $\mathbf{R}_k$ parametrized by arc length s; λ and ν are positive constants; $P(I|\mathbf{R}_k)$ is the probability of the image within region $\mathbf{R}_k$, described by a parametric distribution (model) with parameter vector α_k. In this case,

α_k characterizes the image data within region $\mathbf{R}_k$ and the regions are assumed to differ by their parameters.

From a Bayesian perspective, the first term in (3.35) can be viewed as the image likelihood in terms of the region models, and the last two terms are partition priors. From the MDL perspective, functional (3.35) can be viewed as the continuum limit of the Leclerc's discrete cost function, the first term corresponding to the code length of the specification of the image within region $\mathbf{R}_k$ following $P(I|\mathbf{R}_k)$, and the second term to the code length of the specification of the segmentation boundaries. ν is the code length of the specification of the distribution and the code system of region $\mathbf{R}_k$, assumed the same for all regions.

3.3.1 Optimization

Functional (3.35) depends on two types of variables, the segmentation regions determined by the set of boundaries

$$\Gamma = \cup_{k \in [1 \ldots N]} \partial \mathbf{R}_k \tag{3.36}$$

and the regions parameters. The region competition algorithm iterates two stages, one which optimizes the functional by evolution of the set of boundaries Γ, with the number of regions fixed, and the other which alters the number of regions by merging neighboring regions, each stage decreasing the functional.

First stage: curve evolution

With N fixed, this first stage iterates two steps, each decreasing the functional.

First step: The first step computes the optimal regions parameters with Γ fixed

$$\{\hat{\alpha}_k\} = \arg\min_{\alpha_k} \left(- \int_{\mathbf{R}_k} \log P(I/\mathbf{R}_k) d\mathbf{x} \right) \tag{3.37}$$

For several parametric distributions, the optimal parameters can be computed analytically. For instance, in the case of the Gaussian distribution, each region $\mathbf{R}_k$ is characterized by a mean $\mu_{\mathbf{R}_k}$ and a variance $\sigma^2_{\mathbf{R}_k}$

$$\alpha_k = \{\mu_{\mathbf{R}_k}, \sigma_{\mathbf{R}_k}\} \text{ with } P(I|\mathbf{R}_k) = \frac{1}{\sqrt{2\pi\sigma^2_{\mathbf{R}_k}}} e^{-\frac{(I-\mu_{\mathbf{R}_k})^2}{2\sigma^2_{\mathbf{R}_k}}} \tag{3.38}$$

The optimal parameters correspond to the maximum likelihood estimates (MLE) which verify the necessary conditions, for $k = 1 \ldots N$

$$\begin{aligned} \frac{\partial \left(\int_{\mathbf{R}_k} \log P(I|\mathbf{R}_k) d\mathbf{x} \right)}{\partial \mu_{\mathbf{R}_k}} &= 0 \\ \frac{\partial \left(\int_{\mathbf{R}_k} \log P(I|\mathbf{R}_k) d\mathbf{x} \right)}{\partial \sigma_{\mathbf{R}_k}} &= 0 \end{aligned} \tag{3.39}$$

In the case of the Gaussian distribution, these conditions yield simply the sample mean and variance of the image data within the segmentation regions

$$\begin{aligned}\hat{\mu}_{\mathbf{R}_k} &= \frac{\int_{\mathbf{R}_k} I d\mathbf{x}}{\int_{\mathbf{R}_k} d\mathbf{x}} \\ \hat{\sigma}^2_{\mathbf{R}_k} &= \frac{\int_{\mathbf{R}_k} (I-\hat{\mu}_{\mathbf{R}_k})^2 d\mathbf{x}}{\int_{\mathbf{R}_k} d\mathbf{x}}\end{aligned} \tag{3.40}$$

Second step: With the region parameters fixed, the second step minimizes the functional with respect to Γ by solving the following descent equation

$$\frac{\partial \Gamma}{\partial t} = -\frac{\partial \mathscr{F}_{Zhu-Yuille}}{\partial \Gamma} \tag{3.41}$$

Let $\mathbf{p} = (x(s), y(s))$ be a point on a boundary $\Gamma_k = \partial \mathbf{R}_k$ which encloses region $\mathbf{R}_k$, where s is arc length and $(x(s), y(s))$ the corresponding parametric representation of $\partial \mathbf{R}_k$. Recall from Chapter 2 that the functional derivative with respect to a simple closed plane curve γ of $\int_{\mathbf{R}_\gamma} f d\mathbf{x}$, where f is independent of $\mathbf{R}_\gamma$, is

$$\frac{\partial \int_{\mathbf{R}_\gamma} f d\mathbf{x}}{\partial \gamma} = f\mathbf{n} \tag{3.42}$$

where $\mathbf{n}$ is the outward unit normal function of γ. Recall also that the functional derivative of $\int_\gamma ds$ is

$$\frac{\partial \int_\gamma ds}{\partial \gamma} = \kappa \mathbf{n} \tag{3.43}$$

where κ is the curvature function of γ. The application to (3.41) gives the following evolution equation at each point $\mathbf{p}$ on Γ

$$\frac{\partial \Gamma}{\partial t}(\mathbf{p}) = \sum_{k \in \mathbf{b_p}} \left(\log P(I(\mathbf{p})|\mathbf{R}_k) - \frac{\lambda}{2} \kappa_k(\mathbf{p}) \right) \mathbf{n}_k(\mathbf{p}) \tag{3.44}$$

where $\mathbf{b_p}$ is the set of indices of boundaries Γ_k containing point $\mathbf{p}$

$$\mathbf{b_p} = \{k | \mathbf{p} \text{ is on } \Gamma_k\}, \tag{3.45}$$

$\mathbf{n}_k$ the external unit normal to $\partial \mathbf{R}_k$, and κ_k the curvature of Γ_k.

Alternating the two steps decreases the functional and curve evolution will converge because the functional is lower bounded.

Second stage: region merging

The curve evolution stage does not alter the number of regions. To optimize the functional with respect to the number of regions, one can alternate curve evolution with traditional statistical test region merging [6]. To merge neighboring regions, the merging must decrease the functional. When $P(I|\mathbf{R}_k)$ is Gaussian, merging of two neighboring regions $\mathbf{R}_i$ and $\mathbf{R}_j$ yields a new region $\mathbf{R}_{ij} = \mathbf{R}_i \cup \mathbf{R}_j$ with the following mean and variance:

$$\begin{aligned}
\mu_{\mathbf{R}_{ij}} &= \frac{\mu_{\mathbf{R}_i} A(\mathbf{R}_i) + \mu_{\mathbf{R}_j} A(\mathbf{R}_j)}{A(\mathbf{R}_i) + A(\mathbf{R}_j)} \\
\sigma^2_{\mathbf{R}_{ij}} &= \frac{\sigma_i^2 A(\mathbf{R}_i) + \sigma_j^2 A(\mathbf{R}_j) + \frac{A(\mathbf{R}_i) A(\mathbf{R}_j)(\mu_{\mathbf{R}_i} - \mu_{\mathbf{R}_j})^2}{A(\mathbf{R}_i) + A(\mathbf{R}_j)}}{A(\mathbf{R}_i) + A(\mathbf{R}_j)}
\end{aligned} \tag{3.46}$$

where $A(\mathbf{R})$ is the area of region $\mathbf{R}$:

$$A(\mathbf{R}) = \int_{\mathbf{R}} d\mathbf{x} \tag{3.47}$$

The variation of the functional corresponding to merging $\mathbf{R}_i$ and $\mathbf{R}_j$ is

$$\begin{aligned}
\Delta \mathscr{F}_{Zhu-Yuille} &= -\lambda \int_{\partial \mathbf{R}_i \cap \partial \mathbf{R}_j} ds - \nu \\
&\quad + \int_{\mathbf{R}_i} \log P(I|\mathbf{R}_i) d\mathbf{x} + \int_{\mathbf{R}_j} \log P(I|\mathbf{R}_j) d\mathbf{x} - \int_{\mathbf{R}_{ij}} \log P(I|\mathbf{R}_{ij}) d\mathbf{x} \\
&= -\lambda l(\mathbf{R}_i \cap \partial \mathbf{R}_j) - \nu + \frac{A(\mathbf{R}_i) \log \frac{\sigma^2_{\mathbf{R}_{ij}}}{\sigma^2_{\mathbf{R}_j}} + 1}{2}
\end{aligned} \tag{3.48}$$

Regions $\mathbf{R}_i$ and $\mathbf{R}_j$ are merged when $\Delta \mathscr{F}_{Zhu-Yuille} < 0$.

Algorithm

The region competition algorithm can be summarized as follows:

1. Initialize the segmentation to a partition using seed regions
2. Repeat until convergence
   a) Repeat until the evolution of Γ converges
      i. Fix Γ and compute the optimal regions parameters $\hat{\alpha}_k$
      ii. Evolve Γ according to (3.44)
   b) Merge any two neighboring regions if the merging decreases the functional

Interpretation of the curve evolution equation

The evolution equation in (3.44) contains two velocities along the normal to Γ; one is an image dependent velocity and can be viewed as an image likelihood ratio test, whereas the other depends solely on the curvature of Γ, not on the image, and can be viewed as a curvature regularization velocity.

The image dependent velocity: To examine the effect of the image velocity, consider the case of a point $\mathbf{p}$ on the interface between two neighboring regions $\mathbf{R}_i$ and $\mathbf{R}_j$, as illustrated in Figure 3.1. In this case, we have $\mathbf{n}_i(\mathbf{p}) = -\mathbf{n}_j(\mathbf{p})$ and $\kappa_i(\mathbf{p})\mathbf{n}_i(\mathbf{p}) = \kappa_j(\mathbf{p})\mathbf{n}_j(\mathbf{p})$. Therefore, the evolution equation in (3.44) reads:

$$\begin{aligned}\frac{\partial \mathbf{p}}{\partial t} &= -\lambda \kappa_i(\mathbf{p})\mathbf{n}_i(\mathbf{p}) + (\log P(I(\mathbf{p})/\mathbf{R}_i) - \log P(I(\mathbf{p})/\mathbf{R}_j))\,\mathbf{n}_i(\mathbf{p}) \\ &= -\lambda \kappa_i(\mathbf{p})\mathbf{n}_i(\mathbf{p}) + \log \frac{P(I(\mathbf{p})/\mathbf{R}_i)}{P(I(\mathbf{p})/\mathbf{R}_j)}\mathbf{n}_i(\mathbf{p}) \qquad (3.49)\end{aligned}$$

If we omit the curvature velocity, this evolution of Γ at point $\mathbf{p}$ becomes guided by the following image likelihood ratio test, which amounts to a competition between $\mathbf{R}_i$ and $\mathbf{R}_j$ as to which region should include $\mathbf{p}$:

- If $\frac{P(I(\mathbf{p})/\mathbf{R}_i)}{P(I(\mathbf{p})/\mathbf{R}_j)} > 1$, i.e., the likelihood of the image at pixel $\mathbf{p}$ is higher when $\mathbf{p}$ belongs to $\mathbf{R}_i$, the image dependent velocity is positive. This causes the curve to expand and take pixel $\mathbf{p}$ within region $\mathbf{R}_i$.

- If $\frac{P(I(\mathbf{p})/\mathbf{R}_i)}{P(I(\mathbf{p})/\mathbf{R}_j)} < 1$, i.e., the likelihood of the image at pixel $\mathbf{p}$ is higher when $\mathbf{p}$ belongs to $\mathbf{R}_j$, the image dependent velocity is negative. Therefore, the curve shrinks and leaves $\mathbf{p}$ in region $\mathbf{R}_j$.

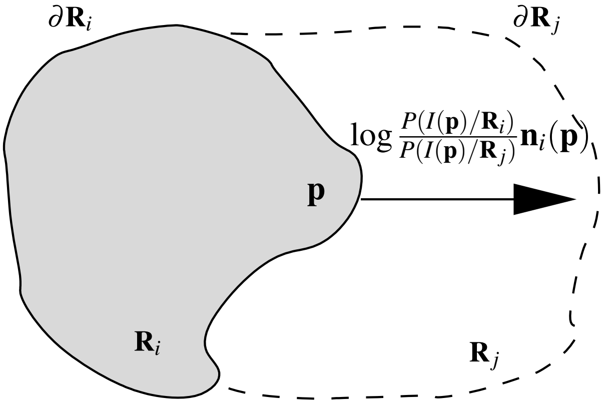


Fig. 3.1. Competition between two regions under the effect of the image dependent velocity.

The curvature dependent velocity: The curvature regularization velocity, also called the Euclidean heat flow [17, 18], has the desirable geometric effect of shortening and smoothing the curve (illustration in Figure 3.2).

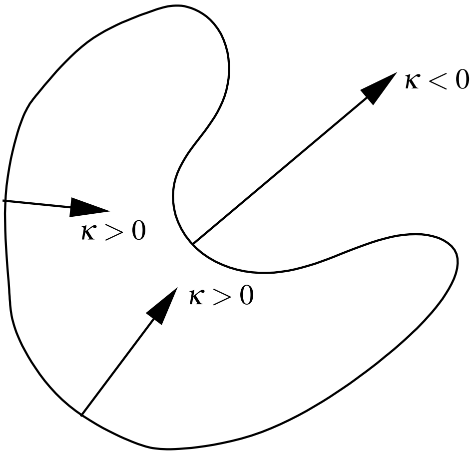


Fig. 3.2. The curvature evolution has the geometric effect of shortening and smoothing the curve.

3.4 A level set formulation of the piecewise constant Mumford-Shah model

The implementation of the region competition algorithm requires an explicit representation of the segmentation boundary curve as a set of points. This implies implementing the curve evolution by explicitly displacing a number of marker points on the curve. With this implementation, errors in the curve position can accumulate during evolution to cause inaccurate and unstable tracking. Moreover, changes in the curve topology, when the curve splits and joins during evolution, can be practically unrealizable [19]. Finally, the region membership of points, needed in various computations, is expensive to determine. Merging regions is also expensive to effect. An efficient, numerically stable alternative which would not be prone to such problems is to represent the evolving curve implicitly by the zero level set of a function. Level sets handle automatically arbitrary topological variations of the active curve, and region membership information is readily available. In this section, we examine the Chan and Vese level set implementation of the piecewise constant Mumford-Shah formulation in the case of two regions [20].

Let $\gamma : s \in [0,1] \to \mathbf{x}(s) = (x(s), y(s)) \in \mathbb{R}^2$ be a simple closed plane curve parametrized by arc length s. γ defines a partition of Ω into two regions,

$$\mathbf{R}_1 = \mathbf{R}_\gamma$$

corresponding to the interior of the curve, and

$$\mathbf{R}_2 = \mathbf{R}_\gamma^c$$

corresponding to the exterior. The problem consists of finding a piecewise constant approximation of image I,

$$\hat{I} = \begin{cases} \mu_1 \text{ in } \mathbf{R}_1 \\ \mu_2 \text{ in } \mathbf{R}_2, \end{cases}$$

following the minimization with respect to γ and $\{\mu_1, \mu_2\}$ of the functional

$$\mathscr{F}_{Chan-Vese} = \lambda_1 \int_{\mathbf{R}_1} (I-\mu_1)^2 \, d\mathbf{x} + \lambda_2 \int_{\mathbf{R}_2} (I-\mu_2)^2 \, d\mathbf{x} + \lambda \int_{\gamma} ds + \nu \int_{\mathbf{R}_1} d\mathbf{x} \quad (3.50)$$

The first two terms are data terms which measure the conformity of the image within each region to the piecewise constant model. These can be viewed as a special case of the data term of the region competition formulation, the case of Gaussian region distributions with variance 1. The last two terms are regularization terms, one measures the curve length and the other the area of the region the curve encloses. These terms promote boundary smoothness and inhibit small and isolated regions. λ_1, λ_2, λ, and ν are positive constants to modulate the contribution of each term. The Chan-Vese functional corresponds to the Mumford-Shah model of piecewise constant images and two-region partitioning. The simple example in Figure 3.3 illustrates the relevance of the Chan-Vese functional. The figure represents a region (the square object) against a background, both with constant intensities, μ_1 and μ_2, respectively. Over all possible positions of the curve, the sum of the data terms is minimized only when the curve coincides with the boundary of the square. If the curve is inside the object, we have $\mathbf{F}_1(\gamma) = \int_{\mathbf{R}_1} (I-\mu_1)^2 \, d\mathbf{x} = 0$ and $\mathbf{F}_2(\gamma) = \int_{\mathbf{R}_2} (I-\mu_2)^2 \, d\mathbf{x} > 0$. If the curve is outside, we have $\mathbf{F}_1(\gamma) > 0$ and $\mathbf{F}_2(\gamma) = 0$. If the curve intersects the boundary of the object, we have $\mathbf{F}_1(\gamma) > 0$ and $\mathbf{F}_2(\gamma) > 0$. These three configurations yield a strictly positive sum of the data terms. When the curve coincides with the boundary of the square, both data terms are zero.

3.4.1 Curve evolution minimization of the Chan-Vese functional

The Chan-Vese functional depends on two types of variables, curve γ and the regions parameters μ_1 and μ_2. Its minimization can be obtained by alternating two steps, one to optimize the functional with respect to γ by curve evolution with the parameters fixed and the other to determine the optimal parameters with γ fixed. The optimal parameters can be computed analytically, and turn out to be the sample means within the regions. With the parameters fixed, the curve evolution equation minimizing $\mathscr{F}_{Chan-Vese}$ with respect to γ is obtained by embedding γ in a one-parameter family of curves $\gamma : s,t \in [0,1] \times \mathbb{R}^+ \rightarrow \gamma(s,t) = \mathbf{x}(s,t) = (x(s,t), y(s,t), t) \in \Omega \times \mathbb{R}^+$ indexed by algorithmic time t, and solving the following descent equation:

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{F}_{Chan-Vese}}{\partial \gamma} \quad (3.51)$$

To compute the functional derivatives with respect to γ of the region integrals in the Chan-Vese functional, we use the basic formula in Equation (3.42), which yields

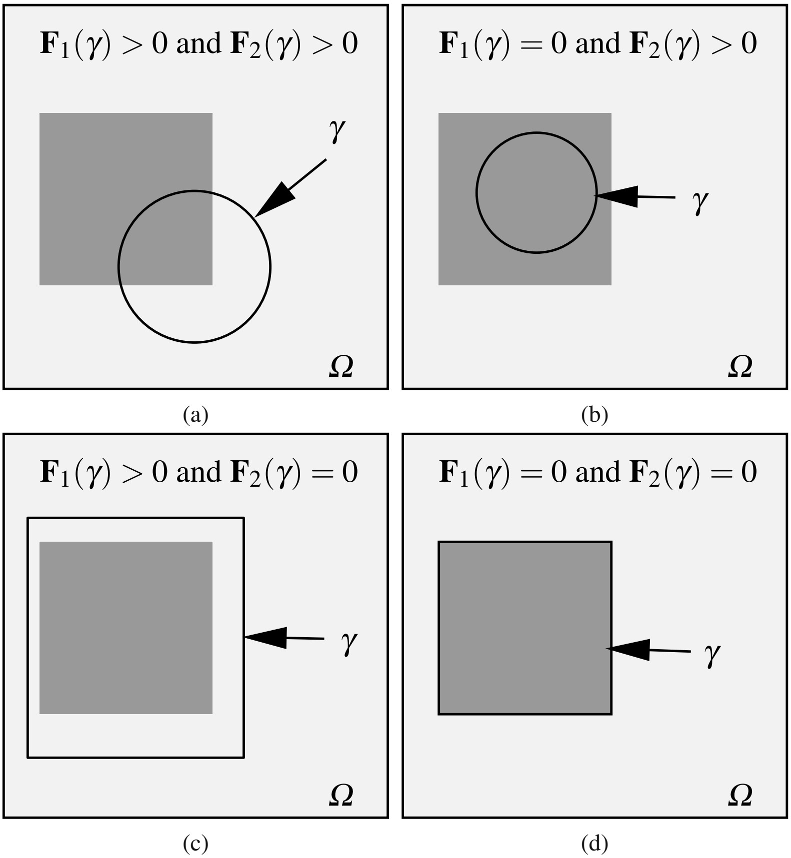


Fig. 3.3. Over all possible positions of the curve, the Chan-Vese data term is minimized only when the curve coincides with the boundary of the square object.

$$
\begin{aligned}
\frac{\partial \int_{\mathbf{R}_1} (I-\mu_1)^2 \, d\mathbf{x}}{\partial \gamma} &= (I-\mu_1)^2 \, \mathbf{n} \\
\frac{\partial \int_{\mathbf{R}_2} (I-\mu_2)^2 \, d\mathbf{x}}{\partial \gamma} &= -(I-\mu_2)^2 \, \mathbf{n} \\
\frac{\partial \int_{\mathbf{R}_1} d\mathbf{x}}{\partial \gamma} &= \mathbf{n}
\end{aligned}
\tag{3.52}
$$

where $\mathbf{n}$ is the outward unit normal to γ. The minus sign in the second equation of the system above is due to the fact that $\mathbf{n}$ being the external unit normal to $\mathbf{R}_1$, the external unit normal to its complement $\mathbf{R}_2$ is $-\mathbf{n}$. The derivatives above and the derivative of the length term in Equation (3.43) give the final evolution equation

$$\frac{\partial \gamma}{\partial t} = -\left(\lambda_1 (I-\mu_1)^2 - \lambda_2 (I-\mu_2)^2 + \lambda \kappa + \nu\right) \mathbf{n} \tag{3.53}$$

Note that in this alternating minimization with respect to the curve and the region parameters, we omitted the dependence of the integrands in the region integrals on the integration domains (the regions). Recall that these integrands depend on the region means and, therefore, on the regions. It is possible to compute accurate curve evolution, which takes into account such dependence [21], but this is inconsequential in the case of the Chan-Vese functional and will lead exactly to the same evolution equation we obtained above.

3.4.2 Level set representation of curve evolution

We recall that with the level set representation, an active curve γ is recovered at all times t as the zero level set of a function $\phi : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$

$$\forall s,t \;\; \phi(\mathbf{x}(s,t),t) = 0, \tag{3.54}$$

and region membership is readily determined by the sign of the level set function. For instance, one can choose the convention that ϕ is positive inside its zero level set and negative outside, i.e.,

$$\begin{aligned} \mathbf{R}_1 &= \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) > 0\} \\ \mathbf{R}_2 &= \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) < 0\} \end{aligned} \tag{3.55}$$

Recall (Chapter 2) that if curve γ evolves according to

$$\frac{\partial \gamma}{\partial t} = V\mathbf{n}, \tag{3.56}$$

ϕ evolves on its zero level according to

$$\frac{\partial \phi}{\partial t} = V\|\nabla \phi\| \tag{3.57}$$

This leads to the level set evolution equation minimizing the Chan-Vese functional

$$\frac{\partial \phi}{\partial t} = -\left(\lambda_1 (I-\mu_1)^2 - \lambda_2 (I-\mu_2)^2 + \lambda \kappa + \nu\right) \|\nabla \phi\| \tag{3.58}$$

where curvature κ is given by

$$\kappa = -\mathrm{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right) = -\frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{\frac{3}{2}}} \tag{3.59}$$

The segmentation is reached at convergence, i.e., when $t \to +\infty$. A priori, the level set evolution is specified for points on the level set function zero level. Therefore, one must define extension velocities to evolve the level set function elsewhere [22].

As mentioned earlier, the extension velocity at any point can be set equal to the velocity of the closest evolving curve point. Extension velocities can also be set so that the level set function is the distance function from the evolving curve. Both of these definitions, which can be implemented via updates of the level set function in a narrow band around the zero level set, require that the initial curves intersect the regions they segment. In this book, we use an alternative whereby the extension velocities are computed everywhere according to the same generic expression as for the points on the active curve.

It is possible to introduce the level set representation of curve γ directly in the segmentation functional [20] with the approximation

$$\begin{aligned}\mathscr{F}_{Chan-Vese} \approx \lambda_1 \int_\Omega (I-\mu_1)^2 H_\varepsilon(\phi) d\mathbf{x} + \lambda_2 \int_\Omega (I-\mu_2)^2 (1-H_\varepsilon(\phi)) d\mathbf{x} \\ + \lambda \int_\Omega \delta_\varepsilon(\phi) \|\nabla\phi\| d\mathbf{x} + \nu \int_\Omega H_\varepsilon(\phi) d\mathbf{x},\end{aligned} \tag{3.60}$$

where H_ε and $\delta_\varepsilon = H'_\varepsilon$ are regularized approximations of the Heaviside function and Dirac measure:

$$H_\varepsilon(z) = \begin{cases} 1 \text{ if } z > \varepsilon \\ 0 \text{ if } z < -\varepsilon \\ \frac{1}{2}[1 + \frac{z}{\varepsilon} + \frac{1}{\pi}\sin(\frac{\pi z}{\varepsilon})] \text{ if } |z| \le \varepsilon \text{ with } \varepsilon > 0 \end{cases}$$

The evolution equation minimizing (3.60) can be computed directly by applying the Euler-Lagrange descent equation with respect to ϕ

$$\frac{\partial \phi}{\partial t} = -\delta_\varepsilon(\phi)\left(\lambda_1 (I-\mu_1)^2 - \lambda_2 (I-\mu_2)^2 - \lambda \operatorname{div}\left(\frac{\nabla\phi}{\|\nabla\phi\|}\right) + \nu\right) \tag{3.61}$$

Note that due to the Dirac δ_ε, this evolution equation amounts to updating the level set function in a narrow band around the zero level set, which requires that the initial curves intersect the regions they segment. An alternative robust to initialization is to extend the evolution to the hole image domain by replacing δ_ε by $\|\nabla\phi\|$ [23]. This would lead to the same evolution equation we obtained earlier in (3.58).

3.4.3 Algorithm summary

The Chan and Vese algorithm can be summarized as follows:

1. Initialize the level set function ϕ
2. Repeat until the evolution of ϕ converges
   a) Update the region parameters as function of ϕ and the image
   $\mu_1 = \frac{\int_{\phi>0} I d\mathbf{x}}{\int_{\phi>0} I d\mathbf{x}}$ and $\mu_2 = \frac{\int_{\phi<0} I d\mathbf{x}}{\int_{\phi<0} I d\mathbf{x}}$
   b) Update ϕ according to evolution equation (3.58)

Note that, in the implementation, the region parameter updates do not need the expressions in step 2.a at each iteration. They can be updated more efficiently using their previous value and pixels corresponding to a change in the sign of the level set function.

3.4.4 Numerical implementation details of the level set evolution equation

The Chan-Vese functional, as well as various other functionals that we will examine in subsequent chapters, result in a level set evolution equation in the general from

$$\frac{\partial \phi}{\partial t} = (V - \lambda \kappa) \|\nabla \phi\| \tag{3.62}$$

where $V : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$ is a scalar function which depends on position and time. In the case of the Chan-Vese functional, V is given by

$$V = -\left(\lambda_1 (I - \mu_1)^2 - \lambda_2 (I - \mu_2)^2 + \nu\right) \tag{3.63}$$

Discretization of evolution equations of the form (3.62) can be written using the finite difference method in [22]:

$$\phi_{ij}^{k+1} = \phi_{ij}^{k} + \triangle t \left(\max(V_{ij}^k, 0)\nabla^+ + \min(V_{ij}^k, 0)\nabla^- - \lambda \kappa_{ij}^k \left((D_{ij}^{0x})^2 + (D_{ij}^{0y})^2 \right)^{\frac{1}{2}} \right) \tag{3.64}$$

where i, j are indices on the discretization grid of Ω, and k is the iteration index. ∇^+ and ∇^- are defined by

$$\begin{aligned}
\nabla^+ = {} & \Big(\max(D_{ij}^{-x}, 0)^2 + \min(D_{ij}^{+x}, 0)^2 \\
& + \max(D_{ij}^{-y}, 0)^2 + \min(D_{ij}^{+y}, 0)^2 \Big)^{\frac{1}{2}} \\
\nabla^- = {} & \Big(\max(D_{ij}^{+x}, 0)^2 + \min(D_{ij}^{-x}, 0)^2 \\
& + \max(D_{ij}^{+y}, 0)^2 + \min(D_{ij}^{-y}, 0)^2 \Big)^{\frac{1}{2}}
\end{aligned} \tag{3.65}$$

where finite difference $x-$derivative operators D^{+x} (forward scheme), D^{-x} (backward scheme), and D^{0x} (central scheme), are applied to ϕ at i, j and iteration k, i.e., $D_{ij}^{+x}, D_{ij}^{-x}, D_{ij}^{0x}$ in (3.64) stand for $D^{+x}(\phi^k)_{ij}, D^{-x}(\phi^k)_{ij}, D^{0x}(\phi^k)_{ij}$ and are given by

$$\begin{aligned}
D_{ij}^{+x} &= \phi_{i+1,j}^k - \phi_{ij}^k \\
D_{ij}^{-x} &= \phi_{ij}^k - \phi_{i-1,j}^k \\
D_{ij}^{0x} &= \frac{1}{2}(\phi_{i+1,j}^k - \phi_{i-1,j}^k)
\end{aligned}$$

Similar comments and expressions apply to the $y-$derivative operators D^{+y}, D^{-y}, and D^{0y}.

The curvature is discretized by computing the gradient of ϕ and its partial derivatives at the current iteration using central differences:

$$\kappa_{i,j}^{k} = -\frac{D_{ij}^{xx}(D_{ij}^{0y})^2 - 2D_{ij}^{0x}D_{ij}^{0y}D_{ij}^{xy} + D_{ij}^{yy}(D_{ij}^{0x})^2}{((D_{ij}^{0x})^2 + (D_{ij}^{0y})^2)^{\frac{3}{2}}} \tag{3.66}$$

where operators D^{xy}, D^{xx}, and D^{yy} are applied to ϕ at i, j and iteration k:

$$\begin{aligned} D_{ij}^{xy} &= \frac{\phi_{i+1,j+1}^{k} - \phi_{i+1,j-1}^{k} - \phi_{i-1,j+1}^{k} + \phi_{i-1,j-1}^{k}}{4} \\ D_{ij}^{xx} &= \phi_{i+1,j}^{k} - 2\phi_{ij}^{k} + \phi_{i-1,j}^{k} \\ D_{ij}^{yy} &= \phi_{i,j+1}^{k} - 2\phi_{ij}^{k} + \phi_{i,j-1}^{k} \end{aligned} \tag{3.67}$$

3.5 Edge-based approaches

In this section, we examine *edge based* approaches, where curve evolution is guided by a description of the image data on the active curve. Whereas region based approaches seek a partition of the image domain using statistics of the image data within the segmentation regions, edge based methods detect the boundaries of the regions by referring to properties along the curve. Generally, edge guided schemes use the assumption that region boundaries correspond to high image transitions (significant edges), in addition to geometric constraints. The *Snakes* [24] and the *geodesic active contours* [25] have been the first developments along this vein.

3.5.1 The Kass-Witkin-Terzopoulos Snakes model

The study in [24] has been a strong precursor of edge detection by curve evolution. In this model, the evolution of a simple closed plane parametric curve, $\gamma(q) = (x(q), y(q)) : [a,b] \subset \mathbb{R} \to \Omega \subset \mathbb{R}^2$, is sought by minimizing the following functional:

$$\mathscr{F}_{Snakes}(\gamma) = \alpha \int_a^b \|\gamma'\|^2 dq + \beta \int_a^b \|\gamma''\|^2 dq - \lambda \int_a^b \|\nabla I(\gamma(q))\|^2 dq \tag{3.68}$$

where

$$\gamma'(q) = \left(\frac{dx(q)}{dq}, \frac{dy(q)}{dq} \right)$$

and a corresponding notation for γ''. Factors α, β, λ are positive constants to weigh the contribution of each term.

The first two terms, referred to as the *internal* energy, enforce curve smoothness. The third term is image data dependent and is called the *external* energy; it drives the curve to places of high gradient magnitude (significant edges) which is thought to characterize the region boundaries.

This functional has shortcomings. First, it is parametrization dependent; different curve parameterizations can produce different results. Second, there is no mechanism to allow the curve to change topology during evolution. As a result, only a single connected region against a background can be segmented. Third, the numerical implementation requires updates in time of a set of marker points on the curve following approximations of the evolution equation, which can lead to numerical instability. In some cases, unwanted effects, such as overlap of the marker points, may occur. Further discussions can be found in [19].

3.5.2 The Geodesic active contour

The geodesic active contour uses the following intrinsic edge based functional [25, 26]:

$$\mathscr{F}_{GAC} = \int_a^b g\left(\|\nabla I(\gamma(q))\|\right) \|\gamma'(q)\| dq \tag{3.69}$$

where g is a positive monotonically decreasing edge indicator function verifying $lim_{t \to +\infty} g(t) = 0$. A common choice of g is

$$g(t) = \frac{1}{1+t^2} \tag{3.70}$$

As with the Snakes functional, the minimization of this functional attracts the active contour toward high transitions of the image data (significant edges) while smoothing and shortening it. However, $\mathscr{F}_{GAC}$ has important advantages. First, it is intrinsic, i.e., parametrization invariant. To see this, consider another parametrization

$$q = \varphi(r), \varphi : [c,d] \to [a,b], \varphi' > 0$$

we have

$$\mathscr{F}_{GAC} = \int_a^b g\left(\|\nabla I(\bar{\gamma}(r))\|\right) \|\bar{\gamma}'(r)\| dr \tag{3.71}$$

with

$$\bar{\gamma}(r) = \gamma(\varphi(r)),$$

which shows that the functional is parametrization invariant.

Parameterizing the curve by the Euclidean arc length s, which verifies

$$ds = \|\gamma'(q)\| dq, \tag{3.72}$$

gives

$$\mathscr{F}_{GAC} = \int_0^{l(\gamma)} g\left(\|\nabla I(\gamma(s))\|\right) ds \tag{3.73}$$

where $l(\gamma)$ is the Euclidean length of the curve. $\mathscr{F}_{GAC}$ can be interpreted as a generalized length, where the Euclidean element of length ds is weighed by the edge significance function g. The problem, then, consists of finding a minimal length curve (a geodesic) using this image induced length measure (hence the name geodesic active contour [25]).

Another important advantage of this functional is that, as we will see subsequently, the ensuing curve evolution equation is amenable to the level set representation, which affords a stable and efficient numerical implementation. Furthermore, the functional does not require *ad hoc* parameters to weight the contribution of different terms, unlike the Snakes functional.

The curve evolution equation minimizing $\mathscr{F}_{GAC}$ is obtained by embedding γ in a one-parameter family of curves $\gamma : s,t \in [0,l] \times \mathbb{R}^+ \to \gamma(s,t) = (x(s,t), y(s,t), t) \in \Omega \times \mathbb{R}^+$ indexed by algorithmic time t, and solving the descent equation

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{F}_{GAC}}{\partial \gamma} \tag{3.74}$$

The functional derivative with respect to γ of the integral over γ of a positive scalar function $h = h(x(s), y(s))$ is given by (Chapter 2)

$$\frac{\partial \oint_\gamma h ds}{\partial \gamma} = \left(\langle \nabla h, \mathbf{n} \rangle + h\kappa\right) \mathbf{n} \tag{3.75}$$

where $\langle . \rangle$ indicates the scalar product. Applied to $\mathscr{F}_{GAC}$, this result leads to the geodesic active contour evolution:

$$\frac{\partial \gamma}{\partial t} = -\left(g\left(\|\nabla I(\gamma)\|\right) \kappa + \langle \nabla g\left(\|\nabla I(\gamma)\|\right), \mathbf{n} \rangle\right) \mathbf{n} \tag{3.76}$$

The first velocity is the curvature regularization velocity (called the Euclidean heat flow) weighted by the edge indicator function g. This velocity allows curve smoothing and shortening modulated by g, which inhibits the contour motion at significant edges. Theoretically, motion stops at ideal edges where $g = 0$ (hence its name of *stopping function*). This is not the case with digital edges. The second velocity attracts the contour towards significant edges because the gradient vector ∇g points toward the image high transitions. This velocity refines edge detection when the stopping function is not sufficiently low everywhere on the boundary of interest, in which case its effect becomes important.

Because the geodesic active contour is based on image information only on the active contour, it may evolve slowly and in the wrong direction when the initialization is far from the edge of interest. One way to speed up curve evolution and control its direction is to add a constant velocity, often referred to as a *balloon force velocity* [25]:

$$- \nu g\left(\|\nabla I(\gamma)\|\right)\mathbf{n} \tag{3.77}$$

This constant velocity can be viewed as the flow optimizing the following weighted area of the region enclosed within the active curve:

$$\nu \int_{\mathbf{R}_\gamma} g\left(\|\nabla I\|\right) d\mathbf{x} \tag{3.78}$$

Coefficient ν can be of any sign, depending on the initial position of the curve with respect to the desired boundary. If the initial curve encloses the desired boundary, ν should be positive to speed up curve shrinking. If the initial curve is inside the desired edge, ν should be negative to speed up curve expansion. Note that region-based flows, such as the Chan-Vese flow, are, generally, more robust to the initial positioning of the active curve [9] because they are guided by image information in all the image domain, rather than only along the curve.

The level set evolution equation corresponding the geodesic active contour, with a balloon velocity added, is

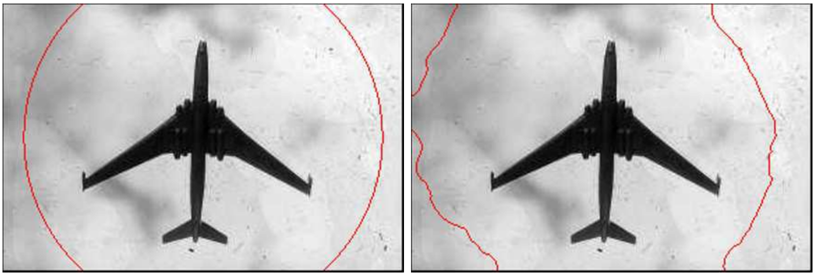
$$\frac{\partial \phi}{\partial t} = g\left(\|\nabla I\|\right)\|\nabla \phi\| \mathrm{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right) + < \nabla g\left(\|\nabla I\|\right), \nabla \phi > - \nu g\left(\|\nabla I(\gamma)\|\right)\|\nabla \phi\| \tag{3.79}$$

Unlike Snakes, the geodesic active contour model accommodates the level set representation, thereby handling implicitly variations in the topology of the active curve. Therefore, it does not require prior knowledge as to the topology of the solution; it can detect the boundaries of an arbitrary number of disconnected objects.

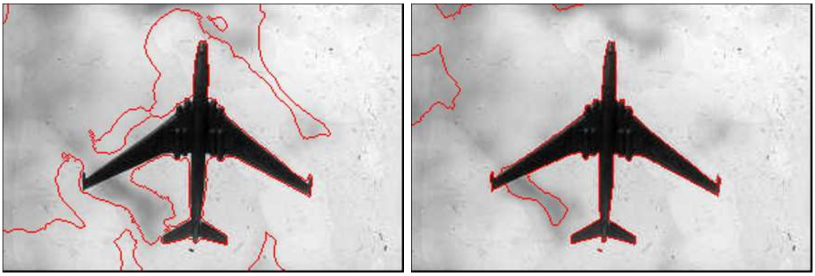
3.5.3 Examples

The example in Figure 3.4 illustrates the segmentation of an aeroplane image with the Chan-Vese model. The initial curve is shown in (a), intermediate curves in (b), (c) and (d), the final curve in (e) and the segmentation represented with the region mean parameters obtained at convergence in (f).

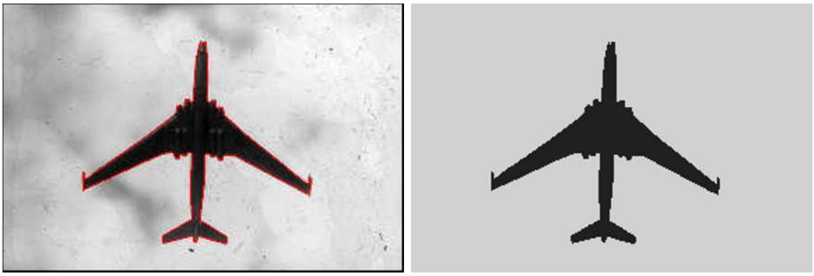
The example in Figure 3.5 illustrates edge detection in a lung image with the geodesic active contour. The initial curve is shown in (a), intermediate curves in (b), (c) and (d), the final curve in (e) and the region within the final curve in (f). This example illustrates how the level set representation handles implicitly arbitrary variations in the curve topology.



(a) Initial curve (b) iteration 125



(c) iteration 250 (d) iteration 375



(e) Final curve (iteration 600) (f) mean parameters

Fig. 3.4. Segmentation example with the Chan-Vese model: (a) initial curve; (b), (c) and (d) intermediate positions of the curve; (e) final position of the curve; (f) segmentation represented with the region mean parameters obtained at convergence.

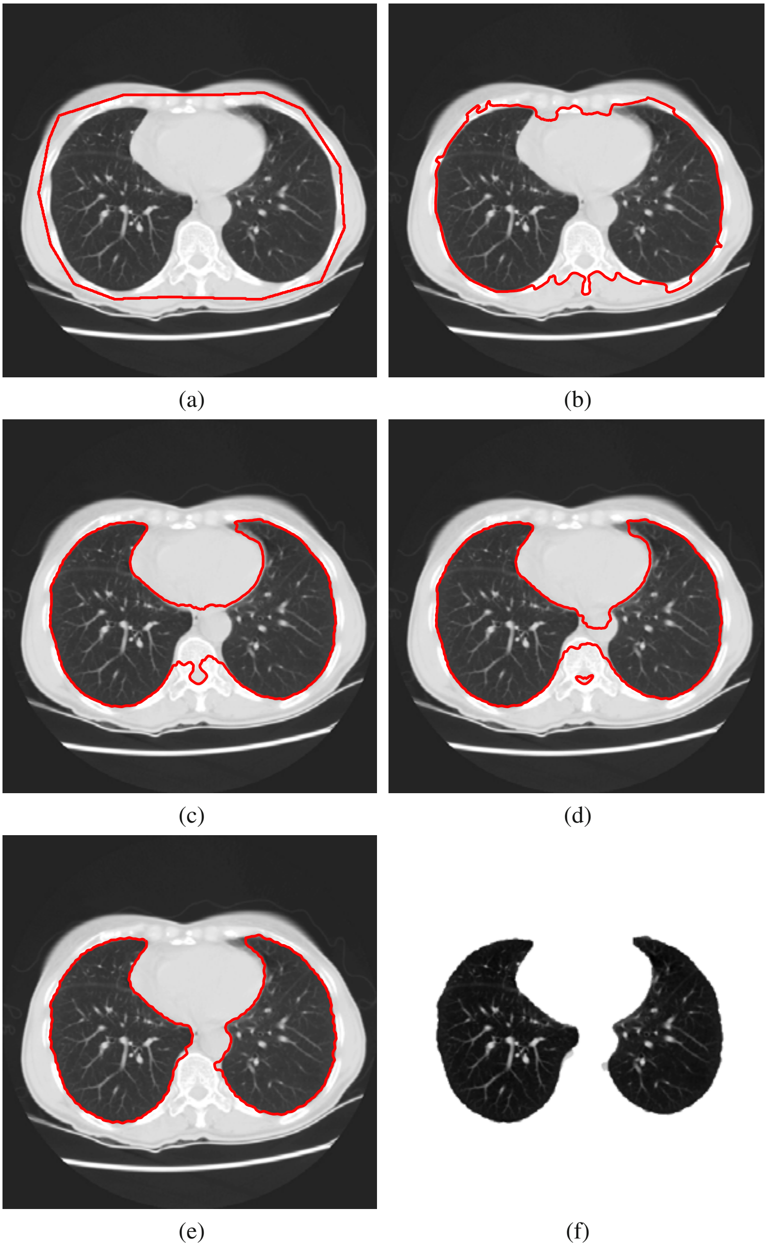


Fig. 3.5. Lung image segmentation with a geodesic active contour: (a) the initial curve; (b), (c), and (d) intermediate positions of the curve; (e) the final position of the curve; (f) the region within the final curve.

References

1. D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.
2. J. M. Morel and S. Solimini, *Variational methods in image segmentation*. Birkhauser, 1995.
3. G. Aubert and P. Kornpbrost, *Mathematical problems in image processing: Partial differential equations and the calculus of variations*. Springer Verlag, 2006.
4. L. Ambrosio and V. M. Tortorelli, “Approximation of functional depending on jumps by elliptic functional via t-convergence,” *Communications on Pure and Applied Mathematics*, vol. 43, no. 8, pp. 999–1036, 1990.
5. T. J. Richardson, “Scale independent piecewise smooth segmentation of images via variational methods,” Ph.D. dissertation, Lab. for Information and Decision Systems, MIT, Cambridge MA 02139, 1990.
6. S. Zhu and A. Yuille, “Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 118, no. 9, pp. 884–900, 1996.
7. Y. G. Leclerc, “Constructing simple stable descriptions for image partitioning,” *International Journal of Computer Vision*, vol. 3, no. 1, pp. 73–102, 1989.
8. 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, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, vol. 13, no. 1/2, pp. 249–268, 2002.
9. D. Cremers, M. Rousson, and R. Deriche, “A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape,” *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, 2007.
10. M. Rousson and R. Deriche, “A variational framework for active and adaptative segmentation of vector valued images,” in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 56–61.
11. I. Ben Ayed, A. Mitiche, and Z. Belhadj, “Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.
12. A. Mansouri, A. Mitiche, and C. Vázquez, “Multiregion competition: A level set extension of region competition to multiple region partioning,” *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
13. A. Blake and A. Zisserman, *Visual reconstruction*. MIT Press, 1987.
14. J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
15. ——, “A universal prior for integers and estimation by minimum description length,” *The Annals of Statistics*, vol. 11, no. 2, pp. 416–431, 1983.
16. G. Dahlquist and A. Bjork, *Numerical methods*. Prentice Hall, 1974.
17. M. Gage and R. S. Hamilton, “The heat equation shrinking convex plane curves,” *Journal of Differential Geometry*, vol. 23, no. 1, pp. 69–96, 1986.
18. M. A. Grayson, “The heat equation shrinks embedded plane curves to round points,” *Journal of Differential Geometry*, vol. 26, no. 2, pp. 285–314, 1987.
19. J. A. Sethian, “Recent numerical algorithms for hypersurfaces moving with curvature-dependent speed: Hamilton-jacobi equations and conservation laws,” *Journal of Differential Geometry*, vol. 31, pp. 131–161, 1990.

20. T. Chan and L. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
21. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, “Image segmentation using active contours: Calculus of variations or shape gradients?” *SIAM Journal of Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.
22. J. A. Sethian, *Level set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
23. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, “A variational level set approach to multiphase motion,” *Journal of Computational Physics*, vol. 127, no. 1, pp. 179–195, 1996.
24. M. Kass, A. P. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
25. V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
26. S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi, “Gradient flows and geometric active contour models,” in *International Conference on Computer Vision (ICCV)*, 1995, pp. 810–815.

4

MULTIREGION SEGMENTATION

4.1 Introduction

In Chapter 3, level set image segmentation into two regions was stated as the evolution of a single regular closed plane curve whose interior and exterior unambiguously define a partition of the image domain [1]. The general case of multiple regions uses several curves which can intersect. Therefore, a two-region formulation cannot be generalized directly by assigning a region to the interior of each curve because region membership becomes ambiguous when the curves intersect.

Let $I : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^n$ be an image function. Partitioning Ω into N regions, with $N > 2$, requires more than one curve. In this chapter, we examine the fundamental problem of evolving multiple curves so that, at convergence, the curves define a partition of the image domain following the minimization of the general functional

$$\mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}\right) = \mathscr{D} + \mathscr{R} \tag{4.1}$$

where

- $\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}$ is a partition, i.e., $\cup_{i\in[1\ldots N]}\mathbf{R}_i = \Omega$ and $\mathbf{R}_i \cap \mathbf{R}_j = \emptyset$ for $i \neq j$.
- $\mathscr{D}$ is a data term measuring how well the data fits a specified description, usually given through statistical models:

$$\mathscr{D} = \sum_{i=1}^{N} \int_{\mathbf{R}_i} e_i d\mathbf{x} \tag{4.2}$$

where e_i is a function which evaluates the conformity of the image data within region $\mathbf{R}_i$ to a statistical model. In Chapter 3, we examined the data terms of fundamental formulations, to which it was possible to give a Bayes/MDL interpretation.

- $\mathscr{R}$ is the length regularization term for smooth segmentation boundaries and to avoid small segmentation fragments:

$$\mathscr{R} = \lambda \sum_{i=1}^{N} \oint_{\partial\mathbf{R}_i} ds \tag{4.3}$$

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_4,

where $\partial \mathbf{R}_i$ is the boundary of region $\mathbf{R}_i$ and λ is a positive real constant to weigh the relative contribution of the two terms in the functional.

Associating a curve to the boundary of each region and minimizing $\mathscr{F}$ with respect to the curves may lead to an ambiguous result if the curves intersect at convergence, causing the regions they define to overlap. From an initial image domain partition into regions bounded by N closed simple plane curves $\{\gamma_i\}$, the region competition algorithm in [2] evolves $\Gamma = \cup_i \gamma_i$. The evolution is determined by the minimization of a functional which references Γ as a curve defining a partition. The implementation of the algorithm requires a careful initialization and uses an explicit representation of Γ as a set of points, a representation which does not accommodate changes in the curve topology during evolution. The formulation precludes the use of level sets because a single level set cannot represent a partition into more than two regions. As mentioned in previous chapters, the level set representation of curves is quite important because it allows changes in their topology during evolution, and it can be implemented by efficient stable numerical schemes [3]. In the level set formalism, a curve is represented implicitly as the zero level set of a function over the image domain. The advantages of such a representation over an explicit description of a curve as a set of points, as with the Snakes [4] formulation, has been generally acknowledged and well documented.

In this chapter, we examine level set multiregion segmentation methods aimed at obtaining a partition of the image domain at completion. There are several such methods:

(1) *Adding a partition constraint to the segmentation functional*

This is the method advocated in [5, 6, 7]. Each region is mapped to the interior of a curve represented implicitly by a level set function, and a term is added to the segmentation functional to draw the solution toward a partition. This does not guarantee a partition. Curve evolution will likely give an ambiguous segmentation if the partition constraint is not sufficiently enforced; if it is too strongly enforced, the curves will evolve mainly as a result of the partition constraint, not of the image data.

(2) *Defining a correspondence between the regions of segmentation and the active curves*

We will examine two closely related methods [8, 9]. The method in [8] establishes a correspondence between the characteristic functions of regions and the different combinations of level set signs so as to guarantee, by construction, a partition all the time during curve evolution. The method in [9] establishes a simpler, systematic general mapping between the segmentation regions and the regions defined by the curves and their intersections. As with the method in [8], the mapping guarantees, by definition, a partition at all times during curve evolution. We will give the level set evolution equations of each method and discuss their advantages and limitations.

(3) *Image segmentation as regularized clustering*

The method investigated in [10, 11, 12] views image segmentation as spatially regularized image data clustering, or quantization [13], leading to the simultaneous minimization of $N-1$ functionals to segment an image into N regions, each minimization involving a single region and its complement. The resulting curve evolution equations produce a partition at convergence provided the curves are initialized so as to define a partition. We will give the equations and show that curves evolving according to these will not intersect if they do not initially.

(4) *Embedding a partition constraint directly in the minimization equations*

This method [14, 15] embeds directly a simple partition constraint in the curve evolution equations. Starting from an arbitrary initial partition, the minimization is carried out following a first order expansion of the data term with an embedded partition constraint which implements the rule that if a point leaves a region, it is claimed by a single other region. The scheme results in a significant reduction of the computational load.

This chapter contains experimental demonstrations of the methods and comparisons between their computational requirements and convergence times.

4.2 Multiregion segmentation using a partition constraint functional term

In their work devoted to the level set motion of junctions and boundaries of multiple phases, Zhao et *al.* proposed a constraint to bias the solution toward a partition [5]. The scheme was later used in image segmentation [6, 7]. Each region (phase) $\mathbf{R}_i$ is mapped to the interior of a curve represented implicitly by the zero level of a level set function ϕ_i, $i \in [1,\ldots,N]$:

$$\mathbf{R}_i = \{\mathbf{x} \in \Omega, \phi_i(\mathbf{x}) > 0\} \tag{4.4}$$

The partition constraint is expressed as follows:

$$\beta \int_\Omega \left(\sum_{i=1}^{N} H(\phi_i) - 1 \right)^2 d\mathbf{x} \tag{4.5}$$

where H is the Heaviside function

$$H(z) = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ if } z < 0 \end{cases} \tag{4.6}$$

and β is a positive constant to balance the effect of the partition constraint against the other functional terms. Figure 4.1 illustrates how this constraint ensures coupling multiple level set functions by enforcing the condition

$$\sum_{i=1}^{N} H(\phi_i) = 1, \tag{4.7}$$

to discourage both empty regions and overlap between regions.

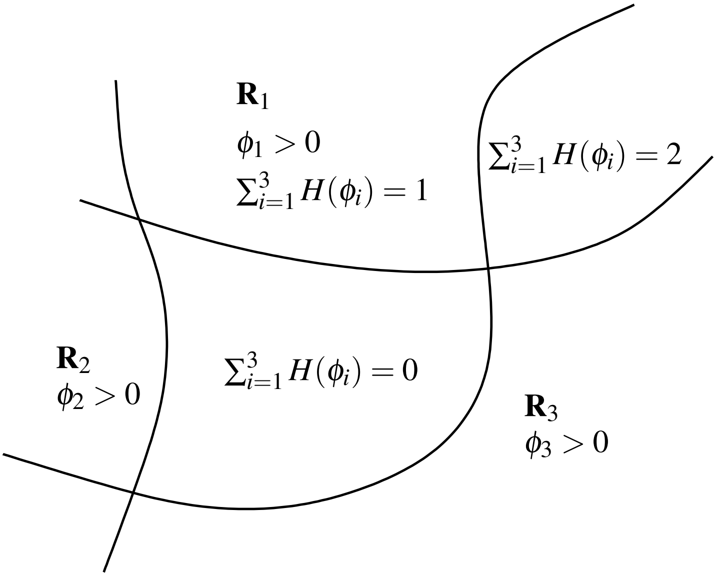


Fig. 4.1. Illustration of the partition constraint proposed by Zhao et *al.* [5]

This method requires N level set functions for a segmentation into N regions. Assigning a value to β can be problematic because curve evolution will produce overlapping regions if the partition constraint is not sufficiently enforced, and if it is too strongly enforced the curves will evolve more as a result of the partition constraint than of the image data. It is not clear how this parameter should be fixed so as to produce a partition without occulting the role of the data term.

4.3 Multiphase level set image segmentation

We will examine the method referred to as multiphase level set image segmentation [8]. This method establishes an explicit correspondence between N regions of segmentation and a partition defined using $\log_2 N$ level set functions. Let $\phi_i : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}, i \in [1,\ldots,M]$, be M level set functions with $M = \log_2 N$. Consider the vector level set function $\phi = (\phi_1,\ldots,\phi_M)$ and the vector-valued image $H(\phi) = (H(\phi_1),\ldots,H(\phi_M)) : \Omega \rightarrow \{0,1\}^M$ with H being the Heaviside function. The binary vectors in image $H(\phi)$ can have up to $N = 2^M$ values, $H_i, i \in [1,\ldots,N]$.

One can define the segmentation regions as follows:

Two pixels $\mathbf{x}$ *and* $\mathbf{y}$ *are in the same region if and only if* $H(\phi)(\mathbf{x}) = H(\phi)(\mathbf{y})$.

With this definition, each set

$$\mathbf{R}_i = \{\mathbf{x} \in \Omega | H(\phi)(\mathbf{x}) = H_i\}, i \in [1, \ldots, N], \tag{4.8}$$

defines a region, and the corresponding $N = 2^M$ regions define, by construction, a partition of the image domain. At all times during the evolution of the level set functions, each pixel belongs to one and only one region, i.e., by definition, there is neither vacuum nor overlap. This is an important advantage over the formulation in [5, 6, 7], which encourages but does not guarantee a partition. The multiphase functional can be written as follows:

$$\mathscr{F}\left(\{\mathbf{R}_i\}_{i \in [1,\ldots,N]}\right) = \sum_{i=1}^{N} \int_{\Omega} e_i \chi_{\mathbf{R}_i} d\mathbf{x} + \lambda \sum_{i=1}^{N} \int_{\Omega} \|\nabla \chi_{\mathbf{R}_i}\| d\mathbf{x} \tag{4.9}$$

where ∇ denotes the spatial gradient, $\chi_{\mathbf{R}} : \Omega \to \{0, 1\}$ is the characteristic function of region $\mathbf{R}$:

$$\chi_{\mathbf{R}}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathbf{R} \\ 0 & \text{if } \mathbf{x} \in \mathbf{R}^c, \end{cases} \tag{4.10}$$

and $\mathbf{R}^c$ is the complement of $\mathbf{R}$. For simplicity, the regularization term can be replaced by the sum of the lengths of the curves [8]:

$$\lambda \sum_{i=1}^{N} \int_{\Omega} \|\nabla H(\phi_i)\| d\mathbf{x} \tag{4.11}$$

Although it causes parts of the segmentation boundaries to be counted more than once, this slight modification does not, in general, have a significant effect on the segmentation outcome [8].

For an illustration, consider the case of a segmentation into at most four regions. This requires two level set functions ϕ_1 and ϕ_2:

$$\begin{cases} \mathbf{R}_1 = \{\mathbf{x} \in \Omega, \phi_1(\mathbf{x}) > 0, \phi_2(\mathbf{x}) > 0\} \\ \mathbf{R}_2 = \{\mathbf{x} \in \Omega, \phi_1(\mathbf{x}) > 0, \phi_2(\mathbf{x}) < 0\} \\ \mathbf{R}_3 = \{\mathbf{x} \in \Omega, \phi_1(\mathbf{x}) < 0, \phi_2(\mathbf{x}) > 0\} \\ \mathbf{R}_4 = \{\mathbf{x} \in \Omega, \phi_1(\mathbf{x}) < 0, \phi_2(\mathbf{x}) < 0\} \end{cases} \tag{4.12}$$

This is illustrated in Fig. 4.2. In the following, we use the piecewise constant image model for simplicity, i.e.,

$$e_i = (I - c_i)^2, \tag{4.13}$$

where c_i is the parameter of region $\mathbf{R}_i$. However, the formulation can be extended to other parametric image models. Using a regularized version H_ε of H (Chapter 3), the multiregion segmentation functional becomes:

$$
\begin{aligned}
\mathscr{F}_{\varepsilon,4} = & \int_{\Omega} (I-c_1)^2 H_{\varepsilon}(\phi_1) H_{\varepsilon}(\phi_2)\, d\mathbf{x} \\
& + \int_{\Omega} (I-c_2)^2 H_{\varepsilon}(\phi_1)\,(1-H_{\varepsilon}(\phi_2))\, d\mathbf{x} \\
& + \int_{\Omega} (I-c_3)^2\,(1-H_{\varepsilon}(\phi_1))\, H_{\varepsilon}(\phi_2)\, d\mathbf{x} \\
& + \int_{\Omega} (I-c_4)^2\,(1-H_{\varepsilon}(\phi_1))\,(1-H_{\varepsilon}(\phi_2))\, d\mathbf{x} \\
& + \lambda \int_{\Omega} \|\nabla H_{\varepsilon}(\phi_1)\| d\mathbf{x} + \lambda \int_{\Omega} \|\nabla H_{\varepsilon}(\phi_2)\| d\mathbf{x}
\end{aligned}
\tag{4.14}
$$

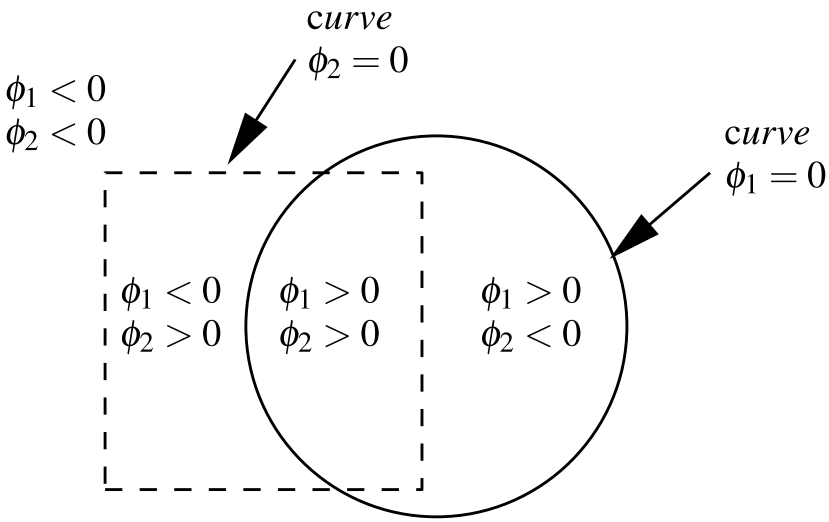


Fig. 4.2. Correspondence between $2 = \log_2 4$ level set functions and 4 segmentation regions with the method in [8]

The functional depends on two types of variables: the level set functions $\phi_i, i = 1,2$, and the region parameters $c_i, i = 1,\ldots,4$. Its minimization can be carried out by iterating two steps:

Step 1: With ϕ_1 and ϕ_2 fixed, minimization of $\mathscr{F}_{\varepsilon,4}$ with respect to $c_i, i = 1,\ldots,4$, yields the following update equations for the region parameters:

$$\begin{cases} c_1 = \frac{\int_{\phi_1>0,\phi_2>0} I d\mathbf{x}}{\int_{\phi_1>0,\phi_2>0} d\mathbf{x}} \\ c_2 = \frac{\int_{\phi_1>0,\phi_2<0} I d\mathbf{x}}{\int_{\phi_1>0,\phi_2<0} d\mathbf{x}} \\ c_3 = \frac{\int_{\phi_1<0,\phi_2>0} I d\mathbf{x}}{\int_{\phi_1<0,\phi_2>0} d\mathbf{x}} \\ c_4 = \frac{\int_{\phi_1<0,\phi_2<0} I d\mathbf{x}}{\int_{\phi_1<0,\phi_2<0} d\mathbf{x}} \end{cases} \tag{4.15}$$

Step 2: With c_i fixed, $i = 1,\ldots,4$, minimization of $\mathscr{F}_{\varepsilon,4}$ with respect to the level set functions is performed by embedding $\phi_i, i = 1,2$, into a one-parameter family of level set functions $\phi_i(\mathbf{x},t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$, indexed by algorithmic time $t > 0$, and solving the corresponding Euler-Lagrange descent equations:

$$\begin{cases} \frac{\partial \phi_1}{\partial t} = \delta_\varepsilon(\phi_1)\{\lambda \operatorname{div}\left(\frac{\nabla\phi_1}{\|\nabla\phi_1\|}\right) + \left((I-c_4)^2 - (I-c_2)^2\right)(1 - H_\varepsilon(\phi_2)) \\ \quad + \left((I-c_3)^2 - (I-c_1)^2\right) H_\varepsilon(\phi_2)\} \\ \frac{\partial \phi_2}{\partial t} = \delta_\varepsilon(\phi_2)\{\lambda \operatorname{div}\left(\frac{\nabla\phi_2}{\|\nabla\phi_2\|}\right) + \left((I-c_4)^2 - (I-c3)^2\right)(1 - H_\varepsilon(\phi_1)) \\ \quad + \left((I-c_2)^2 - (I-c_1)^2\right) H_\varepsilon(\phi_1)\} \end{cases} \tag{4.16}$$

where δ_ε is given by

$$\delta_\varepsilon = H'_\varepsilon \tag{4.17}$$

To extend the evolution to all the level sets of $\phi_i, i \in [1,\ldots,N]$, one can simply replace $\delta_\varepsilon(\phi_i)$ by $\|\nabla\phi_i\|$ in the equations.

The multiphase level set segmentation approach is often used [16, 17, 18]. It has several desirable properties:

- By construction, it guarantees a partition at all times during curve evolution.
- It requires only $\log_2 N$ level set functions for a segmentation into N regions.
- It can detect and represent triple junctions and other complex topologies.

However, it also has disadvantages:

- The complexity of the evolution equations increases exponentially with the number of regions.
- The method evaluates an expensive point membership function, which involves checking the signs of all level set intersections.

- Rather than fixing N, the method seeks a segmentation into up to a power of 2 regions. Some of these regions can vanish. However, there is no clear indication on the actual number of regions the method will produce. In some instances, unwanted divisions of regions can occur, for example with images of planar intensity variations.

4.4 Level set multiregion competition

In this section, we examine the method in [9] based on a representation of an image domain partition by a general, explicit correspondence between the N segmentation regions and the regions of a partition defined by $N-1$ simple closed plane curves and their intersections. As with the method in [8], this formulation maintains a partition at all times by construction. However, it determines a segmentation into a *fixed* but arbitrary number of regions N, rather than a power of 2 regions as in [8].

4.4.1 Representation of a partition into a fixed but arbitrary number of regions

The method uses the following mapping between the family $\{\mathbf{R}_{\gamma_i}\}$ of regions enclosed by the curves $\{\gamma_i\}$ and the segmentation regions $\{\mathbf{R}_i\}$:

$$\begin{cases} \mathbf{R}_1 = \mathbf{R}_{\gamma_1} \\ \mathbf{R}_2 = \mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2} \\ \dots \\ \mathbf{R}_i = \mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \dots \cap \mathbf{R}_{\gamma_{i-1}}^c \cap \mathbf{R}_{\gamma_i} \\ \dots \\ \mathbf{R}_N = \mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \dots \cap \mathbf{R}_{\gamma_{N-1}}^c \end{cases}$$

where $\mathbf{R}^c$ denotes the complement of region $\mathbf{R}$. The family $\{\mathbf{R}_{\gamma_1}, \mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}, \mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \mathbf{R}_{\gamma_3}, \dots\}$ thus obtained is, by construction, a partition of the image domain. The representation is illustrated in Figure 4.3.

With this choice of representation of an N-region partition, and using the length prior, the segmentation functional in (4.1) is written as:

$$\begin{aligned} \mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,\dots,N]}\right) = & \int_{\mathbf{R}_{\gamma_1}} e_1 d\mathbf{x} + \int_{\mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}} e_2 d\mathbf{x} + \dots \\ & + \int_{\mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \mathbf{R}_{\gamma_3}} e_3 d\mathbf{x} + \dots \\ & + \int_{\mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \mathbf{R}_{\gamma_3}^c \dots \cap \mathbf{R}_{\gamma_{N-1}}} e_{N-1} d\mathbf{x} + \dots \\ & + \int_{\mathbf{R}_{\gamma_1}^c \cap \mathbf{R}_{\gamma_2}^c \cap \mathbf{R}_{\gamma_3}^c \dots \cap \mathbf{R}_{\gamma_{N-1}}^c} e_N d\mathbf{x} + \dots \\ & + \lambda \sum_{i=1}^{N-1} \int_{\gamma_i} ds \end{aligned} \tag{4.18}$$

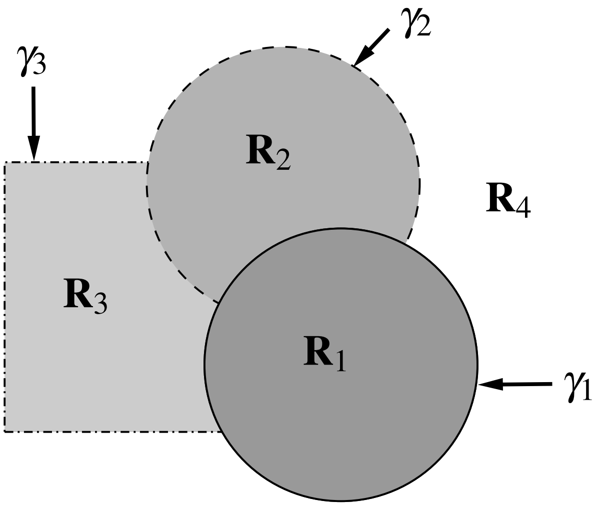


Fig. 4.3. Representation of a region partition into four regions (three curves).

For simplicity, the regularization term in (4.1) is taken to be curve length.

4.4.2 Curve evolution equations

Minimization of functional $\mathscr{F}$ with respect to curves $(\gamma_i)_{i\in[1,\dots,N-1]}$ is performed by embedding each curve into a one-parameter family $\gamma_i(s,t) : [0,1]\times\mathbb{R}^+ \to \Omega, \quad i = 1,\dots,N-1$ of plane curves, indexed by algorithmic time t, and solving the following system of evolution equations:

$$\frac{\partial \gamma_i}{\partial t} = -\frac{\partial \mathscr{F}}{\partial \gamma_i}, \quad i = 1,\dots,N-1 \tag{4.19}$$

The functional derivatives $\frac{\partial \mathscr{F}}{\partial \gamma_i}$ can be computed by first rewriting in a suitable form the area integrals in the functional. First with γ_1, $\mathscr{F}$ is rewritten as follows:

$$\mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,\dots,N]}\right) = \int_{\mathbf{R}_{\gamma_1}} e_1 d\mathbf{x} + \int_{\mathbf{R}^c_{\gamma_1}} \Phi_1 d\mathbf{x} + \lambda \int_{\gamma_1} ds + \lambda \sum_{i=2}^{N-1} \int_{\gamma_i} ds \tag{4.20}$$

where Φ_1 is

$$\begin{aligned}\Phi_1 &= e_2 \chi_{\mathbf{R}_{\gamma_2}} \\ &+ e_3 \chi_{\mathbf{R}^c_{\gamma_2}} \chi_{\mathbf{R}_{\gamma_3}} \\ &+ \cdots \\ &+ e_{N-1} \chi_{\mathbf{R}^c_{\gamma_2}} \chi_{\mathbf{R}^c_{\gamma_3}} \cdots \chi_{\mathbf{R}^c_{\gamma_{N-2}}} \chi_{\mathbf{R}_{\gamma_{N-1}}} \\ &+ e_N \chi_{\mathbf{R}^c_{\gamma_2}} \chi_{\mathbf{R}^c_{\gamma_3}} \cdots \chi_{\mathbf{R}^c_{\gamma_{N-2}}} \chi_{\mathbf{R}^c_{\gamma_{N-1}}}\end{aligned}$$

Both Φ_1 and $\sum_{i=2}^{N-1} \int_{\gamma_i} ds$ are independent of γ_1. Therefore, the functional derivative $\frac{\partial \mathscr{F}}{\partial \gamma_1}$ is (Chapter 2):

$$\frac{\partial \mathscr{F}}{\partial \gamma_1} = (e_1 - \Phi_1 + \lambda \kappa_1)\,\mathbf{n}_1 \tag{4.21}$$

where $\mathbf{n}_1$ is the outward unit normal to γ_1, and κ_1 its curvature function. For the functional derivative $\frac{\partial \mathscr{F}}{\partial \gamma_2}$, $\mathscr{F}$ is rewritten as follows:

$$\begin{aligned} \mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}\right) = & \int_{\mathbf{R}_{\gamma_1}} e_1 d\mathbf{x} + \int_{\mathbf{R}_{\gamma_2}} \chi_{\mathbf{R}_1^c} e_2 d\mathbf{x} + \\ & + \int_{\mathbf{R}_{\gamma_2}^c} \chi_{\mathbf{R}_1^c} \Phi_2 d\mathbf{x} + \lambda \int_{\gamma_2} ds + \\ & + \lambda \sum_{i\neq 2} \int_{\gamma_i} ds \end{aligned} \tag{4.22}$$

where Φ_2 is

$$\begin{aligned} \Phi_2(\mathbf{x}) = & \; e_3 \chi_{\mathbf{R}_{\gamma_3}} \\ & + e_4 \chi_{\mathbf{R}_{\gamma_3}^c} \chi_{\mathbf{R}_{\gamma_4}} \\ & + \ldots \\ & + e_{N-1} \chi_{\mathbf{R}_{\gamma_3}^c} \chi_{\mathbf{R}_{\gamma_4}^c} \cdots \chi_{\mathbf{R}_{\gamma_{N-2}}^c} \chi_{\mathbf{R}_{\gamma_{N-1}}} \\ & + e_N \chi_{\mathbf{R}_{\gamma_3}^c} \chi_{\mathbf{R}_{\gamma_4}^c} \cdots \chi_{\mathbf{R}_{\gamma_{N-2}}^c} \chi_{\mathbf{R}_{\gamma_{N-1}}^c} \end{aligned}$$

Here again, $\int_{\mathbf{R}_{\gamma_1}} e_1 d\mathbf{x}$, Φ_2 and $\sum_{i\neq 2} \int_{\gamma_i} ds$ have no dependence on γ_2. Therefore,

$$\frac{\partial \mathscr{F}}{\partial \gamma_2} = \left(\chi_{\mathbf{R}_1^c}(e_2 - \Phi_2) + \lambda \kappa_2\right)\mathbf{n}_2 \tag{4.23}$$

where $\mathbf{n}_2$ is the outward unit normal to γ_2, and κ_2 its curvature function.

The data term in (4.23) is taken into consideration only outside $\mathbf{R}_1$ where, according to (4.19), the evolution of γ_2 is

$$\frac{\partial \gamma_2}{\partial t} = -\frac{\partial \mathscr{F}}{\partial \gamma_2} = -(e_2 - \Phi_2 + \lambda \kappa_2)\,\mathbf{n}_2 \tag{4.24}$$

Because the evolution of γ_2 within $\mathbf{R}_1$ will not affect the energy functional, we can simply use the evolution equation (4.24) everywhere for γ_2, which lends to a uniform implementation.

Proceeding similarly to compute functional derivatives $\frac{\partial \mathscr{F}}{\partial \gamma_i}$ for all i, the minimization of the multiregion functional is performed by the following system of coupled curve evolution equations:

$$\begin{cases} \frac{\partial \gamma_1}{\partial t} = -\left(e_1 - \Phi_1 + \lambda\,\kappa_1\right)\mathbf{n}_1 \\ \frac{\partial \gamma_2}{\partial t} = -\left(e_2 - \Phi_2 + \lambda\,\kappa_2\right)\mathbf{n}_2 \\ \quad\vdots \\ \frac{\partial \gamma_i}{\partial t} = -\left(e_i - \Phi_i + \lambda\,\kappa_i\right)\mathbf{n}_i \\ \quad\vdots \\ \frac{\partial \gamma_{N-1}}{\partial t} = -\left(e_{N-1} - \Phi_{N-1} + \lambda\,\kappa_{N-1}\right)\mathbf{n}_{N-1} \end{cases} \tag{4.25}$$

where $\mathbf{n}_i$ is the outward unit normal to γ_i, κ_i its curvature function, and Φ_i is, for $i = 1, \ldots, N-1$,

$$\begin{aligned} \Phi_i &= e_{i+1}\chi_{\mathbf{R}_{\gamma_{i+1}}} \\ &+ e_{i+2}\chi_{\mathbf{R}^c_{\gamma_{i+1}}}\chi_{\mathbf{R}_{\gamma_{i+2}}} \\ &+ \cdots \\ &+ e_{N-1}\chi_{\mathbf{R}^c_{\gamma_{i+1}}}\cdots\chi_{\mathbf{R}^c_{\gamma_{N-2}}}\chi_{\mathbf{R}_{\gamma_{N-1}}} \\ &+ e_N\chi_{\mathbf{R}^c_{\gamma_{i+1}}}\cdots\chi_{\mathbf{R}^c_{\gamma_{N-2}}}\chi_{\mathbf{R}^c_{\gamma_{N-1}}} \end{aligned}$$

4.4.3 Level set implementation

The implicit representation of each curve γ_i, $i = 1, \ldots, N-1$, as the zero level set of a function $\phi_i(\mathbf{x},t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ gives the following level set evolution equations:

$$\begin{cases} \frac{\partial \phi_1}{\partial t} = -\left(e_1 - \Phi_1 + \lambda\,\kappa_1\right)\|\nabla\phi_1\| \\ \quad\vdots \\ \frac{\partial \phi_i}{\partial t} = -\left(e_i - \Phi_i + \lambda\,\kappa_i\right)\|\nabla\phi_i\| \\ \quad\vdots \\ \frac{\partial \phi_{N-1}}{\partial t} = -\left(e_{N-1} - \Phi_{N-1} + \lambda\,\kappa_{N-1}\right)\|\nabla\phi_{N-1}\| \end{cases} \tag{4.26}$$

where Φ_i is given by

$$\begin{aligned} \Phi_i &= e_{i+1}\chi_{\{\phi_{i+1}>0\}} \\ &+ e_{i+2}\chi_{\{\phi_{i+1}\le 0\}}\chi_{\{\phi_{i+2}>0\}} \\ &+ \cdots \\ &+ e_{N-1}\chi_{\{\phi_{i+1}\le 0\}}\cdots\chi_{\{\phi_{N-2}\le 0\}}\chi_{\{\phi_{N-1}>0\}} \\ &+ e_N\chi_{\{\phi_{i+1}\le 0\}}\cdots\chi_{\{\phi_{N-2}\le 0\}}\chi_{\{\phi_{N-1}\le 0\}} \end{aligned}$$

and

$$\chi_{\{\phi \leq 0\}} = \begin{cases} 1 & \text{if } \phi \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\chi_{\{\phi > 0\}} = \begin{cases} 1 & \text{if } \phi > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.27}$$

Recall that κ_i is given as a function of ϕ_i by

$$\kappa_i = -\text{div}\left(\frac{\nabla\phi_i}{\|\nabla\phi_i\|}\right) \tag{4.28}$$

The velocities can be evaluated everywhere and used to evolve the level sets everywhere, i.e, the equations (4.26) can be generalized to evolve the level set functions over the whole domain Ω.

4.5 Multiregion level set segmentation as regularized clustering

This method looks at image segmentation from the viewpoint of spatially regularized image data clustering [10, 11]. This leads to segment N regions by the simultaneous minimization of $N-1$ functionals, each involving a single region and its complement. First, consider the minimization of (4.1) without the regularization term. The problem consists of finding a partition which minimizes the data term $\mathscr{D}$, which can be rewritten as follows:

$$\mathscr{D} = \int_{\mathbf{R}_i} e_i \, d\mathbf{x} + \sum_{\substack{j=1 \\ j \neq i}}^{N} \int_{\mathbf{R}_j} e_j \, d\mathbf{x} \tag{4.29}$$

We have the following inequalities on the second term of the right hand side of (4.29):

$$\begin{aligned} \sum_{\substack{j=1 \\ j \neq i}}^{N} \int_{\mathbf{R}_j} e_j \, d\mathbf{x} &\geq \int_{\mathbf{R}_i^c} \sum_{\substack{j=1 \\ j \neq i}}^{N} e_j \chi_{\mathbf{R}_j} \, d\mathbf{x} \\ &\geq \int_{\mathbf{R}_i^c} \min_{j \neq i}(e_j) \, d\mathbf{x} \end{aligned} \tag{4.30}$$

The equality in the first line of (4.30) occurs for a partition, and in the second line for a partition when

$$\sum_{j \neq i} e_j \chi_{\mathbf{R}_j} = \min_{j \neq i}(e_j) \tag{4.31}$$

Therefore, the problem of minimizing (4.29) can be viewed as image data clustering, or quantization [13], and formulated as the simultaneous minimization of the following functionals, each corresponding to a two-region problem involving one region and its complement:

$$\mathscr{F}_{\mathbf{R}_i} = \int_{\mathbf{R}_i} e_i \, d\mathbf{x} + \int_{\mathbf{R}_i^c} \psi_i d\mathbf{x} \quad i = 1, \ldots, N-1 \tag{4.32}$$

where ψ_i is

$$\psi_i = \min_{j \neq i} (e_j) \tag{4.33}$$

and $\mathbf{R}_N$ is the background region, i.e.,

$$\mathbf{R}_N = \bigcap_{i=1}^{N-1} \mathbf{R}_i^c \tag{4.34}$$

The simultaneous minimization of the $N-1$ functionals in (4.32) produces, as in standard clustering, a partition of the image domain with $\sum_{j \neq i} e_j \chi_{\mathbf{R}_j} = \psi_i$. For smooth segmentation boundaries and to avoid small, isolated segmentation fragments, one can add to each of the functionals in (4.32) the length of the boundary of $\mathbf{R}_i$:

$$\mathscr{F}_{\mathbf{R}_i,\lambda} = \int_{\mathbf{R}_i} e_i \, d\mathbf{x} + \int_{\mathbf{R}_i^c} \psi_i d\mathbf{x} + \lambda \, l(\partial \mathbf{R}_i), \quad i = 1, \ldots, N-1 \tag{4.35}$$

The problem consists of finding the following set of regions:

$$\begin{cases} \widetilde{\mathbf{R}}_i & = \min\limits_{\mathbf{R}_i} \left(\mathscr{F}_{\mathbf{R}_i,\lambda} \right), \; i = 1, \ldots, N-1 \\ \widetilde{\mathbf{R}}_N & = \bigcap\limits_{i=1}^{N-1} \widetilde{\mathbf{R}}_i^c \end{cases} \tag{4.36}$$

In the next section, we will derive the corresponding curve evolution equations, and show that curves evolving according to these equations will not intersect if they do not initially.

4.5.1 Curve evolution equations

Let $\gamma_i : [0,1] \to \Omega, \; i = 1, \ldots, N-1$, be simple closed plane curves parameterized by arc parameter $s \in [0,1]$. Let $\mathbf{R}_i$ be $\mathbf{R}_{\gamma_i}$, the interior of γ_i, $i = 1, \ldots, N-1$. The length of the boundary of $\mathbf{R}_i$ is

$$l(\partial \mathbf{R}_i) = \int_{\gamma_i} ds \tag{4.37}$$

For each γ_i, $i = 1, \ldots, N-1$, the curve evolution equation for the minimization of $\mathscr{F}_{\mathbf{R}_i,\lambda}$ is obtained by embedding γ_i in a family of one-parameter curves $\gamma_i(s,t) : [0,1] \times \mathbb{R}^+ \to \Omega$, and solving the following Euler-Lagrange descent equation:

$$\frac{\partial \gamma_i}{\partial t} = -\frac{\partial \mathscr{F}_{\mathbf{R}_i,\lambda}}{\partial \gamma_i} \tag{4.38}$$

which gives

$$\frac{\partial \gamma_i}{\partial t} = -(e_i - \psi_i + \lambda \kappa_i)\,\mathbf{n}_i, \quad i = 1, \ldots, N-1 \tag{4.39}$$

We will now show that curves which move according to these equations will not intersect if they do not initially.

Let γ_i and γ_j be two distinct curves. If γ_i and γ_j do not intersect initially, they must first meet at a set of points before they can intersect. The curve velocity, on the right-hand side of (4.39), has two terms, the curvature term $(-\lambda\kappa_i\mathbf{n}_i)$ and the data term $(-(e_i - \psi_i)\mathbf{n}_i)$. We will study the effect of each of these terms at a point of contact.

Effect of the curvature term

Figures 4.4 and 4.5 show the two possible configurations at a point of contact (up to an exchange of the curve names). Let $\mathbf{x}_0$ be the point of contact.

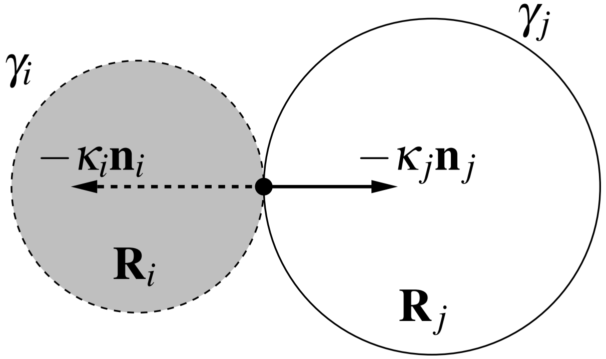


Fig. 4.4. Case 1

In the first case, the curvatures are both positive ($\kappa_i(\mathbf{x}_0) \geq 0$, $\kappa_j(\mathbf{x}_0) \geq 0$), which means that both curves retract and, therefore, will not cross at $\mathbf{x}_0$. In the second case, the curvatures are of opposite signs ($\kappa_i(\mathbf{x}_0) \geq 0$, $\kappa_j(\mathbf{x}_0) \leq 0$), which means that the curves move in the same direction. However, because $|\kappa_i(\mathbf{x}_0)| \geq |\kappa_j(\mathbf{x}_0)|$, curve γ_i retracts faster than γ_j advances and, therefore, the two curves will not cross at $\mathbf{x}_0$.

Effect of the data term

Let $d_i = -(e_i - \psi_i)$. There are two cases to consider:

1. $e_i(\mathbf{x}_0) = \min\limits_{l \in [1,\ldots,N]} (e_l(\mathbf{x}_0))$ or $e_j(\mathbf{x}_0) = \min\limits_{l \in [1,\ldots,N]} (e_l(\mathbf{x}_0))$

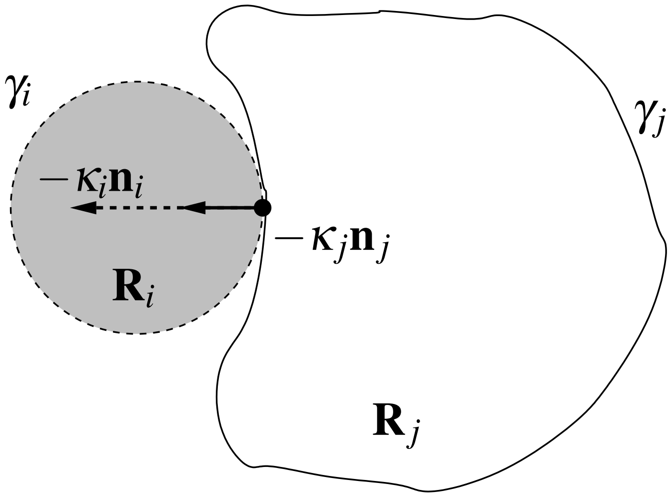


Fig. 4.5. Case 2

2. $e_k(\mathbf{x}_0) = \min\limits_{l\in[1,\ldots,N]} (e_l(\mathbf{x}_0)),\ \ k \neq i,\ k \neq j$

In the first case, let us assume, without loss of generality, that

$$e_j(\mathbf{x}_0) = \min_{l\in[1,\ldots,N]} (e_l(\mathbf{x}_0))$$

This means that $d_i(\mathbf{x}_0) \leq 0$ and $d_j(\mathbf{x}_0) = e_j(\mathbf{x}_0) - e_i(\mathbf{x}_0) \geq 0$. Therefore, γ_i and γ_j move in the same direction, with γ_i advancing and γ_j retracting. However, because $|d_i(\mathbf{x}_0)| \leq |d_j(\mathbf{x}_0)|$, γ_j retracts faster than γ_i advances and the two curves cannot cross at $\mathbf{x}_0$. In the second case, we have $d_i(\mathbf{x}_0) \geq 0$ and $d_j(\mathbf{x}_0) \geq 0$ and, therefore, the curves retract and will not cross at $\mathbf{x}_0$.

4.5.2 Level set implementation

As usual, we represent each curve implicitly as the zero level of a function $\phi_i(\mathbf{x},t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$. As a result, the level set evolution equations corresponding to (4.39) are

$$\frac{\partial \phi_i}{\partial t} = -(e_i - \psi_i + \lambda\kappa_i)\,\|\nabla \phi_i\| \tag{4.40}$$

As with the method in the preceding section, (4.40) can be used over the whole domain Ω.

4.6 Embedding a partition constraint directly in the minimization equations

We will describe a minimization scheme which embeds directly a simple partition constraint in the evolution equations of multiple curves [14, 15]. To expose the method, we first revisit the two-region segmentation to rewrite the standard Euler-Lagrange curve evolution equation using a first-order expansion of the data term. This first-order analysis gives an interpretation of curve evolution strictly from the point of view of *pixel-membership variations*, and leads to an extension to multiregion segmentation where the functional minimization embeds directly a simple, efficient partition constraint that maintains implicitly a partition of the image domain at all times. Basically, the partition constraint states that *if a point leaves a region, it goes to a single other region*, i.e., it is claimed by a region without transiting through others as could be the case in [10, 11].

4.6.1 Two-region segmentation: first order analysis

We use a single curve γ in the case of two regions. Let $\mathbf{R}_1 = \mathbf{R}_\gamma$ and $\mathbf{R}_2$ its complement: $\mathbf{R}_2 = \mathbf{R}_1^c$. The minimization of $\mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,2]}\right)$ with respect to γ follows the descent equation

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{F}\left(\{\mathbf{R}_i\}_{i\in[1,2]}\right)}{\partial \gamma} = -\frac{\partial \mathscr{D}}{\partial \gamma} - \frac{\partial \mathscr{R}}{\partial \gamma} \tag{4.41}$$

with

$$\frac{\partial \mathscr{D}}{\partial \gamma} = (e_1 - e_2)\mathbf{n} \tag{4.42}$$

In the following, we use a first-order expansion of the data term [14, 15] to give a pixel membership variation interpretation to this curve evolution equation, leading to a simple, efficient partition maintaining generalization to multiple curve evolution.

Consider an elementary unit local deformation $\delta\gamma$ of γ along its normal $\mathbf{n}$:

$$\begin{cases} \delta\gamma = (\delta x, \delta y) = ||\delta\gamma||\mathbf{n} \\ ||\delta\gamma|| = 1 \end{cases} \tag{4.43}$$

The first order expansion of $\mathscr{D}$ gives

$$\mathscr{D}(\gamma + \delta\gamma) = \mathscr{D}(\gamma) + < \frac{\partial \mathscr{D}}{\partial \gamma}, \delta\gamma > \tag{4.44}$$

Because $\frac{\partial \mathscr{D}}{\partial \gamma}$ is in the direction of $\mathbf{n}$ and $||\delta\gamma|| = 1$, (4.44) gives

$$\left(\mathscr{D}(\gamma + \delta\gamma) - \mathscr{D}(\gamma)\right)\delta\gamma = \frac{\partial \mathscr{D}}{\partial \gamma} \tag{4.45}$$

Consider a discrete elementary variation of γ involving a single pixel $\mathbf{p}$. Then the elementary variation $\Delta a_{\mathbf{R}}$ of the area of region $\mathbf{R}$ is

$$\begin{cases} \Delta a_{\mathbf{R}} = 1 & \text{if the curve is expanding} \\ \Delta a_{\mathbf{R}} = -1 & \text{if the curve is shrinking} \end{cases} \tag{4.46}$$

Let $\Delta\xi_i(\mathbf{p})$ and $\Delta\mathscr{D} = \mathscr{D}(\gamma+\delta\gamma) - \mathscr{D}(\gamma)$ be the corresponding elementary variation of $\xi_i = \int_{\mathbf{R}_i} e_i d\mathbf{x}$, $i = 1,2$, and of the data term. We have

$$\begin{aligned} \Delta\mathscr{D} &= \Delta\xi_1(\mathbf{p}) + \Delta\xi_2(\mathbf{p}) \\ &= \Delta a_{\mathbf{R}_1} e_1(\mathbf{p}) + \Delta a_{\mathbf{R}_2} e_2(\mathbf{p}) \end{aligned} \tag{4.47}$$

Suppose, without loss of generality, that γ is expanding to contain $\mathbf{p}$, i.e.,

$$\begin{cases} \Delta a_{\mathbf{R}_1} = 1 \\ \Delta a_{\mathbf{R}_2} = -\Delta a_{\mathbf{R}_1} = -1 \end{cases} \tag{4.48}$$

then

$$\begin{aligned} \Delta\xi_1 &= \Delta\xi_1^+ = e_1 \\ \Delta\xi_2 &= \Delta\xi_1^- = -e_2 \end{aligned} \tag{4.49}$$

where the superscript + (−) indicates the increase (decrease) of the data term due to the gain (loss) of pixel $\mathbf{p}$ by a region. Using (4.49), (4.47), and (4.45) gives

$$\begin{aligned} \frac{\partial\mathscr{D}}{\partial\gamma} &= \left(\Delta\xi_1^+ - \Delta\xi_2^+\right)\delta\gamma \\ &= (e_1 - e_2)\,\delta\gamma \end{aligned} \tag{4.50}$$

With $\delta\gamma = \mathbf{n}$, the functional derivative of $\mathscr{D}$ with respect to γ is approximated by

$$\frac{\partial\mathscr{D}}{\partial\gamma} = \left(\Delta\xi_1^+ - \Delta\xi_2^+\right)\mathbf{n} \tag{4.51}$$

When γ is shrinking at $\mathbf{p}$, rather than expanding, we have

$$\begin{cases} \Delta\xi_1 = -\Delta\xi_1^+ = -e_1 \\ \Delta\xi_2 = \Delta\xi_2^+ = e_2 \\ \delta\gamma = -\mathbf{n} \end{cases} \tag{4.52}$$

Thus, when γ is shrinking at $\mathbf{p}$ we obtain the same expression of the velocity as in (4.51). Accounting for the length regularization, we have the approximation

$$\frac{\partial\gamma}{\partial t} = -\left(\Delta\xi_1^+ - \Delta\xi_2^+ + \lambda\kappa\right)\mathbf{n} \tag{4.53}$$

The corresponding level set evolution equation is

$$\frac{\partial\phi}{\partial t} = -\left(\Delta\xi_1^+ - \Delta\xi_2^+ + \lambda\kappa\right)\|\nabla\phi\| \tag{4.54}$$

4.6.2 Extension to multiregion segmentation

In this section, we generalize the evolution equation in (4.54) to an arbitrary, fixed number of regions, so that the evolution of multiple curves produces a partition. Consider a family of simple closed curves $\gamma_k|_{k=1,\ldots,N-1}$ and let $\mathbf{R}_k$ be the region enclosed by γ_k:

$$\mathbf{R}_k = \mathbf{R}_{\gamma_k}, \quad k = 1,\ldots,N-1 \tag{4.55}$$

Let

$$\mathbf{R}_N = \bigcap_{k=1}^{N-1} \mathbf{R}_k^c \tag{4.56}$$

One can impose simple, efficient partition constraints directly on the multiregion curve evolution as follows [14, 15]:

Partition constraints:

Constraint 1

Start from an initial partition $\mathscr{P}^0 = \{\mathbf{R}_k^0\}_{k\in[1\ldots N]}$.

Constraint 2

Suppose we have a partition $\mathscr{P}^t = \{\mathbf{R}_k^t\}_{k\in[1,\ldots,N]}$ at iteration t, and let $\mathbf{p} \in \Omega$. If $\mathbf{p} \in \mathbf{R}_i^t, i \in [1,\ldots,N]$ and $\mathbf{p}$ leaves region $\mathbf{R}_i^t$, it must move to another region $\mathbf{R}_j, j \in [1,\ldots,N], j \neq i$, and only one other region, i.e.,

$$\begin{cases} \mathbf{p} \in \mathbf{R}_j^{t+1} \\ \forall k \neq j, \mathbf{x} \notin \mathbf{R}_k^{t+1} \end{cases} \tag{4.57}$$

To satisfy *constraint* (2), the curve evolution equations at $\mathbf{p}$ must involve at most two curves, i.e., only two regions: region $\mathbf{R}_i$ which contains pixel $\mathbf{p}$ currently and another region $\mathbf{R}_j, j \neq i$. To obtain the multiregion curve evolution equations satisfying *constraint* (2), we fix curves $\gamma_k, k \notin \{i,j\}$, and minimize the functional with respect to the variation of γ_i if $i \neq N$, and γ_j if $j \neq N$, i.e.,

$$\begin{cases} \text{if } \ i \neq N, \ \ \frac{\partial \mathscr{F}}{\partial \gamma_i} = \frac{\partial(\xi_i+\xi_j)}{\partial \gamma_i} + \frac{\partial \lambda \int_{\gamma_i} ds}{\partial \gamma_i} \\ \text{if } \ j \neq N, \ \ \frac{\partial \mathscr{F}}{\partial \gamma_j} = \frac{\partial(\xi_i+\xi_j)}{\partial \gamma_j} + \frac{\partial \lambda \int_{\gamma_j} ds}{\partial \gamma_j} \end{cases} \tag{4.58}$$

Therefore, multiregion segmentation reduces to a two-region segmentation problem corresponding to the variation $\Delta\xi_i + \Delta\xi_j$ of $\mathscr{D}$ in domain $\mathbf{R}_i \cup \mathbf{R}_j$. In this case, the level set curve evolution equations corresponding to the minimization of $\mathscr{F}$ with respect to γ_i, if $i \neq N$, and with respect to γ_j, if $j \neq N$, are given by

$$\begin{cases} \quad\quad \text{if} \quad i \neq N, \\ \frac{\partial \phi_i}{\partial t}(\mathbf{p},t) = -\left(\Delta\xi_i^+(\mathbf{p}) - \Delta\xi_j^+(\mathbf{p}) + \lambda\,\kappa_i(\mathbf{p})\right)\|\nabla\phi_i(\mathbf{p},t)\| \\ \\ \quad\quad \text{if} \quad j \neq N, \\ \frac{\partial \phi_j}{\partial t}(\mathbf{p},t) = -\left(\Delta\xi_j^+(\mathbf{p}) - \Delta\xi_i^+(\mathbf{p}) + \lambda\,\kappa_j(\mathbf{p})\right)\|\nabla\phi_j(\mathbf{p},t)\| \end{cases} \tag{4.59}$$

where ϕ_k is the level set function corresponding to $\gamma_k, k \in [1,\dots,N-1]$, and κ_k the curvature of the zero level set of ϕ_k. It is clear that the coupled curve evolution equations (4.59) satisfy partition *constraint (2)*. There are two cases to consider:

Case 1: $i = N$ or $j = N$
System (4.59) reduces to a single curve evolution equation corresponding to a two-region segmentation in the domain $\mathbf{R}_i \cup \mathbf{R}_j$.

Case 2: $i \neq N$ and $j \neq N$
Omitting the contribution of the curvature term, the two evolving curves γ_i and γ_j have opposite velocities at $\mathbf{p}$. Thus, if γ_i shrinks at $\mathbf{p}$, γ_j expands to contain it and vice versa. If the contribution of the curvature term is important, both evolving curves shrink and $\mathbf{p}$ leaves the interior of one curve to enter the background region $\mathbf{R}_N$.

Definition of region $\mathbf{R}_j$

The problem now is to define which of the regions $\mathbf{R}_j, j \in [1,\dots,N], j \neq i$, will be involved in system (4.59). Suppose $\mathbf{p}$ leaves $\mathbf{R}_i$ to enter $\mathbf{R}_j, j \in [1,\dots,N], j \neq i$. The resulting variation of the data term $\mathscr{D}$ is

$$\Delta\mathscr{D} = \Delta\xi_j^+(\mathbf{p}) - \Delta\xi_i^+(\mathbf{p}) \tag{4.60}$$

Since the purpose is to minimize $\mathscr{F}$, the sharpest variation is given by

$$\begin{aligned} j_0 &= \arg \min_{\{j \in [1,\dots,N],\ \mathbf{p} \notin \mathbf{R}_j\}} (\Delta\xi_j^+(\mathbf{p}) - \Delta\xi_i^+(\mathbf{p})) \\ &= \arg \min_{\{j \in [1,\dots,N],\ \mathbf{p} \notin \mathbf{R}_j\}} \Delta\xi_j^+(\mathbf{p}) \end{aligned} \tag{4.61}$$

This leads to the following multiregion level set equations for all $\mathbf{p}$:

$\forall i \in [1,\dots,N]$, *if* $\mathbf{p} \in \mathbf{R}_i$, do

$$\begin{cases} \qquad\qquad \text{if} \quad i \neq N, \\ \frac{\partial \phi_i}{\partial t}(\mathbf{p},t) = -\left(\Delta\xi_i^+(\mathbf{p}) - \Delta\xi_{j_0}^+(\mathbf{p}) + \lambda\kappa_i\right)\|\nabla\phi_i(\mathbf{p},t)\| \\ \\ \qquad\qquad \text{if} \quad j_0 \neq N, \\ \frac{\partial \phi_{j_0}}{\partial t}(\mathbf{p},t) = -\left(\Delta\xi_{j_0}^+(\mathbf{p}) - \Delta\xi_i^+(\mathbf{p}) + \lambda\kappa_{j_0}\right)\|\nabla\phi_{j_0}(\mathbf{p},t)\| \end{cases} \tag{4.62}$$

where $i \in [1,\ldots,N]$ is the index of the region containing $\mathbf{p}$ and j_0 is given by the best variation (4.61).

This multiregion method has a computational advantage over the other methods we examined in this chapter. It activates at most two evolution equations at each iteration, i.e., the number of level set updates is independent of the number of regions. On the contrary, the methods in [8, 9, 5, 11] activate a number of level set evolution equations which grows with the number of regions. For instance, the method in [8] activates $\log_2 N$ equations, and the schemes in [9, 11] require $N-1$ instantiations. Recall that the methods in [8, 9] evaluate an expensive point membership function. For the method in [8], this involves checking the signs of all the level set intersections. For the method in [9], this involves, for a given level set, checking the sign of all lower numbered level sets. In the next section, we will illustrates some of these computational aspects with an experimental example.

4.6.3 Example

We show several experiments with the brain image in Figure 4.6 to illustrate the convergence speed and computational load of the four multiregion segmentation methods we examined in this chapter. Figure 4.6 depicts the segmentation of the brain image into three regions with the implicit partition method in [15] using the piecewise constant image model. Figure 4.6 (a), (b), and (c) show, respectively, the position of the initial curves, the final curves, and the final segmentation represented by the regions means. We ran a similar experiment for each of the methods discussed in this chapter, and plotted the evolution of the functional as a function of the number of iterations in Figure 4.7. The implicit partition [15] and regularized clustering methods [11] led to decreases in the functional much faster than the explicit partition methods in [8, 9]. For instance the method in [15] needs approximately 30 iterations to converge, whereas the method in [9] needs more than 400. The good performance of the method in [15] is due to its stepwise optimality in Equation (4.61), i.e., it effects the maximum decrease in the functional at each curve evolution step.

Figure 4.8 shows the computation time spent at an iteration as a function of the number of regions for the implicit partition in [15], the regularized clustering in [11], and the explicit partition in [9]. The implicit partition method has the slowest curve growth. This was to be expected because the number of level set updates with this method is independent of the number of regions, unlike the other methods we examined in this chapter.

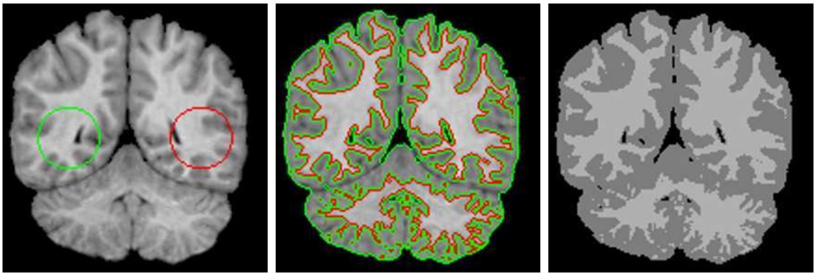

(a) (b) (c)

Fig. 4.6. Multiregion segmentation example with the partition constrained minimization scheme in [15] and the piecewise constant model: (a) initial curves; (b) final positions of the curves; (c) segmentation represented with the mean parameters obtained at convergence.

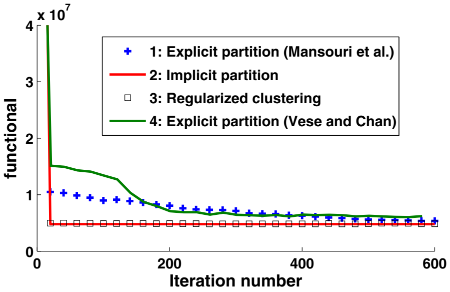


Fig. 4.7. Evolution of the functional as a function of the iteration number for the implicit partition method in [15], the regularized clustering in [11], and the explicit partition methods in [8, 9].

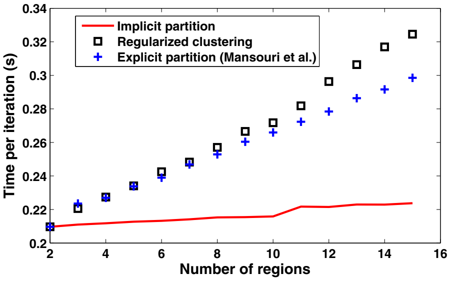


Fig. 4.8. The computation time spent at an iteration as a function of the number of regions for the implicit partition in [15], the regularized clustering in [11], and the explicit partition in [9].

References

1. T. Chan and L. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
2. S. Zhu and A. Yuille, “Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 118, no. 9, pp. 884–900, 1996.
3. J. A. Sethian, *Level set Methods and Fast Marching Methods.* Cambridge University Press, 1999.
4. M. Kass, A. P. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
5. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, “A variational level set approach to multiphase motion,” *Journal of Computational Physics*, vol. 127, no. 1, pp. 179–195, 1996.
6. N. Paragios and R. Deriche, “Coupled geodesic active regions for image segmentation: A level set approach,” in *Europeean Conference on Computer Vision (ECCV)*, 2000, pp. 224–240.
7. C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia, “A level set model for image classification,” *International Journal of Computer Vision*, vol. 40, no. 3, pp. 187–197, 2000.
8. L. Vese and T. Chan, “A multiphase level set framework for image segmentation using the Mumford and Shah model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
9. A. Mansouri, A. Mitiche, and C. Vázquez, “Multiregion competition: A level set extension of region competition to multiple region partioning,” *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
10. C. Vázquez, A. Mitiche, and I. Ben Ayed, “Image segmentation as regularized clustering: A fully global curve evolution method,” in *International Conference on Image Processing (ICIP)*, 2004, pp. 3467–3470.
11. C. Vázquez, A. Mitiche, and R. Laganière, “Joint segmentation and parametric estimation of image motion by curve evolution and level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 782–793, 2006.
12. T. Brox and J. Weickert, “Level set segmentation with multiple regions,” *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213–3218, 2006.
13. A. Gersho and R. M. Gray, *Vector quantization and signal compression.* Kluwer Academic Publishers, 1991.
14. I. Ben Ayed and A. Mitiche, “A partition constrained minimization scheme for efficient multiphase level set image segmentation,” in *International Conference on Image Processing (ICIP)*, 2006, pp. 1641–1644.
15. I. Ben Ayed, A. Mitiche, and Z. Belhadj, “Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.
16. D. Cremers, M. Rousson, and R. Deriche, “A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape,” *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, 2007.
17. M. Rousson and R. Deriche, “A variational framework for active and adaptative segmentation of vector valued images,” in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 56–61.

18. J. Kim, J. W. F. III, A. J. Yezzi, M. Cetin, and A. S. Willsky, “A nonparametric statistical method for image segmentation using information theory and curve evolution,” *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1486–1502, 2005.

5

IMAGE MODELS

5.1 Introduction

To partition an image into regions, statistical segmentation uses a characterization of the image data distribution within each region, i.e., image descriptions which distinguish each region from the others. The descriptions validity largely affects the segmentation accuracy, i.e., the descriptions must fit the class of images to segment. In this chapter, we will examine various statistical descriptions in active curve image segmentation and map each to a curve evolution equation.

Let $I : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^n$ be an image. A two-region active curve segmentation method is specified by the evolution equation of a simple closed plane curve $\gamma : [0,1] \to \Omega$, which defines a partition of the image domain Ω into two regions,

$$\mathbf{R}_1 = \mathbf{R}_\gamma$$

corresponding to the interior of the curve, and

$$\mathbf{R}_2 = \mathbf{R}_\gamma^c$$

corresponding to the exterior. The evolution equations we will derive under various image models result from the minimization of functionals of the following general form:

$$\mathscr{D} + \lambda \int_\gamma ds \tag{5.1}$$

where $\mathscr{D}$ is a data term which measures how well the image fits a specified statistical description and λ a positive constant balancing the contribution of the curve length term against the data term. In general, the data term depends on the statistics of the image within the regions inside and outside the curve. We will examine various terms, each corresponding to a different statistical image description and leading to a different evolution equation. Some of these descriptions are derived from Bayesian estimation, whereas others are built on information theoretic concepts or measures of discrepancy (or affinity) between distributions. For each statistical description,

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_5,

we will determine the curve evolution equation to minimize the data term. These equations are all of the form:

$$\frac{\partial \gamma}{\partial t} = V\mathbf{n} \tag{5.2}$$

where γ is a one-parameter family of curves $\gamma(s,t) : [0,1] \times \mathbb{R}^+ \to \mathbb{R}^2 \times \mathbb{R}^+$, indexed by algorithmic time t; $\mathbf{n}$ is the outward unit normal to γ, and $V : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$ is a velocity function with coordinates of position and time, and dependent on the image and image statistics within the segmentation regions. Accounting for the length term of the segmentation functional, we have (Chapter 2)

$$\frac{\partial \gamma}{\partial t} = (V - \lambda \kappa)\,\mathbf{n} \tag{5.3}$$

where κ is the curvature of γ. Each method will be implemented via level sets, i.e., by defining γ as the zero level set of a function $\phi : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$. The corresponding evolution of ϕ is (Chapter 2)

$$\frac{\partial \phi}{\partial t} = (V - \lambda \kappa)\,\|\nabla \phi\| \tag{5.4}$$

with κ given by

$$\kappa = -\mathrm{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right) \tag{5.5}$$

and the region membership by

$$\begin{aligned} \mathbf{R}_1 &= \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) > 0\} \\ \mathbf{R}_2 &= \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) < 0\} \end{aligned} \tag{5.6}$$

The image models will be described in the two-region segmentation case. There are several ways to generalize the corresponding formulations to multiregion segmentation (Chapter 4) and Section 5.2.7 recalls one such way. The two-region image-model dependent evolution equations will be given explanatory statistical interpretations and verified by experimentation.

5.2 Segmentation by maximizing the image likelihood

In Chapter 3, we have seen that image segmentation can be stated as a Bayesian maximum *a posteriori* estimation [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11], which results in a data term of the following general form:

$$\mathscr{L}(\mathscr{P}^2_\Omega | I, P) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} -\log P(I|\mathbf{R}_i)\,d\mathbf{x} \tag{5.7}$$

where $\mathscr{P}_{\Omega}^{2}$ denotes a partition of the image domain Ω into 2 regions $\mathbf{R}_i$, $i = 1,2$. Minimization of this data term yields a partition which maximizes the image likelihood given the region image models. It is crucial that the models capture the image characteristics which distinguish one region from the other.

Models can be determined before segmentation (supervised segmentation) or during segmentation (unsupervised segmentation). In supervised segmentation, models are generally learned from segmented training images of the same class as the observed images to segment subsequently [8]. They can also be learned interactively from user-specified areas in the image to segment [10]. When region models are known beforehand, one can use the basic formula (Chapter 2)

$$\frac{\partial \int_{\mathbf{R}_\gamma} f d\mathbf{x}}{\partial \gamma} = f\mathbf{n} \tag{5.8}$$

valid when f is a function independent of γ. For (5.7), this gives the following minimization descent equation:

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{L}(\mathscr{P}_{\Omega}^{2}|I,P)}{\partial \gamma} = \log\left(\frac{P(I|\mathbf{R}_1)}{P(I|\mathbf{R}_2)}\right)\mathbf{n} \tag{5.9}$$

This image dependent velocity performs a hypothesis testing by an image likelihood ratio test. It alters the region membership of each point on the curve by evaluation of the hypotheses that the image at the point is drawn from the model of $\mathbf{R}_1$ or $\mathbf{R}_2$. If $\frac{P(I(\mathbf{p})|\mathbf{R}_1)}{P(I(\mathbf{p})|\mathbf{R}_2)} > 1$, i.e., the image likelihood at point $\mathbf{p}$ is higher when $\mathbf{p}$ belongs to $\mathbf{R}_1$, the speed is positive, i.e., the velocity vector is directed as the outward normal, causing the curve to expand and take $\mathbf{p}$ within $\mathbf{R}_1$; otherwise, the curve shrinks, causing $\mathbf{p}$ to belong to $\mathbf{R}_2$.

The prevalent way to define the data term is to use a statistical parametric image model, where each segmentation region is characterized by a set of parameters which distinguish the region from the other. In this case, the region model parameters become variables dependent on the segmentation (or on the active curve), and have to be optimized concurrently with curve evolution minimization. In the following, we first revisit active curve segmentation with the Gaussian distribution by taking into account the dependence between the region model parameters and the active curve. Then, we will examine segmentation with a number of useful parametric distributions, including the Gamma, Bernoulli, Poisson, Weibull, and Wishart. We further discuss a minimum description length (MDL) interpretation of the functional length term, which shows how to determine its weight systematically when used in conjunction with a data term of the type in (5.7).

5.2.1 The Gaussian model

Consider a model which assumes that the image data in each region $\mathbf{R}_i, i \in 1,2$, is Gaussian with parameters $\mu_{\mathbf{R}_i}$ and $\sigma_{\mathbf{R}_i}$:

$$P_{Gaussian}(I|\mathbf{R}_i) = \frac{1}{\sqrt{2\pi\sigma^2_{\mathbf{R}_i}}} e^{-\frac{(I-\mu_{\mathbf{R}_i})^2}{2\sigma^2_{\mathbf{R}_i}}} \tag{5.10}$$

Because the data term measures the conformity of the image within each region to a Gaussian model, we have

$$\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{Gaussian}) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} G(I,\mu_{\mathbf{R}_i},\sigma_{\mathbf{R}_i}) d\mathbf{x} \tag{5.11}$$

with

$$\begin{aligned} G(I,\mu_{\mathbf{R}_i},\sigma_{\mathbf{R}_i}) &= -\log P_{Gaussian}(I|\mathbf{R}_i) \\ &= \frac{1}{2}\log(2\pi\sigma^2_{\mathbf{R}_i}) + \frac{(I-\mu_{\mathbf{R}_i})^2}{2\sigma^2_{\mathbf{R}_i}} \end{aligned} \tag{5.12}$$

This data term depends on two types of variables, the parameters of regions and the segmentation. We recall that in the case of the Gaussian distribution, the optimal parameters can be computed analytically, and turn out to be the sample means and variances of the image within the regions (Chapter3)

$$\begin{aligned} \hat{\mu}_{\mathbf{R}_i} &= \frac{\int_{\mathbf{R}_i} I d\mathbf{x}}{\int_{\mathbf{R}_i} d\mathbf{x}} \\ \hat{\sigma}^2_{\mathbf{R}_i} &= \frac{\int_{\mathbf{R}_i} (I-\hat{\mu}_{\mathbf{R}_i})^2 d\mathbf{x}}{\int_{\mathbf{R}_i} d\mathbf{x}} \end{aligned} \tag{5.13}$$

These expressions show that each parameter is a function of the corresponding region domain, but the segmentation is unknown. The integrands in the region integrals in (5.11) depend on the region parameters and, therefore, on the integration domains, which means a dependence of the region parameters on the curve. Formally, this dependence must be accounted for in the derivation of the functional with respect to the curve [9, 12, 13, 14], a derivation which results in the curve evolution equation. In many studies, however, such a dependence is ignored as the curve evolution equation is deduced by assuming the distribution parameters to be independent of the curve, which results in a greedy, possibly non-optimal algorithm which iterates two consecutive steps: one to evolve the curve with the region parameters fixed and the other to compute the parameters within the curve defined regions. In this case, the curve evolution equation in the minimization of $\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{Gaussian})$ is

$$\frac{\partial\gamma}{\partial t} = -\frac{\partial\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{Gaussian})}{\partial\gamma} = V_{Gaussian}\mathbf{n} \tag{5.14}$$

with

$$V_{Gaussian} = -\left(G(I,\hat{\mu}_{\mathbf{R}_1},\hat{\sigma}_{\mathbf{R}_1}) - G(I,\hat{\mu}_{\mathbf{R}_2},\hat{\sigma}_{\mathbf{R}_2})\right) \tag{5.15}$$

For the particular case of the Gaussian, the dependence of the parameters on the curve does not actually affect the curve evolution equation, which is not the case in general. This is shown next.

Derivation by time-dependent parameters

To take into account the dependence of the region parameters on the curve, one can use the derivative with respect to algorithmic time of region integrals having the following general form [14]:

$$\mathscr{J} = \int_{\mathbf{R}_\gamma} \psi(\mathbf{x},t)d\mathbf{x} \tag{5.16}$$

This derivative is given by (refer to [14], Th. 4.2, p. 352)

$$\frac{\partial \mathscr{J}}{\partial t} = \int_{\mathbf{R}_\gamma} \frac{\partial \psi}{\partial t} d\mathbf{x} + \int_\gamma \langle \psi \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \tag{5.17}$$

where $\langle , \rangle$ denotes the scalar product. Now, assuming the region parameters $\mu_{\mathbf{R}_1}$ and $\sigma_{\mathbf{R}_1}$ depend on the segmentation and, therefore, on time, the derivative with respect to time of the first region integral in functional (5.11) can be derived using (5.17):

$$\frac{\partial \int_{\mathbf{R}_1} G(I,\mu_{\mathbf{R}_1},\sigma_{\mathbf{R}_1}) d\mathbf{x}}{\partial t} = \int_{\mathbf{R}_1} \frac{\partial G(I,\mu_{\mathbf{R}_1},\sigma_1)}{\partial t} d\mathbf{x} + \int_\gamma \langle G(I,\mu_{\mathbf{R}_1},\sigma_{\mathbf{R}_1})\mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \tag{5.18}$$

The integrand in the first term of the above derivative, which stems from the dependence of the region parameters on the curve (dependence on the curve induces dependence on time) can be expressed as follows:

$$\frac{\partial G(I,\mu_{\mathbf{R}_1},\sigma_{\mathbf{R}_1})}{\partial t} = \frac{\partial G}{\partial \mu_{\mathbf{R}_1}} \frac{\partial \mu_{\mathbf{R}_1}}{\partial t} + \frac{\partial G}{\partial \sigma_{\mathbf{R}_1}} \frac{\partial \sigma_{\mathbf{R}_1}}{\partial t} \tag{5.19}$$

We have

$$\frac{\partial G}{\partial \mu_{\mathbf{R}_1}} = -\frac{I-\mu_{\mathbf{R}_1}}{\sigma^2_{\mathbf{R}_1}} \tag{5.20}$$

and

$$\frac{\partial G}{\partial \sigma_{\mathbf{R}_1}} = \frac{1}{\sigma_{\mathbf{R}_1}} - \frac{(I-\mu_{\mathbf{R}_1})^2}{\sigma^3_{\mathbf{R}_1}} \tag{5.21}$$

Embedding these derivatives in the first term in (5.18) yields

$$\begin{aligned} \int_{\mathbf{R}_1} \frac{\partial G(I,\mu_{\mathbf{R}_1},\sigma_{\mathbf{R}_1})}{\partial t} d\mathbf{x} = &-\frac{1}{\sigma^2_{\mathbf{R}_1}} \frac{\partial \mu_{\mathbf{R}_1}}{\partial t} \left(\int_{\mathbf{R}_1} I d\mathbf{x} - \mu_{\mathbf{R}_1} \int_{\mathbf{R}_1} d\mathbf{x} \right) \\ &+ \frac{1}{\sigma_{\mathbf{R}_1}} \frac{\partial \sigma_{\mathbf{R}_1}}{\partial t} \left(\int_{\mathbf{R}_1} d\mathbf{x} - \int_{\mathbf{R}_1} \frac{(I-\mu_{\mathbf{R}_1})^2}{\sigma^2_{\mathbf{R}_1}} d\mathbf{x} \right) \end{aligned} \tag{5.22}$$

We use (5.13) for the parameters of region $\mathbf{R}_1$ in (5.22) to eliminate the first term on the right hand side of (5.18) and obtain the derivative of $\int_{\mathbf{R}_1} G(I, \mu_{\mathbf{R}_1}, \sigma_{\mathbf{R}_1}) d\mathbf{x}$ with respect to time

$$\frac{\partial \int_{\mathbf{R}_1} G(I, \mu_{\mathbf{R}_1}, \sigma_{\mathbf{R}_1}) d\mathbf{x}}{\partial t} = \int_\gamma \langle G(I, \mu_{\mathbf{R}_1}, \sigma_{\mathbf{R}_1}) \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \tag{5.23}$$

Therefore, the curve evolution equation which decreases $\int_{\mathbf{R}_1} G(I, \mu_{\mathbf{R}_1}, \sigma_{\mathbf{R}_1}) d\mathbf{x}$ most rapidly is

$$\frac{\partial \gamma}{\partial t} = -G(I, \mu_{\mathbf{R}_1}, \sigma_{\mathbf{R}_1}) \mathbf{n} \tag{5.24}$$

Application to $\mathscr{L}(\mathscr{P}_\Omega^2 | I, P_{Gaussian})$ leads to (5.15). Therefore, in the case of the Gaussian distribution, the derivative of the region integrals with respect to time behave as for an integrand independent of γ, i.e., the dependence of the parameters on the curve does not give extra terms in the evolution equation. *S*hape derivatives [9, 12, 13] are an alternative to handle region dependent integrands.

Algorithm summary

The algorithm can be summarized as a two-step iteration until convergence:

1. Initialize the level set function ϕ
2. Iterate until convergence
   a) Update the region parameters according to (5.13) for $\mathbf{R} = \mathbf{R}_1, \mathbf{R}_2$ with the regions recovered from ϕ according to (5.6)
   b) Update ϕ according to (5.4) with V given by $V_{Gaussian}$ in (5.15)

Generalization to multivariate Gaussian images

If the image is modeled as multivariate (vectorial) Gaussian within each of the two regions $\mathbf{R}_i$, with mean vector $\mathbf{m}_{\mathbf{R}_i}$ and covariance matrix $\Sigma_{\mathbf{R}_i}$, the data term becomes

$$\sum_{i=1}^{2} \int_{\mathbf{R}_i} G^{vect}(I, \mathbf{m}_{\mathbf{R}_i}, \Sigma_{\mathbf{R}_i}) d\mathbf{x} \tag{5.25}$$

with

$$G^{vect}(I, \mathbf{m}_{\mathbf{R}_i}, \Sigma_{\mathbf{R}_i}) = \frac{n}{2} \log(2\pi) + \frac{1}{2} \log(\det(\Sigma_{\mathbf{R}_i})) + \frac{1}{2} \left((\mathbf{I} - \mathbf{m}_{\mathbf{R}_i})^T \Sigma_{\mathbf{R}_i}^{-1} (\mathbf{I} - \mathbf{m}_{\mathbf{R}_i}) \right) \tag{5.26}$$

where I is of dimension n and det denotes the determinant. In this case, the optimal parameters can also be computed analytically, jointly with the segmentation, and turn out to be the region sample mean vectors and covariance matrices [7, 9, 15]

$$\hat{\mathbf{m}}_{\mathbf{R}_i}^{(l)} = \frac{\int_{\mathbf{R}_i} I^{(l)} d\mathbf{x}}{\int_{\mathbf{R}_i} d\mathbf{x}}$$

$$\hat{\Sigma}_{\mathbf{R}_i}^{(l,k)} = \frac{\int_{\mathbf{R}_i} (I^{(l)} - \hat{\mathbf{m}}_{\mathbf{R}_i}^{(l)})(I^{(k)} - \hat{\mathbf{m}}_{\mathbf{R}_i}^{(k)})\, d\mathbf{x}}{\int_{\mathbf{R}_i} d\mathbf{x}}$$

where $\Sigma^{(l,k)}$ denotes coefficient (l,k) of matrix Σ and $\mathbf{m}^{(l)}$ is the element l of vector $\mathbf{m}$, with $l,k \in [1,\dots,n]$. Proceeding as in the case of a scalar image, we obtain [9, 15]

$$\frac{\partial \gamma}{\partial t} = V_{Gaussian}^{vect} \mathbf{n} \tag{5.27}$$

with

$$V_{Gaussian}^{vect} = -\left(G^{vect}(I, \hat{\mathbf{m}}_{\mathbf{R}_1}, \hat{\Sigma}_{\mathbf{R}_1}) - G^{vect}(I, \hat{\mathbf{m}}_{\mathbf{R}_2}, \hat{\Sigma}_{\mathbf{R}_2})\right) \tag{5.28}$$

5.2.2 The Gamma image model

The Gamma distribution is an accurate model for the important images of the Synthetic Aperture Radar (SAR) family [11]. These images are quite poorly described by the Gaussian model. Due in part to the presence of speckle, a strong multiplicative noise, SAR image segmentation is generally acknowledged as a difficult problem. A radar acquires a complex signal that is the result of several elementary scatters within a resolution cell. For a single look SAR image, the image $I : \Omega \to \mathbb{R}$ is the squared norm of the complex signal. In the case of multi-look SAR images, the L-look image is obtained by averaging L acquired images. A SAR image can be described in each region $\mathbf{R}$ by a Gamma distribution of L looks and parameter $\mu_{\mathbf{R}}$ [16]:

$$P_{Gamma}^{L}(I|\mathbf{R}) = \frac{L^L}{\mu_{\mathbf{R}}(L-1)!} \left(\frac{I}{\mu_{\mathbf{R}}}\right)^{L-1} e^{-\frac{LI}{\mu_{\mathbf{R}}}} \tag{5.29}$$

Therefore, each region is characterized by its parameter $\mu_{\mathbf{R}}$ and the number of looks L, which we assume does not vary with the region. The distribution in (5.29) reduces to the exponential distribution when $L = 1$, a model that has often been used [17, 18, 19]. The data term for the Gamma model is (for two regions)

$$\mathscr{L}(\mathscr{P}_{\Omega}^{2} | I, P_{Gamma}^{L}) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} -\log P_{Gamma}^{L}(I|\mathbf{R}_i) d\mathbf{x} \tag{5.30}$$

As with the Gaussian model, the optimal parameter for a region has an analytic expression, which turns out to be the sample mean within the region (Eq. 5.13).

Embedding the optimal region parameters in the data term in (5.30) yields the following expression [11]:

$$\mathscr{L}(\mathscr{P}^2_{\Omega}|I,P^L_{Gamma}) = L\sum_{i=1}^{2}\mathbf{A}(\mathbf{R}_i)\log\hat{\mu}_{\mathbf{R}_i} + c(L,I) \tag{5.31}$$

where $\mathbf{A}(\mathbf{R})$ denotes the area of region $\mathbf{R}$

$$\mathbf{A}(\mathbf{R}) = \int_{\mathbf{R}} d\mathbf{x} \tag{5.32}$$

and $c(L,I)$ is a constant independent of the segmentation. Omitting this constant, and with the expression of each region parameter as a function of the corresponding region domain, the problem amounts to minimizing the following functional with respect to the segmentation:

$$\sum_{i=1}^{2}\mathbf{A}(\mathbf{R}_i)\log\hat{\mu}_{\mathbf{R}_i} \tag{5.33}$$

Curve evolution equation

The derivative of (5.33) with respect to γ can be written as follows:

$$\begin{aligned}\frac{\partial\left(\mathbf{A}(\mathbf{R}_1)\log\hat{\mu}_{\mathbf{R}_1} + \mathbf{A}(\mathbf{R}_2)\log\hat{\mu}_{\mathbf{R}_2}\right)}{\partial\gamma} &= \log\hat{\mu}_{\mathbf{R}_1}\frac{\partial\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} + \frac{\mathbf{A}(\mathbf{R}_1)}{\hat{\mu}_{\mathbf{R}_1}}\frac{\partial\hat{\mu}_{\mathbf{R}_1}}{\partial\gamma} \\ &+ \log\hat{\mu}_{\mathbf{R}_2}\frac{\partial\mathbf{A}(\mathbf{R}_2)}{\partial\gamma} + \frac{\mathbf{A}(\mathbf{R}_2)}{\hat{\mu}_{\mathbf{R}_2}}\frac{\partial\hat{\mu}_{\mathbf{R}_2}}{\partial\gamma}\end{aligned} \tag{5.34}$$

By the basic formula in (5.8), we have

$$\begin{aligned}\frac{\partial\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} &= \mathbf{n}; \quad \frac{\mathbf{A}(\mathbf{R}_2)}{\partial\gamma} = -\mathbf{n} \\ \frac{\partial\mathbf{S}(\mathbf{R}_1)}{\partial\gamma} &= I\mathbf{n}; \quad \frac{\mathbf{S}(\mathbf{R}_2)}{\partial\gamma} = -I\mathbf{n}\end{aligned} \tag{5.35}$$

where $\mathbf{S}(\mathbf{R})$ is

$$\mathbf{S}(\mathbf{R}) = \int_{\mathbf{R}} I d\mathbf{x} \tag{5.36}$$

Using these derivatives, we compute the functional derivative of each region mean as follows:

$$\begin{aligned}\frac{\partial\hat{\mu}_{\mathbf{R}_1}}{\partial\gamma} &= \frac{\partial\left(\frac{\mathbf{S}(\mathbf{R}_1)}{\mathbf{A}(\mathbf{R}_1)}\right)}{\partial\gamma} = \frac{\mathbf{A}(\mathbf{R}_1)\frac{\partial\mathbf{S}(\mathbf{R}_1)}{\partial\gamma} - \mathbf{S}(\mathbf{R}_1)\frac{\partial\mathbf{A}(\mathbf{R}_1)}{\partial\gamma}}{\mathbf{A}(\mathbf{R}_1)^2} = \frac{(I-\hat{\mu}_{\mathbf{R}_1})}{\mathbf{A}(\mathbf{R}_1)}\mathbf{n} \\ \frac{\partial\hat{\mu}_{\mathbf{R}_2}}{\partial\gamma} &= \frac{\partial\left(\frac{\mathbf{S}(\mathbf{R}_2)}{\mathbf{A}(\mathbf{R}_2)}\right)}{\partial\gamma} = \frac{\mathbf{A}(\mathbf{R}_2)\frac{\partial\mathbf{S}(\mathbf{R}_2)}{\partial\gamma} - \mathbf{S}(\mathbf{R}_2)\frac{\partial\mathbf{A}(\mathbf{R}_2)}{\partial\gamma}}{\mathbf{A}(\mathbf{R}_2)^2} = -\frac{(I-\hat{\mu}_{\mathbf{R}_2})}{\mathbf{A}(\mathbf{R}_2)}\mathbf{n}\end{aligned} \tag{5.37}$$

Substitution of (5.35) and (5.37) in (5.34) yields the functional derivative of the data term (5.33) and, therefore, the descent equation to minimize $\mathscr{L}(\mathscr{P}_\Omega^2|I, P_{Gamma}^L)$

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{L}(\mathscr{P}_\Omega^N|I, P_{Gamma}^L)}{\partial \gamma} = V_{Gamma}\mathbf{n} \tag{5.38}$$

with V_{Gamma} given by

$$V_{Gamma} = -\left(\log \hat{\mu}_{\mathbf{R}_1} + \frac{I}{\hat{\mu}_{\mathbf{R}_1}} - \log \hat{\mu}_{\mathbf{R}_2} - \frac{I}{\hat{\mu}_{\mathbf{R}_2}}\right) \tag{5.39}$$

Algorithm summary

The algorithm can be summarized as a two-step iteration until convergence:

1. Initialize the level set function ϕ
2. Iterate until convergence
   a) Update the optimal region means according to the first equation in (5.13) for $\mathbf{R} = \mathbf{R}_1, \mathbf{R}_2$, with the regions recovered from ϕ according to (5.6)
   b) Update ϕ according to (5.4) with V given by V_{Gamma} in (5.39)

5.2.3 Generalization to distributions of the exponential family

In this section, we examine a generalization of the Gamma-driven curve evolution method studied in the previous section to distributions of the exponential family [2]. This family includes Gamma, Poisson, Bernoulli, and other distributions [20] that can be used to describe the image within the segmentation regions. Consider the class of active curve data terms following a general expression which measures the conformity of the image within each region to a distribution of the exponential family. The data term is

$$\mathscr{L}(\mathscr{P}_\Omega^2|I, P_{EF}) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} -\log P_{EF}(I|\mathbf{R}_i) d\mathbf{x} \tag{5.40}$$

where $P_{EF}(I|\mathbf{R}_i)$ is given by the following general expression

$$P_{EF}(I|\mathbf{R}_i) = k(I)e^{\{a(\alpha_{\mathbf{R}_i})h(I)+b(\alpha_{\mathbf{R}_i})\}} \tag{5.41}$$

$\alpha_{\mathbf{R}}$ is the vector of parameters of region $\mathbf{R}$; $k(I)$ and $h(I)$ are functions of the image independent of the region parameters; $a(\alpha_{\mathbf{R}})$ and $b(\alpha_{\mathbf{R}})$ are functions of the region parameters but independent of the image. For several distributions, the Gamma, Bernoulli, and Poisson, for instance, the optimal region parameters for a given region have an analytic expression determined by the necessary conditions, for $i = 1, 2$,

$$\frac{\partial\left(\int_{\mathbf{R}_i}\log P_{EF}(I|\mathbf{R}_i)d\mathbf{x}\right)}{\partial\alpha_{\mathbf{R}_i}}=0 \tag{5.42}$$

Embedding the optimal parameters in the data term (5.40) yields the following general expression [2]

$$\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{EF})=-K_1\sum_{i=1}^{2}\mathbf{A}(\mathbf{R}_i)H(\theta_{\mathbf{R}_i})+K_2 \tag{5.43}$$

where $\theta_{\mathbf{R}_i}$ and H depend on the distribution. The first row of Table 5.1 gives H for the Bernoulli, Poisson, and Gamma distributions. For these distributions, $\theta_{\mathbf{R}_i}$ turns out to be the mean of the image within region $\mathbf{R}_i$

$$\theta_{\mathbf{R}_i}=\frac{\int_{\mathbf{R}_i}Id\mathbf{x}}{\int_{\mathbf{R}_i}d\mathbf{x}} \tag{5.44}$$

K_1 and K_2 depend on the distribution, but not on the segmentation.

Curve evolution equation

The functional derivative of the data term (5.40) with respect to γ is

$$\frac{\partial\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{EF})}{\partial\gamma}=\sum_{i=1}^{2}\left(\frac{\partial\mathbf{A}(\mathbf{R}_i)}{\partial\gamma}H(\theta_{\mathbf{R}_i})+\mathbf{A}(\mathbf{R}_i)H'(\theta_{\mathbf{R}_i})\frac{\partial\theta_{\mathbf{R}_i}}{\partial\gamma}\right) \tag{5.45}$$

Using the derivatives of region areas and means (5.35) and (5.37) in the descent equation minimizing $\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{EF})$, we arrive at the following evolution equation

$$\frac{\partial\gamma}{\partial t}=-\frac{\partial\mathscr{L}(\mathscr{P}^2_\Omega|I,P_{EF})}{\partial\gamma}=V_{EF}\mathbf{n} \tag{5.46}$$

with V_{EF} given by

$$V_{EF}=-\left(H(\theta_{\mathbf{R}_1})-H(\theta_{\mathbf{R}_2})+H'(\theta_{\mathbf{R}_1})(I-\theta_{\mathbf{R}_1})-H'(\theta_{\mathbf{R}_2})(I-\theta_{\mathbf{R}_2})\right) \tag{5.47}$$

and $H^{'}$ given in the second row of Table 5.1.

Distribution	Bernoulli	Poisson	Gamma
$H(z)$	$-z\log z-(1-z)\log(1-z)$	$-z\log z$	$\log z$
$H'(z)$	$\log(1-z)-\log z$	$-1-\log z$	$\frac{1}{z}$

Table 5.1. Function H that defines the image likelihood data term for different distributions of the exponential family.

5.2.4 The Weibull image Model

In this section, we examine a curve evolution segmentation with a data term which measures the conformity of the image within each region to a Weibull distribution [5]:

$$P_{Weibull}(I|\mathbf{R}_i) = \frac{\beta_{\mathbf{R}_i}}{\eta_{\mathbf{R}_i}} \left(\frac{I}{\eta_{\mathbf{R}_i}} \right)^{\beta_{\mathbf{R}_i}-1} e^{-\left(\frac{I}{\eta_{\mathbf{R}_i}} \right)^{\beta_{\mathbf{R}_i}}} \tag{5.48}$$

Each region $\mathbf{R}_i$ is characterized by a shape parameter $\beta_{\mathbf{R}_i} > 0$ and a scale parameter $\eta_{\mathbf{R}_i} > 0$. Variation of the Weibull parameters yields a spectrum of models which includes the exponential and approximations of the Gaussian and Raleigh. $\beta_{\mathbf{R}} = 1$ describes the image in $\mathbf{R}$ by the exponential distribution. $\beta_{\mathbf{R}} = 3$ gives an approximation of the normal distribution and $\beta_{\mathbf{R}} = 2$ of the Rayleigh (Figure 5.1). Therefore, the Weibull distribution is so versatile as to represent a wide variety of images. It can be useful in applications where the image model is uncertain or when the segmentation regions require different models. Vision studies have used the Weibull distribution to model a variety of signals such as radar [21], sonar [22], and medical images [23, 24], as well as video shot duration [25], and stochastic textures [26].

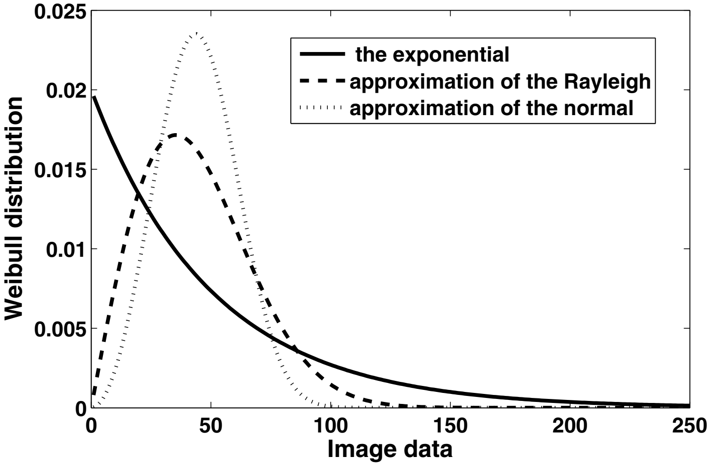


Fig. 5.1. Effect of the shape parameter on the Weibull distribution. $\beta_{\mathbf{R}} = 1$ describes the image in $\mathbf{R}$ by the exponential distribution. $\beta_{\mathbf{R}} = 3$ gives an approximation of the normal distribution and $\beta_{\mathbf{R}} = 2$ of the Rayleigh.

The Weibull model data term is

$$\mathscr{L}(\mathscr{P}_{\Omega}^{2}|I, P_{Weibull}) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} w_{\mathbf{R}_i}(I) d\mathbf{x} \tag{5.49}$$

where

$$\begin{aligned} w_{\mathbf{R}_i}(I) &= -\log P_{Weibull}(I|\mathbf{R}_i) \\ &= -\log\left(\frac{\beta_{\mathbf{R}_i}}{\eta_{\mathbf{R}_i}}\right) - (\beta_{\mathbf{R}_i}-1)\log\left(\frac{I}{\eta_{\mathbf{R}_i}}\right) + \left(\frac{I}{\eta_{\mathbf{R}_i}}\right)^{\beta_{\mathbf{R}_i}} \end{aligned} \tag{5.50}$$

Minimization

The Weibull model optimal region parameters do not have an analytic expression. However, the minimization of (5.49) can proceed by iterative two-stage greedy descent, one to minimize with respect to the region parameters with the curve fixed, and the other with respect to the curve keeping the parameters fixed.

Stage 1: We fix curve γ and minimize the data term (5.49) with respect to the region parameters. After some algebraic manipulations, we have, for $i = 1, 2$,

$$\frac{\partial \mathscr{L}(\mathscr{P}_\Omega^2|I, P_{Weibull})}{\partial \beta_{\mathbf{R}_i}} = -\frac{\mathbf{A}(\mathbf{R}_i)}{\beta_{\mathbf{R}_i}} - \int_{\mathbf{R}_i}\left(\log\frac{I}{\eta_{\mathbf{R}_i}} - \left(\frac{I}{\eta_{\mathbf{R}_i}}\right)^{\beta_{\mathbf{R}_i}}\log\frac{I}{\eta_{\mathbf{R}_i}}\right)d\mathbf{x} \tag{5.51}$$

and the corresponding descent equations are

$$\begin{aligned} \frac{\partial \beta_{\mathbf{R}_i}}{\partial t} &= -\frac{\partial \mathscr{L}(\mathscr{P}_\Omega^2|I, P_{Weibull})}{\partial \beta_{\mathbf{R}_i}} \\ &= \frac{\mathbf{A}(\mathbf{R}_i)}{\beta_{\mathbf{R}_i}} + \int_{\mathbf{R}_i}\left(\log\frac{I}{\eta_{\mathbf{R}_i}} - \left(\frac{I}{\eta_{\mathbf{R}_i}}\right)^{\beta_{\mathbf{R}_i}}\log\frac{I}{\eta_{\mathbf{R}_i}}\right)d\mathbf{x} \end{aligned} \tag{5.52}$$

The derivatives of (5.49) with respect to the scale parameters $\eta_{\mathbf{R}_i}, i = 1, 2$, are

$$\frac{\partial \mathscr{L}(\mathscr{P}_\Omega^2|I, P_{Weibull})}{\partial \eta_{\mathbf{R}_i}} = \frac{\beta_{\mathbf{R}_i}\mathbf{A}(\mathbf{R}_i)}{\eta_{\mathbf{R}_i}} - \beta_{\mathbf{R}_i}\int_{\mathbf{R}_i}\frac{I^{\beta_{\mathbf{R}_i}}}{\eta_{\mathbf{R}_i}^{\beta_{\mathbf{R}_i}+1}}d\mathbf{x} \tag{5.53}$$

which give the following expressions of scale parameters

$$\eta_{\mathbf{R}_i} = \left(\frac{\int_{\mathbf{R}_i} I^{\beta_{\mathbf{R}_i}}d\mathbf{x}}{A(\mathbf{R}_i)}\right)^{\frac{1}{\beta_{\mathbf{R}_i}}} \tag{5.54}$$

When $\beta_{\mathbf{R}_i} = 1$, these updates correspond to the means of the image within the regions.

Stage 2: With the parameters fixed, the derivative of (5.49) with respect to γ can be obtained by application of the basic formula in (5.8). This leads to

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{L}(\mathscr{P}_{\Omega}^{2}|I, P_{Weibull})}{\partial \gamma} = V_{Weibull}\mathbf{n} \tag{5.55}$$

with $V_{Weibull}$ given by

$$V_{Weibull} = -\left(w_{\mathbf{R}_1} - w_{\mathbf{R}_2}\right) \tag{5.56}$$

Algorithm summary

The algorithm can be summarized as follows:

1. Initialize the level set function ϕ
2. Iterate until convergence
    a) Iterate updates of the shape parameters according to (5.52)
    b) Update the scale parameters according to (5.54)
    c) Update ϕ according to (5.4) with V given by $V_{Weibull}$ in (5.56)

Step 2.a can be iterated to convergence. In practice, however, a few iterations generally suffice.

5.2.5 The Complex Wishart Model

In this section, we examine active curve segmentation with a data term based on the Wishart model. This model applies to polarimetric images common in remote sensing [27] and medical imaging [28]. A polarimetric image sensor applies a wave scattering mechanism and different transmission and reception wave polarizations to acquire a signal which consists, at each point of the image domain positional array, of an $n \times n$ (typically 3×3) complex matrix containing both amplitude and phase information. Segmentation of such images is complicated by the complexity of the data, the occurrence of multiplicative random speckle noise due to signal interference, and overlapped region data distributions.

A polarimetric image is a multivalued image consisting at each pixel $\mathbf{x} \in \Omega$ of an $n \times n$ complex Hermitian positive definite matrix $D(\mathbf{x})$. With the Wishart model, we have [27, 29, 30]

$$P_{Wishart}^{L}(D|\mathbf{R}_i) = \frac{\det(D)^{L-n} e^{-Ltr(\Sigma_{\mathbf{R}_i}^{-1} D)}}{K(L,n)\det(\Sigma_{\mathbf{R}_i})^{L}}, \tag{5.57}$$

where det denotes the determinant and tr the trace; L is a fixed integer, and $K(L,n)$ is a constant which depends on L and n. Therefore, each region $\mathbf{R}_i$ is characterized by its covariance matrix $\Sigma_{\mathbf{R}_i}$. The accuracy of the Wishart distribution to model polarimetric images has been verified in several studies [6, 27, 29, 30].

The Wishart data term measures the conformity of D within each region to a Wishart distribution representation [6]:

$$\mathscr{L}(\mathscr{P}_{\Omega}^{2}|I,P_{Wishart}^{L}) = \sum_{i=1}^{2} \int_{\mathbf{R}_i} -\log P_{Wishart}^{L}(D|\mathbf{R}_i)d\mathbf{x} \tag{5.58}$$

The maximum likelihood estimate (MLE) of region covariance matrix, $\hat{\Sigma}_{\mathbf{R}_i}$, can be derived analytically from the necessary condition

$$\frac{\partial \int_{\mathbf{R}_i} \log P_{Wishart}^{L}(D|\mathbf{R}_i)d\mathbf{x}}{\partial \Sigma_{\mathbf{R}_i}} = 0 \tag{5.59}$$

This gives [27]

$$\hat{\Sigma}_{\mathbf{R}_i} = \frac{\int_{\mathbf{R}_i} D d\mathbf{x}}{\mathbf{A}(\mathbf{R}_i)} \tag{5.60}$$

Coefficient (l,k) of this matrix, $l \in [1,\ldots,n]$ and $k \in [1,\ldots,n]$, is given by

$$\hat{\Sigma}_{\mathbf{R}_i}^{(l,k)} = \frac{\int_{\mathbf{R}_i} D^{(l,k)} d\mathbf{x}}{\mathbf{A}(\mathbf{R}_i)} \tag{5.61}$$

Using the expression of the optimal covariance matrices in the data term yields, after some manipulations [6]:

$$\mathscr{L}(\mathscr{P}_{\Omega}^{2}|D,P_{Wishart}^{L}) = L\sum_{i=1}^{2} \mathbf{A}(\mathbf{R}_i)\log\det(\hat{\Sigma}_{\mathbf{R}_i}) + C(L,n,D) \tag{5.62}$$

where $C(L,n,D)$ is a constant (independent of the segmentation) which can be discarded. With the expression of each optimal covariance matrix as a function of the corresponding region domain, the problem amounts to minimizing the following functional with respect to segmentation:

$$\sum_{i=1}^{2} \mathbf{A}(\mathbf{R}_i)\log\det(\hat{\Sigma}_{\mathbf{R}_i}) \tag{5.63}$$

Curve evolution equation

The functional derivative of the $\mathbf{R}_1$ term in (5.63) with respect to γ can be written as

$$\frac{\partial \mathbf{A}(\mathbf{R}_1)\log\det(\hat{\Sigma}_{\mathbf{R}_1})}{\partial\gamma} = \log\det(\hat{\Sigma}_{\mathbf{R}_1})\frac{\partial \mathbf{A}(\mathbf{R}_1)}{\partial\gamma} + \frac{\mathbf{A}(\mathbf{R}_1)}{\det(\hat{\Sigma}_{\mathbf{R}_1})}\frac{\partial\det(\hat{\Sigma}_{\mathbf{R}_1})}{\partial\gamma} \tag{5.64}$$

We also have

$$\begin{aligned}\frac{\partial\det(\hat{\Sigma}_{\mathbf{R}_1})}{\partial\gamma} &= \sum_{l,k}\frac{\partial\det(\hat{\Sigma}_{\mathbf{R}_1})}{\partial\hat{\Sigma}_{\mathbf{R}_1}^{(l,k)}}\frac{\partial\hat{\Sigma}_{\mathbf{R}_1}^{(l,k)}}{\partial\gamma} \\ &= \sum_{l,k} Cof_{\hat{\Sigma}_{\mathbf{R}_1}}^{(l,k)}\frac{\partial\hat{\Sigma}_{\mathbf{R}_1}^{(l,k)}}{\partial\gamma}\end{aligned} \tag{5.65}$$

where Cof_{Σ} denotes the cofactor matrix of Σ. The functional derivative of $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)}$ can be written

$$\frac{\partial \hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)}}{\partial \gamma} = \frac{\partial \left(\frac{\mathbf{S}_D^{(l,k)}(\mathbf{R}_1)}{\mathbf{A}(\mathbf{R}_1)} \right)}{\partial \gamma} = \frac{\mathbf{A}(\mathbf{R}_1) \frac{\partial \mathbf{S}_D^{(l,k)}(\mathbf{R}_1)}{\partial \gamma} - \mathbf{S}_D^{(l,k)}(\mathbf{R}_1) \frac{\partial \mathbf{A}(\mathbf{R}_1)}{\partial \gamma}}{\mathbf{A}(\mathbf{R}_1)^2} \tag{5.66}$$

where $\mathbf{S}_D^{(l,k)}(\mathbf{R}_1)$ is given by

$$\mathbf{S}_D^{(l,k)}(\mathbf{R}_1) = \int_{\mathbf{R}_1} D^{(l,k)} d\mathbf{x} \tag{5.67}$$

Notice that the integrand in this region integral is independent of the curve. Therefore, it suffices to apply the basic formula in (5.8) to obtain the corresponding functional derivative

$$\frac{\partial \mathbf{S}_D^{(l,k)}(\mathbf{R}_1)}{\partial \gamma} = D^{(l,k)} \mathbf{n} \tag{5.68}$$

Embedding in (5.66) this derivative and the derivative of the region area, which we computed earlier in (5.35), we obtain the functional derivative of coefficient (l,k) of the region covariance matrix

$$\frac{\partial \hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)}}{\partial \gamma} = \frac{D^{(l,k)} - \hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)}}{\mathbf{A}(\mathbf{R}_1)} \mathbf{n} \tag{5.69}$$

Using this result in (5.65) yields

$$\frac{\partial \det(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1})}{\partial \gamma} = \frac{1}{\mathbf{A}(\mathbf{R}_1)} \left(\sum_{l,k} Cof_{\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}}^{(l,k)} D^{(l,k)} - \sum_{l,k} Cof_{\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}}^{(l,k)} \hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)} \right) \mathbf{n} \tag{5.70}$$

Now notice the following equalities:

$$\begin{cases} \sum_{l,k} Cof_{\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}}^{(l,k)} D^{(l,k)} = tr(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{-1} D) \det(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}) \\ \sum_{l,k} Cof_{\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}}^{(l,k)} \hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}^{(l,k)} = n \det(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1}) \end{cases} \tag{5.71}$$

Using these equalities in (5.70) gives the functional derivative of the determinant of the region covariance matrix

$$\frac{\partial \det(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1})}{\partial \gamma} = \frac{\det(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}_1})}{\mathbf{A}(\mathbf{R}_1)} \left(tr(\hat{\boldsymbol{\Sigma}}_{\mathbf{R}}^{-1} D) - n \right) \mathbf{n} \tag{5.72}$$

Finally, using in (5.64) this derivative and the derivative of the region area yields, after some manipulations

$$\frac{\partial \mathbf{A}(\mathbf{R}_1) \log \det(\hat{\Sigma}_{\mathbf{R}_1})}{\partial \gamma} = \left(\log \det(\hat{\Sigma}_{\mathbf{R}_1}) + tr(\hat{\Sigma}_{\mathbf{R}_1}^{-1} D) - n \right) \mathbf{n} \tag{5.73}$$

Similarly,

$$\frac{\partial \mathbf{A}(\mathbf{R}_2) \log \det(\hat{\Sigma}_{\mathbf{R}_2})}{\partial \gamma} = -\left(\log \det(\hat{\Sigma}_{\mathbf{R}_2}) + tr(\hat{\Sigma}_{\mathbf{R}_2}^{-1} D) - n \right) \mathbf{n} \tag{5.74}$$

Using (5.73) and (5.74) in the descent equation minimizing the data term in (5.58) leads to the curve evolution equation:

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{L}(\mathscr{P}_\Omega^2 | D, P_{Wishart}^L)}{\partial \gamma} = V_{Wishart} \mathbf{n} \tag{5.75}$$

with $V_{Wishart}$ given by

$$V_{Wishart} = -\left(\log \det(\hat{\Sigma}_{\mathbf{R}_1}) + tr(\hat{\Sigma}_{\mathbf{R}_1}^{-1} D) - \log \det(\hat{\Sigma}_{\mathbf{R}_2}) - tr(\hat{\Sigma}_{\mathbf{R}_2}^{-1} D) \right) \tag{5.76}$$

Algorithm summary

The algorithm can be summarized as a two-step iteration until convergence:

1. Initialize the level set function
2. Iterate until convergence
   a) Update the optimal region covariance matrices according to (5.61) with the regions recovered from ϕ according to (5.6)
   b) Update ϕ according to evolution equation (5.4) with V given by $V_{Wishart}$ in (5.76)

5.2.6 MDL interpretation of the smoothness term coefficient

The image likelihood data terms we examined in this section are often used in conjunction with a smoothness term proportional to the length of the active curve γ. The role of this term is to smooth the segmentation boundaries and prevent the occurrence of small, isolated regions. The choice of the coefficient of this term is important because it can affect the results in a significant way [2]. In this section, we will see that an MDL interpretation [31] of the objective functional prescribes a coefficient value of approximately 2. We consider segmentation functionals of the following general form:

$$\sum_{i=1}^{2} \int_{\mathbf{R}_i} -\log P(I|\mathbf{R}_i) d\mathbf{x} + \lambda \int_{\gamma} ds \tag{5.77}$$

Consider a digital approximation of the image, i.e., a discrete image on a discrete image domain. The code length of its specification is the number of bits to encode γ

and the image within the regions. To encode γ, one needs to encode a starting point $\mathbf{p}$ on γ and its chain code, i.e., the chain of directions to travel from one point to the next on it. For an eight point neighborhood, i.e., each pixel is connected to its eight immediate neighbors, the number of bits to encode the chain code of γ is

$$\log_2 8^{\hat{\mathscr{L}}(\gamma)} = \hat{\mathscr{L}}(\gamma) \log_2 8 \;\text{ bits} \tag{5.78}$$

Because the number of possible locations of the starting point $\mathbf{p}$ is $\hat{\mathbf{A}}(\Omega)$, the number of pixels in the image domain Ω, encoding the curve takes [32]

$$\log_2 \hat{\mathbf{A}}(\Omega) + \hat{\mathscr{L}}(\gamma) \log_2 8 \;\text{ bits} \tag{5.79}$$

If γ consists of $\mathscr{C}$ closed curves, the number of bits to encode it is

$$\mathscr{C} \log_2 \hat{\mathbf{A}}(\Omega) + \hat{\mathscr{L}}(\gamma) \log_2 8 \;\text{ bits} \tag{5.80}$$

To encode the image within the regions, information theory instructs us that the minimum number of bits required to describe the discrete set of observations is the negative base 2 logarithm of the probability of the observations [33]. Therefore, the number of bits to encode the image within the regions is, assuming P is sampled and then discretized,

$$\sum_{i=1}^{2} \sum_{\mathbf{R}_i} -\log_2 P(I|\mathbf{R}_i) d\mathbf{x} \tag{5.81}$$

In nats, the total code length is

$$\sum_{i=1}^{2} \sum_{\mathbf{R}_i} -\log P(I|\mathbf{R}_i) d\mathbf{x} + \log 8 \hat{\mathscr{L}}(\gamma) + \mathscr{C} \log \hat{\mathbf{A}}(\Omega) \tag{5.82}$$

Note that the last term of (5.82) is independent of the segmentation. The first term is a data term and the second a length term. Therefore, this MDL interpretation of (5.77) informs us to use a smoothness term coefficient of approximately 2 ($\log 8$). It is interesting that in the experimental simulations in [2] this value corresponds to the minimum of the mean number of misclassified pixels. We will be able to verify this later with an experimental example.

5.2.7 Generalization to multiregion segmentation

There are several ways of generalizing the two-region active curve segmentation formulations we have examined to an arbitrary but fixed number of regions [6, 15, 34, 35, 36]. These give coupled evolution equations of multiple curves so as to guarantee a partition of the image domain, and were reviewed in Chapter 4. In the experimental examples we give in this chapter, we use the multiregion generalization in [35] which, we recall, represents an image domain partition into N regions by the following mapping between the family $\{\mathbf{R}_{\gamma_i}\}$ of regions enclosed by $N-1$ curves $\{\gamma_i\}$, $i = 1,\ldots,N-1$, and the segmentation regions $\{\mathbf{R}_i\}$, $i = 1,\ldots,N$:

$$\begin{cases} \mathbf{R}_1 = \mathbf{R}_{\gamma_1} \\ \mathbf{R}_2 = \mathbf{R}^c_{\gamma_1} \cap \mathbf{R}_{\gamma_2} \\ \dots \\ \mathbf{R}_i = \mathbf{R}^c_{\gamma_1} \cap \mathbf{R}^c_{\gamma_2} \cap \dots \cap \mathbf{R}^c_{\gamma_{i-1}} \cap \mathbf{R}_{\gamma_i} \\ \dots \\ \mathbf{R}_N = \mathbf{R}^c_{\gamma_1} \cap \mathbf{R}^c_{\gamma_2} \cap \dots \cap \mathbf{R}^c_{\gamma_{i-1}} \cap \mathbf{R}^c_{\gamma_{N-1}} \end{cases} \tag{5.83}$$

By writing the region integrals of the general form $\int_{\mathbf{R}_i} e_i d\mathbf{x}$, $i = 1, \dots, N$, such as those in the functionals we examined earlier, as functions of the characteristic functions of the regions inside and outside the curves, one can use the same derivation as in the two-region formulation to write the functional derivative of each integral (Chapter 4, [35]). Taking the regularization term proportional to the sum of all curve lengths and using the implicit representation of each curve γ_i, $i = 1, \dots, N-1$, as the zero level set of a function $\phi_i(\mathbf{x}, t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$, lead to a system of evolution equations in the following general form

$$\begin{cases} \frac{\partial \phi_1}{\partial t} = -\left(e_1 - \Phi_1 + \lambda \kappa_1\right) \|\nabla \phi_1\| \\ \qquad \vdots \\ \frac{\partial \phi_i}{\partial t} = -\left(e_i - \Phi_i + \lambda \kappa_i\right) \|\nabla \phi_i\| \\ \qquad \vdots \\ \frac{\partial \phi_{N-1}}{\partial t} = -\left(e_{N-1} - \Phi_{N-1} + \lambda \kappa_{N-1}\right) \|\nabla \phi_{N-1}\| \end{cases} \tag{5.84}$$

where e_i is given in Table 5.2 for the Gaussian, Gamma, Weibull, and Wishart distributions.

Φ_i, we recall, is given as a function of ϕ_i by

$$\begin{aligned} \Phi_i = \; & e_{i+1} \chi_{\{\phi_{i+1} > 0\}} \\ & + e_{i+2} \chi_{\{\phi_{i+1} \le 0\}} \chi_{\{\phi_{i+2} > 0\}} \\ & + \dots \\ & + e_{N-1} \chi_{\{\phi_{i+1} \le 0\}} \cdots \chi_{\{\phi_{N-2} \le 0\}} \chi_{\{\phi_{N-1} > 0\}} \\ & + e_N \chi_{\{\phi_{i+1} \le 0\}} \cdots \chi_{\{\phi_{N-2} \le 0\}} \chi_{\{\phi_{N-1} \le 0\}} \end{aligned}$$

with

$$\begin{aligned} \chi_{\{\phi \le 0\}} &= \begin{cases} 1 & \text{if } \phi \le 0 \\ 0 & \text{otherwise} \end{cases} \\ \chi_{\{\phi > 0\}} &= \begin{cases} 1 & \text{if } \phi > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{5.85}$$

and κ_i by

$$\kappa_i = -\mathrm{div}\left(\frac{\nabla \phi_i}{\|\nabla \phi_i\|}\right) \tag{5.86}$$

Model	e_i
Gaussian (scalar images)	$G(I, \mu_{\mathbf{R}_i}, \sigma_{\mathbf{R}_i})$
Gaussian (multivariate images)	$G^{vect}(I, \mathbf{m}_{\mathbf{R}_i}, \Sigma_{\mathbf{R}_i})$
Gamma	$\log \hat{\mu}_i + \frac{I}{\hat{\mu}_i}$
Weibull	$w_{\mathbf{R}_i}$
Wishart	$\log\det(\hat{\Sigma}_{\mathbf{R}_i}) + tr(\hat{\Sigma}_{\mathbf{R}_i}^{-1} D)$

Table 5.2. Expressions of e_i for different distributions.

5.2.8 Examples

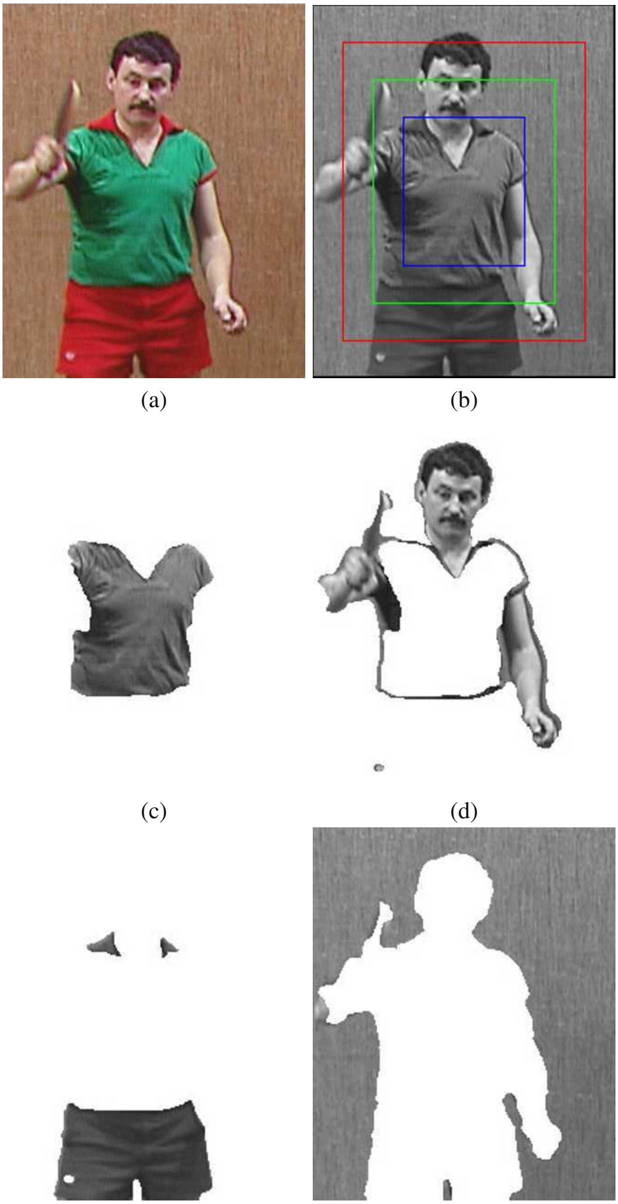
Color image segmentation with the multivariate Gaussian model

Figure 5.2 depicts the segmentation of a real color image into four regions with the multivariate Gaussian model and the multiregion generalization in [35]. The image we used in Figure 5.2 has been converted from its original RGB color space to CIElab [37] and the vector composed by the chrominance components has been used as input to the algorithm. The luminance information has not been included. The original color image is shown in Figure 5.2(a). The initial level set contours are shown in Figure 5.2(b) superimposed on the luminance image. The final four segmented regions are depicted in Figure 5.2 (c), (d), (e), and (f).

For luminance images, sampled by standard cameras, the Gaussian model is often assumed to be applicable. Although effective in several cases, the Gaussian model is not adequately descriptive in general. Images acquired by sensors other than conventional cameras often do not follow a Gaussian distribution, as in medical [23, 24], sonar [22], synthetic aperture radar [11, 21], and polarimetric images [6, 27]. In the following, we examine image segmentation with parametric models other than Gaussian.

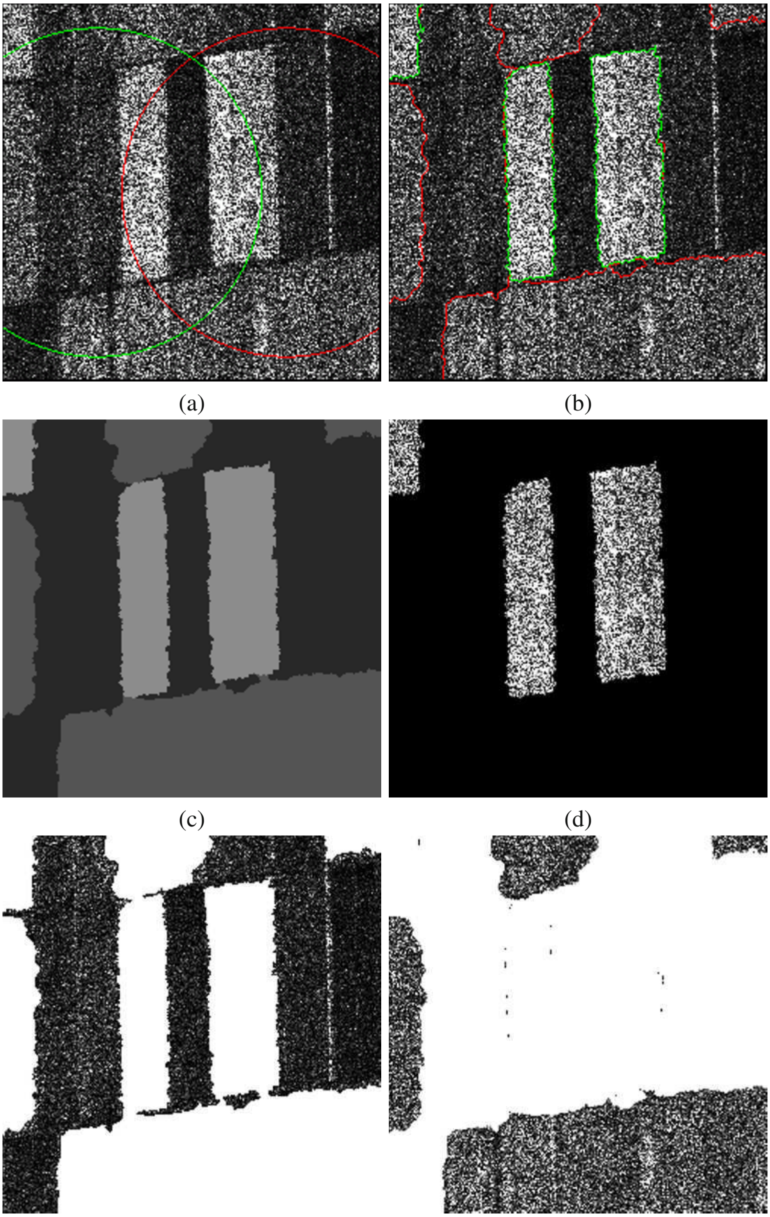
Segmentation of a SAR image with the Gamma model

Figure 5.3 depicts the segmentation into three regions of a 1-look real SAR image of an agricultural land area. Using the Gamma model and the multiregion method in [35], two curves were evolved from the initial positions depicted in Figure 5.3 (a). Figures 5.3 (b) and (c) show, respectively, the final position of the curves and the corresponding segmentation represented by the region means at convergence. Figures (d), (e), and (f) depict the segmentation regions separately.



(a) (b) (c) (d) (e) (f)

Fig. 5.2. Color image segmentation into four regions: (a) original color image; (b) the three initial curves superimposed on the luminance image; (c), (d), (e) and (f) the four segmented regions.



(a) (b)

(c) (d)

(e) (f)

Fig. 5.3. Segmentation of a SAR image into three regions: (a) two initial curves; (b) final curves; (c) the segmentation at convergence represented by the computed region mean parameters; (d), (e), and (f) the segmentation regions displayed separately.

Example of segmentation with the Weibull model

The image shown in Figure 5.4 (a) with two initial curves (red and green) is a synthetic image of three regions of different distributions. The image data in the lighter region is generated from the Gaussian distribution. The gray region is derived from the Rayleigh distribution, and the darker from the Poisson. The segmentation of this image into three regions used the Weibull model and the multiregion generalization in [35]. The final position of the curves and the corresponding segmentation are depicted in, respectively, Figure 5.4 (b) and (c). The results in Figure 5.4 are obtained with the weight of the smoothness term set equal to ones ($\lambda = 1$).

Effect of the weight of the smoothness term

As an illustration of the effect of the weight of the smoothness term, Figure 5.5 shows the segmentations obtained for $\lambda = 0.01, 2$, and 5. The partition boundaries are smoother for $\lambda = 5$ but the segmentations are good for both $\lambda = 2$ and $\lambda = 5$, as are those for the several values between 1 and 5 which we experimented with. In the range between 1 and 5, the differences are quite small. For $\lambda = 0.01$, however, small islands fragment the segmentation quite noticeably. Table 5.3 lists the percentage of correctly classified pixels for various values of λ. The results in this experiment are consistent with the MDL interpretation of the objective functional which prescribes a value of parameter λ approximately equal to 2 [2].

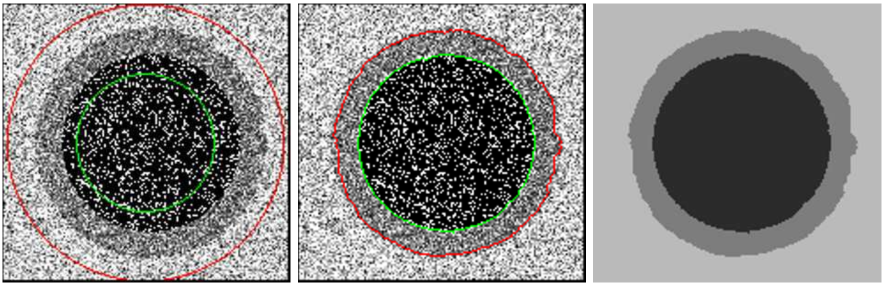

(a) (b) (c)

Fig. 5.4. Segmentation into three regions of a synthetic image with different noise models using the Weibull model and the multiregion generalization in [35]: (a) two initial curves; (b) final position of the curves; (c) segmentation.

5.3 Maximization of the mutual information between the segmentation and the image

In this section, we examine curve evolution by maximization of the mutual information between the segmentation and the image [38]. The ensuing image segmentation

Fig. 5.5. Effect of the weight of the length term: segmentation with $\lambda = 0.01$ in (a), $\lambda = 2$ in (b), and $\lambda = 5$ in (c).

Weight of the length term	Percentage
$\lambda = 0.01$	94.93%
$\lambda = 1$	96.49%
$\lambda = 2$	97.53%
$\lambda = 5$	97.07%

Table 5.3. Portion of correctly classified pixels for the synthetic image with different noise models.

algorithm uses unknown, non-parametric distributions of the image within the segmentation regions. These distributions have to be recovered jointly with the segmentation and, therefore, prior knowledge of their form is not required. Recall that an active curve γ defines a partition of the image domain Ω into two regions, $\mathbf{R}_1$ corresponding to the interior of γ and $\mathbf{R}_2$ to the exterior. This partition yields a binary labeling, $L_\gamma : \Omega \to \{L_1, L_2\}$, which maps the image domain to a set of two labels

$$\forall\, \mathbf{x} \in \Omega \quad L_\gamma(\mathbf{x}) = \begin{cases} L_1 \text{ if } \mathbf{x} \in \mathbf{R}_1 \\ L_2 \text{ if } \mathbf{x} \in \mathbf{R}_2 \end{cases} \tag{5.87}$$

Conversely, this labeling yields a segmentation into two regions. Therefore, segmentation can be viewed as a labeling problem stated as the maximization of the mutual information between the image and the labeling. Generally, the mutual information measures the amount of dependence between two random variables. The labeling L_γ can be viewed as a binary random variable with probabilities

$$\begin{cases} Pr(L_\gamma = L_1) = \frac{\mathbf{A}(\mathbf{R}_1)}{\mathbf{A}(\Omega)} \\ Pr(L_\gamma = L_2) = \frac{\mathbf{A}(\mathbf{R}_2)}{\mathbf{A}(\Omega)} \end{cases} \tag{5.88}$$

The mutual information between image I and labeling L_γ is [38]

$$\mathscr{I}\left(I \| L_\gamma\right) = H - P(L_\gamma = L_1) h(I | L_\gamma = L_1) - P(L_\gamma = L_2) h(I | L_\gamma = L_2) \tag{5.89}$$

where H is a constant and h denotes the conditional entropy, given by

$$h(I|L_\gamma = L_i) = -\int_{\mathbf{R}_i} P(I|\mathbf{R}_i)\log P(I|\mathbf{R}_i)d\mathbf{x}, \;\; i = 1,2 \tag{5.90}$$

The entropy of a distribution measures the amount of *uncertainty* associated with it [39]. As such, each of the conditional entropies in (5.90) measures the degree of randomness (or heterogeneity) of the image within a segmentation region. The more homogeneous the segmentation regions, the lower the conditional entropies and the higher the mutual information. This interpretation supports the use of the mutual information between the labeling and the image as a criterion for segmentation. The conditional entropy $h(I|L_\gamma = L_i)$ corresponds to the expected value of the logarithm of the image distribution in region $\mathbf{R}_i$. Therefore, it can be approximated by the following sample mean [38]

$$h(I|L_\gamma = L_i) \approx -\frac{\int_{\mathbf{R}_i} \log P(I|\mathbf{R}_i)d\mathbf{x}}{\mathbf{A}(\mathbf{R}_i)}, \;\; i = 1,2 \tag{5.91}$$

Using kernel density estimation [40], we can express the image distributions in the segmentation regions as follows

$$P(I|\mathbf{R}_i) = \frac{\varepsilon_{\mathbf{R}_i}}{\mathbf{A}(\mathbf{R}_i)} \tag{5.92}$$

where $\varepsilon_{\mathbf{R}_i} : \Omega \to \mathbb{R}$ is given by

$$\varepsilon_{\mathbf{R}_i}(\mathbf{x}) = \int_{\mathbf{R}_i} K(I(\mathbf{x}) - I(\mathbf{y}))d\mathbf{y} \tag{5.93}$$

with K being the Dirac function or the Gaussian kernel of width σ, given by

$$K(h) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{h^2}{2\sigma^2}} \tag{5.94}$$

When K is the Dirac function, the estimate in (5.92) reduces to the normalized histogram of the image within the corresponding region.

In the following, we examine the maximization of the mutual information by curve evolution, which comes to minimizing the following sum of region integrals

$$-\sum_{i=1}^{2} \int_{\mathbf{R}_i} \log \frac{\varepsilon_{\mathbf{R}_i}}{\mathbf{A}(\mathbf{R}_i)} d\mathbf{x} \tag{5.95}$$

5.3.1 Curve evolution equation

The first region term in functional (5.95) can be written as follows

$$-\int_{\mathbf{R}_1} \log \frac{\varepsilon_{\mathbf{R}_1}}{\mathbf{A}(\mathbf{R}_1)} d\mathbf{x} = \mathbf{A}(\mathbf{R}_1)\log \mathbf{A}(\mathbf{R}_1) - \int_{\mathbf{R}_1} \log \varepsilon_{\mathbf{R}_1} d\mathbf{x} \tag{5.96}$$

In the second term on the right-hand side, the integrand is the logarithm of an integral over the region the curve encloses. As such, it depends on the active curve. To take into account this dependence in the evolution equation, we first compute the equation minimizing functionals in the following general form

$$\mathscr{E} = \int_{\mathbf{R}_\gamma} f(\varepsilon) d\mathbf{x} \tag{5.97}$$

where $\varepsilon : \Omega \to \mathbb{R}$ is given by

$$\varepsilon(\mathbf{x}) = \int_{\mathbf{R}_\gamma} k_{\mathbf{y}}(\mathbf{x}) d\mathbf{y} \tag{5.98}$$

and $k_{\mathbf{y}}$ is a scalar function independent of γ. To do so, we embed ε into a family of functions $\varepsilon(\mathbf{x},t) : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ indexed by algorithmic time t, and compute the derivatives with respect to time of $\mathscr{E}$ and ε using the formula in (5.17)

$$\begin{cases} \frac{\partial \mathscr{E}}{\partial t} = \int_{\mathbf{R}_\gamma} f'(\varepsilon) \frac{\partial \varepsilon}{\partial t} d\mathbf{x} + \int_\gamma \langle f(\varepsilon)\mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \\ \\ \frac{\partial \varepsilon}{\partial t} = \int_\gamma \langle k_\gamma \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \end{cases} \tag{5.99}$$

Note that the region integral in $\frac{\partial \varepsilon}{\partial t}$ is null because the integrand in ε is independent of the curve and, therefore, of time. Embedding the second equation in (5.99) into the first leads to

$$\begin{aligned} \frac{\partial \mathscr{E}}{\partial t} &= \int_{\mathbf{R}_\gamma} f'(\varepsilon) \int_\gamma \langle k_\gamma \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds d\mathbf{x} + \int_\gamma \langle f(\varepsilon)\mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \\ &= \int_{\mathbf{R}_\gamma} \int_\gamma \langle f'(\varepsilon) k_\gamma \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds d\mathbf{x} + \int_\gamma \langle f(\varepsilon)\mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds \\ &= \int_\gamma \langle \left(f(\varepsilon) + \int_{\mathbf{R}_\gamma} f'(\varepsilon) k_\gamma d\mathbf{x} \right) \mathbf{n}, \frac{\partial \gamma}{\partial t} \rangle ds. \end{aligned} \tag{5.100}$$

Therefore, the curve evolution equation which decreases $\mathscr{E}$ most rapidly is

$$\frac{\partial \gamma}{\partial t} = -\left(f(\varepsilon) + \int_{\mathbf{R}_\gamma} f'(\varepsilon) k_\gamma d\mathbf{x} \right) \mathbf{n} \tag{5.101}$$

To apply this to the second term on the right hand side of (5.96), we take $f = -\log$ and (5.98) with

$$k_y(\mathbf{x}) = K(I(\mathbf{x}) - I(y)) \tag{5.102}$$

This gives, after some manipulations, the equation minimizing $\left(-\int_{\mathbf{R}_1} \log \varepsilon_{\mathbf{R}_1} d\mathbf{x} \right)$

$$\frac{\partial \gamma}{\partial t} = \left(\log \mathbf{A}(\mathbf{R}_1) + \log P(I|\mathbf{R}_1) + \frac{1}{\mathbf{A}(\mathbf{R}_1)} \int_{\mathbf{R}_1} \frac{k_\gamma}{P(I|\mathbf{R}_1)} d\mathbf{x} \right) \mathbf{n} \tag{5.103}$$

The curve evolution equation minimizing the first term in (5.96) can be obtained by the following descent equation and the derivative of the region area in (5.35):

$$\begin{aligned}\frac{\partial\gamma}{\partial t} &= -\frac{\partial\mathbf{A}(\mathbf{R}_1)\log\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} = -\log\mathbf{A}(\mathbf{R}_1)\frac{\partial\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} - \mathbf{A}(\mathbf{R}_1)\frac{\partial\log\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} \\ &= -(1+\log\mathbf{A}(\mathbf{R}_1))\frac{\partial\mathbf{A}(\mathbf{R}_1)}{\partial\gamma} = -(1+\log\mathbf{A}(\mathbf{R}_1))\mathbf{n}\end{aligned} \tag{5.104}$$

Finally, using (5.103) and (5.104) gives the curve evolution equation minimizing the first region term in functional (5.95). The equation minimizing the second region integral has a similar form up to the sign. This finally gives the curve evolution equation to minimize functional (5.95)

$$\frac{\partial\gamma}{\partial t} = V_{MI}\mathbf{n} \tag{5.105}$$

with V_{MI} given by

$$V_{MI} = \left(\log\frac{P(I|\mathbf{R}_1)}{P(I|\mathbf{R}_2)} + \frac{1}{\mathbf{A}(\mathbf{R}_1)}\int_{\mathbf{R}_1}\frac{k_\gamma}{P(I|\mathbf{R}_1)}d\mathbf{x} - \frac{1}{\mathbf{A}(\mathbf{R}_2)}\int_{\mathbf{R}_2}\frac{k_\gamma}{P(I|\mathbf{R}_2)}d\mathbf{x}\right) \tag{5.106}$$

and k_y is given by (5.102).

5.3.2 Statistical interpretation

The first term in V_{MI} performs a hypothesis test using an image likelihood ratio which alters the region membership of each point on the active curve so as to increase the homogeneity of each region according to the distribution of the image within the region. It has a form similar to the image velocity in (5.9) we obtained earlier by maximizing the image likelihood when the region models are assumed fixed. Here, however, the region models are variables corresponding to the kernel density estimates of the distributions of the image within the regions inside and outside the active curve, which gives rise to the second and third terms in V_{MI}. These additional terms come from the dependence of the region models on the curve. Their role is to refine and reinforce the effect of the image likelihood ratio which, alone, does not account for the actual variations of the region models induced by curve evolution.

5.3.3 Algorithm summary

The algorithm can be summarized as a two-step iteration until convergence:

1. Initialize the level set function ϕ
2. Iterate until convergence
   a) Compute the kernel density estimates of the image within the regions according to (5.92) as well as the region areas and integrals in the second and third terms in V_{MI}; the regions are recovered from ϕ according to (5.6)
   b) Update ϕ according to (5.4) with V given by V_{MI} in (5.76)

5.4 Segmentation by maximizing the discrepancy between the regions image distributions

In this section, we study segmentation by maximizing the discrepancy (minimizing the similarity) between the distributions of the image in the regions inside and outside the curve. The data term measures, via the Bhattacharyya coefficient, the amount of overlap between the distributions of the image within the region the curve encloses and the region outside [41]:

$$\mathscr{B}(P(.|\mathbf{R}_1)\|P(.|\mathbf{R}_2)) \tag{5.107}$$

where distributions $P(.|\mathbf{R})$ are given by the following kernel density estimates [40]:

$$\forall z \in \mathscr{Z} \quad P(z|\mathbf{R}) = \frac{\int_{\mathbf{R}} K(z-I)d\mathbf{x}}{\mathbf{A}(\mathbf{R})}, \quad \mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2\} \tag{5.108}$$

$\mathscr{Z}$ is the set of values of the image, K the Dirac function or the Gaussian kernel given by (5.94), and $\mathscr{B}(p\|q)$ the Bhattacharyya coefficient measuring the amount of overlap between two distributions p and q:

$$\mathscr{B}(p\|q) = \sum_{z\in\mathscr{Z}} \sqrt{p(z)q(z)} \tag{5.109}$$

Note that the values of $\mathscr{B}$ are in $[0,1]$, where 0 indicates that there is no overlap, and 1 a perfect match. The Bhattacharyya coefficient is an affinity measure. As such, its minimization yields a segmentation into two regions of minimal information similarity.

The functional derivative of $\mathscr{B}(P(.|\mathbf{R}_1)\|P(.|\mathbf{R}_2))$ with respect to γ can be written as follows

$$\begin{aligned}\frac{\partial \mathscr{B}(P(.|\mathbf{R}_1)\|P(.|\mathbf{R}_2))}{\partial \gamma} &= \frac{1}{2}\sum_{z\in\mathscr{Z}} \sqrt{\frac{P(z|\mathbf{R}_2)}{P(z|\mathbf{R}_1)}} \frac{\partial P(z|\mathbf{R}_1)}{\partial \gamma} \\ &+ \frac{1}{2}\sum_{z\in\mathscr{Z}} \sqrt{\frac{P(z|\mathbf{R}_1)}{P(z|\mathbf{R}_2)}} \frac{\partial P(z|\mathbf{R}_2)}{\partial \gamma}\end{aligned} \tag{5.110}$$

The integrands in the region integrals in $P(.|\mathbf{R}_1)$ (Equation 5.108), namely the region area and the integral of $K(z-I)$ over $\mathbf{R}_1$, are independent of the curve. Consequently, applying the basic formula in (5.8) to these integrals gives

$$\begin{aligned}\frac{\partial P(z|\mathbf{R}_1)}{\partial \gamma} &= \frac{\mathbf{A}(\mathbf{R}_1)\frac{\partial \int_{\mathbf{R}_1} K(z-I)d\mathbf{x}}{\partial \gamma} - \int_{\mathbf{R}_1} K(z-I)d\mathbf{x}\frac{\partial \mathbf{A}(\mathbf{R}_1)}{\partial \gamma}}{\mathbf{A}(\mathbf{R}_1)^2} \\ &= \frac{1}{\mathbf{A}(\mathbf{R}_1)}\left(K(z-I) - \mathbf{P}(z|\mathbf{R}_1)\right)\mathbf{n}\end{aligned} \tag{5.111}$$

Similarly,

$$\frac{\partial P(z|\mathbf{R}_2)}{\partial \gamma} = -\frac{1}{\mathbf{A}(\mathbf{R}_2)}\left(K(z-I) - \mathbf{P}(z|\mathbf{R}_2)\right)\mathbf{n} \tag{5.112}$$

Embedding (5.111) and (5.112) into (5.110), and after some algebraic manipulations, we obtain the curve evolution equation to minimize the overlap between the distributions of the image within the segmentation regions:

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{B}(P(.|\mathbf{R}_1)\|P(.|\mathbf{R}_2))}{\partial \gamma} = V^{\mathscr{B}}_{Discrepancy}\mathbf{n} \tag{5.113}$$

with $V^{\mathscr{B}}_{Discrepancy}$ given by

$$\begin{aligned} V^{\mathscr{B}}_{Discrepancy} = & \frac{\mathscr{B}(P(.|\mathbf{R}_1)\|P(.|\mathbf{R}_2))}{2}\left(\frac{1}{\mathbf{A}(\mathbf{R}_1)} - \frac{1}{\mathbf{A}(\mathbf{R}_2)}\right) \\ & + \sum_{z\in\mathscr{Z}} \frac{K(z-I)}{2}\left(\frac{1}{\mathbf{A}(\mathbf{R}_2)}\sqrt{\frac{P(z|\mathbf{R}_1)}{P(z|\mathbf{R}_2)}} - \frac{1}{\mathbf{A}(\mathbf{R}_1)}\sqrt{\frac{P(z|\mathbf{R}_2)}{P(z|\mathbf{R}_1)}}\right) \end{aligned} \tag{5.114}$$

5.4.1 Statistical interpretation

$V_{Discrepancy}$ contains two terms; the first is independent of the spatial coordinates and vanishes with the Bhattacharyya coefficient, i.e., at maximal discrepancy. The second is spatial coordinate dependent and amounts to an image likelihood ratio test which alters the region membership of the points on the active curve so as to increase the homogeneity of each region according to its current image distribution. For a clear interpretation of the second term, assume that K is the Dirac function. The coordinate-dependent velocity at a point $\mathbf{p}$ on the curve is

$$\frac{1}{2}\left(\frac{1}{\mathbf{A}(\mathbf{R}_2)}\sqrt{\frac{P(I(\mathbf{p})|\mathbf{R}_1)}{P(I(\mathbf{p})|\mathbf{R}_2)}} - \frac{1}{\mathbf{A}(\mathbf{R}_1)}\sqrt{\frac{P(I(\mathbf{p})|\mathbf{R}_2)}{P(I(\mathbf{p})|\mathbf{R}_1)}}\right)\mathbf{n} \tag{5.115}$$

The speed is positive when $P(I(\mathbf{p})|\mathbf{R}_1) > P(I(\mathbf{p})|\mathbf{R}_2)$. Therefore, the curve expands to take $\mathbf{p}$ within $\mathbf{R}_1$; this amounts to a likelihood ratio test which rejects the hypothesis that $\mathbf{p}$ belongs to $\mathbf{R}_2$. Otherwise, a negative speed shrinks the curve, causing $\mathbf{p}$ to be in $\mathbf{R}_2$.

5.4.2 The kernel width

Kernel width selection for density estimation has been intensively studied in statistics [42, 43], and there are several ways of computing the optimal value of parameter σ which occurs as a constant in (5.94) and, therefore, in (5.108). Generally, the optimal kernel width minimizes a given error between the original density and its estimate, and depends on the size and statistics of the data [42]. It can be computed accurately via iterative procedures in [42]. However, embedding such procedures into the segmentation process results a heavy computation load. An easier, more efficient alternative is to consider the following approximation of the optimal kernel width for region $\mathbf{R}$ [41, 43]:

$$h = \mathbf{A}(\mathbf{R})^{-\frac{1}{5}} \hat{\sigma}_{\mathbf{R}} \tag{5.116}$$

where $\hat{\sigma}_{\mathbf{R}}$ is the sample standard deviation of the data within region $\mathbf{R}$.

5.4.3 Algorithm summary

The algorithm can be summarized as a two-step iteration until convergence

1. Initialize the level set function ϕ
2. Iterate until convergence
   a) Compute the kernel density estimates of the image within the regions according to (5.108) as well as the region areas; the regions are recovered from ϕ according to (5.6)
   b) Update ϕ according to evolution equation (5.4) with V given by $V_{Discrepancy}$ in (5.114).

5.4.4 Example

The example of Figure 5.6 depicts the two-region segmentation of a real zebra image by the maximization via the Bhattacharyya coefficient of the discrepancy between the image distributions within the two regions. The initial curve is shown in Figure 5.6 (a), two intermediate steps in (b) and (c), and the final curve in (d).

5.5 Image segmentation using a region reference distribution

In this section, we study segmentation using a reference distribution of the image in a region of interest [13, 44, 45]. The problem consists of finding a region within the image domain, so that the distribution of the image within the region most closely matches a model distribution learned *a priori*. Image segmentation using a reference distribution of the image in a region of interest serves various computer vision tasks such as tracking a moving object in an image sequence, co-segmentation, a process which consists of finding a common region in image pairs, and content-based image retrieval [46]. The model distribution can be learned *a priori* from a pre-segmented training image, for instance an initial frame in an image sequence, or from an image having similar contents as the image of interest. It can also be learned interactively from user-specified regions in the image of interest.

In the following, we will examine two criteria for matching the image distribution within the region the active curve encloses, $\mathbf{R}_1$, and the model distribution $\mathbf{M}$. One

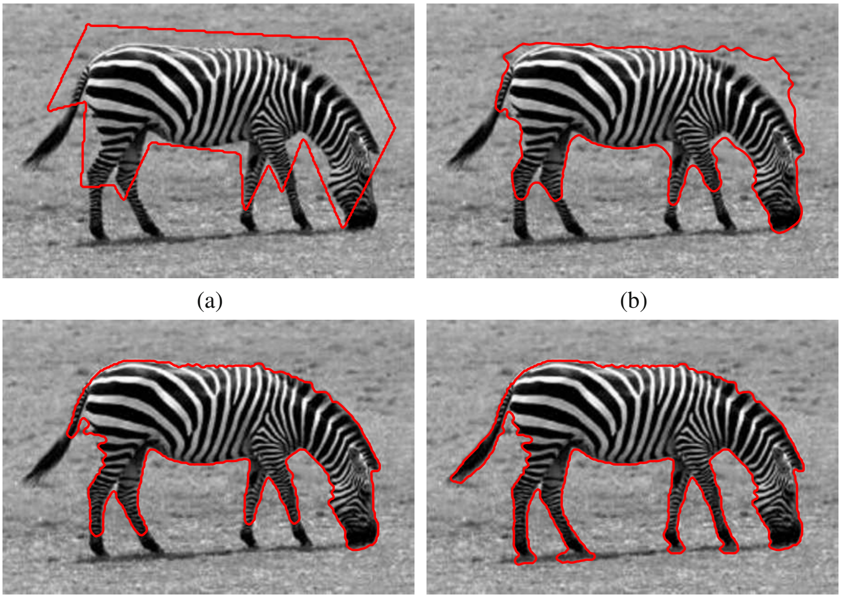


Fig. 5.6. Two-region segmentation of a real zebra image by the maximization via the Bhattacharyya coefficient of the discrepancy between the image distributions: (a) initial curve; (b) and (c) intermediate steps; (d) final curve.

criterion is based on the Bhattacharyya coefficient and the other on the Kullback-Leibler divergence [44]. Since the Bhattacharyya coefficient is an affinity measure, we shall seek its maximum rather than minimum. With the Bhattacharyya coefficient, we maximize

$$\mathscr{B}(P(.|\mathbf{R}_1)\|\mathbf{M}) \tag{5.117}$$

where $P(z|\mathbf{R}_1)$ is the kernel density estimate in (5.108). The functional derivative of the data term in (5.117) with respect to γ can be written as follows

$$\frac{\partial \mathscr{B}(P(.|\mathbf{R}_1)\|\mathbf{M})}{\partial \gamma} = \frac{1}{2}\sum_{z\in\mathscr{Z}} \sqrt{\frac{\mathbf{M}(z)}{P(z|\mathbf{R}_1)}}\frac{\partial P(z|\mathbf{R}_1)}{\partial \gamma} \tag{5.118}$$

Combining this equation and the derivative of the region distribution in (5.111) gives the ascent curve evolution equation to maximize (5.117)

$$\frac{\partial \gamma}{\partial t} = \frac{\partial \mathscr{B}(P(.|\mathbf{R}_1)\|\mathbf{M})}{\partial \gamma} = V^{\mathscr{B}}_{Matching}\mathbf{n} \tag{5.119}$$

with $V^{\mathscr{B}}_{Matching}$ given by

$$V^{\mathscr{B}}_{Matching} = \frac{1}{2\mathbf{A}(\mathbf{R}_1)} \left(\sum_{z \in \mathscr{Z}} K(z-I) \sqrt{\frac{\mathbf{M}(z)}{P(z|\mathbf{R}_1)}} - \mathscr{B}(P(.|\mathbf{R}_1)\|\mathbf{M}) \right) \tag{5.120}$$

Another way is to minimize the Kullback-Leibler divergence between model **M** an the distribution of the image within $\mathbf{R}_1$:

$$\mathscr{K}(P(.|\mathbf{R}_1)\|\mathbf{M}) = \sum_{z \in \mathscr{Z}} \mathbf{M}(z) \log \frac{\mathbf{M}(z)}{P(z|\mathbf{R}_1)} \tag{5.121}$$

The Kullback-Leibler divergence measures the discrepancy between two distributions. As such, its minimization yields a region whose image distribution most closely matches **M**. The functional derivative of the data term in (5.121) with respect to γ is

$$\frac{\partial \mathscr{K}(P(.|\mathbf{R}_1)\|\mathbf{M})}{\partial \gamma} = - \sum_{z \in \mathscr{Z}} \frac{\mathbf{M}(z)}{P(z|\mathbf{R}_1)} \frac{\partial P(z|\mathbf{R}_1)}{\partial \gamma} \tag{5.122}$$

Combining this equation and the derivative of the region distribution in (5.111) leads to the following descent curve evolution equation to minimize (5.121):

$$\frac{\partial \gamma}{\partial t} = - \frac{\partial \mathscr{K}(P(.|\mathbf{R}_1)\|\mathbf{M})}{\partial \gamma} = V^{\mathscr{K}}_{Matching} \mathbf{n} \tag{5.123}$$

with $V^{\mathscr{K}}_{Matching}$ given by

$$V^{\mathscr{K}}_{Matching} = \frac{1}{\mathbf{A}(\mathbf{R}_1)} \left(\sum_{z \in \mathscr{Z}} K(z-I) \frac{\mathbf{M}(z)}{P(z|\mathbf{R}_1)} - 1 \right) \tag{5.124}$$

5.5.1 Statistical interpretation

For a clear interpretation of the Kullback-Leibler divergence velocity, let K be the Dirac function. The velocity at a point **p** on the curve is

$$\frac{1}{\mathbf{A}(\mathbf{R}_1)} \left(\frac{\mathbf{M}(I(\mathbf{p}))}{P(I(\mathbf{p})|\mathbf{R}_1)} - 1 \right) \mathbf{n} \tag{5.125}$$

It is clear that this velocity amounts to a hypothesis test by an image likelihood ratio. The test evaluates the hypothesis that the image at point **p** is drawn from model **M** against the hypothesis that it is instead drawn from the image distribution in the interior of the curve. If the image distribution in the interior of the curve at $I(\mathbf{p})$ is lower than for model **M**, the likelihood ratio $\frac{\mathbf{M}(I(\mathbf{p}))}{P(I(\mathbf{p})|\mathbf{R}_1)}$ is greater than 1 and causes the curve to expand to include **p** in $\mathbf{R}_1$. This make sense because it results in increasing the image distribution within the curve at value $I(\mathbf{p})$, which means a better match with the model at that value.

When K is the Dirac function, the velocity at a point **p** on the curve using the Bhattacharyya coefficient is

$$\frac{1}{2\mathbf{A}(\mathbf{R}_1)}\left(\sqrt{\frac{\mathbf{M}(I(\mathbf{p}))}{P(I(\mathbf{p})|\mathbf{R}_1)}}-\mathscr{B}(P(.|\mathbf{R}_1)||\mathbf{M})\right)\mathbf{n} \tag{5.126}$$

This velocity has a similar meaning as (5.125). The difference is that the likelihood ratio is compared to a variable dependent on the current Bhattacharyya coefficient rather than to 1. Notice that when $P(.|\mathbf{R}_1)$ matches perfectly $\mathbf{M}$, the likelihood ratio and Bhattacharyya coefficient equal 1. This means that the velocities in (5.125) and (5.126) become zero, which forces the curve to stop.

5.5.2 Summary of the algorithms

The algorithms can be summarized as follows:

1. Initialize the level set function ϕ
2. Iterate until convergence
    a) Update the area of region $\mathbf{R}_1$,
    b) Update the kernel density estimate of the image distribution in region $\mathbf{R}_1$ according to (5.108),
    c) If the Bhattacharyya coefficient is used, update (5.117)
    d) Update ϕ according to Equation (5.4) with V given by $V^{\mathscr{K}}_{Matching}$ in (5.124) for the Kullback–Leibler divergence, or $V^{\mathscr{B}}_{Matching}$ in (5.120) for the Bhattacharyya coefficient.

5.5.3 Example

Figure 5.7 illustrates the segmentation of a person against a background in a color image, using a model learned from the same image and the Bhattacharyya coefficient. The image used for segmentation was obtained from the original RGB representation by

$$I = R + 16G + 256B \tag{5.127}$$

The initial curve is shown in Figure 5.7 (a), an intermediate position in (b), and the final in (c).

Figure 5.8 plots the evolution of the maximized affinity measure, the Bhattacharyya coefficient in this example, as a function of the number of iterations. It shows how the Bhattacharyya coefficient approaches 1 (the maximum value in this example where the model has been learned from the image itself; in actual situations, of course, the training and test images are distinct) during curve evolution.

5.6 Segmentation with an overlap prior

Methods using a data term which maximizes the discrepancy between the image distributions in the segmentation regions may perform poorly in situations, for instance,

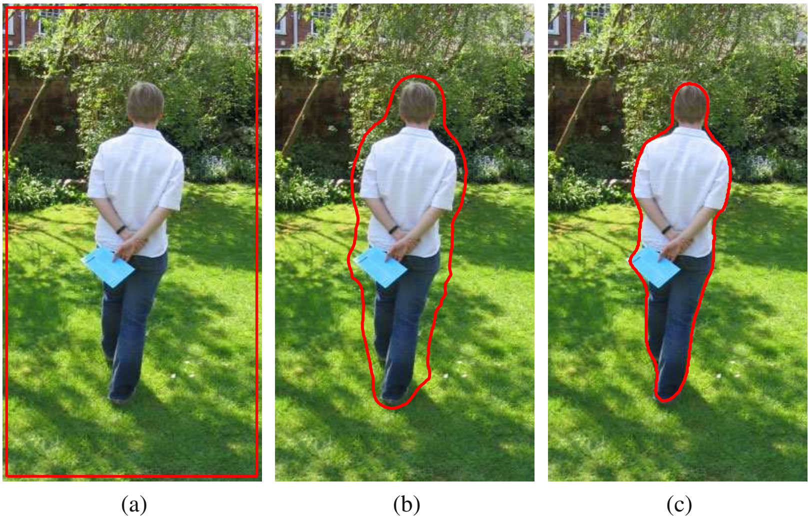

(a) (b) (c)

Fig. 5.7. Color image segmentation with a reference distribution: initial, intermediate, and final curves are superimposed on the image in, respectively, (a), (b), and (c).

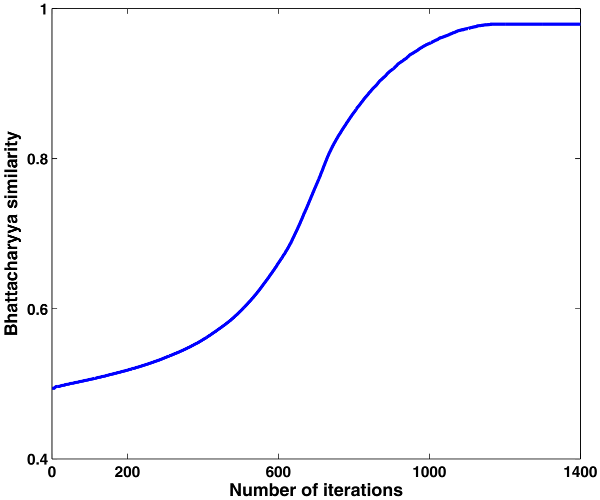


Fig. 5.8. Color image segmentation using a reference distribution of the region of interest: evolution of the Bhattacharyya coefficient as a function of the number of iterations.

where parts of the segmentation regions have the same image statistics, i.e., when there is a significant overlap between the image distributions in the segmentation regions. This occurs frequently in medical image segmentation. Figure 5.9 depicts a typical instance, where parts of the desired regions in a cardiac image have approximately the same image profile. This results in a significant overlap between the distributions of these regions as depicted in Figure 5.9 (b). In this case, prior information about such an overlap is useful to recover the desired segmentation. In the following, we examine curve evolution minimization of a data term which measures the conformity to a learned description of the overlap between the image distributions in the segmentation regions [47].

Again, let us assume that the target region to be delineated by the active curve is characterized by a model **M** learned *a priori* from a segmented training image, for instance an initial frame in an image sequence. Consider the following Bhattacharyya measure of overlap between the distribution of the image outside the curve and the model of the region inside:

$$\mathscr{O} = \mathscr{B}(\mathbf{M} \| P(.|\mathbf{R}_2)) \tag{5.128}$$

Further assume that a prior estimate $\mu_{\mathscr{O}}$ of $\mathscr{O}$ is learned beforehand from a pre-segmented training image. In order to incorporate information about the overlap between the image distributions in the segmentation regions, one can minimize the following overlap prior which measures the conformity of $\mathscr{O}$ to the learned overlap $\mu_{\mathscr{O}}$ [47]:

$$|\mathscr{O} - \mu_{\mathscr{O}}| \tag{5.129}$$

where $|.|$ denotes the absolute value. The functional derivative of this data term with respect to γ can be written as follows:

$$\frac{\partial |\mathscr{O} - \mu_{\mathscr{O}}|}{\partial \gamma} = \frac{\mathscr{O} - \mu_{\mathscr{O}}}{|\mathscr{O} - \mu_{\mathscr{O}}|} \frac{\partial \mathscr{O}}{\partial \gamma} \tag{5.130}$$

The derivative of $\mathscr{O}$ with respect to γ can be written as follows:

$$\frac{\partial \mathscr{O}}{\partial \gamma} = \frac{1}{2} \sum_{z \in \mathscr{Z}} \sqrt{\frac{\mathbf{M}(z)}{P(z|\mathbf{R}_2)}} \frac{\partial P(z|\mathbf{R}_2)}{\partial \gamma} \tag{5.131}$$

Combining this equation and the derivative of the region distribution in (5.112), and embedding the result in (5.130), we obtain the following descent curve evolution equation minimizing (5.129):

$$\frac{\partial \gamma}{\partial t} = -\frac{\partial \mathscr{O}}{\partial \gamma} = V_{Overlap} \mathbf{n} \tag{5.132}$$

with $V_{Overlap}$ given by

$$V_{Overlap} = \frac{\mathscr{O} - \mu_{\mathscr{O}}}{2\mathbf{A}(\mathbf{R}_2)|\mathscr{O} - \mu_{\mathscr{O}}|} \left(\sum_{z \in \mathscr{Z}} K(z - I) \sqrt{\frac{\mathbf{M}(z)}{P(z|\mathbf{R}_2)}} - \mathscr{O} \right) \tag{5.133}$$

5.6.1 Statistical interpretation

For a clearer interpretation of the overlap prior velocity, assume that K is the Dirac function. The velocity at a point $\mathbf{p}$ on the curve is

$$\frac{\mathcal{O}-\mu_{\mathcal{O}}}{2\mathbf{A}(\mathbf{R}_2)|\mathcal{O}-\mu_{\mathcal{O}}|}\left(\sqrt{\frac{\mathbf{M}(I(\mathbf{p}))}{P(I(\mathbf{p})|\mathbf{R}_2)}}-\mathcal{O}\right)\mathbf{n} \tag{5.134}$$

This velocity is the product of two terms; one, which is coordinate dependent, can be viewed as performing a pointwise hypothesis test by an image likelihood ratio; the other is independent of the spatial coordinates and can be viewed as performing a global overlap test. Notice that the hypothesis testing term has a form similar to the velocity we examined earlier in (5.126). As such, its role is to increase or decrease the similarity, i.e., the overlap between the image distribution outside the curve and model $\mathbf{M}$. To examine how the overlap test controls curve evolution, consider the following two cases.

Case 1: The curve is initialized so as to define a positive overlap test, i.e., $\mathcal{O} > \mu_{\mathcal{O}}$.

If the model at $\mathbf{p}$ is such that $\mathbf{M}(I(\mathbf{p})) > P(I(\mathbf{p})|\mathbf{R}_2)\mathcal{O}^2$, the velocity (along the outward normal) is positive, thereby excluding $\mathbf{p}$ from the region outside the curve. Otherwise, the pixel is included. Therefore, the pointwise hypothesis test guides curve evolution so as to decrease the overlap between the model and the distribution of the image outside the curve. This makes sense because the current overlap is higher than its expected value. When the overlap reaches its expected value $\mu_{\mathcal{O}}$, overlap test term becomes zero, forcing the curve evolution to end.

Case 2: The curve is initialized so as to define a negative overlap test, i.e., $\mathcal{O} < \mu_{\mathcal{O}}$.

In this case, the pointwise hypothesis test guides curve evolution so as to increase the overlap between the model and the image distribution outside the curve. Similar to case 1, the overlap test term ends curve evolution when the overlap reaches its expected value.

In summary, the velocity obtained from the overlap prior can be viewed as a pointwise hypothesis test controlled by a global overlap test as follows.

1. The overlap test ends hypothesis testing when the expected overlap measure $\mu_{\mathcal{O}}$ is reached.
2. Given the initial overlap, the sign of the overlap test term defines the direction of the hypothesis testing velocity, i.e., decides whether the hypothesis testing term should increase or decrease the discrepancy between the model and the image distribution outside the curve.

5.6.2 Example

The cardiac magnetic resonance (MR) image segmentation example in Figure 5.9 illustrates the use of an overlap prior. The purpose is to delineate the heart cavity in a MR cardiac sequence given a user provided delineation in an initial frame. This task, of interest in automating the diagnosis of cardiovascular diseases [48], is difficult because the intensity profiles of the cavity and the nearby background are similar. Figure 5.9 depicts the expected (manually performed) delineation of the cavity by the red curve in (a), a region in the nearby background enclosed within a blue curve and intersecting the right ventricle in (a), and the significant overlap between the image distributions in these two regions in (b). The ground truth cavity region corresponding to the manual delineation is shown in (c). We ran three experiments to segment the image according to three different data terms, one containing an overlap prior and the two others not. For the three terms, we used the same curve initialization and region models learned *a priori* from a previous frame. We considered (1) segmentation with a reference distribution (SRD), i.e., the data term in (5.117) whose optimization seeks a region consistent with a learned model; (2) segmentation with a reference distribution and an overlap prior (SRDOP), i.e., the data term uses both (5.117) and the overlap prior in (5.129); and (3) segmentation with a likelihood prior (SLP), i.e., the data term is the logarithm of the image likelihood in (5.7) with the region models fixed (learned a *priori*). Figure 5.9 (d) and (e) depict the regions obtained with, respectively, SRD and SLP. With these methods, part of the background, which has an intensity profile similar to the cavity, is included in the final region, whereas adding an overlap prior, i.e., using SRDOP, biases the solution accurately toward the cavity, giving the result in (f) which is very similar to the ground truth in (c). The curve initialization, an intermediate step with SRDOP, and the final curve with SRDOP are depicted, respectively, in (g), (h), and (i).

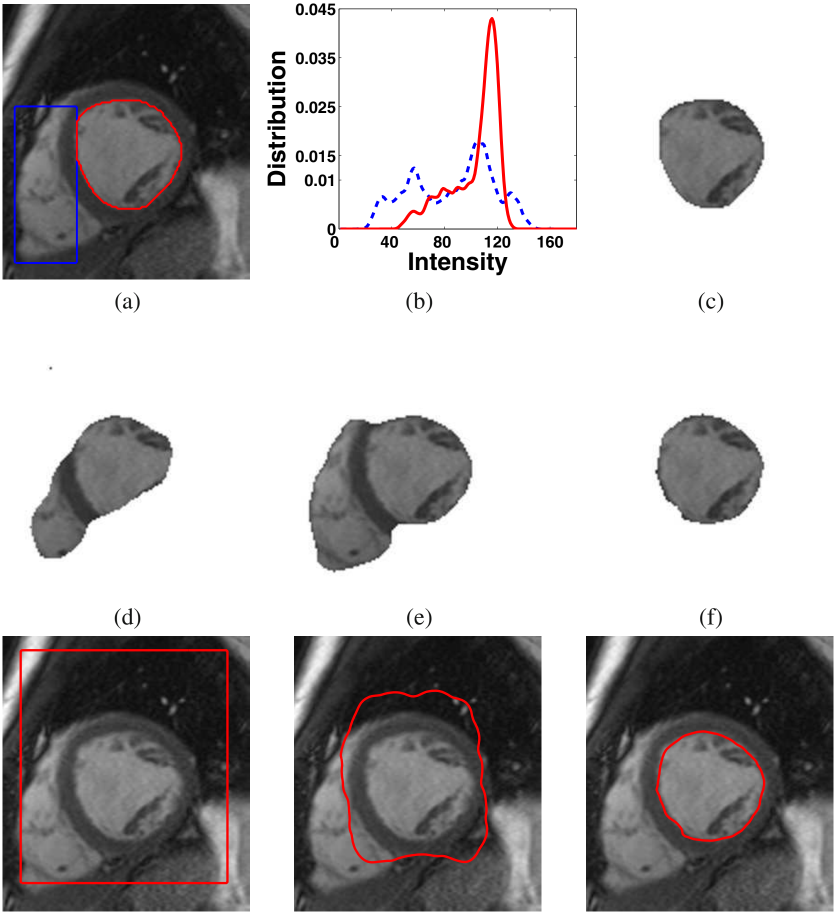
To illustrate quantitatively the effect of the overlap prior, Table 5.4 gives the following image statistics of the segmentations obtained with SRD, SLP, and SRDOP.

- *Similarity at convergence*: The Bhattacharyya measure of similarity between the distribution of the image within the curve at convergence and the model.
- *Overlap at convergence*: the Bhattacharyya measure of overlap between the regions inside and outside the curve at convergence.

For SRD and SRDOP, we obtained approximately the same similarity to the model (refer to the first row of Table 5.4), although the corresponding regions are different (refer to the second row in Figure 5.9). The measure of overlap obtained with SRDOP is different from those obtained with SRD and SLP, and is approximately equal to the learned prior (refer to the second row of Table 5.4). Consequently, matching the distribution inside the curve to a model is not sufficient in such situations where different regions in the image have similar distributions. The overlap prior allowed a correct segmentation. This typical example occurs frequently in medical image segmentation, where parts of different regions have approximately the same image profile. In such cases, SRD and SLP may not be able to recover accurately the target segmentation and the overlap information can be useful.

Method	SRD	SLP	SRDOP
Similarity at convergence	0.98	0.95	0.98
Overlap at convergence	0.50	0.45	0.54

Table 5.4. Cardiac image segmentation example: image statistics corresponding to the segmentations obtained with SRD, SLP, and SRDOP. The learned overlap measure is equal to $\mathbf{0}.53$.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 5.9. A typical cardiac magnetic resonance image: (a) target delineation (manual) of the left ventricle cavity (red curve); (b) overlap between the distributions of the cavity and the nearby background (region inside the blue curve in (a)); (c) target region (the left ventricle cavity); (d), (e), and (f) depict the regions obtained with, respectively, SRD, SLP, and SRDOP; (g), (h), and (i) depict the initial curve, an intermediate step, and the final curve with SRDOP. For the three energies, we used the same initialization in (g) and the same learning frame.

References

1. C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia, "A level set model for image classification," *International Journal of Computer Vision*, vol. 40, no. 3, pp. 187–197, 2000.
2. P. Martin, P. Réfrégier, F. Goudail, and F. Guérault, "Influence of the noise model on level set active contour segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 799–803, 2004.
3. T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
4. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, 2007.
5. I. Ben Ayed, N. Hennane, and A. Mitiche, "Unsupervised variational image segmentation/classification using a weibull observation model," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3431–3439, 2006.
6. I. Ben Ayed, A. Mitiche, and Z. Belhadj, "Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.
7. S. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 118, no. 9, pp. 884–900, 1996.
8. 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, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, vol. 13, no. 1/2, pp. 249–268, 2002.
9. M. Rousson and R. Deriche, "A variational framework for active and adaptative segmentation of vector valued images," in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 56–61.
10. Y. Boykov and G. Funka Lea, "Graph cuts and efficient n-d image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
11. I. Ben Ayed, A. Mitiche, and Z. Belhadj, "Multiregion level set partitioning on synthetic aperture radar images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 793–800, 2005.
12. S. Jehan-Besson, M. Barlaud, and G. Aubert, "Dream2s: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation," *International Journal of Computer Vision*, vol. 53, no. 1, pp. 45–70, 2003.
13. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients?" *SIAM Journal of Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.
14. M. C. Delfour and J. P. Zolesio, *Shapes and Geometries: Analysis, Differential Calculus and Optimization*. SIAM series on Advances in Design and Control, 2001.
15. C. Vázquez, A. Mitiche, and I. Ben Ayed, "Image segmentation as regularized clustering: A fully global curve evolution method," in *International Conference on Image Processing (ICIP)*, 2004, pp. 3467–3470.
16. J. W. Goodman, "Laser speckle and related phenomena," in *Statistical Properties of Laser Speckle Patterns*, 1974, pp. 9–75.
17. O. Germain and P. Réfrégier, "Statistical active grid for segmentation refinement," *Pattern Recognition Letters*, vol. 22, no. 10, pp. 1125–1132, 2001.

18. F. Galland, N. Bertaux, and P. Réfrégier, "Minimum description length synthetic aperture radar image segmentation," *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 995–1006, 2003.
19. O. Germain and P. Réfrégier, "Edge location in sar images: performance of the likelihood ratio filter and accuracy improvement with an active contour approach," *IEEE Transactions on Image Processing*, vol. 10, no. 1, pp. 72–78, 2001.
20. A. Azzalini, *Statistical inference based on the likelihood*. Chapman and Hall, 1996.
21. M. di Bisceglie and C. Galdi, "Cfar detection of extended objects in high-resolution sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 833–843, 2005.
22. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical mrf model," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216–1231, 2000.
23. R. Archibald, J. Hu, A. Gelb, and G. E. Farin, "Improving the accuracy of volumetric segmentation using pre-processing boundary detection and image reconstruction," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 459–466, 2004.
24. J. Hu, A. Razdan, G. M. Nielson, G. E. Farin, D. P. Baluch, and D. G. Capco, "Volumetric segmentation using weibull e-sd fields," *IEEE Transactions Visualization and Computer Graphics*, vol. 9, no. 3, pp. 320–328, 2003.
25. N. Vasconcelos and A. Lippman, "Statistical models of video structure for content analysis and characterization," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 3–19, 2000.
26. J. M. Geusebroek and A. W. M. Smeulders, "A six-stimulus theory for stochastic texture," *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 7–16, 2005.
27. J. M. Beaulieu and R. Touzi, "Segmentation of textured polarimetric sar scenes by likelihood approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 2063–2072, 2004.
28. S. L. Jacques, J. C. Ramella-Roman, and K. Lee, "Imaging skin pathology with polarized light," *Journal for Biomedical Optics*, vol. 7, pp. 329–340, 2002.
29. K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex wishart distribution and its application to change detection in polarimetric sar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, 2003.
30. P. R. Kersten, J.-S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and em clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 519–527, 2005.
31. J. Rissanen, *Stochastic Complexity in Statistical Inquiry*. Singapore: World Scientific, 1989.
32. H. Freeman, "On the encoding of arbitrary geometric configurations," *IRE Transactions on Electronic Computers*, vol. 10, no. 2, pp. 260–268, 1961.
33. J. Rissanen, "A universal prior for integers and estimation by minimum description length," *The Annals of Statistics*, vol. 11, no. 2, pp. 416–431, 1983.
34. L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
35. A. Mansouri, A. Mitiche, and C. Vázquez, "Multiregion competition: A level set extension of region competition to multiple region partioning," *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
36. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of Computational Physics*, vol. 127, no. 1, pp. 179–195, 1996.

37. C. Poynton, *A technical introduction to digital video*. John Wiley & Sons, Inc., 1996.
38. J. Kim, J. W. F. III, A. J. Yezzi, M. Cetin, and A. S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1486–1502, 2005.
39. C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.
40. E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, p. 1065–1076, 1962.
41. O. V. Michailovich, Y. Rathi, and A. Tannenbaum, "Image segmentation using active contours driven by the bhattacharyya gradient flow," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2787–2801, 2007.
42. M. C. Jones, J. S. Marron, and S. Sheather, "A brief survey of bandwidth selection for density estimation," *Journal of the American Statistical Association*, vol. 91, no. 433, p. 401–407, 1996.
43. P. Deheuvels, "Estimation nonparamétrique de la densite par histogrammes généralisés," *Revue de Statistique Appliquée*, vol. 25, pp. 5–42, 1977.
44. D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 518–526, 2004.
45. T. Zhang and D. Freedman, "Improving performance of distribution tracking through background mismatch," *IEEE Transacions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 282–287, 2005.
46. C. Rother, T. P. Minka, A. Blake, and V. Kolmogorov, "Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2006, pp. 993–1000.
47. I. Ben Ayed, S. Li, and I. Ross, "A statistical overlap prior for variational image segmentation," *International Journal of Computer Vision*, vol. 85, no. 1, pp. 115–132, 2009.
48. ——, "Embedding overlap priors in variational left ventricle tracking," *IEEE Transactions on Medical Imaging*, vol. 28, no. 12, pp. 1902–1913, 2009.

6

REGION MERGING PRIORS

6.1 Introduction

Although the actual number of segmentation regions is not known in several applications, current active curve methods assume that it is given beforehand. In the methods we examined so far, the number of regions occurs as a constant in the objective functional and its optimization. How to allow it to vary is an important question which has been generally avoided. A few studies have considered determining the number of regions automatically, either during curve evolution [1] or as a process external to curve evolution optimization [2, 3, 4]. In these studies, a *region merging prior* is added to the segmentation functional to alter the number of regions. Region merging priors cause the objective functional to decrease when regions are merged. A maximum number of regions, which is available in most applications, is used as an initial number of regions. Under the effect of a region merging prior, the effective number of regions, equal to the maximum number of regions initially, decreases automatically during the optimization of the objective functional to be, ideally, the desired number of regions. In this chapter, we examine two region merging priors. One is related to the regions logarithmic area and has an entropic interpretation [1]; the other is proportional to the number of regions and has a minimum description length (MDL) interpretation [2, 3, 4]. There are fundamental differences between the two priors. The MDL prior is independent of the curves and, therefore, does not affect the curve evolution equations. To alter the number of regions with the MDL prior, the studies [2, 3] alternate local region-merging operations[1] and curve evolution with a fixed number of regions. In this case, the objective functional optimization consists of several stages, each stage consisting of curve evolution optimization followed by local region-merging operations. By contrast, the entropic region merging prior encourages region merging to occur intrinsically with curve evolution. As a result, it allows the number of regions to vary automatically during curve evolution, thereby optimizing the objective functional implicitly with respect to the number of regions. In this

---

[1] Region merging techniques use the repeated application of a statistical test on neighboring regions to decide whether to merge them [5, 6].

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_6,

case, a maximum (initial) number of regions is used as a constant in the definition of the segmentation functional and the optimization consists of a single stage of curve evolution during which the effective number of regions can decrease automatically.

We will study in detail an objective functional containing the entropic prior, a data term which measures conformity of data within each region to a piecewise constant model, and the length term for smooth region boundaries. An interpretation of the curve evolution equations minimizing this functional will show the link between region merging and multiple curve evolution: the entropic prior can cause some curves to vanish and others to expand, thereby effectively leading to region merging by curve evolution, although not in the traditional sense of one-step merging of two regions. Instead, regions are merged progressively by curve evolution.

Variational region merging raises two important issues which we will discuss. One issue concerns determining the coefficient of the region merging prior. In general, this coefficient affects the number of regions produced at convergence and it can be quite problematic to fix its value. For the entropic prior, we will be able to determine good values of the coefficient via a statistical interpretation. The other issue concerns the influence of the initial number of regions on the segmentation.

Let us first recall the main steps of active curve segmentation into a fixed number of regions and give a simple example to illustrate the usefulness of a region merging prior. To segment an image $I : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ into N regions is to determine a partition $\mathscr{P} = \{\mathbf{R}_i\}_{i \in [1,\dots,N]}$ of the image domain so that the image is homogeneous with respect to some prescribed characteristics in each region. Active curve segmentation follows the minimization of a functional generally containing a data term, to measure the conformity of the image data within each region to a given model, and a regularization term for smooth region boundaries. Consider a partition $\mathscr{P} = \{\mathbf{R}_i\}_{i \in [1,\dots,N]}$ of Ω defined by a family of simple closed plane curves $\{\gamma_i\}_{i \in [1,\dots,N-1]}$. Using the piecewise constant image model and the boundary length regularization, the problem is to find the set of curves which minimizes the following functional:

$$\mathscr{F}(\{\gamma_i\}_{i \in [1,\dots,N-1]}) = \sum_{i=1}^{N} \int_{\mathbf{R}_i} (I - \mu_{\mathbf{R}_i})^2 d\mathbf{x} + \lambda \sum_{i=1}^{N-1} \int_{\gamma_i} ds \tag{6.1}$$

where $\mu_{\mathbf{R}_i}$ is the region $\mathbf{R}_i$ parameter and λ is a positive real constant. Optimization of the functional with respect to the region parameters determines these as the intensity mean values inside the regions:

$$\mu_{\mathbf{R}_i} = \frac{\int_{\mathbf{R}_i} I d\mathbf{x}}{\int_{\mathbf{R}_i} d\mathbf{x}} \tag{6.2}$$

Optimization with respect to $\{\gamma_i\}_{i \in [1,\dots,N-1]}$ yields the curve evolution equations (Chapter 4).

The simple example in Fig. 6.1 illustrates the usefulness of a region merging prior. Consider an image composed of two disjoint regions $\mathbf{R}_1$ and $\mathbf{R}_2$ against a background $\mathbf{R}_3 = (\mathbf{R}_1 \cup \mathbf{R}_2)^c$. Let the intensities in $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$ be constant and equal, respectively, to $\mu_{\mathbf{R}_1}$, $\mu_{\mathbf{R}_2}$, and $\mu_{\mathbf{R}_3}$. In the special case where $\mu_{\mathbf{R}_1} = \mu_{\mathbf{R}_2}$, there is no incentive in model (6.1) to merge $\mathbf{R}_1$ and $\mathbf{R}_2$ because

$$\sum_{i=1}^{3} \int_{\mathbf{R}_i} (I - \mu_{\mathbf{R}_i})^2 d\mathbf{x} = 0, \tag{6.3}$$

which corresponds to the minimum of the data term. More generally, one can show that, for two non intersecting regions $\mathbf{R}_1$ and $\mathbf{R}_2$, we have [7]:

$$\begin{cases} \int_{\mathbf{R}_1} (I - \mu_{\mathbf{R}_1})^2 d\mathbf{x} + \int_{\mathbf{R}_2} (I - \mu_{\mathbf{R}_2})^2 d\mathbf{x} \leq \int_{\mathbf{R}_1 \cup \mathbf{R}_2} (I - \mu_{\mathbf{R}_1,\mathbf{R}_2})^2 d\mathbf{x} \\ \lambda(\partial\mathbf{R}_1 + \partial\mathbf{R}_2) = \lambda\partial(\mathbf{R}_1 \cup \mathbf{R}_2) \end{cases} \tag{6.4}$$

where $\mu_{\mathbf{R}_1}$, $\mu_{\mathbf{R}_2}$, and $\mu_{\mathbf{R}_1,\mathbf{R}_2}$ are the mean of $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_1 \cup \mathbf{R}_2$, respectively. Minimization of (6.1) would not favor merging $\mathbf{R}_1$ and $\mathbf{R}_2$, even when $\mu_{\mathbf{R}_1} = \mu_{\mathbf{R}_2}$. Therefore, it may lead to an over-segmentation when N is larger than the actual number of regions. A region merging prior in (6.1) which can merge regions, as when $\mu_{\mathbf{R}_1} = \mu_{\mathbf{R}_2}$, would be beneficial. In the following, we will examine two such terms.

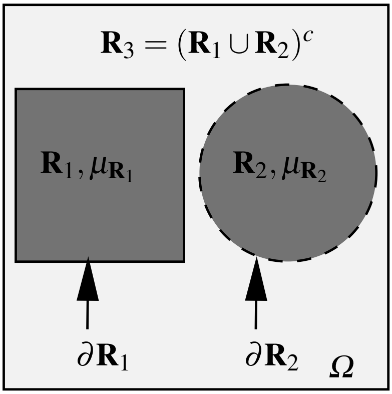


Fig. 6.1. Minimization of functional (6.1) with $N = 3$ may lead to a segmentation into 3 regions of an image containing 2 regions.

6.2 Definition of a region merging prior

A region merging prior $\mathscr{P}_{\mathscr{RM}}$ is a function from the set of partitions of Ω to $\mathbb{R}$ which satisfies the following condition:

For each partition $\mathscr{P} = \{\mathbf{R}_i\}_{i\in[1,\dots,N]}$ *of* Ω, *and for each subset* $\mathscr{J}$ *of* $[1,\dots,N]$:

$$\mathscr{P}_{\mathscr{RM}}\left(\{\cup_{j\in\mathscr{J}} \mathbf{R}_j, \{\mathbf{R}_k\}_{k\in[1,\dots,N],\ k\notin\mathscr{J}}\}\right) < \mathscr{P}_{\mathscr{RM}}\left(\{\mathbf{R}_i\}_{i\in[1,\dots,N]}\right) \tag{6.5}$$

This condition means that any region merging must decrease the prior term.

6.3 A minimum description length prior

Several studies used a prior having a minimum description length (MDL) interpretation to merge segmentation regions; this prior is expressed as a positive constant multiplied by the number of regions [2, 3, 4]:

$$\mathscr{P}_{MDL} = \nu N \tag{6.6}$$

It is independent of the curves and, as such, does not affect the curve evolution equations. To alter the number of regions, it requires an additional process external to curve evolution optimization. For instance, the region competition algorithm alternates local region merging operations and curve evolution with a fixed number of regions [2, 3]. Recall that classical region merging rests on the repeated application of local operations, such as search, and statistical testing, to decide whether to merge neighboring regions [5, 6]. The process is computationally costly [2, 3] and the outcome dependent upon

- initial conditions [3],
- *ad hoc* parameters [3, 5, 6], such as the scale of local operations and thresholds and,
- the order of the local operations [3, 6].

An alternative process to avoid such problems would use a prior which allows region merging to occur intrinsically during curve evolution [1]. In the following we study such a prior.

6.4 An entropic region merging prior

Consider the following region merging prior related to the regions logarithmic area[2]:

$$\mathscr{P}_{Entropy}(\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}) = -\beta \sum_{i=1}^{N} \mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i), \tag{6.7}$$

where $\mathbf{A}(\mathbf{R}_i)$ is the area of region $\mathbf{R}_i$,

$$\mathbf{A}(\mathbf{R}_i) = \int_{\mathbf{R}_i} d\mathbf{x}, \tag{6.8}$$

and β is a positive real constant to weigh the relative contribution of the region merging term in the segmentation functional. Prior (6.7) satisfies the region merging condition (6.5). It suffices to verify this in the case of two non empty regions $\mathbf{R}_i$ and $\mathbf{R}_j$. Merging the regions causes the following variation of $\mathscr{P}_{Entropy}$:

$$\begin{aligned}\Delta\mathscr{P}_{Entropy} = & -\beta\,(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j))\log(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j)) \\ & + \beta\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i) + \beta\mathbf{A}(\mathbf{R}_j)\log \mathbf{A}(\mathbf{R}_j)\end{aligned} \tag{6.9}$$

---

[2] For $\mathbf{A}(\mathbf{R}_i) = 0$, we pose $\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i) = 0$.

Because

$$\begin{cases} \mathbf{A}(\mathbf{R}_i)\log\mathbf{A}(\mathbf{R}_i) < \mathbf{A}(\mathbf{R}_i)\log\left(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j)\right) \\ \mathbf{A}(\mathbf{R}_j)\log\mathbf{A}(\mathbf{R}_j) < \mathbf{A}(\mathbf{R}_j)\log\left(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j)\right), \end{cases} \tag{6.10}$$

we have:

$$\mathbf{A}(\mathbf{R}_i)\log\left(\mathbf{A}(\mathbf{R}_i)\right)+\mathbf{A}(\mathbf{R}_j)\log\left(\mathbf{A}(\mathbf{R}_j)\right) < \left(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j)\right)\log\left(\mathbf{A}(\mathbf{R}_i)+\mathbf{A}(\mathbf{R}_j)\right) \tag{6.11}$$

Therefore, merging $\mathbf{R}_i$ and $\mathbf{R}_j$ decreases $\mathscr{P}_{Entropy}$.

6.4.1 Entropic interpretation

The prior (6.7) has an entropic interpretation if we define the probability of a region $\mathbf{R}_i$ by

$$p_i = \frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)}, \tag{6.12}$$

where $\mathbf{A}(\Omega)$ is the area of the image domain, then

$$\begin{aligned} -\beta\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log\mathbf{A}(\mathbf{R}_i) &= -\beta\mathbf{A}(\Omega)\sum_{i=1}^{N}\frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)}\log\frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)} - \beta\mathbf{A}(\Omega)\log\mathbf{A}(\Omega) \\ &= \beta\mathbf{A}(\Omega)H(\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}) - \beta\mathbf{A}(\Omega)\log\mathbf{A}(\Omega) \end{aligned} \tag{6.13}$$

where

$$H(\{\mathbf{R}_i\}_{i\in[1,\ldots,N]}) = -\sum_{i=1}^{N} p_i\log p_i \tag{6.14}$$

is the partition entropy. Note that $-\beta\mathbf{A}(\Omega)\log\mathbf{A}(\Omega)$ is a constant independent of the partition which can be ignored.

6.4.2 Segmentation functional

We will study the effect of the entropic region merging prior in a functional which also contains the piecewise constant model data term and the length regularization term. Let N be the maximum number of regions, i.e., a number such that the actual number of regions is inferior or equal to N. Such a number is generally available in applications. Segmentation into a number of regions less or equal to N is:

$$
\begin{aligned}
\mathscr{F}_{\mathscr{R}\mathscr{M}}(\{\gamma_i\}_{i\in[1,\ldots,N-1]}) = & \underbrace{\sum_{i=1}^{N}\int_{\mathbf{R}_i}(I-\mu_{\mathbf{R}_i})^2 d\mathbf{x}}_{\textit{Data term}} \\
& - \underbrace{\beta\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log\mathbf{A}(\mathbf{R}_i)}_{\textit{Region merging prior}} \\
& + \underbrace{\lambda\sum_{i=1}^{N-1}\int_{\gamma_i} ds}_{\textit{regularization}}
\end{aligned}
\tag{6.15}
$$

In the following, we will derive the curve evolution equations minimizing $\mathscr{F}_{\mathscr{R}\mathscr{M}}$, and show how the effective number of active curves can decrease automatically during curve evolution as a result of the entropic region merging prior.

6.4.3 Minimization equations

Minimization of $\mathscr{F}_{\mathscr{R}\mathscr{M}}$ with respect to curves $\{\gamma_i\}_{i\in[1,\ldots,N-1]}$ is carried out by embedding these into a family of one-parameter curves $\gamma_i(s,t) : [0,1]\times\mathbb{R}^+ \rightarrow \Omega$ and solving the partial differential equations:

$$
\frac{\partial\gamma_i}{\partial t} = -\frac{\partial\mathscr{F}_{\mathscr{R}\mathscr{M}}}{\partial\gamma_i}, i = 1,\ldots,N-1 \tag{6.16}
$$

Multiregion segmentation uses several active curves. When the interior of each curve defines a region, curve evolution equations can lead to overlapping regions. Chapter 4 reviewed several methods of defining regions from simple closed curves so as to guarantee a partition [8, 9, 10, 11]. Here, we use the partition constrained minimization scheme of [11, 12] which embeds a partition constraint directly in the level set evolution equations: if a point leaves a region $\mathbf{R}_i$, it goes to a single other region $\mathbf{R}_j$ without transiting through others. Starting from an arbitrary initial partition, this constraint leads to a partition. At each iteration, the scheme involves two regions for each pixel $\mathbf{x}$, the region $\mathbf{R}_k$ which contains $\mathbf{x}$ currently, and a region $\mathbf{R}_j$, $j \neq k$, which corresponds to the largest decrease in the functional were $\mathbf{x}$ transferred to this region. Recall that, for a level set implementation of curve evolution equations $\frac{\partial\gamma_i}{\partial t}$, $i = 1,\ldots,N$, each curve γ_i is represented implicitly by the zero level set of a function $\phi_i : \Omega\subset\mathbb{R}^2\rightarrow\mathbb{R}$, with the region inside γ_i ($\mathbf{R}_i = \mathbf{R}_{\gamma_i}$) corresponding to $\phi_i > 0$. For a functional $\mathscr{F}$ of the form:

$$
\mathscr{F}(\{\gamma_i\}_{i\in[1,\ldots,N-1]}) = \sum_{i=1}^{N}\int_{\mathbf{R}_i} e_i d\mathbf{x} + \lambda\sum_{i=1}^{N-1}\int_{\gamma_i} ds, \tag{6.17}
$$

where $e_i, i = 1,\ldots,N$ are scalar functions, the multiregion partition constrained level set equations minimizing $\mathscr{F}$ are given by (Chapter 4):

$$\begin{cases} \text{if } k \neq N, \frac{\partial \phi_k}{\partial t} = -\left(e_k - e_j + \lambda \kappa_k\right) \|\nabla \phi_k\| \\ \text{if } j \neq N, \frac{\partial \phi_j}{\partial t} = -\left(e_j - e_k + \lambda \kappa_j\right) \|\nabla \phi_j\| \end{cases} \tag{6.18}$$

where κ_i is the curvature of the zero level set of ϕ_i, k the index of the region containing $\mathbf{x}$, and j given by:

$$j = \arg \min_{\{i \in [1,\ldots,N],\ \mathbf{x} \notin \mathbf{R}_i\}} e_i(\mathbf{x}) \tag{6.19}$$

To use this result here, we rewrite the functional in Equation 6.15 as follows (we replace $\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i)$ by $\int_{\mathbf{R}_i} \log \mathbf{A}(\mathbf{R}_i)$):

$$\begin{aligned} \mathscr{F}_{\mathscr{R}\mathscr{M}}(\{\gamma_i\}_{i\in[1,\ldots,N-1]}) &= \sum_{i=1}^{N} \int_{\mathbf{R}_i} (I - \mu_{\mathbf{R}_i})^2 d\mathbf{x} - \beta \sum_{i=1}^{N} \int_{\mathbf{R}_i} \log \mathbf{A}(\mathbf{R}_i) d\mathbf{x} \\ &+ \lambda \sum_{i=1}^{N-1} \int_{\gamma_i} ds \\ &= \sum_{i=1}^{N} \int_{\mathbf{R}_i} \underbrace{\left((I - \mu_{\mathbf{R}_i})^2 - \beta \log \mathbf{A}(\mathbf{R}_i)\right) d\mathbf{x}}_{e_i} \\ &+ \lambda \sum_{i=1}^{N-1} \int_{\gamma_i} ds \end{aligned} \tag{6.20}$$

Thus, applying (6.18) and (6.19) to $e_i = (I - \mu_{\mathbf{R}_i})^2 - \beta \log \mathbf{A}(\mathbf{R}_i)$ yields the partition constrained level set evolution equations minimizing $\mathscr{F}_{\mathscr{R}\mathscr{M}}$:

$$\begin{cases} \text{if } \ k \neq N, \frac{\partial \phi_k}{\partial t} = -\{\underbrace{\left(I(\mathbf{x}) - \mu_{\mathbf{R}_k}\right)^2 - \left(I(\mathbf{x}) - \mu_{\mathbf{R}_j}\right)^2}_{\textit{Region data competition}} \\ \qquad - \underbrace{\beta\left(\log \mathbf{A}(\mathbf{R}_k) - \log \mathbf{A}(\mathbf{R}_j)\right)}_{\textit{Region merging}} \\ \qquad + \lambda \kappa_k\} \|\nabla \phi_k\| \\ \text{if } \ j \neq N, \frac{\partial \phi_j}{\partial t} = -\{\left(I(\mathbf{x}) - \mu_{\mathbf{R}_j}\right)^2 - \left(I(\mathbf{x}) - \mu_{\mathbf{R}_k}\right)^2 \\ \qquad - \beta\left(\log \mathbf{A}(\mathbf{R}_j) - \log \mathbf{A}(\mathbf{R}_k)\right) \\ \qquad + \lambda \kappa_j\} \|\nabla \phi_j\| \end{cases} \tag{6.21}$$

where k is the index of the region containing $\mathbf{x}$, and j is given by:

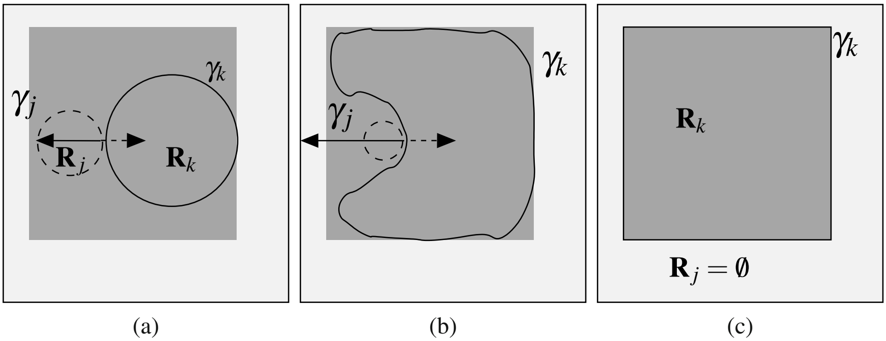
$$j = \arg \min_{\{i \in [1,\ldots,N],\ \mathbf{x} \notin \mathbf{R}_i\}} (I(\mathbf{x}) - \mu_i)^2 - \beta \log \mathbf{A}(\mathbf{R}_i) \tag{6.22}$$

6.4.4 A region merging interpretation of the level set evolution equations

Examination of the level set evolution equations in (6.21) reveals how region merging can occur intrinsically during curve evolution. When two disjoint regions $\mathbf{R}_k = \mathbf{R}_{\gamma_k}$ (the interior of γ_k) and $\mathbf{R}_j = \mathbf{R}_{\gamma_j}$ (the interior of γ_j) have close intensities (Figure 6.2 a), the velocity resulting from the data term is weak:

$$\left(I - \mu_{\mathbf{R}_k}\right)^2 - \left(I - \mu_{\mathbf{R}_j}\right)^2 \approx 0 \tag{6.23}$$

Omitting the curvature term, the evolution of curves γ_k and γ_j is mainly due to the region merging prior velocity. As ϕ_k increases and ϕ_j decreases under the effect of the term $(\log \mathbf{A}(\mathbf{R}_k) - \log \mathbf{A}(\mathbf{R}_j))$, this velocity expands the region having the larger area (Figure 6.2 b), and shrinks the other region until only one curve encloses both regions and the other curve disappears (Figure 6.2 c).



(a) (b) (c)

Fig. 6.2. A region merging interpretation of the curve evolution equations. (a)-(b) the evolution of the two homogeneous active regions under the effect of the entropic region merging prior velocity: when $\mathbf{A}(\mathbf{R}_k) > \mathbf{A}(\mathbf{R}_j)$, $\mathbf{R}_k$ expands and $\mathbf{R}_j$ shrinks; (c) (γ_k) encloses both regions while γ_j vanishes.

6.4.5 The weight of the entropic prior

On the one hand, the data term increases when regions are merged and, on the other hand, the entropic region merging term, $-\sum_{i=1}^{N} \mathbf{A}(\mathbf{R}_i) \log \mathbf{A}(\mathbf{R}_i)$, decreases when regions are merged. The role of the weighting parameter β is to balance the contribution of the region merging term against the other terms so as to yield, ideally, the actual number of regions. The weighting parameter β can be interpreted as a unit conversion factor between the data term unit and the region merging term unit. Therefore, and considering the form of these terms, we can take, for $\mathbf{A}(\Omega) > 1$:[3]

---

[3] Formally, we have: $\mathbf{A}(\Omega) = \int_\Omega d\mathbf{x}$. In practice, $\mathbf{A}(\Omega)$ is an integer corresponding to the number of pixels in the image domain. Consequently, we ignore the cases $\mathbf{A}(\Omega) = 0$ and $\mathbf{A}(\Omega) = 1$.

$$\beta = \alpha \frac{\int_{\Omega}(I-\mu)^2 d\mathbf{x}}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)} \tag{6.24}$$

where μ is the mean intensity over the whole image, $\mathbf{A}(\Omega)$ is the image domain area, and α is a constant without unit. Note that $\frac{\int_{\Omega}(I-\mu)^2}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)}$ is independent of the segmentation and is the ratio of the data term to the region merging term in the case of a partition with a single region (corresponding to the whole image domain). Using (6.24), the sum of the data term and the region merging prior is:

$$\sum_{i=1}^{N}\int_{\mathbf{R}_i}(I-\mu_{\mathbf{R}_i})^2 d\mathbf{x} - \alpha \underbrace{\frac{\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)}}_{close\ to\ 1} \int_{\Omega}(I-\mu)^2 d\mathbf{x} \tag{6.25}$$

Note that the denominator in (6.24) can be viewed as a normalization factor for the prior, and the integral in the numerator as a normalization factor for the data term because minimizing (6.25) is equivalent to minimizing:

$$\underbrace{\frac{\sum_{i=1}^{N}\int_{\mathbf{R}_i}(I-\mu_{\mathbf{R}_i})^2 d\mathbf{x}}{\int_{\Omega}(I-\mu)^2 d\mathbf{x}}}_{\in[0,1]} - \alpha \underbrace{\frac{\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)}}_{\in[0,1]} \tag{6.26}$$

Therefore, a value of α close to 1 would be reasonable. The following analysis and experimentation show that this is the case. We will need the following inequalities which hold for a partition $\mathscr{P} = \{\mathbf{R}_i\}_{i\in[1,\ldots,N]}$ of Ω:

$$1 - \frac{\log N}{\log \mathbf{A}(\Omega)} \leq \frac{\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)} \leq 1 \quad \forall \mathbf{A}(\Omega) > 1 \tag{6.27}$$

Proof: The right hand side inequality in (6.27), $\frac{\sum_{i=1}^{N}\mathbf{A}(\mathbf{R}_i)\log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)\log \mathbf{A}(\Omega)} \leq 1$, is a straightforward application of condition (6.5). For the left hand inequality, we apply

$$\log(z) \leq z-1, \forall z \in [0,+\infty[$$

to $\frac{1}{Np_i}$, where $p_i = \frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)}, \forall i \in [1,\ldots,N]$:

$$\log(\frac{1}{Np_i}) \leq \frac{1}{Np_i} - 1, \forall i \in [1,\ldots,N] \tag{6.28}$$

Multiplying both sides of (6.28) by p_i gives:

$$p_i \log(\frac{1}{Np_i}) \leq \frac{1}{N} - p_i, \forall i \in [1,\dots,N] \Rightarrow \sum_{i=1}^{N} p_i \log(\frac{1}{Np_i}) \leq \underbrace{1 - \sum_{i=1}^{N} p_i}_{0}$$

$$\Rightarrow -\sum_{i=1}^{N} \frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)} \log \frac{\mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega)} \leq \log(N)$$

$$\Rightarrow \frac{\sum_{i=1}^{N} \mathbf{A}(\mathbf{R}_i) \log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega) \log \mathbf{A}(\Omega)} \geq 1 - \frac{\log N}{\log \mathbf{A}(\Omega)},$$

which completes the proof of (6.27).

In practice, N is generally much smaller than $\mathbf{A}(\Omega)$ and $\frac{\sum_{i=1}^{N} \mathbf{A}(\mathbf{R}_i) \log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega) \log \mathbf{A}(\Omega)}$ is close to 1. For example, for a maximum number of regions equal to 10 and a 256**x**256 image, we have approximately $0.8 \leq \frac{\sum_{i=1}^{N} \mathbf{A}(\mathbf{R}_i) \log \mathbf{A}(\mathbf{R}_i)}{\mathbf{A}(\Omega) \log \mathbf{A}(\Omega)} \leq 1$. The small interval of variation of the normalized prior in (6.27) suggests that α will vary in a small interval centered close to 1. Note that with a value of α close to 1 in (6.25), the sum of the data and the region merging terms will be close to the in-between cluster distance according to a classical data clustering relation in statistical pattern recognition [7]:

$$\underbrace{\sum_{i=1}^{N} \int_{\mathbf{R}_i} (I - \mu_{\mathbf{R}_i})^2 d\mathbf{x}}_{within-cluster\ distance} - \underbrace{\int_{\Omega} (I - \mu)^2 d\mathbf{x}}_{total\ distance} = \underbrace{-\sum_{i=1}^{N} (\mu_{\mathbf{R}_i} - \mu)^2}_{in-between\ cluster\ distance} \tag{6.29}$$

It is also possible to learn an interval of values of coefficient λ applicable to the images of a given class. For a set of training segmentation images, this is done by determining the values of λ which produce the desired number of regions in a relevant set of training images and segmentation examples.

6.5 Example

6.5.1 Segmentation with the entropic region merging prior

Figure 6.3 shows segmentations of a brain image for different initial (maximum) number N of regions (from 3 to 6). The first row of Figure 6.3 depicts the curve initializations corresponding to the different values of N. To illustrate the influence of the entropic region merging prior, we show the results obtained with the entropic region merging prior ($\alpha = 1$) in the second and third rows, and the results obtained without a region merging prior ($\alpha = 0$) in the fourth and fifth rows. With the entropic region merging prior, some curves vanish at convergence (second row), leading to the same segmentation into 3 regions (third row) for N (the initial number of regions) going from 3 to 6 (first row). Without the region merging prior, none of the curves disappears and the functional minimization splits up the image into exactly N regions, leading to four different segmentations (fifth row). Figure 6.4 contains a plot of the data term at convergence versus N, for $\alpha = 0$ and $\alpha = 1$ (without and with the entropic region merging prior). For $\alpha = 0$ (dashed curve), the minimized data term decreases when N increases. The region merging prior, which increases with N, balances the effect of the data term, leading approximately to the same value at convergence for the different values of N (solid curve). This illustrates the robustness of the entropic region merging with respect to the initial (maximum) number of regions.

Table 6.1 shows the effect of the weight of the entropic region merging prior (parameter α) on the number of segmentation regions at convergence, using 7 initial regions ($N = 7$). It lists the number of regions obtained at convergence, and the interval of α values leading to this number. All the values of α in the interval $[0.58, 2.29]$ give the same segmentation into the desired number of regions, i.e., three regions. This is consistent with the statistical interpretation we discussed in Section 6.4.5 of the weight of the entropic region merging prior, and which prescribes a value of α of about 1.

α	$2.31 \leq \alpha \leq 7.10$	$0.58 \leq \alpha \leq 2.29$	$0.36 \leq \alpha \leq 0.56$
regions	2	3	4

Table 6.1. Brain image: effect of the weight of the entropic region merging prior (parameter α) on the number of regions

6.5.2 Segmentation with the MDL region merging prior

By contrast to using the entropic region merging prior, segmentation with the MDL prior is sensitive to the initial number of regions. It cannot lead to the same segmentation of the brain image into three regions for the five different initializations (the initial number of regions varies from 3 to 7). Table 6.2 lists, for each initialization, the intervals of ν values necessary to obtain three regions at convergence with the MDL prior. To end region merging with a 3-region initialization, i.e., to obtain segmentation into 3 regions, the weight of the MDL prior ν should be in $[0, 1.55(10^6)]$. However, all ν values in this interval do not allow region merging with the 6-, 5- and 4-region initializations (refer to Table 6.2), thereby resulting in four different segmentations. With the entropic region merging prior, all α values in $[0.58, 2.29]$ give the same segmentation into 3 regions (refer to table 6.1). Fig. 6.5 plots the number of regions obtained at convergence versus the initial number of regions for both the entropic and the MDL prior. This demonstrates that the entropic prior is less sensitive than the MDL prior to variations of the initial number of regions.

6.5.3 Computation time

Table 6.3 gives the computation time on a 2 GHZ machine for the brain image segmentation with different initializations, for the entropic and the MDL priors. The

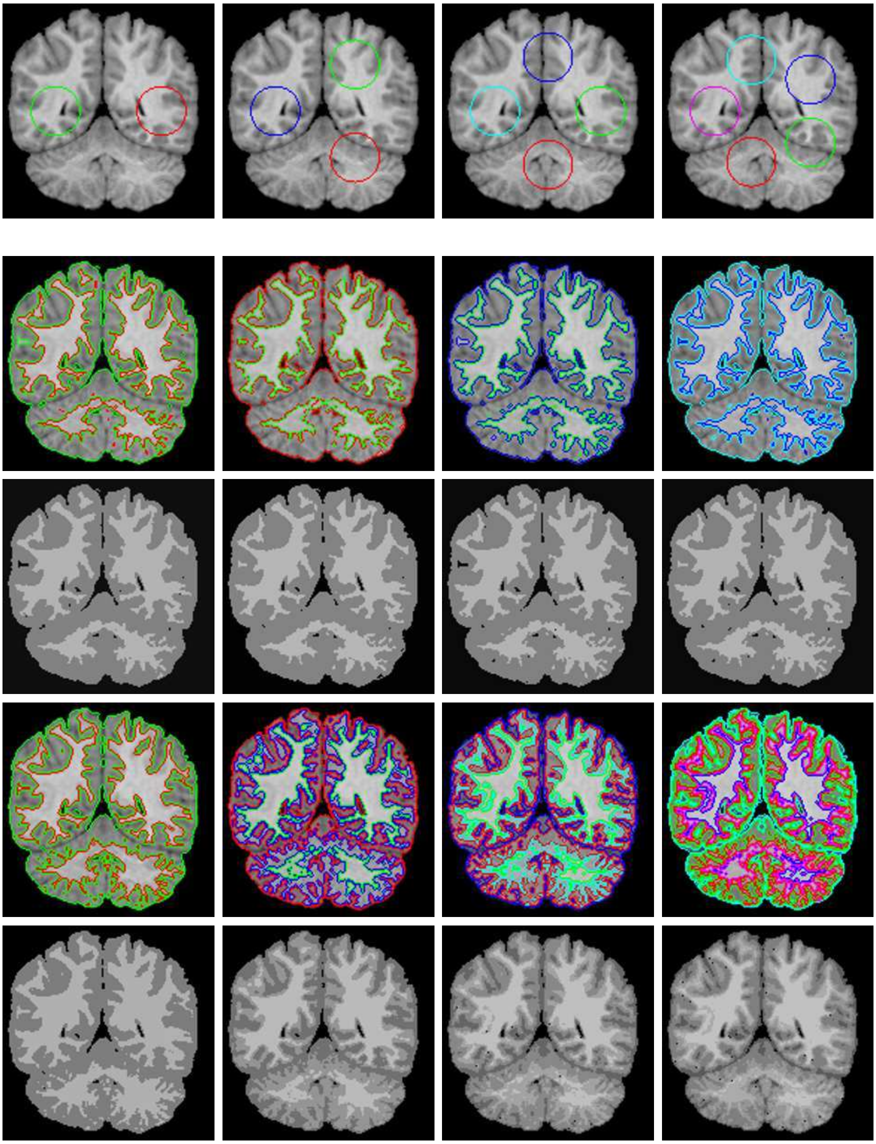


Fig. 6.3. Segmentation of a brain image. Row 1: different number of regions at the initialization; row 2: final curves which remained at convergence *with* the entropic region merging prior ($\alpha = 1$); row 3: final segmentations *with* the entropic region merging prior ($\alpha = 1$); row 4: final curves (N curves) *without* a region merging prior ($\alpha = 0$); row 5: final segmentations *without* a region merging prior ($\alpha = 0$). The displayed results are obtained with $\lambda = 2$.

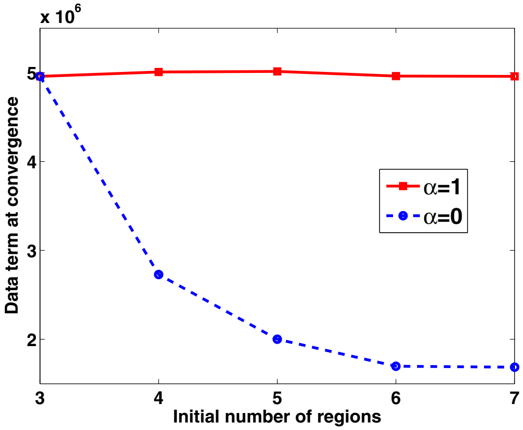


Fig. 6.4. Effect of the entropic region merging prior: graph of the minimized data term versus the initial (maximum) number of regions (N).

Initial number of regions	7	6	5	4	3
$\nu(10^{-16})$	$[0.93, +\infty[$	$[(1.87, +\infty[$	$[2.73, +\infty[$	$[(2.43, +\infty[$	$[0, 1.55[$

Table 6.2. Brain image: effect of the weight of the MDL prior (parameter ν) on the number of regions at convergence. 3-region initialization: ν values stopping the region merging. 4-region to 7-region initializations: ν values allowing the region merging.

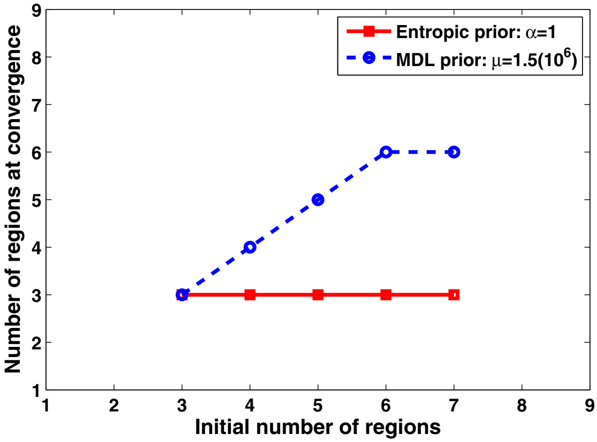


Fig. 6.5. Robustness with respect to variation of the initial number of region: the number of regions obtained at convergence versus the initial number of regions using the MDL and entropic priors.

initial number of regions varies from 3 to 7. Segmentation with the entropic prior needs relatively less computation time, particularly when the initial number of regions increases. This was predictable because segmentation follows one stage of curve evolution whereas segmentation with the MDL prior alternates iteratively two stages: curve evolution and local region-merging operations.

Initial number of regions	3	4	5	6	7
Entropic prior	0.22s	0.25s	0.3s	0.35s	0.4s
MDL prior	0.65s	0.78s	1.04s	1.41s	3.21s

Table 6.3. Computation time for the the brain image segmentation when the initial number of regions varies from 3 to 7.

References

1. I. Ben Ayed and A. Mitiche, “A region merging prior for variational level set image segmentation,” *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2301–2313, 2008.
2. S. Zhu and A. Yuille, “Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 118, no. 9, pp. 884–900, 1996.
3. T. Kadir and M. Brady, “Unsupervised non-parametric region segmentation using level sets,” in *International Conference on Computer Vision (ICCV)*, 2003, pp. 1267–1274.
4. T. Brox and J. Weickert, “Level set segmentation with multiple regions,” *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213–3218, 2006.
5. F. Nielsen and R. Nock, “On region merging: The statistical soundness of fast sorting, with applications,” in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2003, pp. 19–26.
6. R. Nock and F. Nielsen, “Statistical region merging,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452–1458, 2004.
7. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley, 2000.
8. L. Vese and T. Chan, “A multiphase level set framework for image segmentation using the Mumford and Shah model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
9. C. Vázquez, A. Mitiche, and I. Ben Ayed, “Image segmentation as regularized clustering: A fully global curve evolution method,” in *International Conference on Image Processing (ICIP)*, 2004, pp. 3467–3470.
10. A. Mansouri, A. Mitiche, and C. Vázquez, “Multiregion competition: A level set extension of region competition to multiple region partioning,” *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
11. I. Ben Ayed and A. Mitiche, “A partition constrained minimization scheme for efficient multiphase level set image segmentation,” in *International Conference on Image Processing (ICIP)*, 2006, pp. 1641–1644.
12. I. Ben Ayed, A. Mitiche, and Z. Belhadj, “Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.

7

MOTION BASED IMAGE SEGMENTATION

7.1 Introduction

In his book on human visual perception *The perception of the visual world* [1], 1950, J. J. Gibson studied the *flow of projected surfaces* which he called *optical flow*. He characterized optical flow by speed and direction at each point of the retinal image domain. Nakayama and Loomis [2] referred to optical flow later as the *instantaneous velocity field* over the image positional domain and provided an analytic working definition of *optical velocity*, the instantaneous velocity vector. In this book, we will retain the essence of Gibson's and Nakayama's definition and characterization which view optical flow as the field of optical velocities over the image domain. Taking *image* in its most common mathematical sense to mean the result of an application, a projection on the retinal domain in this case, we will also refer to optical flow as *image motion* or, when there is no ambiguity, simply *motion* as in the title of this chapter.

Image motion occurs whenever there is a relative motion of the viewing system and the viewed environmental surfaces. The processing of this ever-present motion by the human visual system plays several functions which result in a rich, effortless perceptual ability [2]. In machine vision, motion analysis has several essential roles as well [3, 4]. The fundamental roles are (motion) *detection*, to separate the image domain into two regions, one corresponding to the moving scene objects and the other to the background; (motion) *segmentation* to partition the image domain into regions of differing motion properties; *three-dimensional interpretation* to explain the image in terms of the structure and the motion of the moving real objects in the viewed environment; and *tracking* to locate and follow objects in their movement throughout an image sequence.

In this chapter, we will be concerned with motion based image segmentation, i.e., image domain partitioning using optical flow. Motion-based segmentation is of broad interest in vision because it serves various useful applications such as video coding [5], video indexing and retrieval [6, 7] video event understanding [4, 8, 9], tracking [10, 11, 12, 13], and three-dimensional interpretation [3, 14, 15, 16].

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_7,

Optical flow is not a sensed variable; it must be estimated. One can estimate it, using a variational method [17], for instance, and then segment it, using, for instance, the Chan-Vese level set method [18] described in a preceding chapter. The issue, in this case, is to compute a boundary preserving estimate of optical flow via proper regularization [17, 19]. The Appendix describes two prevalent image motion estimation methods.

An advantageous alternative to this two-step processing is *joint estimation and segmentation* because estimation and segmentation are in reciprocal dependence. We will describe two such methods, both of which can be seen as applications to optical flow of the Mumford-Shah image segmentation functional [20] which, we recall, is

$$\mathscr{E}(g,K) = \int_{\Omega} (g - g_0)^2 d\mathbf{x} + \rho \int_{\Omega \setminus K} \|\nabla g\|^2 d\mathbf{x} + \lambda \rho l(K) \tag{7.1}$$

We recall that $\mathscr{E}$ is the 'energy', or 'merit', of the partition determined by the image approximation g with boundaries K, g_0 is the sensed input image, ρ (scale) and λ (weight) are positive real constants, and l is the length of K.

The transposition of (7.1) to optical flow will bring in a term of conformity of image motion to data, namely, the image spatial and temporal variations; a term to embed the desired properties of the flow, smoothness, for instance; and a term related to the length of the flow boundaries to produce regular boundaries according to a definition, albeit implicit, of a flow edge.

The two methods we are about to describe differ in their representation of the flow boundaries. One [21] uses the discrete Leclerc minimum description length (MDL) formulation [22]. It references the segmentation individual edge points and, as such, can be viewed as an implicit edge detection method. The other is a level set method, using closed curves to delineate (optical flow) segmentation regions, each described by the parameters of a general linear model of motion [23].

Before detailing these methods, we review briefly others which addressed joint optical flow estimation and segmentation via functional optimization [24, 25, 26, 27].

A set of label variables over the image domain describes the segmentation in [24]. The problem is cast in a Bayesian parametric motion estimation formulation which, under assumptions, is transformed into an energy function containing the classic quadratic terms of smoothness of flow and conformity of flow to data, in addition to two other terms related to the segmentation, one referencing the flow and the other the labels of the segmentation, both terms derived under a Markov/Gibbs field model. Three parameters control the relative contribution of the four terms. Optimization is carried out deterministically to ease the computational burden.

The objective functional in [27] contains five terms. Two terms are related to parametric estimation of the flow field under an affine approximation model. One is a measure of conformity of the flow to data and the other is a smoothness prior. Both terms use a statistically robust function rather than the quadratic function as in [28], and are given a semi-quadratic form via the introduction of two sets of auxiliary variables over the image domain, one set for each term. The other terms are related to the segmentation. One term is to favor low values along the segmentation boundaries of the smoothness related auxiliary variable. Another term is to enforce conformity of the motion field to the parametric model within the segmentation regions. The last term is a region prior proportional to the length of the segmentation boundaries. The relative contribution of the five terms in the objective functional is set by four parameters. Optimization is carried out hierarchically for computational efficiency. Because segmentation boundaries are implicitly defined, an external process locates them after completion of the optimization algorithm.

Joint optical flow estimation and segmentation of a piecewise constant motion field is cast into an active curve evolution formulation in [25]. The objective functional contains a data term according to a normalized version of the Horn-and-Schunck equation and a regularization term related to the length of the segmentation boundaries. The Euler-Lagrange equations result in active curve evolution, implemented via level sets. Joint parametric estimation and segmentation, extended to space-time, was the focus of the study in [26], which represented each component of motion by a linear combination of a set of parameters. However, the coefficients of the combinations were not related to any specific model of motion which would define them as functions of image coordinates. The affine model was given as an example and the constant model was implemented.

Other studies have regarded segmentation and estimation of motion as separate processes and do not, therefore, account for the mutual dependence of the two processes [29, 30, 31, 32, 33, 34].

7.2 Piecewise constant MDL estimation and segmentation of optical flow

Leclerc [22] introduced the minimum description length (MDL) formalism in image segmentation. Let us recall briefly from Chapter 3 the main points of the formulation, which can be regarded as a discrete implementation of the Mumford-Shah functional [35]. The method minimizes a discrete objective functional which assigns to an image partition the code length of its specification according to a predefined description. The energy functional for the piecewise constant model (piecewise constant g) is:

$$\mathscr{E}(g) = a \sum_{i \in D} \left(\frac{g_i - g_{0i}}{\sigma} \right)^2 + \frac{\beta}{2} \sum_{i \in D} \sum_{j \in \mathscr{N}_i} (1 - \delta(g_i - g_j)) \tag{7.2}$$

where g_0 and g are, respectively, the original discrete image and its segmented approximation; indices indicate position of evaluation; $\mathscr{N}_i$ is some fixed neighborhood of i; and

$$\delta(z) = \begin{cases} 1 & \text{for } z = 0 \\ 0 & \text{for } z \neq 0 \end{cases} \tag{7.3}$$

Using entropy coding and under the assumption that g and g_0 differ by white Gaussian noise with variance σ^2, the first term of the objective functional approximates the number of bits to code the difference between g and g_0. The second term approximates the number of bits to code the boundary elements of g and the image values in the corresponding piecewise constant partition. Under the piecewise constant model, edges are points at the border between two regions of different image value. A partition region can be as small as a single pixel or as large as the whole image domain.

The energy functional (7.2) of Leclerc is not differentiable because of the presence of the δ function. To determine the necessary conditions of minimization by differentiation, one substitutes a Gaussian kernel for this δ function and uses *continuation* indexed by the width of the kernel: Starting from an initial large value, the width is gradually lowered and, at each step, a solution to the corresponding problem is computed using as initial approximation the solution to the previous problem.

We will now transpose Leclerc's formulation to joint estimation and segmentation of optical flow.

Digital images are noisy due to various distortions such as sensor noise, sampling, and digitization. Assume that noise causes the Horn and Schunck gradient equation (see Chapter 2) to be verified only up to a Gaussian white noise, i.e.,

$$I_{xi}u_i + I_{yi}v_i + I_{ti} = \mu_i \tag{7.4}$$

where i is the index in the image positional array of the point of evaluation; I_x, I_y and I_t are quantized approximations of the spatial and temporal derivatives of the image sequence; u, v are the (quantized) coordinate functions of optical flow, and μ is quantized white Gaussian noise.

Let conformity of a motion field to image data be described by the field likelihood with respect to the image spatial and temporal derivatives, i.e., with respect to the *residuals* $\{I_{xi}u_i + I_{yi}v_i + I_{ti}\}_{i \in D}$. Using entropy coding, the number of bits to code this description of a motion field W over the image positional array D is approximated by:

$$\begin{aligned} -\log_2 P(\{I_{xi}u_i + I_{yi}v_i + I_{ti}\}_{i\in D} | \{(u_i, v_i)\}_{i\in D}) &= -\log_2 \prod_{i\in D} P(\mu_i = I_{xi}u_i + I_{yi}v_i + I_{ti}) \\ &= c + a \sum_{i\in D} \left(\frac{I_{xi}u_i + I_{yi}v_i + I_{ti}}{\sigma} \right)^2 \end{aligned} \tag{7.5}$$

where σ is the standard deviation of μ_i $\forall i \in D$; a and c are constants. Let $\{\mathbf{R}_k\}$ be a partition of the image domain and let the motion field have a parametric form within each region of this partition.

Then the motion field can be described by its boundaries $\{\partial \mathbf{R}_k\}$ and its parameters $\{\Theta_k\}$. In the discrete case, the number of bits to code this description can be approximated by

$$\frac{b}{2} \sum_k l(\gamma_k) + \sum_k \theta_k \tag{7.6}$$

where γ_k is a discrete approximation of $\partial \mathbf{R}_k$, $l(\gamma_k)$ is its length measured by the number of pixels it threads through; b is the average number of bits to code the direction of a line segment between two adjacent points of γ_k, and θ_k is the number of bits to code the set of parameters Θ_k and the coordinates of a starting point on γ_k. The division of b by 2 is due to the fact that each boundary segment appears twice in the summation. For a piecewise constant parametric model of image motion, we have

$$\sum_k l(\gamma_k) = \sum_{i \in D} \sum_{j \in \mathcal{N}_i} (1 - \delta(u_i - u_j)\delta(v_i - v_j)) \tag{7.7}$$

where $\mathcal{N}_i$ is the set of the 4-neighbors of i, and δ is defined in (7.3). If the number of bits to code the region parameters is the same for all regions, i.e., $\theta_k = \theta \quad \forall k$, the second sum in (7.6) can be approximated by

$$\sum_k \theta_k \approx r\theta \tag{7.8}$$

where r is the average number of regions determined by the total number of boundary pixels divided by the average number of pixels per boundary. Consequently, the total number of bits bit to describe a piecewise constant motion field over a partition of the image positional array can be approximated by

$$c + a \sum_{i \in D} \left(\frac{I_{xi}u_i + I_{yi}v_i + I_{ti}}{\sigma} \right)^2 + \frac{\beta}{2} \sum_{i \in D} \sum_{j \in \mathcal{N}_i} (1 - \delta(u_i - u_j)\delta(v_i - v_j)) \tag{7.9}$$

with $\beta = b + \theta/d$, where d is the average number of pixels per boundary. In practice, β is approximately equal to 2; its exact value is not essential.

The minimum description length estimate of optical flow is the motion field $\tilde{\mathbf{W}}$ over D which corresponds to a minimum in the total bit cost of description:

$$\tilde{W} = \arg \min_{W=(u,v)} \left\{ a \sum_{i \in D} \left(\frac{I_{xi}u_i + I_{yi}v_i + I_{ti}}{\sigma} \right)^2 + \frac{\beta}{2} \sum_{i \in D} \sum_{j \in \mathcal{N}_i} (1 - \delta(u_i - u_j)\delta(v_i - v_j)) \right\} \tag{7.10}$$

7.2.1 Numerical implementation

As in the original Leclerc formulation for intensity images, the objective function (7.10) is not differentiable because of the presence of the delta function. Of course, a stochastic method such as simulated annealing can carry out its minimization. However, such methods are notoriously demanding computationally, even when the procedure parameter updating is properly scheduled. As with the Leclerc original formulation, an advantageous alternative is to embed the minimization (7.10) in a family of minimizations indexed by the parameter of a differentiable approximation of the δ function, and use *continuation* [22, 36].

We will base continuation on the following substitution:

$$\delta(u_i - u_j)\delta(v_i - v_j) \quad \longleftarrow \quad e_{ij}(W,s) = e^{\frac{-(u_i-u_j)^2-(v_i-v_j)^2}{(s\sigma)^2}} \tag{7.11}$$

Without loss of generality, we used $s\sigma$ in (7.11) rather that just s in order to simplify subsequent expressions. However, the parameter of continuation is s. Following the substitution, the objective functional to minimize becomes:

$$\mathscr{E}(W,s) = a\sum_{i\in D}\left(\frac{I_{xi}u_i + I_{yi}v_i + I_{ti}}{\sigma}\right)^2 + \frac{\beta}{2}\sum_{i\in D}\sum_{j\in\mathscr{N}_i}(1 - e_{ij}(W,s)) \tag{7.12}$$

Continuation solves problem (7.10) as follows: Let $s_1, s_2, \ldots$ be a decreasing sequence of s values which tends to zero. We solve the following sequence of problems:

$$\text{Minimize} \quad \mathscr{E}(W,s_l) \tag{7.13}$$

For each value s_l of s, the necessary conditions for a minimum of $\mathscr{E}$ result in two equations for each pixel $i \in D$:

$$\begin{aligned} I_{xi}^2 u_i + I_{xi}I_{yi}v_i + I_{xi}I_{ti} + \beta_l \sum_{j\in\mathscr{N}_i}(u_i - u_j)e_{ij}(W,s_l) &= 0 \\ I_{xi}I_{yi}u_i + I_{yi}^2 v_i + I_{yi}I_{ti} + \beta_l \sum_{j\in\mathscr{N}_i}(v_i - v_j)e_{ij}(W,s_l) &= 0 \end{aligned} \tag{7.14}$$

where $\beta_l = \beta\log 2/s_l^2$. This results in a large scale sparse system of equations, most of which are linear. This instructs us to use the following Jacobi-type iterative scheme:

$$\begin{aligned} u_i^{k+1} &= \frac{-I_{x_i}I_{ti} - I_{x_i}I_{y_i}v_i^k + \beta_l\sum_{j\in\mathscr{N}_i}e_{ij}^k(W,s_l)u_j^k}{I_{x_i}^2 + \beta_l\sum_{j\in\mathscr{N}_i}e_{ij}^k(W,s_l)} \\ v_i^{k+1} &= \frac{-I_{y_i}I_{t_i} - I_{x_i}I_{y_i}u_i^k + \beta_l\sum_{j\in\mathscr{N}_i}e_{ij}^k(W,s_l)v_j^k}{I_{y_i}^2 + \beta_l\sum_{j\in\mathscr{N}_i}e_{ij}^k(W,s_l)} \end{aligned} \tag{7.15}$$

Iterations (7.15) are Jacobi-like because we apply, at each iteration, a Jacobi update to a linear system of equations obtained by evaluating the exponential term using the values of motion of the preceding iteration.

The solution of each problem in (7.13) is the initial approximation for the next. As s tends to zero, we tend to the original problem (7.10) because $e_{ij}(W,s)$ approaches δ. For s tending to infinity, the second term in (7.12) tends to zero. This instructs us to set the first problem with s_1 large and use as initial approximation the motion field $W^{\perp}$ composed of the normal component of image motion determined by the Horn and Schunck optical flow constraint equation (see Chapter 2). The iterations terminate with s_l small. In the experiments we conducted, about 100 iterations of continuation and 5 of (7.15) sufficed.

7.2.2 Example

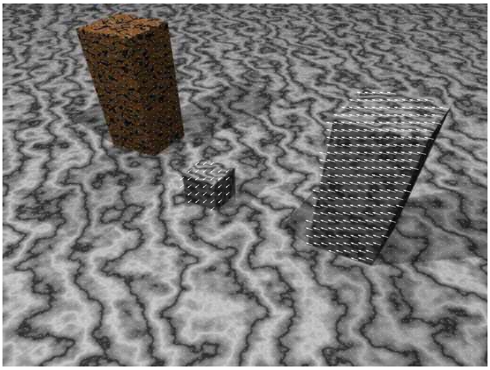
The scheme is illustrated with the *Marbled blocks* synthetic test sequence (KOGS/ IAKS laboratory database, University of Karlsruhe, Germany). There are three textured blocks in the image. The leftmost block is not moving, the other two are. The viewing system is not moving so that the moving blocks background is static. The rightmost block moves to the left, slightly away from the viewing system; the smallest block moves slightly to the left and forward in depth. The images are noised; they are also characterized by very weak texture variation across the top edges of the blocks. The variation of depth, and of the image motion thereof, is strong at the blocks boundaries not in contact with the floor. The scene and the actual motion field are shown in Fig. 7.1 a, and the MDL motion estimate in Fig. 7.1 b. The MDL estimation/segmentation depicted in Fig. 7.1b does not include the delineation of the segmentation regions because the MDL representation uses only an implicit definition of motion edges. According to this representation, there is an edge between any two pixels of different image motion; in practice, this corresponds to a severely over-segmented image. The average motion norm error (over the whole image) is 0.13 pixel and the average direction error (on the two moving blocks) is 4.7°. The standard deviation for the norm is 0.2, which is similar to those of the Horn and Shunck [28] and of the Deriche, Kornprobst, and Aubert [19] methods. The standard deviations for the direction of 5.8 (small block) and 3.5 (large block) are also similar to those of the methods [19, 28].

7.3 Joint segmentation and linear parametric estimation of optical flow

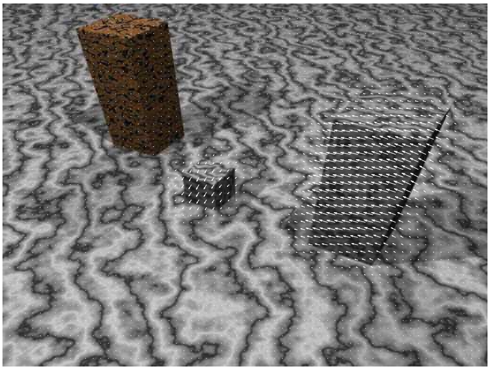
The purpose of using a parametric model of motion is to partition the image into regions of differing model parameters. Therefore, befitting models can not only achieve a good segmentation but will also describe the motion field economically and accurately.

Several studies have used parametric models for motion estimation and segmentation [24, 26, 29, 30, 31, 32, 33, 34, 37, 38, 39, 40] and, in general, the investigations were limited to the constant and affine models. Ideally, the model should only account for the actual motion variations, modulo sensing noise and approximate digital representations. When segmentation is the main purpose, rather than accurate estimation, the constant and affine models are sufficiently descriptive to segment a variety of motion fields, such as those generated by familiar man-made rigid objects. A model which describes finer grades of motion than are present in the underlying true motion field would likely lead to a sketchy segmentation because it can account for complex motion variations within a region. On the other hand, a model too simple to represent the actual variations would likely lead to a fragmented segmentation because the model would apply in regions of smaller extent, rather than larger.

In general, accurate segmentation requires more complex models when the variations of motion are complex within or in between the segmentation regions. More



(a)



(b)

Fig. 7.1. (a) Ground truth image motion superimposed on the first of the two *Marbled block* images used in the experiment; (b) the MDL motion estimate.

complex models are also needed when the accuracy of the estimated motion is important, e.g., for 3D interpretation [3]. Non parametric variational motion methods such as [41, 42], which use a Mumford-Shah type functional [20], can yield very good, boundary preserving estimates of motion but do not address explicit segmentation.

The goal in this section is to formulate joint image motion estimation and segmentation as a variational problem using a general linear model of motion and active curve evolution. The objective functional of such a formulation expands each of the two component functions of motion in a linear space of functions $V = \text{span}\{\theta_i\}_{i=1}^{M}$, and represents image partitions via the interior of regular simple closed plane curves. The functional we investigate has three integral terms. One term, the data term, measures the conformity in each segmentation region of the motion coordinate functions representation in V to the image spatiotemporal first order variations. This is expressed, as usual, via the the optical flow constraint of Horn and Schunck.

Another term biases the segmentation toward a partition with smooth boundaries, as is common in image segmentation. The third term biases the solution toward a partition with high motion contrast at the region boundaries. This last term is not essential but brings in additional information; it does not reference motion explicitly; instead, it uses a characterization of motion boundaries in terms of the image spatiotemporal variations.

The general linear parametric model used here affords a wide choice of models, such as affine, polynomial, exponential polynomial, sines/cosine, among a variety of others. The Euler-Lagrange equations lead to an algorithm which reduces to iterations of curve evolution, via level sets, and least squares estimation of the motion parameters within each segmentation region. Therefore, the representation generality does not overly increase the algorithmic complexity. The dependence of the parameters on the segmentation does not give extra terms in the curve evolution equations [43]. Therefore, it is not necessary to assume that these parameters are independent of the segmentation [26, 34, 40]. Also, it is not necessary to determine the parameters before segmentation [32, 33], or invoke a greedy algorithm in lieu of gradient descent [44] to optimize the objective functional.

7.3.1 Formulation

For greater clarity, we will treat the case of a segmentation into two regions and then generalize to the segmentation into a fixed but arbitrary number of regions.

Let $\gamma : s \in [0,1] \rightarrow \Omega \subset \mathbb{R}^2$ be a closed simple plane curve. Let $\mathbf{R}_1 = \mathbf{R}_\gamma$ be the region interior to γ, and $\mathbf{R}_2 = \mathbf{R}_\gamma^c$ its complement (the region exterior to γ).

Let $W_i(\mathbf{x}) = (u_i(\mathbf{x}), v_i(\mathbf{x}))^T, i = 1,2$, designate the image motion at $\mathbf{x} \in \mathbf{R}_i$. The component functions of image motion in each of the two regions $\mathbf{R}_i$, $i = 1,2$, are taken to be functions in $V = \text{span}\{\theta_j\}_{j=1}^M$:

$$W_i(\mathbf{x}) = \alpha_i^T \theta(\mathbf{x}), \quad i = 1,2, \tag{7.16}$$

where

$$\theta = \begin{pmatrix} \theta_1 \ \theta_2 \cdots \theta_M \end{pmatrix}^T, \tag{7.17}$$

and $\alpha_i, i = 1,2$, are $M \times 2$ matrices of parameters:

$$\alpha_i = \begin{pmatrix} \alpha_{i11} \ \alpha_{i21} \cdots \alpha_{iM1} \\ \alpha_{i12} \ \alpha_{i22} \cdots \alpha_{iM2} \end{pmatrix}^T \tag{7.18}$$

Coefficient α_{ijl} corresponds to function θ_j for the horizontal ($l = 1$) or vertical ($l = 2$) coordinate functions of motion W_i in $\mathbf{R}_i$, $i = 1,2$.

This representation affords a wide variety of models which generalize the currently prevalent constant and affine models [45]. For example, a polynomial function space of dimension one, which has six basis functions, describes affine motion, and a polynomial function space of dimension two, with has twelve basis functions, describes general quadratic motion. Here following are two useful general models [23].

Polynomial basis: The polynomial space representation contains the constant, affine, and quadratic motion models as instances. The basis function of a polynomial space of degree P is described by

$$\theta_j(\mathbf{x}) = x^{k-l} y^l, \tag{7.19}$$

where x, y are the coordinates of $\mathbf{x}$, $j \in [0, (P+1)(P+2)/2]$, and

$$\begin{aligned} k &= \left\lfloor \frac{\sqrt{8j+1}-1}{2} \right\rfloor \\ l &= j - \frac{k(k+1)}{2} \end{aligned} \tag{7.20}$$

The degree of the affine model is $P = 1$ and there are 3 basis functions:

$$\begin{aligned} \theta_0(\mathbf{x}) &= 1 \\ \theta_1(\mathbf{x}) &= x \\ \theta_2(\mathbf{x}) &= y \end{aligned} \tag{7.21}$$

With digital images, x and y are discrete and $x \in [0, W-1]$, $y \in [0, H-1]$, where W and H are, respectively, the width and height of the image in pixels.

DCT basis: The discrete cosine transform (DCT) basis, of degree P, specified over an image positional array, is defined by:

$$\theta_j(\mathbf{x}) = \rho_1 \rho_2 \cos\left(\frac{\pi(2x+1)q}{2W}\right) \cos\left(\frac{\pi(2y+1)p}{2H}\right), j \in [0, P^2 - 1] \tag{7.22}$$

where

$$\begin{aligned} p &= \lfloor j/P \rfloor \\ q &= j - pP \\ \rho_1 &= \begin{cases} \frac{1}{\sqrt{P}} & p = 0 \\ \sqrt{\frac{2}{P}} & p \in [1, P-1] \end{cases} \\ \rho_2 &= \begin{cases} \frac{1}{\sqrt{P}} & q = 0 \\ \sqrt{\frac{2}{P}} & q \in [1, P-1] \end{cases} \end{aligned} \tag{7.23}$$

This variational formulation minimizes an objective functional which is the sum of three integral terms:

$$\mathscr{E} = \mathscr{E}_d + \lambda \mathscr{E}_s + \varepsilon \mathscr{E}_b \tag{7.24}$$

where $\mathscr{E}_d$, $\mathscr{E}_s$, and $\mathscr{E}_b$, described in detail subsequently, are, respectively, the terms of image motion conformity to the image spatio-temporal data, smoothness of the motion segmentation boundaries, and motion contrast at region boundaries. The positive constants λ and ε weigh the contribution of each of the three terms.

The data term: As with the previous MDL formulation, and practically all the variational statements of image motion segmentation, the integrand in each region data term is the squared total time derivative of the image sequence function, i.e., the squared left-hand side of the Horn and Schunck optical flow constraint, or gradient equation [28]:

$$\mathscr{E}_d = \frac{1}{2} \sum_{i=1}^{2} \int_{\mathbf{R}_i} \left(\langle \nabla I, W_i \rangle + I_t\right)^2 d\mathbf{x} \tag{7.25}$$

Substitution of the optical flow component functions parametric expression (7.16) in (7.25) gives the data term, which is a function of active curve γ:

$$\mathscr{E}_d = \frac{1}{2} \sum_{i=1}^{2} \int_{\mathbf{R}_i} \left(\langle \nabla I, \alpha_i^T \theta \rangle + I_t\right)^2 d\mathbf{x} \tag{7.26}$$

Because the regions $\mathbf{R}_i$, $i = 1,2$, are defined by γ, the motion coordinate parameters in (7.26) are functions of γ.

A few studies have investigated data terms which use functions of image motion other than the Horn and Shunck. However, these were similar, or they used optical flow computed by an outside process using the Horn and Schunck constraint. For instance, the data term of [26] uses the angle between the spatio-temporal gradient and the motion vector to separate motions according to direction, rather than speed and direction. As such, it is related to the Horn and Shunck data term. The study in [34] uses the norm of the difference between the least squares solution of the optical flow gradient equation in the two regions of a two-region segmentation formulation [46].

The boundary smoothness regularization term: The common segmentation boundary length term is used to provide smooth region contours and avoid partitions fragmented into small components:

$$\mathscr{E}_s = \frac{1}{2} \int_{\gamma} ds \tag{7.27}$$

The parametric representation will provide a motion field smooth everywhere except at motion boundaries. Therefore, a regularization term with respect to motion is not necessary.

The motion boundary contrast term: This term is not essential in the sense that the problem would remain well posed if it were omitted. However, it brings in additional information because it uses a characterization of motion boundaries in terms of image measurements to further constrain the solution. It has the form, for either region $\mathbf{R}_i$:

$$\mathscr{E}_b = -\frac{1}{2}\int_{\gamma} g\, ds \tag{7.28}$$

where function g is derived as follows [13]. Assuming that the motion field is smooth everywhere (i.e., $\nabla u_i \approx 0$ and $\nabla v_i \approx 0$) except at motion boundaries, differentiation with respect to the image spatial coordinates of the Horn and Schunck optical flow constraint

$$\langle \nabla I, W_i \rangle + I_t = 0 \tag{7.29}$$

gives

$$\begin{aligned} \nabla\left(\langle \nabla I, W_i\rangle + I_t\right) &= \mathbf{H}W_i + \nabla\left(I_t\right) && (7.30)\\ &\approx \mathbf{0} && (7.31)\end{aligned}$$

where $\mathbf{H}$ is the Hessian matrix of the image function I, i.e., the matrix of second order derivatives of I (we assume that I is twice continuously differentiable within regions)

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{pmatrix} \tag{7.32}$$

Then, combining (7.30) with (7.29), gives the following motion boundary characterization function:

$$g = \left|\det(\mathbf{H}) I_t - \langle \nabla I, (\mathbf{H}^* \nabla I_t)\rangle\right| \tag{7.33}$$

where $|\cdot|$ denotes the absolute value and det the determinant; $\mathbf{H}^*$ is the transpose of the matrix of cofactors of $\mathbf{H}$. It has the property

$$\mathbf{H}^*\mathbf{H} = \det(\mathbf{H})\mathbf{I}, \tag{7.34}$$

where $\mathbf{I}$ is the identity matrix. Function g is an indicator of motion boundaries because it takes small values inside motion regions, where motion is smooth, and inconsistent, generally large values at motion boundaries where the optical flow constraint equation is theoretically not valid.

Using (7.26), (7.27), and (7.28), the objective functional is:

$$\mathscr{E} = \frac{1}{2}\sum_{i=1}^{2}\int_{\mathbf{R}_i}\left(\langle \nabla I, \alpha_i^T \theta\rangle + I_t\right)^2 d\mathbf{x} + \frac{1}{2}\lambda\int_{\gamma} ds - \frac{1}{2}\varepsilon\int_{\gamma} g\, ds \tag{7.35}$$

Parametric motion estimation and segmentation consists now of determining a curve γ and coefficients α_1, α_2 which minimize functional (7.35).

This formulation can be viewed as a motion-based active curve application of the Mumford-Shah functional [20]. It generalizes motion segmentation by constant/affine representation [25] to general linear parametric segmentation. It is also a generalization of the Chan and Vese functional [18] to general linear parametric motion-based image segmentation.

7.3.2 Functional minimization

The minimization of (7.35) follows its derivation with respect to γ and motion parameters α_1, α_2 (see Chapter 2 for basic formulas).

Minimization with respect to the motion parameters

Using a shorthand matrix derivative notion, the motion parameters of each region $\mathbf{R}_i$, $i = 1,2$, follow the necessary conditions

$$\frac{\partial}{\partial \alpha_i}\left(\frac{1}{2}\int_{\mathbf{R}_i}\left(\langle \nabla I, \alpha_i^T \theta\rangle + I_t\right)^2 d\mathbf{x}\right) = \mathbf{0}, \qquad i = 1,2 \tag{7.36}$$

For simplicity, the dependence of α_i on γ has been omitted because it is inconsequential at this point. Algebraic manipulations lead to

$$\frac{\partial}{\partial \alpha_i}\left(\frac{1}{2}\int_{\mathbf{R}_i}\left(\langle \nabla I, \alpha_i^T \theta\rangle + I_t\right)^2 d\mathbf{x}\right) = \mathbf{B}_i \beta_i + \mathbf{d}_i, \qquad i = 1,2 \tag{7.37}$$

where:

- $\mathbf{B}_i$ is the (region dependent) $2M \times 2M$ block matrix

  $$\mathbf{B}_i = \begin{pmatrix} \mathbf{B}_i^{11} & \mathbf{B}_i^{12} \\ \mathbf{B}_i^{21} & \mathbf{B}_i^{22} \end{pmatrix} \tag{7.38}$$

  where the element (m,n), $m = 1,\ldots,M$, $n = 1,\ldots,M$ of sub-matrix $\mathbf{B}_i^{rc}$, $r,c = 1,2$ is given by

  $$\mathbf{B}_i^{rc}[m,n] = \int_{\mathbf{R}_i} I_r I_c \theta_m \theta_n d\mathbf{x} \tag{7.39}$$

  and where I_l is the image spatial derivative with respect to the horizontal ($l = 1$) and the vertical ($l = 2$) image coordinates.
- β_i is a $2M \times 1$ vector constructed by (vertical) concatenation of the motion parameters α_{i1} and α_{i2} corresponding to the horizontal (u) and vertical (v) coordinates of optical flow. Its elements are, where $m = 1,\ldots,M$,

$$\beta_i[m] = \alpha_{im1} \tag{7.40}$$

$$\beta_i[M+m] = \alpha_{im2} \tag{7.41}$$

- $\mathbf{d}_i$ is a $2M \times 1$ vector with elements, where $m = 1, \dots, M$,

$$\mathbf{d}_i[m] = \int_{\mathbf{R}_i} I_t I_1 \theta_m d\mathbf{x} \tag{7.42}$$

$$\mathbf{d}_i[M+m] = \int_{\mathbf{R}_i} I_t I_2 \theta_m d\mathbf{x} \tag{7.43}$$

In matrix notation, the necessary conditions with respect to the motion parameters $(\beta_i, i = 1, 2)$ are:

$$\mathbf{B}_i \beta_i = -\mathbf{d}_i \tag{7.44}$$

A least squares estimate of the motion parameters can be computed from the linear system of the equations (7.44) (using singular value decomposition, for instance) For each region $\mathbf{R}_i$, the matrix $\mathbf{B}_i$ (Equation 7.38) and vector $\mathbf{d}_i$ are in terms of the image first order spatio-temporal variations and the given basis functions of the image motion representation (Equations 7.39, 7.42 and 7.43).

Minimization with respect to γ: the curve evolution equation

To write the necessary conditions for a minimum of the objective functional with respect to γ, we need to determine its functional derivative with respect to γ (refer to Chapter 2). The derivation will take into consideration the dependence of α_1, α_2 on γ.

Let $\widetilde{\mathbf{h}} \in \mathbb{R}^{2M}$ have elements, for $m = 1, \cdots, M$,

$$\widetilde{\mathbf{h}}[m] = I_1 \theta_m \tag{7.45}$$

$$\widetilde{\mathbf{h}}[M+m] = I_2 \theta_m \tag{7.46}$$

Recall, from the preceding section, that I_1 and I_2 designate the spatial derivatives of I with respect to the horizontal and vertical image coordinates, respectively. First, we consider the derivation of the region integrals in the data term of the objective functional (7.35). We treat $i = 1$ (i.e., for region $\mathbf{R}_1 = \mathbf{R}_\gamma$), the treatment for $i = 2$ (i.e., for $\mathbf{R}_2 = \mathbf{R}_\gamma^c$) being identical. With the definition of $\widetilde{\mathbf{h}}$, and of $\mathbf{B}_1$, $\mathbf{d}_1$, and β_1 in the preceding Section 7.3.2, and using (7.44) as well as the equalities

$$\int_{\mathbf{R}_1} \widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^T \, d\mathbf{x} = \mathbf{B}_1 \quad \text{and} \quad \int_{\mathbf{R}_1} I_t \widetilde{\mathbf{h}} \, d\mathbf{x} = \mathbf{d}_1$$

we have

$$
\begin{aligned}
&\int_{\mathbf{R}_1} \left(\langle \nabla I, \alpha_1^T \theta\rangle + I_t\right)^2 d\mathbf{x} && (7.47)\\
&= \int_{\mathbf{R}_1} \left(\beta_1^T \widetilde{\mathbf{h}} + I_t\right)^2 d\mathbf{x} && (7.48)\\
&= \beta_1^T \left(\mathbf{B}_1\, \beta_1 + \mathbf{d}_1\right) + \beta_1^T \mathbf{d}_1 + \int_{\mathbf{R}_1} (I_t)^2\, d\mathbf{x} && (7.49)\\
&= \beta_1^T \mathbf{d}_1 + \int_{\mathbf{R}_1} (I_t)^2\, d\mathbf{x}, && (7.50)
\end{aligned}
$$

which means that

$$
\begin{aligned}
&\frac{\partial}{\partial \gamma} \int_{\mathbf{R}_1} \left(\langle \nabla I, \alpha_1^T \theta\rangle + I_t\right)^2 d\mathbf{x} && (7.51)\\
&= \frac{\partial \beta_1^T}{\partial \gamma}\mathbf{d}_1 + \beta_1^T \frac{\partial \mathbf{d}_1}{\partial \gamma} + \frac{\partial}{\partial \gamma}\left(\int_{\mathbf{R}_1} (I_t)^2\, d\mathbf{x}\right) && (7.52)
\end{aligned}
$$

The various derivatives which appear in (7.52) are as follows. Using (7.44),

$$
\frac{\partial \beta_1^T}{\partial \gamma} = \left(-\frac{\partial \mathbf{d}_1^T}{\partial \gamma} - \beta_1^T \frac{\partial \mathbf{B}_1}{\partial \gamma}\right) (\mathbf{B}_1)^{-1} \tag{7.53}
$$

Using the expressions of $\mathbf{d}_1$ and $\mathbf{B}_1$, the functional derivatives of the other terms (refer to Chapter 2) are

$$
\begin{aligned}
\frac{\partial \mathbf{d}_1}{\partial \gamma} &= I_t \widetilde{\mathbf{h}} \otimes \mathbf{n} && (7.54)\\
\frac{\partial \mathbf{B}_1}{\partial \gamma} &= \widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^T \otimes \mathbf{n} && (7.55)\\
\frac{\partial}{\partial \gamma}\left(\int_{\mathbf{R}_1} (I_t)^2\, d\mathbf{x}\right) &= (I_t)^2\, \mathbf{n} && (7.56)
\end{aligned}
$$

where $\otimes$ denotes the tensor product and $\mathbf{n}$ is the external unit normal function of γ.

Using (7.54) and (7.55) in (7.53) gives an equation which we substitute in (7.52); further substitution in (7.52) of (7.56) and (7.54) (which removes the tensor products) allows the following derivation:

$$
\begin{aligned}
&\frac{\partial}{\partial \gamma} \int_{\mathbf{R}_1} \left(\langle \nabla I, \alpha_1^T \theta\rangle + I_t\right)^2 d\mathbf{x} && (7.57)\\
&= \left(\left(-I_t \widetilde{\mathbf{h}}^T - \beta_1^T \widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^T\right) (\mathbf{B}_1)^{-1} \mathbf{d}_1 + I_t \beta_1^T \widetilde{\mathbf{h}} + I_t^2\right) \mathbf{n} && (7.58)\\
&= \left(\beta_1^T \widetilde{\mathbf{h}} + I_t\right)^2 \mathbf{n} && (7.59)\\
&= \left(\langle \nabla I, \alpha_1^T \theta\rangle + I_t\right)^2 \mathbf{n} && (7.60)\\
& && (7.61)
\end{aligned}
$$

Therefore, the functional derivative with respect to γ behaves as for an integrand independent of γ, i.e., the dependence of the motion parameters on the segmentation does not give the derivative additional terms [43]. This result can be derived by a method other than the calculus of variations [43].

Given that the (external) unit normal function to the boundary of $\mathbf{R}_\gamma^c$ is $-\mathbf{n}$, the derivative of the integral on $\mathbf{R}_2 = \mathbf{R}_\gamma^c$ in the data term, i.e., for $i = 2$, is

$$\frac{\partial}{\partial \gamma} \int_{\mathbf{R}_2} \left(\langle \nabla I, \alpha_2^T \theta \rangle + I_t \right)^2 d\mathbf{x} = - \left(\langle \nabla I, \alpha_2^T \theta \rangle + I_t \right)^2 \mathbf{n} \tag{7.62}$$

The functional derivatives with respect to γ of the second and third integrals of the objective function (7.35) are (Chapter 2):

$$\frac{\partial}{\partial \gamma} \int_\gamma ds = \kappa \mathbf{n}$$
$$\frac{\partial}{\partial \gamma} \int_\gamma g\, ds = (\langle \nabla g, \mathbf{n} \rangle + g\kappa)\mathbf{n},$$

which, finally, gives

$$\frac{\partial \mathscr{E}}{\partial \gamma} = \frac{1}{2} \left(e_1 - e_2 + \lambda \kappa - \varepsilon \langle \nabla g, \mathbf{n} \rangle - \varepsilon g \kappa \right) \mathbf{n} \tag{7.63}$$

where κ is the mean curvature function of γ and,

$$e_i = \left(\langle \nabla I, \alpha_i^T \theta \rangle + I_t \right)^2, \;\; i = 1, 2 \tag{7.64}$$

The necessary condition for a minimum of (7.35) with respect to γ follows setting the right-hand side of (7.63) to zero. Embedding γ in a one-parameter family of curves $\gamma(s,t)$, indexed by algorithmic time t, the gradient descent equation corresponding to (7.63) is

$$\frac{\partial \gamma}{\partial t} = -\frac{1}{2} \left(e_1 - e_2 + \lambda \kappa - \varepsilon \langle \nabla g, \mathbf{n} \rangle - \varepsilon g \kappa \right) \mathbf{n} \tag{7.65}$$

The algorithm can be summarized as a two-step iteration until convergence:

1. Initialize γ
2. Iterate until convergence
   a) Compute the regions coefficients by least squares (7.44)
   b) Move γ by performing a descent step (7.65)

7.3.3 Level set implementation

If γ is the zero level of level set function $\phi : \mathbb{R}^2 \to \mathbb{R}$ (see Chapter 2), the evolution of ϕ is

$$\frac{\partial \phi}{\partial \tau} = -\frac{1}{2}\left(e_1 - e_2 + \lambda \kappa - \varepsilon \langle \nabla g, \mathbf{n} \rangle - \varepsilon g \kappa\right) \|\nabla \phi\| \tag{7.66}$$

We recall that

$$\kappa = -\operatorname{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right) \tag{7.67}$$

and

$$\mathbf{n} = -\frac{\nabla \phi}{\|\nabla \phi\|} \tag{7.68}$$

7.3.4 Multiregion segmentation

There are several ways of generalizing a two-region active curve segmentation formulation to an arbitrary but fixed number of regions [44, 46, 47, 48, 49, 50, 51]. These were reviewed in Chapter 4. The example in Section 7.3.5 uses the implementation in [23] of the multiregion extension in [50] which views segmentation as regularized clustering. For a segmentation into N regions, the method uses $N-1$ closed simple plane curves $\gamma_i, i = 1, \ldots, N-1$ which move according to:

$$\frac{\partial \gamma_i}{\partial t} = -\left(e_i - \psi_i + \lambda \kappa_i - \varepsilon \langle \nabla g, \mathbf{n}_i \rangle - \varepsilon g \kappa_i\right) \mathbf{n}_i \tag{7.69}$$

where where κ_i is the curvature function of γ_i, $\mathbf{n}_i$ its external unit normal function, and

$$\psi_i = \min_{j \neq i}(e_j) \tag{7.70}$$

For functionals with a data term and a regularization length term, as those used here, the evolution equation converges to a partition, provided it starts with a partition (Chapter 4). The effect of adding a boundary characterization term has not been studied. The corresponding level set evolution equations are:

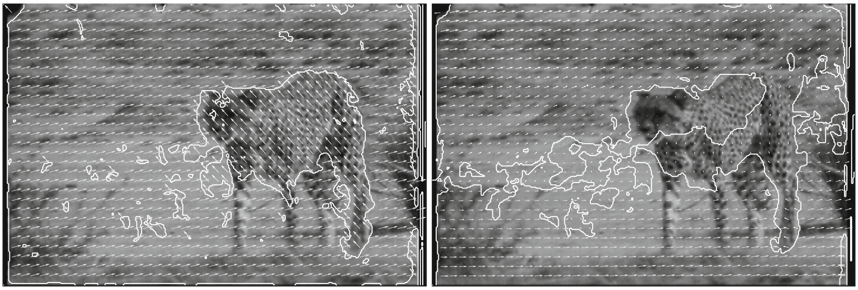
$$\frac{\partial \phi_i}{\partial t} = -\left(e_i - \psi_i + \lambda \kappa_i - \varepsilon \langle \nabla g, \mathbf{n}_i \rangle - \varepsilon g \kappa_i\right) \|\nabla \phi_i\| \tag{7.71}$$

7.3.5 Examples

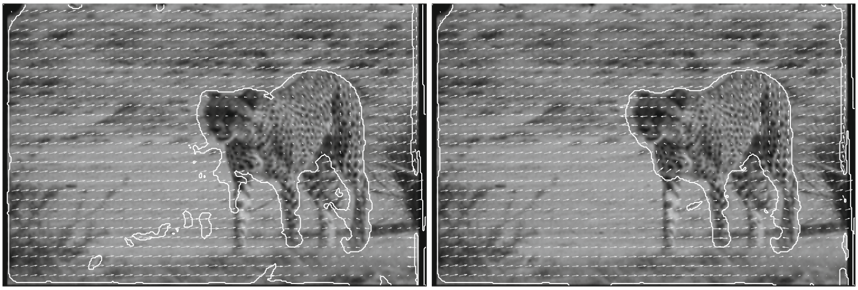
Numerical implementation of the motion based image partitioning methods described in this chapter use discrete approximations of the image spatial and temporal derivatives. These approximations can be computed as described in the Appendix, according to the Horn and Schunck prescription, for instance. Other schemes have been used in [23] for an efficient realization of the formulation.

The following example uses two images of a sequence showing a walking cheetah. The cheetah walks forward and slightly left with a nonrigid articulated movement. Two legs are at rest and the other two move forward, the front leg a little faster than the rest of the body. The camera moves to follow the animal, inducing an image motion to the right over the extent of the image domain.

Results are shown in Fig. 7.2 for a segmentation into two regions (in this case all multiregion extension methods are equivalent). Both the constant and affine models are too simple to represent the image motion and, as a result, lead to fragmented, incorrect segmentations. The DCT model with 16 basis functions gives a good segmentation. The more complex DCT model produces a better segmentation. More examples and evaluations are given in [23].



(a) Constant (b) Affine



(c) DCT with 4 basis functions and (d) DCT with 16 basis functions

Fig. 7.2. Motion segmentation and parametric estimation for the *Cheetah* sequence.

An accurate segmentation allows image motion to be estimated selectively within each of its regions, i.e., within appropriate image supports. Therefore, estimation will preserve motion boundaries and it can be quite accurate. This is illustrated on the *Marbled blocks* synthetic test sequence (KOGS/IAKS laboratory database, University of Karlsruhe, Germany). Windows of the same two images as in the MDL example are used. We recall that two blocks move in a static environment. The larger moves slightly away from the viewing system to the left, and the smaller to the left and forward in depth. The texture remains practically the same across the top edges of the blocks. The change in depth, and in image motion thereof, is sharp across the occlusion boundaries not in contact with the floor. The scene and the ground truth motion field are depicted again in Fig. 7.3 a. The segmentation and the estimated motion obtained using a DCT representation with 16 basis functions, are depicted in Fig. 7.3 b. The motion is accurate, on average, within 0.884° in direction and 0.069 of a pixel in magnitude, with corresponding standard deviations of 1.208° and 0.091 pixel. The corresponding values for the Horn and Schunck method are $5.114°/4.613°$ (direction), 0.404 pixel/0.221 pixel (magnitude). Those of the boundary preserving method in [17] are $4.199°/3.777°$ (direction), 0.480 pixel/0.251 pixel (magnitude).

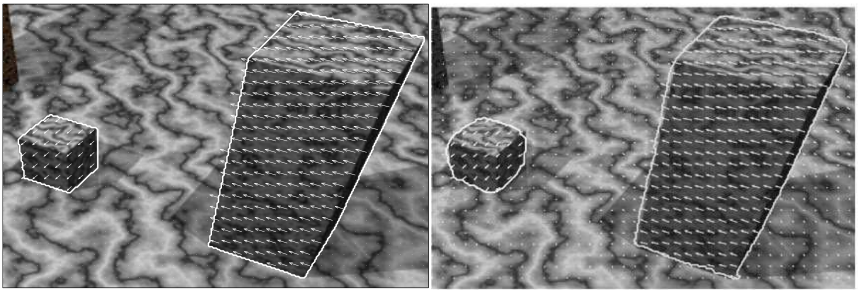

(a) (b)

Fig. 7.3. (a) An image of the *Marbled block* sequence and the ground truth image motion; (b) Computed segmentation and motion using a DCT basis with 16 basis functions.

References

1. J. J. Gibson, *The perception of the visual world*. Houghton Mifflin, 1950.
2. K. Nakayama, "Biological image motion processing: A survey," *Vision Research*, vol. 25, pp. 625–660, 1985.
3. A. Mitiche, *Computational Analysis of Visual Motion*. Plenum Press, New York, 1994.
4. A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: A synopsis of current problems and methods," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.
5. A. Kaup, "Object-based texture coding of moving video in mpeg-4," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, p. 5, Feb. 1999.
6. R. Fablet, P. Bouthemy, and P. Pérez, "Non-parametric motion characterization using causal probabilistic models for video indexing and retrieval," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 393–407, Apr. 2002.
7. P. Bouthemy, C. Garcia, R. Ronfard, G. Tziritas, E. Veneau, and D. Zugaj, "Scene segmentation and image feature extraction for video indexing and retrieval," in *3rd International Conference on Visual Information Systems*, Amsterdam, Jun. 1999, pp. 245–252.
8. J. M. Odobez and P. Bouthemy, "Direct incremental model-based image motion segmentation for video analysis," *Signal Processing*, vol. 6, no. 2, pp. 143–155, 1998.
9. C. Papin, P. Bouthemy, and P. Pérez, "Unsupervized segmentation of low clouds from infrared METEOSTAT images based on a contextual spatio-temporal labeling approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 1, pp. 104–114, 2002.
10. Q. Cai, A. Mitiche, and J. K. Aggarwal, "Tracking human motion in an indoor environment," in *International Conference on Image Processing (ICIP)*, 1995, pp. 215–218.
11. N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 266–280, 2000.
12. A. Mansouri, "Region tracking via level set pdes without motion computation," *IEEE Transacions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 947–961, Jul. 2002.
13. A. Mitiche, R. Feghali, and A. Mansouri, "Motion tracking as spatio-temporal motion boundary detection," *Journal of Robotics and Autonomous Systems*, vol. 43, pp. 39–50, 2003.
14. H. Sekkati and A. Mitiche, "Joint optical flow estimation, segmentation, and 3d interpretation with level sets," *Computer Vision and Image Understanding*, vol. 103, no. 2, pp. 89–100, 2006.
15. ——, "Concurrent 3-d motion segmentation and 3-d interpretation of temporal sequences of monocular images," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 641–653, 2006.
16. A. Mitiche and H. Sekkati, "Optical flow 3d segmentation and interpretation: A variational method with active curve evolution and level sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1818–1829, 2006.
17. G. Aubert and P. Kornpbrost, *Mathematical problems in image processing: Partial differential equations and the calculus of variations*. Springer Verlag, 2006.
18. T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
19. G. Aubert, R. Deriche, and P. Kornprobst, "Computing optical flow via variational techniques," *SIAM Journal of Applied Mathematics*, vol. 60, no. 1, pp. 156–182, 1999.
20. D. Mumford and J. Shah, "Boundary detection by using functionals," *Computer Vision and Image Understanding*, vol. 90, pp. 19–43, 1989.

21. A. Mitiche and S. Hadjres, "Mdl estimation of a dense map of relative depth and 3d motion from a temporal sequence of images," *Pattern Analysis and Applications*, vol. 6, no. 1, pp. 78–87, 2003.
22. Y. G. Leclerc, "Constructing simple stable descriptions for image partitioning," *International Journal of Computer Vision*, vol. 3, no. 1, pp. 73–102, 1989.
23. C. Vázquez, A. Mitiche, and R. Laganière, "Joint segmentation and parametric estimation of image motion by curve evolution and level sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 782–793, 2006.
24. M. Chang, A. Tekalp, and M. Sezan, "Simultaneous motion estimation and segmentation," *IEEE Transactions on Image Processing*, vol. 6, no. 9, pp. 1326–1333, Sep. 1997.
25. D. Cremers and C. Schnörr, "Motion competition: Variational integration of motion segmentation and shape regularization," in *DAGM Symposium on Pattern Recognition*, 2002, pp. 472–480.
26. D. Cremers and S. Soatto, "Variational space-time motion segmentation," in *International Conference on Computer Vision (ICCV)*, 2003, pp. 886–892.
27. É. Mémin and P. Pérez, "Joint estimation-segmentation of optic flow," in *Europeean Conference on Computer Vison (ECCV)*, vol. 2, 1998, pp. 563–577.
28. B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 17, pp. 185–203, 1981.
29. N. Diehl, "Object-oriented motion estimation and segmentation in image sequences," *Signal Processing: Image Communication*, vol. 3, no. 1, pp. 23–56, Feb. 1991.
30. J. V. A. Wang and E. H. Adelson, "Representing moving images with layers," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 625–638, Sep. 1994.
31. G. D. Borshukov, G. Bozdagi, Y. Altunbasak, and A. M. Tekalp, "Motion segmentation by multistage affine classification." *IEEE Transactions on Image Processing*, vol. 6, no. 11, pp. 1591–1594, 1997.
32. A. Mansouri and J. Konrad, "Multiple motion segmentation with level sets," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 201–220, 2003.
33. A. Mansouri, A. Mitiche, and C. Langevin, "Joint Space-Time motion-based segmentation of image sequences with level set PDEs," in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 50–55.
34. A. Mansouri, A. Mitiche, and F. Dolla, "Motion-based figure-ground segmentation by maximum motion separation," in *International Conference on Image Processing (ICIP)*, vol. 3, 2003, pp. 925–928.
35. J. M. Morel and S. Solimini, *Variational methods in image segmentation.* Birkhauser, 1995.
36. G. Dahlquist and A. Bjork, *Numerical methods.* Prentice Hall, 1974.
37. J. M. Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric motion models," *Journal of Visual Communication and Image Representation*, vol. 6, no. 4, pp. 348–365, 1995.
38. Y. Altunbasak, R. M. Mersereau, and A. J. Patti, "A fast parametric motion estimation algorithm with illumination and lens distortion correction," *IEEE Transactions on Image Processing*, vol. 12, no. 4, pp. 395–408, 2003.
39. F. Dufaux, I. Moccagatta, F. Moscheni, and H. Nicolas, "Vector quantization-based motion field segmentation under the entropy criterion," *Journal of Visual Communication and Image Representation*, vol. 5, pp. 356–369, Dec. 1994.
40. D. Cremers, "A multiphase level set framework for motion segmentation," in *Scale Space Theories in Computer Vision*, 2003, pp. 599–614.

41. R. Deriche, P. Kornprobst, and G. Aubert, “Optical-flow estimation while preserving its discontinuities: A variational approach,” in *Asian Conference on Computer Vision (ACCV)*, 1995, pp. 71–80.
42. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” in *European Conference on Computer Vision (ECCV)*, vol. 4, 2004, pp. 25–36.
43. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, “Image segmentation using active contours: Calculus of variations or shape gradients?” *SIAM Journal of Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.
44. S. C. Zhu and A. Yuille, “Region competition: Unifiying snakes, region growing, and Bayes/MDL for multiband image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, Sep. 1996.
45. C. Stiller and J. Konrad, “Estimating motion in image sequences: A tutorial on modeling and computation of 2D motion,” *IEEE Signal Processing Magazine*, vol. 16, no. 4, pp. 70–91, Jul. 1999.
46. A. Mansouri, A. Mitiche, and C. Vázquez, “Multiregion competition: A level set extension of region competition to multiple region partioning,” *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
47. M. Rousson and R. Deriche, “A variational framework for active and adaptative segmentation of vector valued images,” in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 56–61.
48. L. Vese and T. Chan, “A multiphase level set framework for image segmentation using the Mumford and Shah model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
49. A. Yezzi, A. Tsai, and A. Willsky, “A fully global approach to image segmentation via coupled curve evolution equations,” *Journal of Visual Communication and Image Representation*, vol. 13, no. 1, pp. 195–216, Mar. 2002.
50. C. Vázquez, A. Mitiche, and I. Ben Ayed, “Image segmentation as regularized clustering: A fully global curve evolution method,” in *International Conference on Image Processing (ICIP)*, 2004, pp. 3467–3470.
51. I. Ben Ayed, A. Mitiche, and Z. Belhadj, “Polarimetric image segmentation via maximum-likelihood approximation and efficient multiphase level-sets,” *IEEE Transactions on Pattern Analalysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.

8

IMAGE SEGMENTATION ACCORDING TO THE MOVEMENT OF REAL OBJECTS

8.1 Introduction

In the preceding chapter, image segmentation used optical flow. The flow components in each region were described by a linear parametric model. Therefore, the model parameters served to distinguish the segmentation regions. We presented two methods which differed by their representation of the segmentation boundaries. One of the methods (Chapter 7, Section 7.2) can be viewed as an optical flow edge detection method. It used the minimum description length (MDL) formulation [1], and placed the segmentation boundary edges at points between adjacent pixels described by distinct model parameters. The other method we described was a level set method, using closed curves to delineate the segmentation regions each of which is described by the parameters of a general linear model of image motion [2]. In both methods, the parametric models did not express any relationship to the movement of the physical objects from which the optical flow arose. They were strictly statistical, two-dimensional characterizations of the optical flow in the segmentation regions. However, optical flow can be constrained to a three-dimensional interpretation, using, for instance, one of the expressions between optical flow and rigid body motion in Chapter 2, Section 2.4.4. In this case, the segmentation regions will correspond to differently moving real rigid objects.

Image segmentation according to the movement of real objects necessarily brings in variables related to the motion or to the structure of the objects. These three-dimensional (3D) variables are unknowns to be estimated.

The segmentation and the estimation of the 3D variables are in reciprocal dependence: the segmentation requires the 3D variables of each of its regions, and the estimation of the 3D variables must be performed over the extent of each region. Therefore, the segmentation and the estimation of the 3D variables must be concurrent: methods of image segmentation according to the motion of real objects are also methods of 3D interpretation of image sequences. These methods can also be viewed as methods for joint optical flow segmentation and 3D interpretation because optical flow is a function of the 3D variables of structure and motion.

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_8,

The purpose in this chapter is to describe methods of image segmentation according to the movement of real objects. These are variational, level set methods. They all assume objects are rigid and in motion relative to the viewing system, i.e., they allow simultaneous movement of objects and viewing system. They differ in the type of the 3D variables they use in the objective functional. One method [3] uses the optical flow 3D rigid body constraint (Eq. 2.65) and, therefore, eliminates optical flow from the formulation and introduces both the depth and the 3D kinematic screws of each of the moving objects. Another method [4] uses the essential parameters constraint (Eq. 2.67). As a result, optical flow and the 3D essential parameters of each moving object are the unknowns, along with the segmentation of these moving objects. The last method described [5] can be seen as a hybrid method: it segments the image according to optical flow, as in the methods of Chapter 7, but uses additional terms in the objective functional to constrain optical flow to be caused by a rigid 3D motion via the essential parameters.

There are intrinsic limitations in the recovery of 3D structure and motion from optical flow:

- An obvious limitation is that the structure of surfaces which do not move relative to the viewing system cannot be recovered. This is reflected by the viewing system geometry and the rigid body motion models in the optical flow equations (2.64). Both sides of these equations become zero when a surface is not moving with respect to the viewing system, leaving depth undermined. Also, because brightness change is a prerequisite for motion perception, another obvious limitation is that the structure of the parts of a moving surface which is not textured cannot be recovered, except possibly by propagation, via regularization, from the surface textured parts. This is shown in the rigid body constraint (2.65), the left-hand side of which becomes zero when $I_x = I_y = I_t = 0$.

- To a 3D interpretation consistent with the optical flow corresponds a family of scaled interpretations [6], i.e., one cannot distinguish between a surface with a given motion and a scaled version of the surface with a correspondingly scaled motion. This is reflected, for rigid objects, by the invariance of the optical flow equation (2.64) to a (common) scaling of depth (structure) and the translational component of 3D motion. This invariance will also appear in the the 3D interpretation equations of the methods in this chapter.

- A formulation which uses optical flow assumes image motion of small extent in practice. Large extent motion requires some form of multiresolution processing.

There are also limitations specific to the methods described in this chapter:

- The number of segmentation regions, i.e., the number of moving real objects, although arbitrary, is assumed known. This number appears explicitly in the level set functionals of the methods described in this chapter. The possibility of introducing a constraint to vary it during functional optimization as in Chapter 6 has not been studied.

- The second method described, which uses the essential parameters constraint (Section 8.4), cannot distinguish between 3D motions with the same rotational component and parallel translational components [7]. This will become clear in Section 8.4, following the description of the method. This shortcoming can be corrected (Section 8.5).

Before describing the methods, we review briefly the literature on optical flow 3D interpretation and segmentation.

In the general context of 3D interpretation of optical flow, it is customary to distinguish *sparse* from *dense* interpretation. Sparse interpretation computes the 3D variables at a sparse set of feature points over the image positional array. This topic has been addressed in numerous well documented studies [8, 9, 10]. Dense interpretation computes the 3D variables over the extent of the image positional array. This has been relatively little researched, in spite of the many studies on dense estimation of optical flow [11, 12, 13, 14, 15, 16].

It is also customary to distinguish *direct* from *indirect* interpretation. Methods where image motion estimation precedes interpretation are indirect, or two-stage [17, 18, 19]. Those which compute structure without prior explicit image motion estimation are direct methods [20, 21, 22, 23, 24, 25, 26]. In general, direct formulations substitute for image motion its expression in terms of 3D variables.

Most methods assume a viewing system moving in a static environment [23] [25, 24, 26, 27, 28]. This simplifies the problem significantly because segmentation according to the motion of real objects is no longer at issue; the relative motion of camera and environment is the single 3D motion to recover. When viewing system and objects are allowed to move simultaneously and independently, segmentation becomes essential because interpretation must consider separately each real object motion.

Several methods, both variational [3, 4, 5, 29, 30] and non variational [17, 18, 19], allowed the simultaneous movement of viewing system and viewed objects. In this case, the movement of objects is *relative* to the viewing system which is allowed to move.

The discrete method in [30] uses the minimum description length (MDL) encoding under the assumption that the real objects movements are translations, and that the maps of the estimated 3D variables are piecewise constant. This direct method is an immediate generalization to 3D of the MDL optical flow segmentation method described in the preceding chapter. The MDL encoding refers to the 3D variable local edges rather than segmentation boundaries as closed curves. As with optical flow segmentation, this lack of explicit region information generally leads to a fragmented segmentation. The variational method in [29], minimizes a integral functional with a term of conformity of the 3D interpretation to the image sequence spatio-temporal variations, and a term of regularization based on anisotropic diffusion to preserve edges in depth. As with the MDL formulation, segmentation is implicit and generally fragmented.

In the following, we describe the level set methods of [3, 4, 5].

8.2 The functionals

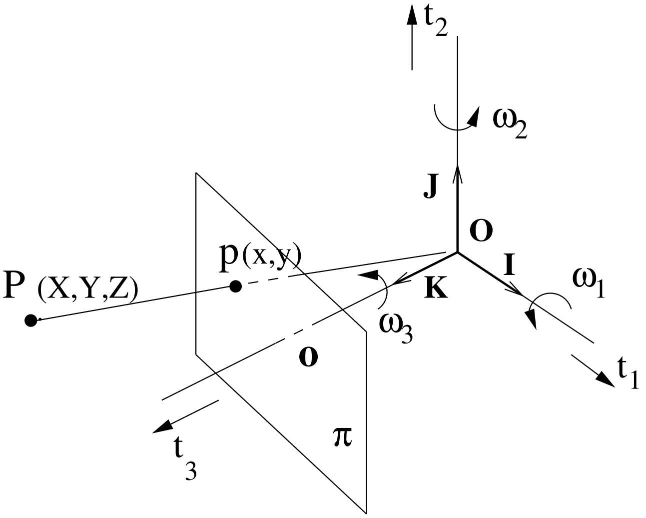


Fig. 8.1. The viewing system is symbolized by an orthonormal direct coordinate system $\mathscr{S} = (\mathbf{O}; \mathbf{I}, \mathbf{J}, \mathbf{K})$ and central projection through **O** on plane π (the image plane) parallel to plane $P_{\mathbf{IJ}}$ and at focal distance f from **O**.

Let Ω be the common domain of an image sequence I, acquired by a possibly moving viewing system. Consider the problem of partitioning Ω into two regions of distinct 3D motions relative to the viewing system. Let γ be a simple closed plane curve and $\mathbf{R}_1 = \mathbf{R}_\gamma$, its interior. Let $\mathbf{R}_2$ be the complement of $\mathbf{R}_1$ ($\mathbf{R}_2 = \mathbf{R}_1^c$), corresponding to the exterior of γ. Assuming that each region corresponds to the motion of a real rigid object, let $(\omega_k, \mathbf{T}_k)$ designate the screw of motion assigned to $\mathbf{R}_k, k = 1, 2$. The reference system has been described in Chapter 2. For the reader's convenience, it is depicted again here (Figure 8.1). Let Z designate the depth function over Ω. The objective functional of the first method we will examine is:

$$\mathscr{E}_1(\gamma, \{\omega_k, \mathbf{T}_k\}_{k=1}^2, Z) = \sum_{k=1}^{2} \int_{\mathbf{R}_k} \left(\psi_k^2 + \mu \|\nabla Z\|^2\right) d\mathbf{x} + \lambda \int_\gamma ds \tag{8.1}$$

where μ and λ are positive real constants, and the expression of $\psi_k, k = 1, 2$, is given by the left-hand side of the optical flow 3D rigid body constraint (Eq. 2.65):

$$\psi_k = I_t + \langle \mathbf{s}, \frac{\mathbf{T}_k}{Z} \rangle + \langle \mathbf{q}, \omega_k \rangle \tag{8.2}$$

with

$$\mathbf{s} = \begin{pmatrix} fI_x \\ fI_y \\ -xI_x - yI_y \end{pmatrix} \qquad \mathbf{q} = \begin{pmatrix} -fI_y - \frac{y}{f}(xI_x + yI_y) \\ fI_x + \frac{x}{f}(xI_x + yI_y) \\ -yI_x + xI_y \end{pmatrix} \tag{8.3}$$

It is clear that the first term, the data term, in the first integral of (8.1) is to evaluate the conformity, within each region, of the 3D interpretation (depth and the screw of motion) to the image spatiotemporal variations. The second term in this integral is to evaluate the smoothness of depth within each region. The last term of the functional is the usual curve length term to bias the interpretation toward a smooth region boundary γ.

The second functional we will study is:

$$\begin{aligned} \mathscr{E}_2(\{\mathbf{e}_k\}_1^2, \gamma, u, v) = & \sum_{k=1}^{2} \int_{\mathbf{R}_k} \left(\langle \mathbf{d}, \mathbf{e}_k \rangle^2 + \mu (I_x u + I_y v + I_t)^2 + \nu (\|\nabla u\|^2 + \|\nabla v\|^2) \right) d\mathbf{x} \\ & + \lambda \int_\gamma ds \end{aligned} \tag{8.4}$$

where (u, v) designates optical velocity, $\mathbf{e}_k$ is the essential parameter vector of the rigid motion of $\mathbf{R}_k$, $k = 1, 2$, and $\mathbf{d} = (x^2, y^2, f^2, 2xy, 2xf, 2yf, -fv, fu, -uy + vx)$. The elements of the essential parameter vector are related to the screw of rigid motion by Eq. 2.68. It is clear that the first term in the integrand of the first integral is to evaluate the conformity to optical flow of the 3D interpretation, via the essential parameters, within each region. The other terms of the integrand are used to estimate optical flow by the Horn and Schunck method. Note that this estimation is done within each region, not across the region boundary γ.

The third, and last, functional we will study is:

$$\mathscr{E}_3(\gamma) = \sum_{k=1}^{2} \int_{\mathbf{R}_k} \left((u^* - \overline{u}_k^*)^2 + (v^* - \overline{v}_k^*)^2 \right) d\mathbf{x} + \lambda \int_\gamma ds \tag{8.5}$$

where (u^*, v^*) is a motion field consistent, in *each* region, with a *single* rigid motion described by essential parameters and computed jointly with the essential parameters by, for $k = 1, 2$,

$$(u^*, v^*, \mathbf{e}_k) = \arg \min_{u, v, \mathbf{e}_k} \int_{\mathbf{R}_k} \left(\langle \mathbf{d}, \mathbf{e}_k \rangle^2 + \mu (I_x u + I_y v + I_t)^2 + \nu (\|\nabla u\|^2 + \|\nabla v\|^2) \right) d\mathbf{x} \tag{8.6}$$

and $\overline{u}_k^*$ $(\overline{v}_k^*)$ is the mean of u^* (v^*) in region $\mathbf{R}_k$, $k = 1, 2$,

$$\begin{aligned} \overline{u}_k^* &= \frac{1}{\int_{\mathbf{R}_k} d\mathbf{x}} \int_{\mathbf{R}_k} u^* d\mathbf{x} \\ \overline{v}_k^* &= \frac{1}{\int_{\mathbf{R}_k} d\mathbf{x}} \int_{\mathbf{R}_k} v^* d\mathbf{x} \end{aligned}$$

$\mathscr{E}_3$ is a piecewise constant optical segmentation functional. However, the regions mean optical velocities, which are the regions parameters, are constrained by (8.6) to correspond to the motion of real rigid objects via rigid motion essential parameters.

It is clear that the first functional, $\mathscr{E}_1$, differs from the other two because it includes depth when the others do not, and does not involve optical flow when the others do. The difference between functionals $\mathscr{E}_2$ and $\mathscr{E}_3$ is more subtle. The segmentation in $\mathscr{E}_2$ depends directly on the essential parameters whereas $\mathscr{E}_3$ is an optical flow based image partitioning functional although the optical flow it uses is constrained, by (8.6), to be consistent in each region with a rigid motion described by the essential parameters.

From a computational point of view, the necessity to regularize depth in $\mathscr{E}_1$ and its non linear occurrence are, *a priori*, disadvantageous. The quadratic data terms in $\mathscr{E}_2$ and $\mathscr{E}_3$ point to a better computational behavior. Also, with both $\mathscr{E}_2$ and $\mathscr{E}_3$, regularized depth can be automatically recovered, *a posteriori*, from the essential parameter vector and optical flow.

Objective functional $\mathscr{E}_2$ appears less computationally demanding than $\mathscr{E}_3$ which involves two minimizations, although each simple. However, we will see that, contrary to $\mathscr{E}_3$, $\mathscr{E}_2$ does not distinguish between screws with the same rotation and collinear translations.

The minimization of each functional follows the Euler-Lagrange descent equations, implemented via level sets.

8.3 Minimization of $\mathscr{E}_1$

We need to derive the necessary conditions for a minimum of $\mathscr{E}_1$ with respect to the screws of motion, depth, and the active curve.

8.3.1 Minimization with respect to the screws of motion

Each term ψ_k, $k = 1,2$, in $\mathscr{E}_1$ depends linearly on $\mathbf{T}_k$ and ω_k. Therefore, minimization of (8.1) with respect to these variables reduces to linear least-squares estimation of the screw parameters within each region. With digital images, let N_k be the number of pixels in region $\mathbf{R}_k$, and let $\mathbf{b}$ be the following vector, evaluated at each point $\mathbf{x}_i$ of $\mathbf{R}_k$:

$$\mathbf{b}(\mathbf{x}_i) = \left(\frac{s_1}{Z}, \frac{s_2}{Z}, \frac{s_3}{Z}, q_1, q_2, q_3\right)_{\mathbf{x}_i} \tag{8.7}$$

where

$$\begin{aligned} (s_1, s_2, s_3)^T &= \mathbf{s} \\ (q_1, q_2, q_3)^T &= \mathbf{q} \end{aligned} \tag{8.8}$$

with $\mathbf{s}$ and $\mathbf{q}$ having the meaning in Eq. 2.66 of Chapter 2. The necessary conditions with respect to the screw of motion for region $\mathbf{R}_k$ are:

$$\mathbf{B}_k \rho_k = \mathbf{c}_k \tag{8.9}$$

where

$$\rho_k = (\mathbf{T}_k, \omega_k)^T \tag{8.10}$$

and matrix $\mathbf{B}_k$ and vector $\mathbf{c}_k$ are:

$$\mathbf{B}_k = \begin{pmatrix} \mathbf{b}(\mathbf{x}_1) \\ \vdots \\ \mathbf{b}(\mathbf{x}_{N_k}) \end{pmatrix} \quad \mathbf{c}_k = \begin{pmatrix} -I_t(\mathbf{x}_1) \\ \vdots \\ -I_t(\mathbf{x}_{N_k}) \end{pmatrix}$$

This generally overdetermined linear system can be solved efficiently by singular value decomposition [31, 32].

8.3.2 Minimization with respect to depth

Note that the data term of $\mathscr{E}_1$ can be rewritten as:

$$\sum_{k=1}^{2} \int_{\Omega} \chi_k \left(\psi_k^2 + \mu \|\nabla Z\|^2 \right) d\mathbf{x} \tag{8.11}$$

where χ_k is the characteristic function of region $\mathbf{R}_k$. Therefore, the Euler-Lagrange descent equation to minimize $\mathscr{E}_1$ with respect to depth is:

$$\frac{\partial Z}{\partial t} = \sum_{k=1}^{2} \chi_k \left(\frac{\langle \mathbf{s}, \mathbf{T}_k \rangle}{Z^2} \psi_k + \mu \nabla^2 Z \right) \tag{8.12}$$

where ∇^2 is the Laplacian operator. With digital images, the Laplacian in this formula is to be estimated within each region. Estimation at a region border point can be done according to a boundary preserving recipe [33].

8.3.3 Minimization with respect to the active curve

Let

$$e_k = \psi_k^2 + \mu \|\nabla Z\|^2, k = 1, 2 \tag{8.13}$$

The Euler-Lagrange descent equation with respect to curve γ (determined, as usual, according to the basic formulas in Chapter 2) is, assuming any region parameter fixed, i.e., independent of the segmentation:

$$\frac{\partial \gamma}{\partial \tau} = -(e_1 - e_2 + \lambda \kappa) \mathbf{n} \tag{8.14}$$

where $\mathbf{n}$ is the external unit normal to γ and κ is its curvature. The level set evolution equation corresponding to (8.14) is (formulas in Chapter 2):

$$\frac{\partial \phi}{\partial \tau} = -(e_1 - e_2 + \lambda \kappa) \|\nabla \phi\| \tag{8.15}$$

where ϕ is the level set function, positive in the interior of γ. Recall that

$$\kappa = -\mathrm{div}\left(\frac{\nabla \phi}{\|\nabla \phi\|}\right) \tag{8.16}$$

8.3.4 Algorithm

The algorithm to minimize $\mathscr{E}_1$ can be summarized has follows:

1. Initialize Z and γ
2. Repeat until convergence
   a) Compute the parameters of motion by least squares (Eq. 8.9)
   b) Update depth using Eq. 8.12 (one or more iterations)
   c) Evolve the level set using one iteration of Eq. 8.15

It is often convenient to initialize depth to correspond to a fronto-parallel plane. The curve can be initialized systematically as a circle which covers about half the image and approximately centered. Note that the objective functional is decreased at each step of the algorithm (greedy algorithm). Therefore, the functional being positive, the algorithm converges to a local minimum (assuming such a minimum exists).

8.3.5 Uncertainty of scale in 3D interpretation

Methods of sparse 3D interpretation [8] can recover depth and the translational component of rigid motion only up to a scale factor. This has been called the uncertainty of scale. We ask whether a similar result holds for dense interpretation by a variational method such as the one investigated here.

Looking at the objective functional $\mathscr{E}_1$ also as a function of the regularization constant μ, we have

$$\mathscr{E}_1(\gamma, \{\omega_k, \mathbf{T}_k\}_{k=1}^2, Z, \mu) = \mathscr{E}_1(\gamma, \{\omega_k, \alpha \mathbf{T}_k\}_{k=1}^2, \alpha Z, \frac{\mu}{\alpha^2}) \tag{8.17}$$

where α is a positive constant. Therefore, to a 3D interpretation corresponds an interpretation scaled by α with μ adjusted. This can be checked directly in the minimization update equations. Let the change of variables $Z \to Z' = \alpha Z$ and $\mathbf{T} \to \mathbf{T}' = \alpha \mathbf{T}$ reflect a change of scale of Z and $\mathbf{T}$, where α is a positive real number. Equation 8.9 is verified by the scaled variables Z' and $\mathbf{T}'$. As to Equation 8.12, we have the following evolution equation for Z', after manipulation and simplification:

$$\frac{\partial Z'}{\partial t} = \frac{\langle \mathbf{s}, \mathbf{T}'_k \rangle}{Z'^2} \psi_k + \mu' \nabla^2 Z', \quad k = 1, 2 \tag{8.18}$$

with $\mu' = \mu/\alpha^2$. Similarly, Equation 8.15 is verified by the scaled variables with μ replaced by μ' in the expression of e_k. Therefore, to any solution obtained with a regularization factor μ and initial condition Z_0, corresponds, with a regularization factor $\mu' = \frac{\mu}{\alpha^2}$ and initial condition αZ_0, and for any positive real number α, a solution with Z and $\mathbf{T}$ scaled by α.

8.3.6 Multiregion segmentation

There are several ways of generalizing a two-region active curve segmentation formulation to an arbitrary but fixed number of regions [34, 35, 36, 37, 38, 39, 40]. These were reviewed in Chapter 4.

8.3.7 Example

This example uses two consecutive images of the *Marbled blocks* synthetic test sequence. This sequence served to test the MDL image segmentation method and was described in the preceding chapter. We describe it again here for the reader's convenience. There are three blocks, two of which are in movement, the small block and the large to its right. The remainder of the imaged scene is static. The larger moving block recedes slightly in depth to the left, relatively to the viewing system; the smaller block moves to the left and forward in depth. The images are characterized by a texture which remains practically the same across the top edges of the blocks. They are also characterized by a sharp discontinuity in depth, and in image motion thereof, at the blocks occluding boundaries not in contact with the floor.

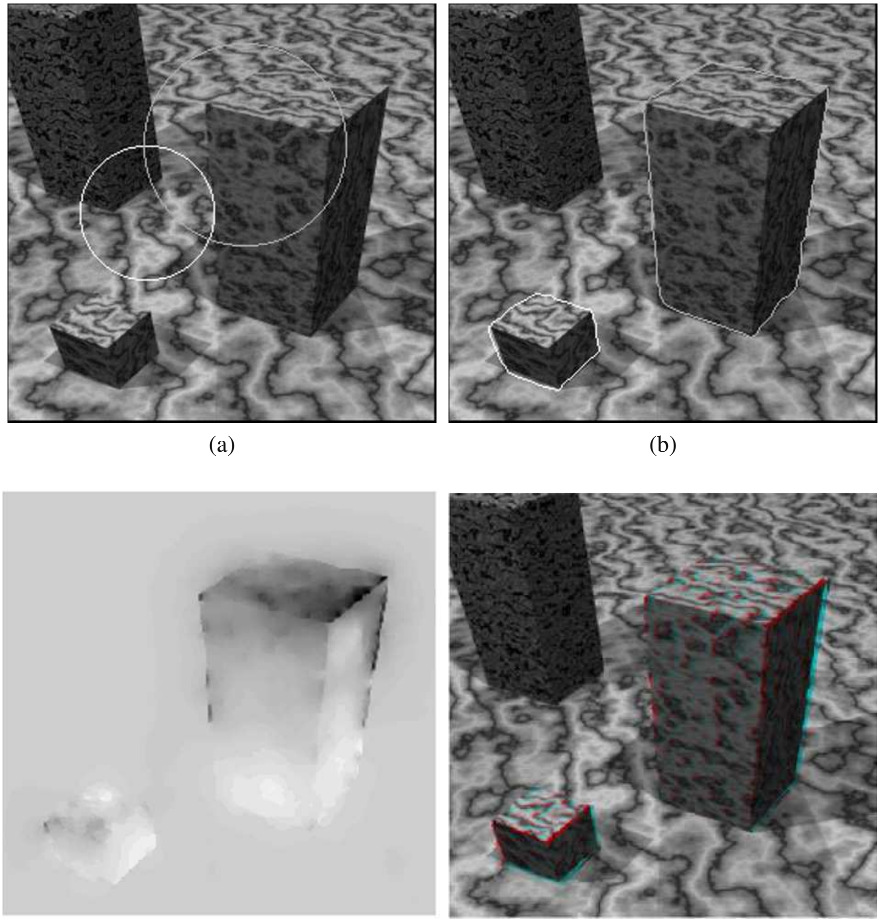
The goal is to segment the images into three regions each of which corresponds to a 3D object with a distinct motion: the small moving block, the large moving block, and the remainder of the imaged scene (i.e., the static background, which is an object with null motion). The scene and the two initial active curves are depicted in Figure 8.2a. The computed segmentation is depicted in Figure 8.2b. Figure 8.2c displays a gray level representation of the reconstructed depth of the blocks and Figure 8.2d an anaglyph of a stereoscopic image reconstructed from the scheme's output. The anaglyph is to be viewed with red/blue glasses.

8.4 Minimization of $\mathscr{E}_2$

Minimization of $\mathscr{E}_2$ follows the necessary conditions for a minimum with respect to the essential parameter vectors $\{\mathbf{e}_k\}_{k=1}^2$, the optical flow field (u,v), and active segmentation curve γ.

8.4.1 Minimization with respect to the essential parameter vectors

As with the screws of motion in $\mathscr{E}_1$, this minimization amounts to a linear least-squares evaluation of the essential parameter vector within each region, which can be



(a) (b)

(c) (d)

Fig. 8.2. (a) The first of the two consecutive *Block* sequence images used in the experiment and the initial curves; (b) the computed 3D motion segmentation; (c) a gray level representation of the moving blocks reconstructed depth and, (d) an anaglyph of a stereoscopic image reconstructed from the scheme's output.

done efficiently via singular value decomposition of the matrix of the corresponding linear system of equations:

$$\mathbf{D}_k \mathbf{e}_k = 0, \quad k = 1, 2, \tag{8.19}$$

where $\mathbf{D}_k$ is the matrix the rows of which are obtained by evaluating vector $\mathbf{d}$ at each point of region $\mathbf{R}_k$. These equations being homogeneous, the essential parameters are determined up to a scale factor, generally by imposing $\|\mathbf{e}_k\| = 1$.

8.4.2 Minimization with respect to optical flow

$\mathscr{E}_2$ can be rewritten as follows:

$$\mathscr{E}_2(u,v) = \sum_{k=1}^{2} \int_{\Omega} \chi_k \left(\langle \mathbf{d}, \mathbf{e}_k \rangle^2 + \mu (I_x u + I_y v + I_t)^2 + \nu (\|\nabla u\|^2 + \|\nabla v\|^2) \right) d\mathbf{x} \tag{8.20}$$

This gives the following Euler-Lagrange equations:

$$\begin{cases} \sum_{k=1}^{2} \chi_k \left((f e_{k,8} - y e_{k,9}) \langle \mathbf{d}, \mathbf{e}_k \rangle + \mu I_x (I_x u + I_y v + I_t) - \nu \nabla^2 u \right) &= 0 \\ \sum_{k=1}^{2} \chi_k \left((-f e_{k,7} + x e_{k,9}) \langle \mathbf{d}, \mathbf{e}_k \rangle + \mu I_y (I_x u + I_y v + I_t) - \nu \nabla^2 v \right) &= 0, \end{cases} \tag{8.21}$$

which amounts to solving for optical flow separately within each region. The Laplacian in these formulas is to be estimated within each region. As with (8.12), estimation of the Laplacian at a region border point can be done according to a boundary preserving recipe [33]. The optical flow estimations equations form a large sparse linear system of equations which can be solved by Jacobi or Gauss-Seidel iterations, as with the Horn and Schunck algorithm (given in the Appendix). The region boundary (which is also a 3D-motion boundary), is preserved because computations are confined to the regions interior.

8.4.3 Minimization with respect to γ

Assuming any region parameter fixed (i.e., independent of the segmentation), the Euler-Lagrange equation with respect to γ gives the following descent equation for the evolution of γ:

$$\frac{\partial \gamma}{\partial t} = -(e_1 - e_2 + \lambda \kappa) \mathbf{n} \tag{8.22}$$

where

$$e_k = \langle \mathbf{d}, \mathbf{e}_k \rangle^2, \quad k = 1,2 \tag{8.23}$$

The corresponding level set evolution equation is:

$$\frac{\partial \phi}{\partial t} = -(e_1 - e_2 + \lambda \kappa) \|\nabla \phi\| \tag{8.24}$$

8.4.4 Recovery of regularized relative depth

For each region, the components of the kinematic screw of rigid 3D motion, $(\mathbf{T}, \omega)$, can be recovered analytically from the region essential parameters using Eq. 2.68 of Chapter 2 [8]. The translation $\mathbf{T}$ is recovered up to a sign and a positive scale factor

[8, 10]. When $\mathbf{T} \neq \mathbf{0}$, we use Eq. 2.64 of Chapter 2 to compute the relative depth as a function of image motion and the components of rigid motion:

$$Z = \left(\frac{(ft_1 - xt_3)^2 + (ft_2 - yt_3)^2}{(u + \frac{xy}{f}\omega_1 - \frac{f^2+x^2}{f}\omega_2 + y\omega_3)^2 + (v + \frac{f^2+y^2}{f}\omega_1 - \frac{xy}{f}\omega_2 - x\omega_3)^2} \right)^{\frac{1}{2}} \tag{8.25}$$

Because the components of $\mathbf{T}$ appear in a ratio with depth in (2.64), translation and depth are recovered up to a common scale factor which is fixed when the essential parameters are computed under a fixed norm constraint (Equation 2.67). Once depth is computed, the sign of $\mathbf{T}$ is adjusted, if necessary, to correspond to positive depth [8, 10]. Note that the recovered depth is regularized because the optical flow and the essential parameters estimates from which it is computed are regularized solutions of the functional minimization.

The recovery up to a scale factor of the 3D translation has two implications. First, the depth in one region is relative to depth in another in the ratio of the norms of their actual translational components of motion. Second, image sequences of moving objects with the same rotational component of motion and collinear translational components may not be properly segmented, although recovery of the scene structure would not be affected.

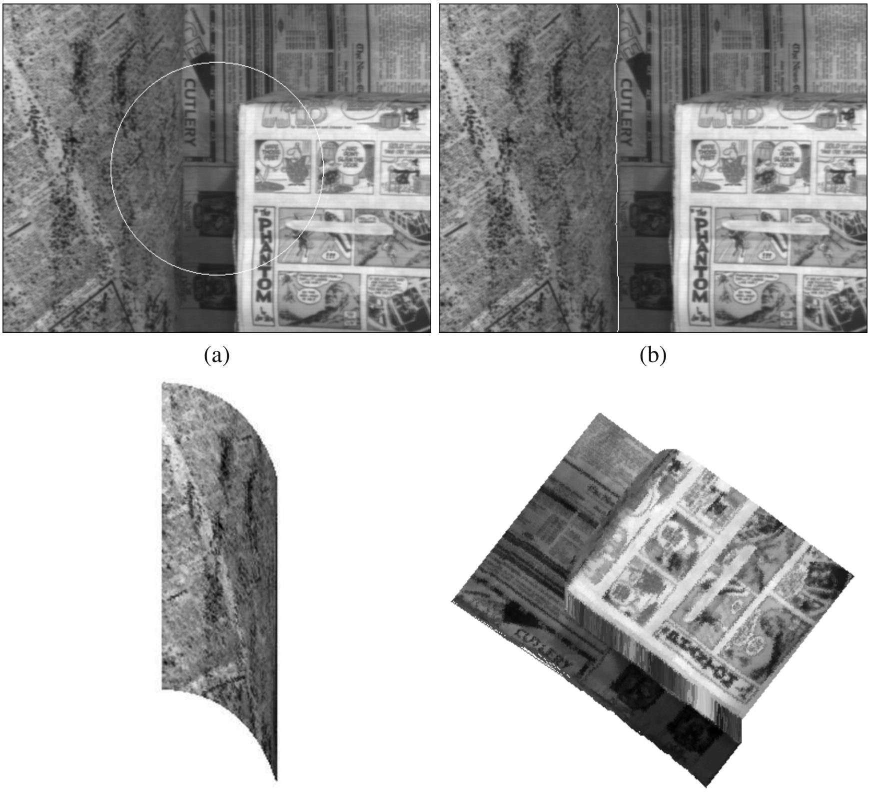
8.4.5 Algorithm

The algorithm to minimize $\mathscr{E}_2$ can be summarized as follows:

1. Initialize the curve and optical flow
2. Repeat until convergence
    a) Estimate the essential parameter vectors
    b) Estimate optical flow
    c) Evolve the curve according to (8.24)

The optical flow can be initialized to zero. More efficiently, it can be initialized with the result of a few iterations of the Horn-and-Schunck algorithm applied to the whole image (see Appendix). As with $\mathscr{E}_1$, the initial curve can be chosen systematically as a circle approximately centered and which covers about half the image. One or two iterations of (8.24) suffice at step 2.a Each step assumes the other variables fixed. Also, optical flow estimation at each step uses the estimate at the previous step. The algorithm converges because each step decreases the objective functional at each iteration.

The method can be generalized to multiple regions (Chapter 4).



(a) (b)

(c) (d)

Fig. 8.3. (a) The first frame of the *Cylinder* sequence and initial active curve; (b) the computed segmentation into two regions, one corresponding to the cylindrical object and the other to the box and background; (c) a view of the recovered cylindrical surface, and (d) a view of the recovered box and background.

8.4.6 Example

This test sequence contains three moving objects (courtesy of C. Debrunner [41]). One object is a cylindrical surface rotating approximately 1^o per frame around a vertical axis. Its 3D translation causes an image motion of about 0.15 pixel/frame. The second object is a flat background also moving right with a corresponding image translation of approximately 0.15 pixel per frame. Finally, there is a box which translates right with a corresponding image translation speed of about 0.30 pixel/frame. The movements of the box and background are parallel (3D) translations. Therefore, the purpose is to apply the method for a 3D-motion segmentation into two regions (refer to the explanation in Section 8.4.4), namely one region corresponding to the imaged surface of the cylindrical object and another to the viewed parts of the box and background together.

Fig. 8.3a shows the first of the two consecutive frames used, on which the initial segmentation curve has been superimposed, and Fig. 8.3b the final (3D motion) segmentation. A view of the reconstructed structure of the cylindrical object (from a viewpoint distinct from the camera), back-shaded with the original texture is displayed in Figure 8.3c. Note that the recovered surface is smooth due to a regularized depth. Views of the reconstructed structures for the box and background are displayed in Figure 8.3d. Fig. 8.4 displays an anaglyph of a stereoscopic image reconstructed from the scheme's output.

Fig. 8.4. An anaglyph of a stereoscopic image reconstructed from the method's output (to be viewed with red/blue filter glasses).

8.5 Minimization of $\mathscr{E}_3$

Functional $\mathscr{E}_3$ is minimized with respect to the segmentation (curve γ), subject to Eq. 8.6. Therefore, the formulation seeks simultaneously a partition of the image domain and an estimate of optical flow which is consistent with a *single* rigid body motion in *each* region of the segmentation. A functional such as (8.5) is often referred to as a piecewise constant segmentation functional [36, 42, 43]. It is the simplest form of the general linear parametric functional [44]. The focus of such formulations is not on computing an accurate parametric representation of an image or a flow field, but on partitioning the image domain into regions which are assumed to differ by the representation parameters. The justification is that a parametric representation can be coarse and still distinguish between the regions of interest. This is the case in this formulation which seeks a segmentation under the assumption that the *desired* regions, i.e., where each region corresponds to a single rigid object in space, differ by their average optical velocity. This assumption is generally valid for optical flow induced by rigid body motion. If necessary, an affine model [43] can be used without affecting the segmentation paradigm. However, 3D interpretation requires both an accurate optical flow estimate [8] and a segmentation where each region corresponds to a single rigid object in space. These requirements are embodied in Eq. 8.6.

As with $\mathscr{E}_2$, the minimization of $\mathscr{E}_3$ proceeds in three iterated steps, following initialization. As with $\mathscr{E}_2$, the optical flow is initialized to zero. Alternatively, it can be initialized with the result of a few iterations of the Horn-and-Schunck algorithm (see Appendix). The initial curve can be chosen systematically as a circle approximately centered and which covers about half the image. The first step estimates the essential parameter vectors $\mathbf{e}_k$, $k = 1,2$, by singular value decomposition, subject to $\|\mathbf{e}_k\| = 1$ [32]. The second step estimates optical flow, given the essential parameters and the segmentation curve. The necessary conditions are the same as those of $\mathscr{E}_2$ (Eq. 8.21):

$$\begin{cases} \sum_{k=1}^{2} \chi_k \left((fe_{k,8} - ye_{k,9})\langle \mathbf{d}, \mathbf{e}_k \rangle + \mu I_x (I_x u + I_y v + I_t) - \nu \nabla^2 u \right) &= 0 \\ \sum_{k=1}^{2} \chi_k \left((-fe_{k,7} + xe_{k,9})\langle \mathbf{d}, \mathbf{e}_k \rangle + \mu I_y (I_x u + I_y v + I_t) - \nu \nabla^2 v \right) &= 0 \end{cases}$$

Assuming any region parameter fixed, i.e., independent of the segmentation, the final step moves γ, via the level set equation:

$$\frac{\partial \phi}{\partial \tau} = -(e_1 - e_2 + \lambda \kappa) \|\nabla \phi\| \tag{8.26}$$

where

$$e_k = (u^* - \overline{u}_k^*)^2 + (v^* - \overline{v}_k^*)^2, \quad k = 1,2 \tag{8.27}$$

Essentially, the algorithms for $\mathscr{E}_2$ and $\mathscr{E}_3$ differ by their curve evolution equation. As with $\mathscr{E}_2$, regularized depth is recovered analytically by Eq. 8.25. The screws of motion are recovered as in Section 8.4.4 [8]. However, this formulation can distinguish between collinear translations because segmentation is based on optical flow rather than the essential parameter vectors as with $\mathscr{E}_2$.

As with $\mathscr{E}_1$ and $\mathscr{E}_2$, the method can be generalized to multiple regions (Chapter 4)

8.5.1 Example

This example uses two consecutive frames of the real image sequence *Berber*. The sculpted head was made to rotate and move forward to the right between the two frames. The object surface was textured by covering it with pasted newspaper. The initial curve of segmentation is shown in Fig. 8.5a, superimposed on the first of the two frames, and Fig. 8.5b displays the computed segmentation. A view of the reconstructed figurine, back-shaded using the original texture, is shown in Figure 8.5c. The incorrect segmentation in Fig. 8.5d was obtained with the 3D interpretation term removed from the functional. Fig. 8.6 displays an anaglyph of a stereoscopic image constructed from the first image and the estimated depth. As mentioned in Section 8.1, $\mathscr{E}_3$ is capable, contrary to $\mathscr{E}_2$, of distinguishing between distinct parallel translations. This is illustrated in Fig. 8.7 with the *Cylinder* sequence, where the box image is distinguished from the background.

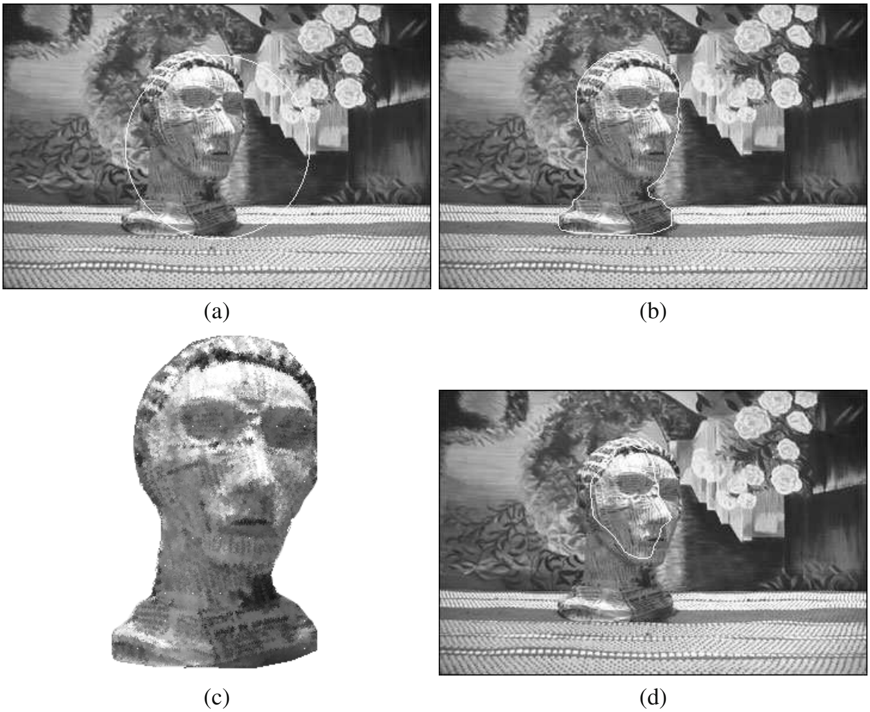

(a) (b) (c) (d)

Fig. 8.5. (a) The first frame of the *Berber* sequence and initial zero level set; (b) the computed segmentation into two regions, one corresponding to the figurine and the other to the background imaged surfaces; (c) a view of the recovered figurine surface, and (d) an incorrect segmentation obtained without the 3D term in the objective functional

Fig. 8.6. *Berber* sequence: An anaglyph of a stereoscopic image reconstructed reconstructed from the first frame and the estimated depth (to be viewed with red/blue filter glasses).

Fig. 8.7. Segmentation of the *Cylinder* sequence by $\mathscr{E}_2$ into three regions: The box is distinguished from the background on the basis of optical flow.

References

1. Y. G. Leclerc, “Constructing simple stable descriptions for image partitioning,” *International Journal of Computer Vision*, vol. 3, no. 1, pp. 73–102, 1989.
2. C. Vázquez, A. Mitiche, and R. Laganière, “Joint segmentation and parametric estimation of image motion by curve evolution and level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 782–793, 2006.
3. H. Sekkati and A. Mitiche, “Concurrent 3-d motion segmentation and 3-d interpretation of temporal sequences of monocular images,” *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 641–653, 2006.
4. A. Mitiche and H. Sekkati, “Optical flow 3d segmentation and interpretation: A variational method with active curve evolution and level sets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1818–1829, 2006.
5. H. Sekkati and A. Mitiche, “Joint optical flow estimation, segmentation, and 3d interpretation with level sets,” *Computer Vision and Image Understanding*, vol. 103, no. 2, pp. 89–100, 2006.
6. S. Ullman, “The interpretation of structure from motion,” *Proceedings of the Royal Society of London, B*, vol. 203, pp. 405–426, 1979.
7. A. Mitiche, *Computational Analysis of Visual Motion*. Plenum Press, New York, 1994.
8. ——, *Computational Analysis of Visual Motion*. Kluwer Academic, 1994.
9. H. Longuet-Higgins and K. Prazdny, “The interpretation of a moving retinal image,” *Proceedings of the Royal Society of London, B*, vol. 208, pp. 385–397, 1981.
10. X. Zhuang and R. Haralick, “Rigid body motion and the optical flow image,” in *International Conference on Artificial Intelligence Applications*, 1984, pp. 366–375.
11. B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 17, pp. 185–203, 1981.
12. A. Mitiche and P. Bouthemy, “Computation and analysis of image motion: A synopsis of current problems and methods,” *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.
13. G. Aubert, R. Deriche, and P. Kornprobst, “Computing optical flow via variational techniques,” *SIAM Journal of Applied Mathematics*, vol. 60, no. 1, pp. 156–182, 1999.
14. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” in *European Conference on Computer Vision (ECCV)*, vol. 4, 2004, pp. 25–36.
15. J. M. Odobez and P. Bouthemy, “Robust multiresolution estimation of parametric motion models,” *Journal of Visual Communication and Image Representation*, vol. 6, no. 4, pp. 348–365, 1995.
16. D. Cremers, “A multiphase level set framework for motion segmentation,” in *Scale Space Theories in Computer Vision*, 2003, pp. 599–614.
17. W. MacLean, A. Jepson, and R. Frecher, “Recovery of egomotion and segmentation of independent object motion using the em algorithm,” in *British Machine Vision Conference (BMVC)*, 1994, pp. 13–16.
18. S. Fejes and L. Davis, “What can projections of flow fields tell us about visual motion,” in *International Conference on Computer Vision (ICCV)*, 1998, pp. 979–986.
19. J. Weber and J. Malik, “Rigid body segmentation and shape description from dense optical flow under weak perspective,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 139–143, 1997.
20. T. Huang, *Image sequence analysis*. Newyrok, Springer-Verlag Berlin Heidelberg, 1981.

21. J. Aloimonos and C. Brown, "Direct processing of curvilinear sensor motion from a sequence of perspective images," in *IEEE Workshop on Computer Vision, Representation and Control*, 1984, pp. 72–77.
22. S. Negahdaripour and B. Horn, "Direct passive navigation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 168–176, 1987.
23. B. Horn and E. Weldon, "Direct methods for recovering motion," *International Journal of Computer Vision*, vol. 2, no. 2, pp. 51–76, 1988.
24. Y. Hung and H. Ho, "A kalman filter approach to direct depth estimation incorporating surface structure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 6, pp. 570–575, 1999.
25. S. Srinivasan, "Extracting structure from optical flow using the fast error search technique," *International Journal of Computer Vision*, vol. 37, no. 3, pp. 203–230, 2000.
26. T. Brodsky, C. Fermuller, and Y. Aloimonos, "Structure from motion: Beyond the epipolar constraint," *International Journal of Computer Vision*, vol. 37, no. 3, pp. 231–258, 2000.
27. N. Gupta and N. Kanal, "3-D motion estimation from motion field," *Artificial Intelligence*, vol. 78, no. 1-2, pp. 45–86, 1995.
28. Y. Xiong and S. Shafer, "Dense structure from a dense optical flow," in *International conference on Computer vision and Image Understanding*, 1998, pp. 222–245.
29. H. Sekkati and A. Mitiche, "Dense 3D interpretation of image sequences: A variational approach using anisotropic diffusion," in *International Conference on Image Analysis and Processing (ICIAP)*. Mantova, Italy, 2003, pp. 424–429.
30. A. Mitiche and S. Hadjres, "Mdl estimation of a dense map of relative depth and 3d motion from a temporal sequence of images," *Pattern Analysis and Applications*, vol. 6, no. 1, pp. 78–87, 2003.
31. P. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, Fifth, 1994.
32. G. E. Forsyth, A. A. Malcolm, and C. B. Moler, *Computer methods for mathematical computations*. Prentice Hall, 1977.
33. R. Feghali and A. Mitiche, "Fast computation of a boundary preserving estimate of optical flow," *SME Vision Quarterly*, vol. 17, no. 3, pp. 1–4, 2001.
34. S. C. Zhu and A. Yuille, "Region competition: Unifiying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, Sep. 1996.
35. M. Rousson and R. Deriche, "A variational framework for active and adaptative segmentation of vector valued images," in *IEEE Workshop on Motion and Video Computing*, 2002, pp. 56–61.
36. L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
37. A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *Journal of Visual Communication and Image Representation*, vol. 13, no. 1, pp. 195–216, Mar. 2002.
38. C. Vázquez, A. Mitiche, and I. Ben Ayed, "Image segmentation as regularized clustering: A fully global curve evolution method," in *International Conference on Image Processing (ICIP)*, 2004, pp. 3467–3470.
39. I. Ben Ayed, A. Mitiche, and Z. Belhadj, "Polarimetric image segmentation via maximum-likelihood approximation and efficient multiphase level-sets," *IEEE Transactions on Pattern Analalysis and Machine Intelligence*, vol. 28, no. 9, pp. 1493–1500, 2006.

40. A. Mansouri, A. Mitiche, and C. Vázquez, “Multiregion competition: A level set extension of region competition to multiple region partioning,” *Computer Vision and Image Understanding*, vol. 101, no. 3, pp. 137–150, 2006.
41. C. Debrunner and N. Ahuja, “Segmentation and factorization-based motion and structure estimation for long image sequences,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 206–211, 1998.
42. T. Chan and L. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
43. D. Cremers and S. Soatto, “Motion cpmetition: A variational approach to piecewise parametric motion segmentation,” *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, 2005.
44. C. Vázquez, A. Mansouri, and A. Mitiche, “Approximation of images by basis functions for multiple region segmentation with level sets,” in *International Conference on Image Processing (ICIP)*, 2004, pp. 549–552.

9

APPENDIX

The Horn and Shunck [1] and the Aubert-Deriche-Kornprobst [2, 3] optical flow formulations and algorithms are fundamental in motion analysis. They have served often as benchmarks to test other optical flow estimation methods. We have used them repeatedly in this book (Chapters 7 and 8). The formulations have been reviewed in Chapter 2. Here following is a summary of the algorithms.

9.1 The Horn and Schunck optical flow estimation algorithm

We recall the Horn and Schunck optical flow estimation functional [1]:

$$\mathscr{E}(u,v) = \int_{\Omega} (I_x u + I_y v + I_t)^2 dxdy + \lambda \int_{\Omega} (\|\nabla u\|^2 + \|\nabla v\|^2) dxdy \tag{9.1}$$

and its corresponding Euler-Lagrange equations

$$\begin{aligned} I_x(I_x u + I_y v + I_t) - \lambda \nabla^2 u &= 0 \\ I_y(I_x u + I_y v + I_t) - \lambda \nabla^2 v &= 0, \end{aligned} \tag{9.2}$$

with boundary conditions:

$$\begin{aligned} \frac{\partial u}{\partial \mathbf{n}} &= 0 \\ \frac{\partial v}{\partial \mathbf{n}} &= 0 \end{aligned} \tag{9.3}$$

where ∇^2 designates the Laplacian and $\frac{\partial}{\partial \mathbf{n}}$ indicates differentiation in the direction of the normal $\mathbf{n}$ to the image domain boundary $\partial\Omega$.

Let D be a regular, unit-spacing grid over Ω, the grid points indexed top-down and left-to-right by the integers $\{1, 2, ..., N\}$. A discrete approximation of the Euler-Lagrange equations (9.2) is, for all grid point indices $i \in \{1, 2, ..., N\}$:

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5_9,

$$\begin{aligned} I_{xi}^2 u_i + I_{xi} I_{yi} v_i + I_{xi} I_{ti} - K \sum_{j \in \mathscr{N}_i} (u_j - u_i) = 0 \\ I_{yi} I_{xi} u_i + I_{yi}^2 v_i + I_{yi} I_{ti} - K \sum_{j \in \mathscr{N}_i} (v_j - v_i) = 0 \end{aligned} \tag{9.4}$$

where $(u_i, v_i) = (u, v)_i$ is the optical flow vector at grid point i; I_{xi}, I_{yi}, I_{ti} are the spatial and temporal derivatives I_x, I_y, I_t evaluated at i; K is proportional to λ; and $\mathscr{N}_i$ is the set of indices of the neighbors of i for some neighborhood system. For the 8-neighborhood, $card(\mathscr{N}_i) < 8$ for points on the boundary of D and $card(\mathscr{N}_i) = 8$ otherwise.

In (9.2), the approximation of the Laplacian of z (where z stands for either u or v) is proportional to $\bar{z} - z$, where $\bar{z}$ is the neighborhood average of z. The coefficient of proportionality is absorbed by K (which replaces the original weighting coefficient λ).

Rewriting (9.4), we have the following system of linear equations, for $i \in \{1, \ldots, N\}$:

$$(S) \left\{ \begin{array}{l} (I_{xi}^2 + Kc_i) u_i + I_{xi} I_{yi} v_i - K \sum_{j \in \mathscr{N}_i} u_j = -I_{xi} I_{ti} \\ \\ I_{xi} I_{yi} u_i + (I_{yi}^2 + Kc_i) v_i - K \sum_{j \in \mathscr{N}_i} v_j = -I_{yi} I_{ti} \end{array} \right.$$

where $c_i = card(\mathscr{N}_i)$. Let $\mathbf{z} = (z_1, \ldots, z_{2N})^t \in \mathbb{R}^{2N}$ be the vector defined by

$$\begin{aligned} \forall i \in \{1, \ldots, N\} \quad z_{2i-1} &= u_i \\ z_{2i} &= v_i \end{aligned} \tag{9.5}$$

Also, let $\mathbf{b} = (b_1, \ldots, b_{2N})^t \in \mathbb{R}^{2N}$ the vector defined by

$$\begin{aligned} \forall i \in \{1, \ldots, N\} \quad b_{2i-1} &= -I_{xi} I_{ti} \\ b_{2i} &= -I_{yi} I_{ti} \end{aligned} \tag{9.6}$$

The system of linear equations (S) is written in matrix form as

$$\mathbf{Az} = \mathbf{b} \tag{9.7}$$

where $\mathbf{A}$ is the $2N \times 2N$ matrix defined by

$$\begin{aligned} \forall i \in \{1, \ldots, N\} \quad \mathbf{A}_{2i-1,2i-1} &= I_{xi}^2 + Kc_i \\ \mathbf{A}_{2i,2i} &= I_{yi}^2 + Kc_i \\ \mathbf{A}_{2i-1,2i} &= \mathbf{A}_{2i,2i-1} = I_{xi} I_{yi} \\ \forall i, j \in \{1, \ldots, N\} \text{ with } j \in \mathscr{N}_i \quad \mathbf{A}_{2i-1,2j-1} &= \mathbf{A}_{2i,2j} = -K, \end{aligned} \tag{9.8}$$

and all other elements are zero.

System (9.7) is a large scale system of linear equations. It is also a sparse system as most entries in matrix $\mathbf{A}$ are zero. Such systems are best solved by iterative algorithms [4, 5]. In the following, we will assume that $\mathbf{A}$ is non-singular.

9.1.1 Iterative resolution by the Jacobi and Gauss-Seidel iterations

Horn and Schunck [1] solve system (9.7) using the Jacobi method for a 2 × 2 block division of matrix **A** [6].The iterations are:

$$
\begin{aligned}
u_i^{k+1} &= \frac{I_{yi}^2 + Kc_i}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \sum_{j \in \mathcal{N}_i} u_j^k - \frac{I_{xi}I_{yi}}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \sum_{j \in \mathcal{N}_i} v_j^k - \frac{I_{xi}I_{ti}}{I_{xi}^2 + I_{yi}^2 + Kc_i} \\
v_i^{k+1} &= \frac{-I_{xi}I_{yi}}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \sum_{j \in \mathcal{N}_i} u_j^k + \frac{I_{xi}^2 + Kc_i}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \sum_{j \in \mathcal{N}_i} v_j^k - \frac{I_{yi}I_{ti}}{I_{xi}^2 + I_{yi}^2 + Kc_i}
\end{aligned}
$$

With the Jacobi method,the update is done for all points of the image domain grid before the updated values are used at the next iteration. With the Gauss-Seidel method, the updated values are used as soon as they are available. For the 2 × 2 block division of matrix **A**, the Gauss-Seidel iterations are [6], for all $i \in \{1, \ldots, N\}$:

$$
\begin{aligned}
u_i^{k+1} = {} & \frac{I_{yi}^2 + Kc_i}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \left(\sum_{j \in \mathcal{N}_i; j<i} u_j^{k+1} + \sum_{j \in \mathcal{N}_i; j>i} u_j^k \right) \\
& - \frac{I_{xi}I_{yi}}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \left(\sum_{j \in \mathcal{N}_i; j<i} v_j^{k+1} + \sum_{j \in \mathcal{N}_i; j>i} v_j^k \right) - \frac{I_{xi}I_{ti}}{I_{xi}^2 + I_{yi}^2 + Kc_i}
\end{aligned} \tag{9.9}
$$

$$
\begin{aligned}
v_i^{k+1} = {} & \frac{-I_{xi}I_{yi}}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \left(\sum_{j \in \mathcal{N}_i; j<i} u_j^{k+1} + \sum_{j \in \mathcal{N}_i; j>i} u_j^k \right) \\
& + \frac{I_{xi}^2 + Kc_i}{c_i(I_{xi}^2 + I_{yi}^2) + Kc_i^2} \left(\sum_{j \in \mathcal{N}_i; j<i} v_j^{k+1} + \sum_{j \in \mathcal{N}_i; j>i} v_j^k \right) - \frac{I_{yi}I_{ti}}{I_{xi}^2 + I_{yi}^2 + Kc_i}
\end{aligned} \tag{9.10}
$$

With the Jacobi iterations (9.9), (u_i^{k+1}, v_i^{k+1}), $i = 1, ..., 2N$ at iteration $k+1$ are computed using the values (u_i^k, v_i^k), $i = 1, ..., 2N$ of the preceding iteration k. Therefore, $4N$ values must be in memory store. With the Gauss-Seidel iterations (9.10), only $2N$ values need storage. However, in contrast with the Gauss-Seidel iterations, Jacobi iterations can be performed in parallel for all domain grid points, which would be a very fast implementation.

One can show that matrix **A** is positive definite [6]. This implies that the Gauss-Seidel iterations converge [4, 5]. It does not imply that Jacobi iterations converge. However, convergence can be shown directly [6].

There is a remarkable block division which makes matrix **A** block tridiagonal [6]. Combined with the property that **A** is symmetric positive definite, this characteristic affords efficient resolution of the corresponding linear system [4]. For an $n \times n$ discrete image, the blocks are $2n \times 2n$. The block tridiagonal form is due to the fact that points with index αn, $1 \leq \alpha \leq n$, do not have a neighbor on the right, and those with index $\beta n + 1$, $0 \leq \beta \leq n - 1$, do not have a neighbor on the left. The readers interested in the details can refer to [6].

9.1.2 Evaluation of derivatives

Horn and Schunck prescribe approximations of the image spatial and temporal derivatives as averages of forward first differences. It is easier to read the formulas by looking at D as a two-dimensional grid indexed *left to right* horizontally and *bottom to top* vertically. Let the horizontal and vertical indices be r and c, respectively. Also, let the two images used be indexed 0 (current) and 1 (next). The Horn and Schunck approximations are:

$$
\begin{aligned}
I_x(r,c) &\approx \tfrac{1}{4}\textstyle\sum_{\Delta r=0}^{1} && \{\, I(r+\Delta r, c+1, 0) - I(r+\Delta r, c, 0) \\
& && +I(r+\Delta r, c+1, 1) - I(r+\Delta r, c, 1) \,\} \\
I_y(r,c) &\approx \tfrac{1}{4}\textstyle\sum_{\Delta c=0}^{1} && \{\, I(r+1, c+\Delta c, 0) - I(r, c+\Delta c, 0) \\
& && +I(r+1, c+\Delta c, 1) - I(r, c+\Delta c, 1) \,\} \\
I_t(r,c) &\approx \tfrac{1}{4}\textstyle\sum_{\Delta r=0}^{1}\sum_{\Delta c=0}^{1} && \{\, I(r+\Delta r, c+\Delta c, 1) - I(r+\Delta r, c+\Delta c, 0) \,\}
\end{aligned}
\tag{9.11}
$$

Alternatively, the spatial derivatives can be estimated using the following averages of central differences:

$$
\begin{aligned}
I_x(r,c) &= \tfrac{1}{6}\textstyle\sum_{\Delta c=-1}^{1}\{\, I(r+1, c+\Delta c) - I(r-1, c+\Delta c) \\
I_t(r,c) &= \tfrac{1}{6}\textstyle\sum_{\Delta r=-1}^{1}\{\, I(r+\Delta r, c+1) - I(r+\Delta r, c-1)
\end{aligned}
\tag{9.12}
$$

Using central differences to compute the temporal derivatives would not be consistent with the in-between consecutive frames velocities to be estimated because it would require using the frames preceding and following the current, rather than consecutive frames.

9.2 The Aubert, Deriche, and Kornprobst algorithm

The Laplacian operator which appears in the Euler-Lagrange equations associated with (9.1) causes smoothing, and blurring thereof, across motion boundaries. To circumvent the problem, the study in [2, 3] investigated the following generalization of the Horn and Schunck functional:

$$
\mathscr{E}(u,v) = \int_\Omega (I_x u + I_y v + I_t)^2 dxdy + K \int_\Omega (g(\|\nabla u\|) + g(\|\nabla v\|))dxdy \tag{9.13}
$$

where g is a function of class C^2. With $g(z) = z^2$, (9.13) reduces to the Horn and Schunck functional (9.1).

The goal of the analysis in [2, 3] was to determine functions g from conditions that would ensure motion boundary preserving anisotropic smoothing of motion. This led to functions such as $g(s) = 2\sqrt{1+s^2} - 2$. A discretization of the the Euler-Lagrange equations corresponding to (9.13),

$$\begin{aligned} I_x(I_x u + I_y v + I_t) &= \frac{K}{2}\mathrm{div}\left(g'(\|\nabla u\|)\frac{\nabla u}{\|\nabla u\|}\right) \\ I_y(I_x u + I_y v + I_t) &= \frac{K}{2}\mathrm{div}\left(g'(\|\nabla v\|)\frac{\nabla v}{\|\nabla v\|}\right), \end{aligned} \tag{9.14}$$

gives a large scale sparse system of nonlinear equations. Instead of solving such a system, the study in [2][3] proposes a more efficient implementation using the half-quadratic minimization algorithm applied to the following functional, the change from the original functional justified under a duality theorem:

$$\begin{aligned} \mathscr{E}(u,v,b_1,b_2) = &\int_\Omega (I_x u + I_y v + I_t)^2 dxdy \\ &+ K\int_\Omega \left(b_1\|\nabla u\|^2 + b_2\|\nabla v\|^2 + \psi(b_1) + \psi(b_2)\right) dxdy \end{aligned} \tag{9.15}$$

Two new functions, $b_1(x,y)$ and $b_2(x,y)$, called auxiliary variables, appear in this functional. Also appears a function ψ, convex and decreasing, related implicitly to g and such that the value of b which minimizes $bs^2 + \psi(b)$ is given by

$$b = \frac{g'(s)}{2s} \tag{9.16}$$

This result is at the heart of the half-quadratic minimization algorithm which, following initialization, iterates two consecutive steps. At each iteration, a minimization with respect to u,v, with b_1,b_2 considered constant, is followed by a minimization with respect to b_1,b_2, with u,v considered constant. The minimization with respect to b_1,b_2 is obtained analytically using (9.16).

Minimization with respect to u,v, with b_1,b_2 considered constant, consists of minimizing the following functional:

$$\int_\Omega (I_x u + I_y v + I_t)^2 + K\left(b_1\|\nabla u\|^2 + b_2\|\nabla v\|^2\right) dxdy \tag{9.17}$$

The Euler-Lagrange equations corresponding to (9.17) are given by:

$$\begin{aligned} I_x(I_x u + I_y v + I_t) &= K\mathrm{div}(b_1\nabla u) \\ I_y(I_x u + I_y v + I_t) &= K\mathrm{div}(b_2\nabla v) \end{aligned} \tag{9.18}$$

The corresponding Neumann boundary conditions are $\partial u/\partial\mathbf{n} = \partial v/\partial\mathbf{n} = 0$. Discretization of the equations results in a large scale sparse system of linear equations which can be solved iteratively by the Jacobi or Gauss-Seidel method as with the Horn and Schunck method [7]. The divergence terms in 9.18 can be discretized as in [8].

Minimization with respect to b_1,b_2, with u,v considered constant, consists of minimizing the functional

$$\int_{\Omega} \left(b_1\|\nabla u\|^2 + b_2\|\nabla v\|^2 + \psi(b_1) + \psi(b_2)\right) dxdy \tag{9.19}$$

The unique solution is given analytically following (9.17):

$$\begin{aligned} b_1 &= \frac{g'(\|\nabla u\|)}{2\|\nabla u\|} \\ b_2 &= \frac{g'(\|\nabla v\|)}{2\|\nabla v\|} \end{aligned} \tag{9.20}$$

The semi-quadratic algorithm to minimize (9.15) can be summarized as follows.

1. Initialize b_1, b_2
2. Repeat until convergence
   a) Minimize with respect to u, v using Jacobi (or Gauss-Seidel) iterations to solve the linear system of equations corresponding to the discretized Eq. (9.18).
   b) Minimize with respect to b_1, b_2 using (9.20) [$b_1 = \frac{g'(\|\nabla u\|)}{2\|\nabla u\|}$, $b_2 = \frac{g'(\|\nabla v\|)}{\|\nabla v\|}$]

The readers interested can find the details in [2, 3].

9.3 Construction of stereoscopic images of a computed 3D interpretation

We can evaluate subjectively a monocular image sequence 3D interpretation by viewing anaglyphs of stereoscopic images constructed using the recovered depth. Anaglyphs use two color channels to code depth on the original image so that viewing with chromatic glasses gives the impression of depth. Common, inexpensive commercial plastic glasses are available. Printed anaglyphs are best viewed on high quality photographic paper. When viewing on a CRT screen, high resolution and display options for high quality color image rendering offer a clearer impression of depth. In all cases, however, anaglyphs offer a good, inexpensive means of viewing 3D interpretation results.

Given an image and the corresponding depth map, we can construct a stereoscopic image using the following rudimentary scheme. Let I_1 be the given image. I_1 will be one of the two images of the stereoscopic pair. We will construct the other image, I_2. Let S_1 be the viewing system representing the camera which acquired I_1, and S_2 that of the (fictitious) camera acquiring I_2. S_2 is placed to differ from S_1 by a translation of amount d along the X-axis.

Let (x_2, y_2) be a point on the image position array of I_2, corresponding to a point **P** in space with coordinates (X_2, Y_2, Z_2) in S_2. The coordinates of **P** in S_1 are

$$\begin{aligned} X_1 &= X_2 + d \\ Y_1 &= Y_2 \\ Z_2 &= Z_1 \end{aligned} \tag{9.21}$$

The image of **P** in the image domain of I_1 are, according to our viewing system model (Figure 2.3):

$$x_1 = f\frac{X_1}{Z_1} \tag{9.22}$$

$$y_1 = y_2 \tag{9.23}$$

Because depth has been estimated, coordinates (x_1, y_1) are known. Image I_2, which will be the second of the stereoscopic pair, is then constructed as follows:

$$I_2(x_2, y_2) = I_1(\tilde{x}_1, y_1) \tag{9.24}$$

where $\tilde{x}_1$ is the x-coordinate of the point on the image positional array of I_1 with x coordinate closest to x_1. Alternatively, one can use interpolation. However, we found it unnecessary for our purpose here. Given the stereoscopic image pair (I_1, I_2), an anaglyph can be generated using [9]. Programs of other methods are freely accessible on the web.

References

1. B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 17, pp. 185–203, 1981.
2. R. Deriche, P. Kornprobst, and G. Aubert, “Optical-flow estimation while preserving its discontinuities: A variational approach,” in *Asian Conference on Computer Vision (ACCV)*, 1995, pp. 71–80.
3. G. Aubert, R. Deriche, and P. Kornprobst, “Computing optical flow via variational techniques,” *SIAM Journal of Applied Mathematics*, vol. 60, no. 1, pp. 156–182, 1999.
4. P. Ciarlet, *Introduction à l’analyse numérique matricielle et à l’optimisation*. Masson, Fifth, 1994.
5. J. Stoer and P.Burlisch, *Introduction to numerical methods*, 2nd ed. Springer, 1993.
6. A. Mitiche and A. Mansouri, “On convergence of the Horn and Schunck optical flow estimation method,” *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 848–852, 2004.
7. G. Aubert and P. Kornpbrost, *Mathematical problems in image processing: Partial differential equations and the calculus of variations*. Springer Verlag, 2006.
8. P. Perona and J. Malik, “Scale space and edge detection using anisotropic diffusion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1981.
9. E. Dubois, “A projection method to generate anaglyph stereo images,” in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, 2001, pp. 1661–1664.

Index

A. Mitiche, I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*,
Springer Topics in Signal Processing, DOI 10.1007/978-3-642-15352-5,