Chapter 2
Deformable Model-Based Medical Image Segmentation

Gavriil Tsechpenakis

Abstract Image data is of immense practical importance in medical informatics. Automated image segmentation, which aims at automated extraction of region boundary features, plays a fundamental role in understanding image content for searching and mining in medical image archives. A challenging problem is to segment regions with boundary insufficiencies, i.e., missing edges and/or lack of texture contrast between regions of interest (ROIs) and background. To address this problem, several segmentation approaches have been proposed in the literature, with many of them providing rather promising results.

In this chapter, we focus on a specific category of image segmentation methods widely used in medical vision, namely the deformable models. We first review two general classes of deformable models, i.e., (1) the parametric deformable models, or active contours, and (2) the geometric or implicit models. Then we describe the feature extraction, i.e., the estimation of image features based on which the segmentation is performed. We show the most common approaches of how the image data is transformed into compact (higher level) numerical representations, which are integrated into the deformable models to play the image-based driving factor for the segmentation. Since these features can be used in a deterministic or probabilistic manner, we describe the basic principles of transforming these numerical representations into probabilities, using two known probabilistic graphical models, namely the Markov Random Fields (MRFs) and the Conditional Random Fields (CRFs). Then we show how these machine learning methods are integrated with deformable models to yield robust segmentation results. To illustrate the latter case, we describe a geometric model, which is integrated with a CRF: the deformable model is driven by probability fields estimated from the images (features), rather than being driven by the image features directly, with the main advantage being the increased robustness in cases of feature ambiguities, i.e., noise. We show different

G. Tsechpenakis (✉)
Department of Computer and Information Science, Indiana University-Purdue University Indianapolis, 723 W. Michigan St, Indianapolis, IN 46202-5132, USA
e-mail: gavriil@cs.iupui.edu

El-Baz, Acharya, Mirmehdi and Suri, *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies: Volume 1*,
DOI 10.1007/978-1-4419-8195-0_2,

examples of medical data deformable model-based segmentation, we draw general conclusions from the methods described in this chapter, and we give future directions for solving challenging and open problems in medical image segmentation.

2.1 Introduction

An important goal of medical image processing is to transform raw images into a numerically symbolic form for better representation, evaluation, and/or content-based search and mining. An essential step in this transformation is the segmentation of the target structures; that is, based on given homogeneity criteria, the task is the image partitioning into regions, which, in medical images, are usually the target anatomic regions (foreground) and their surroundings (background). After this segmentation, the exact shape and appearance features of the targets can be calculated, and, based on the application, they can be used for clinical evaluation, pattern analysis, and/or knowledge discovery. A challenging problem is to segment regions with boundary insufficiencies, i.e., missing edges and/or lack of texture contrast between regions of interest (ROIs) and background.

Here, we focus on a specific category of segmentation methods, namely the deformable models. The main reason why these approaches have been widely used in medical image computing is their robustness, primarily due to the models' constraints, as we explain throughout this chapter. We aim at giving the reader an intuitive but also mathematical description of these model-based methods, we describe their implementation aspects, and then we detail on advanced methods that integrate shape and appearance models for robust yet accurate medical image segmentation.

Deformable models are curves or surfaces, for segmentation in the image domain, or hyper-surfaces, for the segmentation of higher dimensional image data, such as stacks of images, which deform under the influence of internal and external forces to delineate object boundary. The internal forces are defined such that they preserve the shape smoothness of the model, while the external forces are defined by the image features to drive the model toward the desired position/configuration, i.e., to the desired region boundaries. Usually, the core model definition is independent from the features used in the external force terms; in most cases, the image features are application-related, i.e., their choice depends on the image modality. For instance, in ultra-sound images, different regions are determined by region contrast, in terms of the intensity speckle density/distribution, while the edge information, in its definition as the image gradient, is usually too poor to be used. On the other hand, magnetic resonance (MR) and computerized tomography (CT) images have sufficient gradients for edge information to be used in segmentation.

Deformable models are classified into two general approaches, the parametric and the geometric models, depending on how the model is defined in the shape domain.

Intuitively, parametric models, widely known as active contours for the segmentation in the two-dimensional image domain, are curves whose deformations are determined by the displacement of a discrete number of control points along the curve. Apart from active contours, parametric models can be also surfaces, with the control points defining two-dimensional (in the shape domain) deformable grids, for two-dimensional image segmentation, or hyper-surfaces, with the control points defining three-dimensional, intraconnected, clouds of points, for the segmentation of higher-dimensional image data (e.g., image stacks).

The main advantage of parametric models is that they are usually very fast in their convergence, depending on the predetermined number of control points. However, an obvious weakness of these models is that they are topology dependent: a model can only capture a single ROI, and therefore, in images with multiple ROIs, multiple models have to be initialized, one for each ROI.

The second class of deformable models, namely the geometric models, use a distance transformation to define the shape from the n-dimensional to an $n + 1$-dimensional domain, where $n = 1$ for curves, $n = 2$ for surfaces on the image plane, etc. As we explain below, there are three main advantages of such transformation. First, the shape can be defined in a domain with dimensionality similar to the dataset space (for example, for 2D segmentation, a curve is transformed into a 2D surface), which can provide a more mathematically straightforward integration of shape and appearance (image features) in the model definition. Second, the shape can be implicitly defined, with the control/deformation points being at the image pixels' positions. Finally, some methods that we review below and that use such representations are topology independent, i.e., they can capture multiple ROIs with a single model, and therefore they can be robust to initializations.

In the following two sections, we review some of the existing literature on both categories of deformable models, and we explain in detail the mathematical definitions of the most representative models. For simplicity, we will consider the two-dimensional segmentation problem.

2.2 Active Contours

Parametric deformable models or active contours [8, 17, 23, 39] were first introduced in 1988, by Kass et al., under the name "snakes," in a paper-milestone for the computer vision and computer graphics communities [17]. Active contours, or snakes, use parametric curves to represent the model shape, and, during their evolution, the deformations are determined by geometry, kinematics, dynamics, as well as other constraints (if available), such as material properties. Mathematically, active contours are splines, whose state (position and dynamics) is determined by an energy, and their evolution is an energy minimization problem.

Specifically, let us consider the image domain Ω, and a curve C defined by a set of ordered points $p = (x, y) \in \Omega$, where (x, y) are Cartesian coordinates on the image plane. If s represents the parametric domain, i.e., s is a normalized length that

parameterizes the curve, then we can define the curve as $s \rightarrow C(s) = \{x(s), y(s)\}$. This curve is the model that undergoes deformations, based on the states/energies of the individual points $\{x(s), y(s)\}$.

Internal Energy of Active Contours: The internal energy of an active contour can be translated as the summation of forces applied along the curve to preserve its smoothness. Intuitively, C can be seen as a rubber band that can extend and shrink, but it has some deformation limits: excessive stress at a point increases locally its temperature/energy, and when the stress exceeds a limit, the band breaks. Therefore, the internal energy of the rubber band/model determines how smooth the model is locally, and the evolution is toward globally minimizing the local stress, i.e., minimizing the curve's energy, or respectively, forcing the curve to be locally smooth.

Mathematically, and keeping in mind the above example, the internal energy can be described as [17]:

$$E_{\text{int}}(C) = -\int_0^1 e_{\text{int}}(C(s))\,\mathrm{d}s, \tag{2.1}$$

where the individual energies e_{int} represent the local state along the curve, and are defined as,

$$e_{\text{int}}(C(s)) = \alpha(s)\left|\frac{\partial C}{\partial s}\right|_s + \beta(s)\left|\frac{\partial^2 C}{\partial s^2}\right|_s \tag{2.2}$$

The first derivative of C corresponds to the first-order smoothness of the curve, i.e., determines the length (distances between successive points on the curve), while the second derivative determines smoothness in the direction normal to the curve. The parameters α and β regulate the relative importance of the two smoothness terms.

The minimization of the energy in (2.1) and (2.2) forces the curve toward the minimum length and maximum smoothness; intuitively, initializing the curve in any configuration/shape, the minimization of this energy, after some iterations, would yield a shrinking circle.

External Energy of Active Contours: Common active contours use primarily edge (image gradient) information to derive external image forces that drive a shape-based model. In parametric models, a typical formulation for the energy term deriving the external image forces is [17],

$$E_{\text{ext}}(C) = -\int_0^1 |\nabla \hat{I}(C(s))|^2\,\mathrm{d}s, \tag{2.3}$$

where $\hat{I} = G_\sigma * I$ is the image I after smoothing with a Gaussian kernel of standard deviation σ, and $\nabla \hat{I}(C)$ is the image gradient along the curve C. Basically, by minimizing this energy term, the accumulative image gradient along the curve is maximized, which means that the parametric model is attracted by strong edges that correspond to pixels with local-maxima image gradient values.

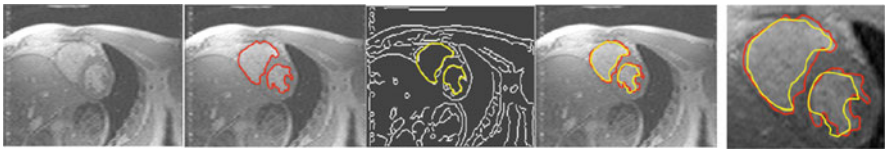

Fig. 2.1 Edge-based active contour [17] for the segmentation of the *left and right ventricles* (LV, RV) in a cardiac MRI. From *left* two *right*: original image, ground-truth boundaries, edge map (Canny edge detection) with the final model solution superimposed (in *yellow*), ground-truth (*red*) and final solution (*yellow*), magnified view of LV and RV along with the ground-truth (*red*) and the final solution (*yellow*)

The dependence on image gradient information, however, makes deformable models sensitive to noise and spurious edges so that they often need to be initialized close to the boundary to avoid getting stuck in local minima.

Figure 2.1 illustrates an example of medical image segmentation using an active contour. The task here is to segment the left and right ventricle (LV and RV) regions in an MR cardiac image. The leftmost image shows the original grayscale image, while the second from left image shows the *ground-truth*, i.e., the actual boundaries of RV and LV, with two red closed lines; these boundaries were obtained by manual segmentation. The third (from left) image shows the image edges obtained with the Canny edge detector; the yellow closed contours superimposed on the edge image show the segmentation result of the deformable model (initialized around RV and LV), which, in this case, uses edge information as external image forces [17]. In the next image, we show both the ground-truth (in red) and the estimated (in yellow) boundaries. In the rightmost image, which shows a magnification of the ventricles, one can observe that the deformable model converges to edges that do not correspond to the actual region boundaries, which is caused by a local minimum of the model's energy.

To overcome the above problem, other parametric methods have been proposed that use region-based external energies or even combinations of both edge- and region-based terms [5, 16, 28, 40]. These methods assume that the parametric model C partitions the image into a ROI and a background region, each one of which has different statistics, often approximated with Gaussian distributions, and this difference locally drives the model evolution. Intuitively, the statistical models for the ROI and the background, I_{ROI} and I_{b} respectively, are compared with the image I, aiming at the minimization of the mean square error [28],

$$E_{\text{MSE}}(\mathbf{x}) = \int \int_{\text{ROI}} \|I_{\text{ROI}}(\mathbf{x}) - I(\mathbf{x})\|^2 \, d\mathbf{x} + \int \int_{\text{b}} \|I_{\text{b}}(\mathbf{x}) - I(\mathbf{x})\|^2 \, d\mathbf{x}, \qquad (2.4)$$

where $\mathbf{x} = (x, y)$, i.e., it represents the location of each pixel in the image domain, in Cartesian coordinates.

In [40], a generalized energy function that combines aspects of active contours and region growing is proposed and the minimization of the objective function is guaranteed to converge to a local minimum. However, this formulation does not

address the problem of unifying shape and appearance, because of the large difference in representation for shape and appearance. While the model shape is represented using a parametric spline curve, the region intensity statistics are captured by parameters of a Gaussian distribution. This representation difference prevents the use of gradient descent methods to update both region parameters and shape parameters in a unified optimization process, so that the two sets of parameters are estimated in separate steps, and the overall energy function is minimized in an iterative way [40]. In other hybrid segmentation frameworks [5, 16], a region-based module is used to get a rough binary mask of the ROI. Then this rough boundary estimation serves as initialization for a deformable model, which deforms to fit edge features in the image using gradient information.

Figure 2.2 illustrates an example of segmentation with a region-based active contour, in a mouse retinal image acquired with Spectral Domain Optical Coherence Tomography (SDOCT). In the first image (A), the slightly brighter region highlighted with the yellow box, corresponds to the ROI, which is a retina tumor to be quantified. One can notice the obvious lack of edges in this kind of images, where different regions can only be determined with the difference in the speckle density (region information).

In this example, the active contour is a probabilistic model that uses intensity distribution as image feature, and its dynamics are defined in a probabilistic framework [29]. Despite the robustness that probabilistic models offer, in all

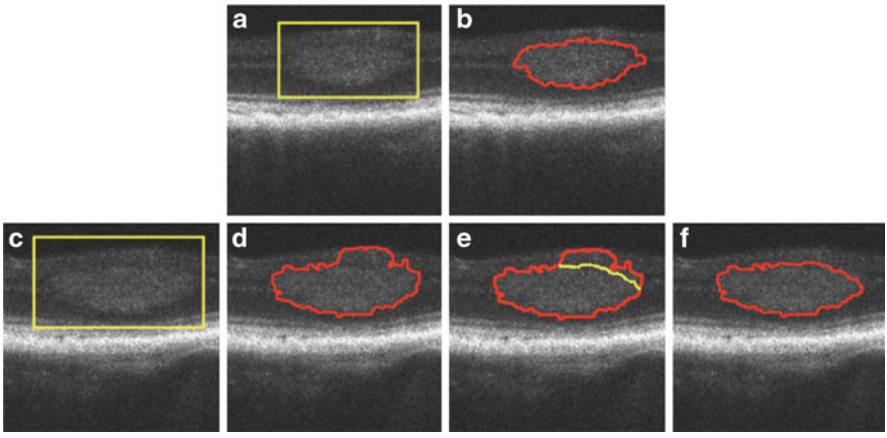

Fig. 2.2 A region-based active contour example for the segmentation of retinal tumors in the mouse model from spectral domain optical coherence tomography (SDOCT) [29]. (**a**) The region of the tumor is selected and highlighted by a *yellow box*; (**b**) the result of the segmentation is displayed and no corrections to the estimated boundary were needed. (**c**–**f**) Example when the tumor was not segmented accurately. (**c**) The tumor region is selected with the *yellow box* on the OCT image; (**d**) the estimated boundary in the upper right region of the tumor is not accurate due to missing contrast between the tumor and the surrounding region; (**e**) boundary correction (*yellow line*), by using a shape prior with high weight in the model's energy; (**f**) the final result after the weight adjustment

existing active contours, the optimization (energy minimization in deterministic models, or maximum a posteriori probability (MAP) estimation in probabilistic models) depends on open parameters that regulate the relative significance between internal and external energy terms, between different terms within the internal energy (as we saw in (2.2)), or even between more energy terms, in models where shape priors, material properties, etc. are considered.

In the model used in this example, due to the anticipated shape of the tumor (ellipsoid-like), a shape prior was also considered in the model's energy. The image (b) in Fig. 2.2 illustrates, in red, the region-based active contour after its convergence; one can observe that the model's solution captures the desired boundaries, however, this was achieved after appropriate tuning of the shape prior and internal energy weights. Panel (c) in the same figure shows another case of retinal tumor, highlighted with a yellow box, while panel (d) shows the model's result. One may observe that part of the resulting contour does not correspond to the actual tumor boundary, mainly due to the lack of contrast between tumor and surroundings. The image in (e) shows how fine tuning of the open parameters can refine the result: the yellow line is the corrected part of the model, after setting a higher weight for the shape prior. Panel (f) shows the final result after this correction.

A very good survey on probabilistic deformable models in medical image analysis can be found in [23].

2.3 Geometric Models

The second class of deformable models is the *geometric* or *implicit* models [2, 21, 24, 25, 30, 38, 42], which use the level-set based shape representation, transforming the curves into higher dimensional scalar functions, as shown in Fig. 2.3: a 1D closed curve, the evolving front, is transformed into a 2D surface, using the scalar distance function that is mathematically defined as follows [33].

The model's shape is embedded in a higher dimensional space of distance transforms, such that the zero-level of a scalar (Euclidean distance) function corresponds to the evolving curve. The interface defines two regions in the image domain Ω, namely the region R_C enclosed by the curve C and the background $\Omega \backslash R_C$. The model shape is represented implicitly by its distance transform,

$$\Phi_C(\mathrm{x}) = \begin{cases} 0, \mathbf{x} \in C \\ +\min_{\mathrm{x}_C \in C} \|\mathbf{x} - \mathbf{x}_\mathrm{C}\|, \mathbf{x} \in R_C \\ -\min_{\mathrm{x}_C \in C} \|\mathbf{x} - \mathbf{x}_C\|, \mathbf{x} \in \Omega \backslash R_C \end{cases}, \tag{2.5}$$

where $\mathbf{x} = (x, y)$ is the image pixel location in Cartesian coordinates. This representation transforms the model's shape into a distance image Φ_C. There are two main advantages for using this definition. First, the integration of the shape with the appearance (feature) based information becomes straightforward, since both the

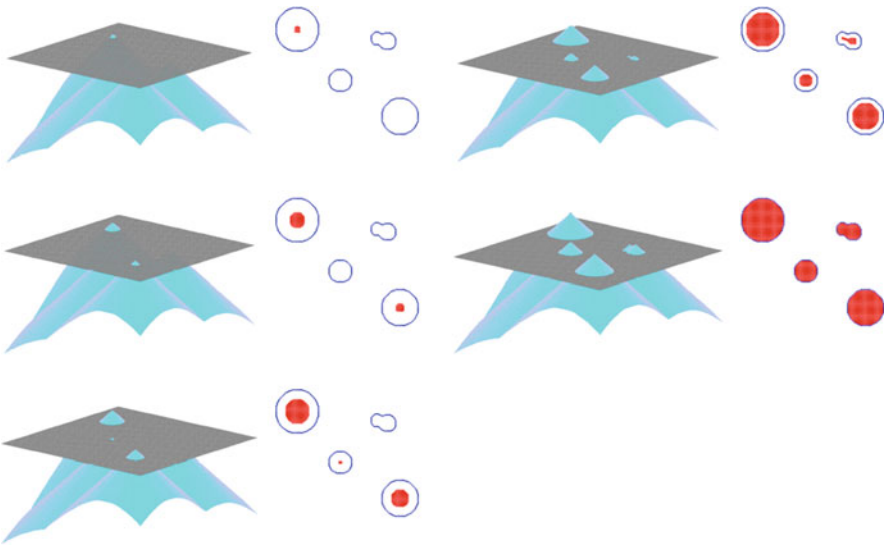

Fig. 2.3 Implicit shape representation using the distance transform. The surfaces are the distance functions, while the *gray planes* indicate the zero level (zero distance). The *blue circles* represent the desired boundaries while the *red disks* correspond to the positive distance values from the current contour in each iteration

image and the shape have the same dimensionality. Second, the shape distance function satisfies the C^1 continuity constraint (first order smoothness): the gradient of Φ_C is defined for all points in the image domain, and its direction is normal to C (i.e., where $\Phi_C = 0$).

In edge-based geometric models, a typical objective function that drives the front propagation of a level set (distance) function is [2]:

$$E(C) = \int_0^1 g(|\nabla \hat{I}(C(s))|)|C'(s)|\mathrm{d}s, \quad \text{where } g(|\nabla \hat{I}|) = \frac{1}{1+|\nabla \hat{I}|^2}. \tag{2.6}$$

Here, C represents the front (i.e., zero level set) curve of the evolving level set function. To minimize the objective function, the front curve deforms along its normal direction $C''(s)$, and its speed is controlled by the speed function $g(|\nabla \hat{I}|)$. The speed function definition, $g(|\nabla \hat{I}|)$, depends on image gradient $\nabla \hat{I}$, and it is positive in homogeneous areas and zero at ideal edges. Hence the curve moves at a velocity proportional to its curvature in homogeneous regions and stops at strong edges. Similarly to active contours, the reliance on image gradient information makes geometric models sensitive to noise and spurious edges, in a sense that, e.g., they may leak through boundary gaps or generate small holes/islands.

Using the above shape representation, and to avoid the limitations of edge-based models, in the Mumford–Shah model [24], the objective function E_{ms} to be

minimized consists of implicit terms for edge and region-driven forces, and the boundary smoothness and length constraints,

$$E_{\text{ms}}(f,\Gamma) = \mu \int \int_{\Omega} \left[f(\mathbf{x}) - I(\mathbf{x})\right]^2 \mathrm{d}\mathbf{x} + \lambda \int \int_{\Omega\setminus\Gamma} ||\nabla f(\mathbf{x})||^2 \mathrm{d}\mathbf{x} + \nu|\Gamma|, \qquad (2.7)$$

where f is the data approximation, i.e., f approximates the input image I, Γ labels the discontinuities in the image domain Ω, i.e., it is the desired boundary approximation, and $\{\mu, \lambda, \nu\}$ are cost weights. According to this objective function, the desired boundary should be the one with the shortest arc-length (third term), it should be the interface between regions that are piecewise smooth (second term), and these regions should fit the input observations (first term).

In the Mumford–Shah model above, one may observe the following. First, it is a region-based model, which becomes apparent from the approximation function f. Then, the integration between the model and the image features comes as a natural consequence from the fact that both the shape and the data (image) domain have same dimensionality. The latter makes the front propagation a two-dimensional fitting problem with shape constraints. Finally, since the optimal boundary estimation problem is transformed into an optimization over the entire image domain, the model is topology independent. Here, we will not discuss the solution of this optimization framework, but we refer the reader to the original work in [24].

Although similar frameworks nicely incorporate gradient and region criteria into a single energy function, no practical globally-optimal solution for the function is available, most notably because of the mathematical difficulties documented, e.g., in [24].

In the recent few years, progress has been made and solutions for several reduced cases of the Mumford–Shah functional have been proposed in the level set framework. For example, the method in [41] is able to segment images that consist of several regions, each one of which is modeled by given statistics, such as the mean intensity and variance. Nevertheless, the algorithm requires a priori knowledge of the number of segments in the image and its performance depends upon the discriminating power of the chosen set of statistics.

Another approach in [30] applies a multiphase level set representation to segmentation assuming piecewise constant intensity within one region. It can be considered as solving a classification problem because it assumes that the mean intensities of all region classes are known a priori, and only the set of boundaries between regions is unknown. In the methods presented in [3, 38], piecewise constant and piecewise smooth approximations of the Mumford–Shah functional are derived for two-phase (i.e., two regions) [3] or multiphase (i.e., multiple regions) [38] cases in a variational level set framework. The optimization of the framework is based on an iterative algorithm that approximates the region mean intensities and level-set shape in separate steps.

Geodesic Active Regions [25] is another method that integrates edge and region-based modules in a level set framework. The algorithm consists of two stages: a modeling stage that constructs a likelihood map of edge pixels and approximates

region/class statistics using Mixture-of-Gaussian components, and a segmentation stage that uses level set techniques to solve for a set of smooth curves that are attracted to edge pixels and partition regions that have the expected properties of the associated classes.

In summary, all the above geometric models solve the image partition problem, which can be computationally expensive when dealing with busy images that contain many objects and clutter. Their assumptions of piecewise constant or piecewise smooth regions, in terms of Gaussian or Mixture-of-Gaussians for the intensity distributions, can also limit their effectiveness in segmenting objects whose interiors have textured appearance and/or complex multimodal intensity distributions.

To tackle the above issues, another class of deformable models is proposed in [11], namely the *Metamorphs*. This framework uses both edge and texture information in a semiparametric model, where the shape is represented similarly as in level-set methods, with the distance function of (2.5). The interior texture is captured using a nonparametric kernel-based approximation of the intensity probability density function (pdf), which is updated in every iteration. The deformations that the model can undergo are defined in a parametric manner using Free Form Deformations (FFD) over a grid in the model interior. The nonparametric region features approximation, and the FFD-constrained model dynamics increase the robustness of Metamorphs, compared to active contours and most of the aforementioned region-based geometric models. However, the parametric, FFD-regulated, shape model constrains the solution to smoother boundaries, and therefore topologically complex region discontinuities cannot be easily captured. Also, an obvious limitation of this method, due to the parametric nature of the shape (grid), is that it is not topology independent: merging different curves on the image plane is formulated as detection of collision of different models; therefore, and in contrast to common level-set based methods, region merging is not a property inherently defined in the model representation.

A family of geometric models that emerged recently is the learning-driven deformable models, which is a general framework for propagating deformable models in a probabilistic manner, by (a) formulating the traditional energy minimization as a MAP estimation problem, while (b) exploiting the Markovian (spatially neighboring) property among the data, by using graphical learning models. In the work of [14], the integration of probabilistic active contours with Markov Random Fields (MRFs) in a graphical framework was proposed to overcome the limitations of edge-based probabilistic active contours. Deformable models with MRFs were also used in [13] to achieve smoother image likelihoods for the model evolution. In this work, although the MRFs were loosely integrated with the deformable model, the results show that the use of MRFs outperforms methods that do not exploit spatial neighborhoods for probability smoothing. More recently, a framework that tightly couples 3D MRFs with deformable models was proposed in [12] for the segmentation of 3D medical data.

To exploit the superiority of Conditional Random Fields (CRFs) [19] compared to common first-order MRFs in image segmentation problems, a coupling framework was proposed in [33, 37], where a CRF and a geometric model are integrated

for the segmentation of textured regions in clutter. Extensions of this coupling have been used in medical imaging [34–36], for 2D and 3D data segmentation. These coupling methodologies handle the classification label field and the deformable model configuration as two different solutions that are met in a joint optimization framework, while this coupling does not avoid some parametric assumptions. However, these methods provide increased robustness in cluttered images and complex textures, while they utilize the deformable model evolution as a way to update the region statistics, and dynamically adapt the graphical model (CRF) parameters.

Figure 2.4 illustrates an example of segmentation using the geometric model in [34]. The leftmost image (a) shows an *en face* fundus image of the human retina obtained with SDOCT [15]; the bright region in the center of the image is clinically called *Geographic Atrophy* (GA) [32], which corresponds to the atrophy of the retinal pigment epithelium (RPE), common in dry age-related macular degeneration (AMD). Figure 4b shows the result of the GA segmentation (in red); (c) and (d) illustrate the distance function (colored surface) as shape representation of the deformable model, for the initialization and the final configuration of the model, respectively. The cross section of the surface with the image plane (zero level) is the evolving boundary. Panel (e) shows eight instances of the deformable model evolution: the red grid points correspond to the model interior during the evolution (the leftmost image corresponds to the initialization shown in (c) and the rightmost image shows the final model interior corresponding to (d)).

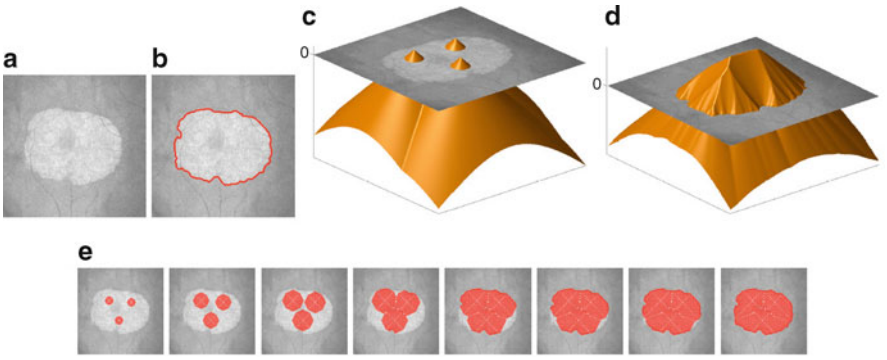

Fig. 2.4 Segmentation of an *en face* fundus image of the human retina [34], obtained with SDOCT [15]: the bright region in the center is clinically called *Geographic Atrophy* (GA) [32], which corresponds to the atrophy of the retinal pigment epithelium (RPE), common in dry age-related macular degeneration. (**a**) Original *en face* image; (**b**) final position of the deformable model capturing the GA boundaries; (**c**–**d**) model shape representation of the initialization and the final solution: the cross section of the surfaces with the image plane (zero plane) correspond to the model boundary in the image domain; (**e**) eight instances of the model interior during the evolution. *The OCT data are courtesy of G. Gregori, B. Lujan, and P.J. Rosenfeld, Bascom Palmer Eye Institute, University of Miami*

Although there is an extensive literature on both active contours and geometric models, applied to medical imaging problems in deterministic and probabilistic approaches, here we will focus on the learning-driven geometric models, as a robust and accurate segmentation methodology, which has been proven appropriate for specific challenging image data.

2.4 Region-Based Models Driven by Probability Fields

The integration of deformable models with learning-based classification is a recently introduced framework for propagating deformable models in a probabilistic manner, by formulating the *traditional* energy minimization as a *MAP* estimation problem. The main advantages of such integration are:

- The model evolution provides a framework for updating the region statistics in a learning-based region classification.
- The probabilistic formulation can provide the desired robustness to data (region) ambiguities, especially when considering markovianity among the data (spatial correlations on the image plane).
- The final solution is a locally smooth boundary around the ROI, due to the deformable model formulation, and therefore the segmentation is more robust compared to traditional segmentation-by-classification methods.

Learning-based pixel and region classification is among the popular approaches for image segmentation. This kind of methods exploit the advantages of supervised learning (training from examples) to assign probabilities of belonging to the ROI to image sites.

2.4.1 Image Likelihoods Using Graphical Models

Graphical models are commonly used to incorporate neighborhood interactions and contextual information, and they can be characterized as either *generative* or *discriminative*. Generative models are commonly used in segmentation/recognition problems where the neighboring property is well defined among the data, and they are robust to compositionality (variations in the input features), without having to see all possibilities during training. However, generative models can be computationally intractable since they require representations of multiple interacting features or long-range dependencies. On the other hand, discriminative learning infers model parameters from training data and directly calculate the class posterior given the data (mapping); such methods are usually very fast at making predictions, since they adjust the resulting classification boundary or function approximation accuracy, without the intermediate goal of forming a *generator* that models the underlying distributions during testing. However, discriminative models often need large

training sets in order to make accurate predictions, and therefore they cannot be used for data with relatively high rates of ambiguities in a straightforward way. To address this problem, some approaches integrate discriminative with generative models, where the parameters of a generative approach are modeled and trained in a discriminative manner. Also, for the same purpose, discriminative methods are used in active learning frameworks, to select the most descriptive examples for labeling, in order to minimize the model's entropy without increasing the size of the training set.

Among the most popular graphical models used in 2D and 3D image segmentation are the MRFs. Intuitively, the common MRF formulation assumes that neighboring image sites should have similar labels, and this (markovian) property results to smooth probability fields. To obtain better probability smoothing, CRFs were introduced in computer vision by Lafferty et al. [19]. Although CRFs were first used to label sequential data, extensions of them are used for image segmentation [10, 18, 34–37]. The main advantage of CRFs is that they handle the known label bias problem [19], avoiding the conditional independence assumption among the features of neighboring sites (the labels neighboring property is driven by the corresponding features). In [18] the Discriminative Random Fields (DRFs) are presented, which allow for computationally efficient MAP inference. Also, in [10], CRFs are used in different spatial scales to capture the dependencies between image regions of multiple sizes. A potential limitation of CRFs is that they do not provide robustness to unobserved or partially observed features, which is a common problem in most discriminative learning models.

Figure 2.5 illustrates the idea of learning-based classification of image sites using a common MRF. Panel (a) shows the original image, namely a SDOCT example of a mouse retinal tumor [29], and panel (b) shows in magnification the region indicated by the yellow box in (a). The yellow patches in (b) indicate the MRF sites, which, in this case, correspond to single pixels. Panel (c) shows

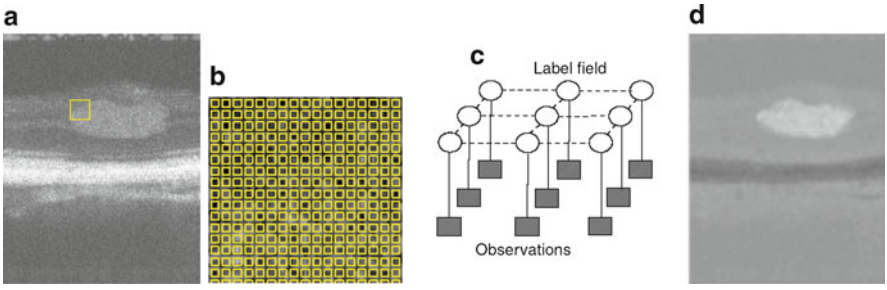

Fig. 2.5 Segmentation of mouse retinal tumor in the SDOCT image [29] in (**a**), similar to the examples in Fig. 2.2: the probability field estimated by an MRF. (**b**) Shows a magnification of the region inside the *yellow box* in (**a**): the *yellow grid* shows the image sites, which in this case correspond to pixels. (**c**) Shows the graphical representation of the MRF, where each feature (*gray box*) corresponds to the intensity value of a pixel. (**d**) Shows the probability field for the entire image: the bright regions indicate high probability of belonging to the ROI (tumor); by thresholding these probabilities we obtain the ROI

the graphical representation of the MRF. The upper level is the label field to be calculated, where each node corresponds to the (unknown) label of each pixel, i.e., "tumor" or "background" ("not tumor"). The lower level is the observation set, where each node (usually indicated with a box) corresponds to the feature vector of each site. Here, the feature vector contains a single value, which is the grayscale value of the pixel. Specifically in MRFs, the label of each site depends on (1) the corresponding observation and (2) the labels of its neighboring sites (pairwise cliques); we illustrate these dependencies with the solid and dashed lines, respectively. The segmentation result is obtained as a global optimization problem, i.e., estimating the optimal label field, given the observations. Panel (d) illustrates the probabilities of all pixels belonging to the ROI, i.e., the tumor: bright regions correspond to high probabilities, while darker regions denote the least likely sites to belong to the tumor. The label field of the image (tumor vs. background) is derived by thresholding these probability values.

In contrast to traditional deformable models that follow deterministic energy minimization approaches, learning-based classification methods are usually based on a probabilistic solution, i.e., they are driven by the maximization of a probability. Note that some discriminative classifiers, such as Support Vector Machines (SVMs), do not offer probabilistic solution, however, there are Maximum Likelihood-based methods that convert the classifier's decision boundaries into class posteriors [27].

Integrating MRFs with Deformable Models: The coupling between learning-based classification and deformable models is a framework for propagating deformable models in a probabilistic manner, by formulating the traditional energy minimization as a MAP problem, while exploiting the Markovian property among the data.

In the work of [14], the integration of probabilistic active contours with MRFs in a graphical framework was proposed to overcome the limitations of edge-based probabilistic active contours. The active contour model and the MRF are considered separately, as two different modules, which are integrated in a simple graphical model, and the solution is given by a Bayesian decomposition.

Specifically, if C is the parametric curve of the active contour, L is the desired label field, i.e., the pixel/region labels assigned by the MRF-based classification, and I is the input image, then the integration of the two models yields the joint MAP estimation task [14],

$$\langle C^*, L^* \rangle = \arg\max_{(C,L)} P(C, L|I) \tag{2.8}$$

where the joint posterior is decomposed into,

$$P(C, L|I) \propto P(I|L)P(L|C)P(C) \tag{2.9}$$

The likelihood $P(I|L)$ is the term typically estimated with the MRF, the term $P(L|C)$ corresponds to the correlation between the active contour state and the

classification solution, while the model prior $P(C)$ represents the "internal state" of the contour, i.e., it is constructed using the internal energy in a softmax (sigmoid) functional.

Intuitively, the above formulation can be seen as a deformable model that is driven by the image likelihood, i.e., the probability field estimated with the MRF. Although this coupling does not provide a solid probabilistic model, the solution is constrained in terms of shape and markovianity among the image data. Considering edge-based features, the above formulation provides increased robustness compared to deterministic active contours.

2.4.2 CRF-Driven Geometric Models

Here, we detail on a deformable model integrated with a graphical learning model [33, 37]. It is a topology-independent solution for segmenting ROIs with texture patterns of any scale, using the implicit distance transform-based shape representation of (2.5) in a geometric model driven by CRFs.

This model integrates region and edge information as image-driven terms, whereas the probabilistic shape and internal (smoothness) terms use representations similar to common geometric models. The evolution of the model is solved as a MAP estimation problem, where the target conditional probability is decomposed into the internal term and the image-driven term, in a framework similar to the decomposition in [14], briefly described in (2.8) and (2.9). For the external/image term of the model, i.e., for estimating the probability fields that drive the evolution, a discriminative CRF is used, in two scales, pixel- and patch-based, to obtain smooth probability fields based on the corresponding image features. The main advantages of this model are:

- The use of the distance transform for the shape representation, to achieve robustness to the model initializations
- The integration of edge and region information, which is being updated during the model evolution, to handle local feature variations, contrary to methods that assume constant or smooth region features
- It avoids the problem of the model getting trapped in local minima, which most of the energy minimization-driven models suffer from
- It exploits the superiority of CRFs compared to MRFs, in their first-order formulations, for image segmentation
- It captures higher scale dependencies, using pixel- and patch-based terms in the CRF

However, although this framework can handle local feature ambiguities, by updating the region statistics dynamically, the model initialization for the target ROIs and the background must include indicative samples of both classes of regions. This is because of the discriminative nature of the driving CRF, i.e., the requirement of sufficient samples for the positive and negative hypotheses, for

training. Also, the computational cost of this method is in average higher than the cost of most recent deformable models and learning-based segmentation approaches. This is mainly due to the inference process and the updating of the region statistics in every iteration of the evolution.

Before we detail on the integration of the deformable model with the probabilistic classification, i.e., the model dynamics, we describe the shape, topology, and internal energy terms below.

Model Shape and Topology: In a similar way as in most geometric models, e.g., the Metamorphs [11], the model's shape is represented by the distance transform of (2.5), such that the zero distance values correspond to the evolving interface (Fig. 2.3). The model M (interface) defines two region classes in the image domain Ω, namely the region R_M enclosed by the model M and the background $\Omega \backslash R_M$. Keeping in mind that the initialization of the model must be inside the target ROI, and therefore the evolution is in a balloon-like manner, i.e., the model expands from the currently estimated R_M outward, we can incorporate the topology of the edges in $\Omega \backslash R_M$ in the implicit shape description. That is, we assume that edges near, and outside, the currently estimated front (in each iteration), may belong to the desired ROI boundaries. Then, the implicit shape-topology description of the model is,

$$\Phi_M(\mathbf{x}) = \frac{1}{1 + \exp\{-D(\mathbf{x})\}}, \quad D(\mathbf{x}) = D_M + D_{\text{edge}}, \tag{2.10}$$

where D_M is the distance transform of the front M, as defined in (2.5), and D_{edge} is the (unsigned) distance image of the "informative edges," i.e., the edges outside M. Specifically, let G denote the image edge map extracted using the Canny edge detector. Then we can define the informative edge set as $\hat{G} = [1 - H(D_M)]G$, where H is the Heaviside function: $y \geqslant 0 \Leftrightarrow H(y) = 1, y < 0 \Leftrightarrow H(y) = 0$. The distance image of the informative edges is then given by,

$$D_{\text{edge}}(\mathbf{x}) = \min_{\mathbf{x}_g \in \hat{G}} ||\mathbf{x} - \mathbf{x}_g||, \forall \mathbf{x} \in \Omega \tag{2.11}$$

This formulation is used to define two regions, namely $\{ R_+ : \Phi(\mathbf{x}) \geqslant 0.5 \}$ and $R_- \equiv \Omega \backslash R_+$, which are different than R_M and $\Omega \backslash R_M$ defined in the representation D_M. The topology map $\Phi_M(\mathbf{x})$ inherently carries two types of information: the shape and the proximity to salient edges. Intuitively, the effect of D_{edge} is to control the evolution using the nearest edges in the exterior of the estimated object region, in a balloon-like manner: (2.10) implies that any point $a \in M$ is displaced due to its closest edge point $b \in \hat{G}$, toward the $\overrightarrow{ab}$ direction by the distance of $\frac{|\overrightarrow{ab}|}{2}$.

Figure 2.6 explains the above, using as an example the segmentation of the left ventricle (LV endocardium) from a CT heart image. On the left, the original image is shown, while in the middle, an instance of the evolving front M (green closed curve) and the edges estimated with the Canny edge detector (white lines) are shown superimposed on the MR image. The right most image shows in magnification the region highlighted with a magenta-colored box in the image in the middle.

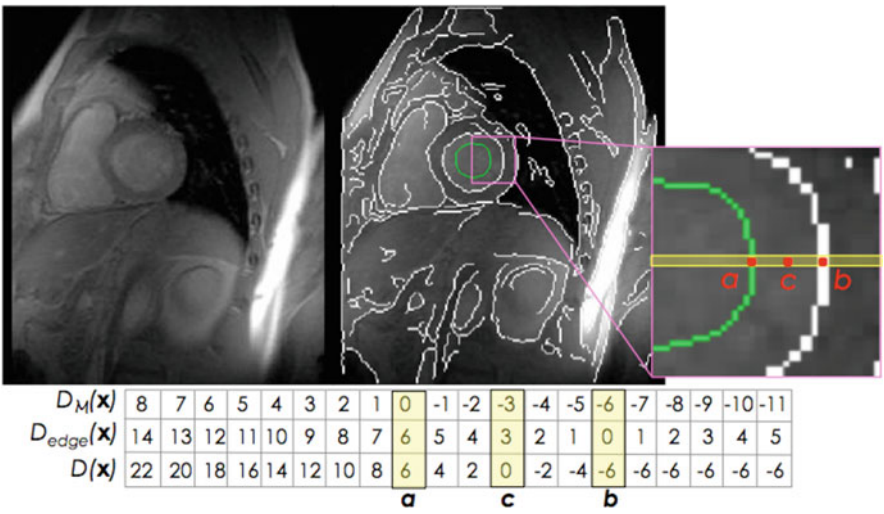

Fig. 2.6 The model shape-topology that incorporates the implicit shape representation of the evolving front and the topology of the closest informative edges to the front

We use the pixels highlighted with the yellow line to show the distances. The points a, b, and c correspond to the point on the evolving front, the closest informative edge pixel, and the point with equal distance between the first two, respectively. The table at the bottom shows the values of D_M, D_{edge}, and D.

At any instance of the model evolution, spurious edges in the model interior are not considered, as imposed by (2.11) and the use of $\hat{G}$ (informative edges). Spurious edges in the model exterior compete with the model shape, as shown in Fig. 2.6, which sometimes creates ambiguities in the model shape-topology Φ_{M}. Let us consider the two possible worst-case scenarios. (a) The front M is close to the desired object boundary, and the background is highly cluttered around the object. (b) M is still far from the desired object boundaries, and the object is highly textured. In both cases, spurious edges would yield noisy Φ_M; then, the model evolution depends on the "goodness" of the utilized region-based features, as described below.

Internal Energy: The internal energy of the model consists of three individual terms, namely the smoothness constraint E_{smooth}, the distance from the target shape E_{shape}, and a partitioning energy term E_{part},

$$E_{\text{int}}(\Phi_M) = E_{\text{smooth}}(\Phi_M) + E_{\text{part}}(\Phi_M) + E_{\text{shape}}(\Phi_M). \tag{2.12}$$

In similar manner as in active contours and geometric models, this energy is responsible for preserving the model's topology, boundary and region-based, in terms of smoothness. In the following, we describe the individual terms and their role to the model evolution.

The smoothness term: The energy term that enforces smoothness along the model boundary is defined as,

$$E_{\text{smooth}}(\Phi_M) = \varepsilon_1 A(R_+) + \varepsilon_2 \int \int_{\partial R_-} ||\nabla \Phi_M(\mathbf{x})|| \mathrm{d}\mathbf{x}, \tag{2.13}$$

where ε_1 and ε_2 are weighting constants, ∂R_-denotes a narrow band around the boundary in R_-, and $A(R_+)$denotes the area of the model interior, calculated as

$$A(R_+) = \int \int_{\Omega} H[\Phi_M(\mathbf{x}) - 0.5] \, \mathrm{d}\mathbf{x} \tag{2.14}$$

with $H(.)$ being the Heaviside function, and the value 0.5 normalizes the topology map, based on the value of the sigmoid function in (2.10). The minimization of this energy forces the model to the position with the minimum area enclosed and the maximum first-order smoothness along the model boundary; $\nabla \Phi_M$ is defined $\forall \mathbf{x} \in \Omega$, and is used similarly as in the Mumford–Shah formulation [24].

The partitioning energy: The partitioning energy $E_{\text{part}}(\Phi_M)$ forces the region $\Phi_M \geqslant 0.5$ toward a connected form. It can be also seen as a term that minimizes the entropy of a set of particles, where the particles are assumed to be the connected components of $\Phi_M \geqslant 0.5$, or equivalently the connected regions of $H(\Phi_M - 0.5)$.

Specifically, let $\{\varphi_M^{(i)}\}_{i=1}^N \subseteq H(\Phi_M - 0.5)$ be the set of N connected regions, and $e_{\text{part}}(\varphi_M^{(i)})$ be the energy of each connected region $\phi_M^{(i)}$. This energy is defined in terms of the distances between $\phi_M^{(i)}$ and the rest of the connected regions in the set $H(\Phi_M - 0.5)$,

$$e_{\text{part}}(\varphi_M^{(i)}) = \sum_{j=1, j \neq i}^{N} \tilde{d}(\varphi_M^{(j)}, \phi_M^{(i)}), \tag{2.15}$$

where

$$\tilde{d}(\varphi_M^{(j)}, \varphi_M^{(i)}) = \min_{\mathrm{x} \in \varphi_M^{(j)}} ||\varphi_M^{(i)} - \mathbf{x}|| \tag{2.16}$$

Then, the partitioning energy of the model is expressed as,

$$E_{\text{part}}(\Phi_M) = \frac{1}{2} \sum_{i=1}^{N} e_{\text{part}}(\varphi_M^{(i)}) \tag{2.17}$$

The minimization of this energy forces the model toward the minimum distances between the connected components of $H(\Phi_M - 0.5)$, i.e., forces different regions (curves) on the image plane to merge. Figure 2.4 illustrates an example of the effect

of this energy term: the model is initialized in three different positions on the image plane, and during the evolution, the model is forced toward a connected form. Intuitively, merging different regions (model interior) competes the data-driven terms, and therefore, sometimes the image features do not allow for this merging.

The shape-based energy: The role of the shape-based energy term, the third term in (2.12), is to force the evolving model toward a desired shape, in cases where we have a prior knowledge about the ROI in terms of its shape. This is a common approach in medical image analysis, since in many applications the goal is to segment ROIs with specific anatomic features. Introducing this term in the model energy, we combine the bottom-up approach (use image features to find a region) with a top-down methodology (use prior knowledge about the ROI to detect it in an image). This shape-based term is defined in a similar way as in [4, 26], in terms of the distance between the model and the target shape. For complexity purposes, i.e., to avoid the estimation of the rotation and scaling parameters for the target shape [4], we can first estimate the normalized shape histograms, which are translation, rotation and scale invariant, and measure the histogram similarity.

Let $h_M(r,v)$ and $h_{\text{shape}}(r,v)$ be the normalized 2D shape histograms of the Heaviside functions $H(\Phi_M - 0.5)$ and $H(\Phi_{\text{shape}})$, which correspond to the binary images of the model and target shape interiors, respectively; r and v (radius and angle, respectively) parameterize the position of each point of the shapes in the image domain Ω, with respect to any chosen fixed point (origin). Then the shape energy term is the similarity between the two histograms, which is given by the Bhattacharyya distance,

$$E_{\text{shape}}(\Phi_M) = -\log\left\{\int_r \int_v \sqrt{h_M(k,l)h_{\text{shape}}(k,l)}\,\mathrm{d}k\,\mathrm{d}l\right\} \tag{2.18}$$

Low values of this energy indicate similarity between the model and target shape; therefore, the minimization of this term forces the model evolution toward the desired shape. This definition is an alternative to the distance $\left\|\Phi_M - \tilde{\Phi}_{\text{shape}}\right\|$ commonly used in shape-driven geometric deformable models [4, 26], where $\tilde{\Phi}_{\text{shape}}$ denotes the prototype Φ_{shape} after the transformation (translation, rotation, and scaling) $T : \Phi_{\text{shape}} \xrightarrow{\text{T}} \tilde{\Phi}_{\text{shape}}$. Figure 2.7 illustrates a toy example of the model evolution driven only by the smoothness and shape energy terms.

The Model Dynamics: The CRF-driven deformable model is a coupling framework, where the goal is to find the optimal model position on the image plane, as well as the optimal probability field of all pixels/regions. Similarly to the MRF-driven active contours [14], the deformable model evolution is formulated as a joint MAP estimation problem for the model position and the image label field,

$$\langle \Phi_M^*, L^* \rangle = \arg\max_{(\Phi_M, L)} P(\Phi_M, L|F), \tag{2.19}$$

where Φ_M is the deformable model configuration, L is the sites' (pixels or image patches) labels, i.e., $L = \{-1,1\}$, with -1 and 1 denoting *background* and *model*

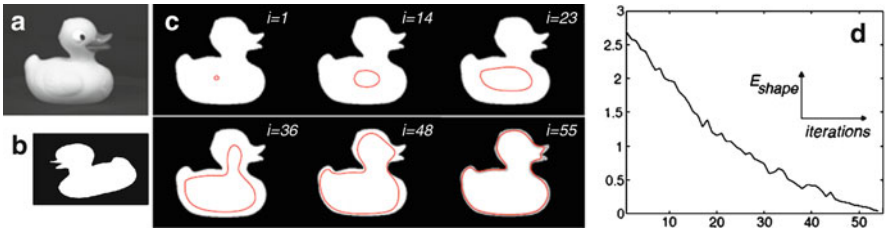

Fig. 2.7 Model evolution towards the minimization of the shape energy term. (**a**) Original toy image; (**b**) prototype; (**c**) six instances (*in red*) of the model evolution, where the number of iterations is indicated with *i*; (**d**) the values of the shape energy term during the evolution: the shape histograms of the prototype (**b**) and the model at each iteration are used in (2.18) and the value of E_{shape} is plotted

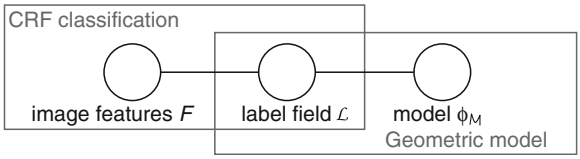

Fig. 2.8 Graphical model for the integration of the CRF with the geometric model

interior respectively, and F is the observations set: without loss of generality, we use here intensities and intensity distributions as image features, although one can integrate different texture formulations (e.g., Gabor filter responses) to describe spatially high-scale texture patterns [37].

To solve the problem of (2.19), the simple graphical model in Fig. 2.8 is used, which integrates the deformable model with learning-based classification, namely a CRF. The posterior probability $P(\Phi_M, L|F)$ is decomposed as,

$$\begin{aligned} P(\Phi_M, L|F) &\propto P(F|L)P(L|\Phi_M)P(\Phi_M) \\ &\propto P(\Phi_M)P(F)P(L|\Phi_M)P(L|F), \end{aligned} \tag{2.20}$$

where $P(\Phi_M)$ is the model prior, which corresponds to the model internal energy, and $P(F)$ is the data prior; $P(L|\Phi_M)$ is defined below as a softmax function of Φ_M, and represents the uncertainty between the classification and the deformable model position; finally, $P(L|F)$ represents the pixel/region classification in a discriminative manner (CRF framework), and we describe its estimation in the next section.

The data prior: In a pixel-wise probability field estimation, where sites indicate pixels and the observations are the pixel intensities, the data prior can be defined in terms of a Gaussian distribution,

$$P(F(\mathbf{x})) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-F(\mathbf{x})^2/\sigma_0^2\right\} \tag{2.21}$$

where σ_0^2 specifies the width of the Gaussian kernel, and $F(\mathbf{x})$ corresponds to the intensity value at the position $\mathbf{x}$ on the image plane. In a patch-wise probability field

estimation, where instead of pixels we use spatially high-scale sites, we can define the data prior using the pixel intensity distribution inside each site.

The model prior: As explained above, a common approach in probabilistic deformable models is to use the internal energy in a softmax function [14], to obtain probability instead of energy values. Using the model energy definition of (2.12), the model prior in defined in terms of the Gibbs functional,

$$P(\Phi_M) = \frac{1}{Z_{\text{int}}} \exp\{-E_{\text{int}}(\Phi_M)\}, \tag{2.22}$$

where the individual terms of $E_{\text{int}}(\Phi_M)$ are calculated using the definitions of (2.13), (2.17), and (2.18); Z_{int} is a normalization constant. The maximization of this prior forces the model toward a position with the minimum enclosed area and maximum smoothness along the boundary, with the smallest distance to the target shape, and the minimum *entropy* as defined in (2.15)–(2.17).

The model-labels likelihood: In this framework, and according to (2.20), the likelihood $P(L|\Phi_M)$ is introduced; this is the uncertainty between the classification results (ROI vs. background) obtained using the deformable model configuration (Φ_M) and the learning-based site labeling (label field L), at each instance of the model evolution. This uncertainty can be formulated as a sigmoid function of D_M (distance transform of M), as shown in [33] and in Fig. 2.9, using $1/[1 + \exp(-\kappa D_M)]$, with κ being a regulating factor; alternatively, and under the shape-topology assumption of (2.10), we can define $P(l|\Phi_M) = \Phi_M$.

According to [33], this term indicates that the probability of a site belonging to the model interior rapidly increases as $D_M(\mathbf{x})>0$ increases, and converges to zero as $D_M(\mathbf{x})<0$ decreases; also $P(l_i|D_M) = 0.5 \forall \mathbf{x} \in \Omega : D_M(\mathbf{x}) = 0$. In Fig. 2.9, we describe this uncertainty term and we show how this conditional probability varies

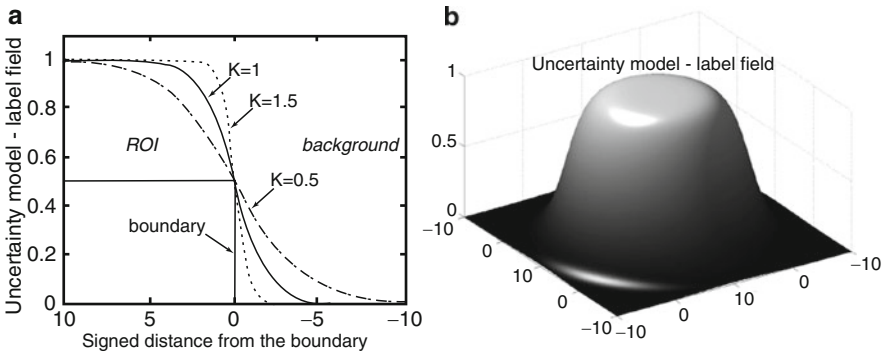

Fig. 2.9 Uncertainty between the deformable model position and the classification: (**a**) the likelihood $P(l|\Phi_M)$, for three different values of the slope parameter ($k = \{0.5, 1, 1.5\}$), using an 1D distance function, and (**b**) the same likelihood, with $k = 1$, for the distance transform of a circle of radius 10

with the values of the distance function D_M: in panel (a) we illustrate the case where $P(l_i|\Phi_M)$ is derived from an 1D distance function D_M, for a regulating $\kappa = \{0.5, 1, 1.5\}$ while in panel (b) we show the same likelihood functional (for $\kappa = 1$) calculated using the shape representation of a circle in the image domain.

The calculation of the remaining term $P(L|F)$ in (2.20) is described below.

The Driving CRF: To calculate the probability field $P(L|F)$, i.e., the image likelihood that drives the deformable model according to (2.19) and (2.20), a CRF framework is used [33]. In the "collaborative" formulation of this CRF, *interaction* terms that enforce similar class labels ("ROI" or "background") between spatially neighboring sites (cliques) containing similar features are considered. Moreover, the idea of *correlative* information between neighboring sites is introduced, by extracting complimentary features for classification in instances of local region ambiguities. Complimentary features are considered to be features from neighboring sites, which when considered jointly, instead of individually, can better describe the appearance of a single ROI.

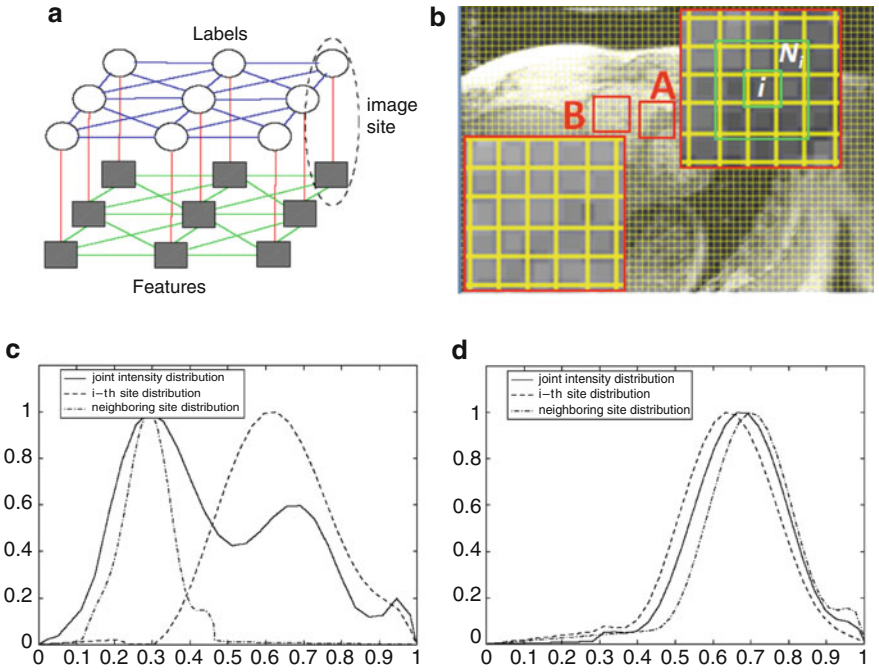

Fig. 2.10 (**a**) Graphical representation of the collaborative CRF. (**b**) Two different site neighborhoods for the example of Fig. 2.1(**a**): some sites in A belong to the ROI and some belong to the background, while B belongs entirely to the ROI. The plots of (**c** and **d**) illustrate the intensity distributions (*dashed lines*) and joint intensity distributions (*solid lines*) for two sites inside the neighborhoods of A and B, respectively (noncomplimentary and complimentary intensities, respectively – see text)

Figure 2.10 illustrates the idea of this collaborative CRF (CoCRF) and the complimentary features, using the cardiac MR example of Fig. 2.1. In (a) the graphical representation of the CoCRF is shown. The label (ROI or background) of each site of the examined image is associated with its feature set, namely its intensity value (site = pixel) or its intensity distribution (site = image patch); such associations are shown in red lines. Also, the unknown label node of each site interacts with the corresponding label nodes of its neighboring sites; these interactions are shown in blue lines. Finally, the neighboring (markovian) property of the image sites is also applied at the observation level, i.e., between the features of neighboring sites.

Panel (b) in Fig. 2.10 shows two different regions (highlighted with the red boxes) in magnification: the neighborhood N_i of the ith site (outlined with green, in region A) consists of the eight immediate (1st order) neighbors. In region A, some sites belong to the same ROI (RV) and some belong to the background and the LV, while region B belongs entirely to the ROI RV. The plots in (c) and (d) illustrate the intensity distributions (features) of the sites of A and B, respectively: the dashed lines show the distributions of the central site and one of its neighbors; the solid lines illustrate the joint distributions between these two sites. In (c) the two sites belong to different regions, while in (d) the two sites are part of the same region (since B is chosen from the ROI). Intuitively, and in contrast to the case in (c), the joint distribution in (d) supports the assumption that the two sites are complimentary, based on the resulting intensity distribution.

CoCRF Formulation: Let $F = \{f_i\}_{i \in S}$ be the feature set extracted from the input image, where f_i is the intensity distribution from a site i, and S is the set of all sites. Also, let $L = \{l_i\}_{i \in S}$ be the set of corresponding labels for all sites in S, with $l_i = \{-1, 1\}$ (background, ROI).

If N_i denotes the spatial neighborhood of each site i, we can then say that conditioned on the discrete observations F, the distribution over the labels $p(L|F)$ can be written as a first order CRF of the form [33],

$$p(L|F) = \frac{1}{z} \exp\left\{ \sum_{i \in S} \mathbf{A}(l_i, f_i) + \sum_{i \in S} \sum_{j \in N_i} [\mathbf{I}(l_i, l_j, f_i, f_j, K_i, K_j) + \mathbf{C}(l_i, l_j, f_{ij})] \right\}, \tag{2.23}$$

where z is a normalization constant.

Site associations: The unary association potential $\mathrm{A}(y_i, x_i)$ is estimated using a discriminative classifier, to directly calculate the class posterior as mapping between the distribution f_i and the class l_i [18],

$$\mathbf{A}(l_i, f_i) = \log P(l_i|f_i), \tag{2.24}$$

In the work of [18], the class posterior $P(l_i|f_i)$ is estimated locally using a Generalized Linear Model (GLM) [22], applied to an embedded feature space, after transforming the features with a nonlinear kernel set, to obtain nonlinear

decision boundaries. Alternatively, as we describe below, we can use a SVM to directly formulate this probability.

The SVM calculates the decision approximation function B in the feature space so that sign$(B(f_i))$can be used for classification of each site with feature vector f_i. However, this mapping does not inherently offer a probabilistic output since the value of $B(f_i)$ does not correspond to a posterior probability value. Adopting the approach proposed by Platt et al. [27], we can calculate the value of posterior probabilities based on the present labeled samples, by the use of a sigmoid function of the form,

$$P(l_i|f_i) = P(l_i|B(f_i)) = \frac{1}{1+\exp\{e_1 B(f_i) + e_2\}}, \tag{2.25}$$

where $\{l_i, f_i\}$ is the label and feature vector of the ith site, and e_1 and e_2 are calculated by solving a negative log maximum likelihood problem, which is a cross entropy error function of the following form:

$$\min_{e_1,e_2} D(e_1, e_2) = -\sum_{i \in S} [c_i \log(P_i) + (1 - c_i) \log(1 - P_i)], \tag{2.26}$$

where $P_i = P(l_i|B(f_i))$, and c_i denotes the true labels of the form $c_i = (y_i + 1/2)$ (for the binary problem, i.e., $l_i = \{-1, 1\}$).

Finally, if we assume a set of different features that describe the image sites, as in [37], and to generalize the definition of (2.24), we can rewrite the association potential as,

$$\mathbf{A}(l_i, f_i) = \sum_{q=1}^{Q} \log P(l_i|f_i^{(q)}), \tag{2.27}$$

where q indexes the features, i.e., $f_i^{(q)}$ is the qth feature of the set f_i. Note that with this definition we assume conditional independence between the features.

Site interactions: A common notion in MRFs is the potential that models the interactions between neighboring sites, namely the interaction potential. The simplest way to form such a potential is to penalize neighboring sites with different labels by a constant factor β,

$$\mathbf{I}(l_i, l_j) = \exp\left\{\frac{\delta(l_i - l_j)}{\sigma^2}\right\}\beta, \quad j \in N_i \tag{2.28}$$

Neighboring sites with different labels result to lower values of this potential, controlled by σ^2 and the cost parameter β; therefore, for the Markovian property to be enforced, the optimal label configuration of a neighborhood is obtained with the pairwise maximization of $\mathbf{I}$. One significant weakness of the above formulation is that it forces pairwise smoothness among the labels, without considering

discontinuities in the features. As a natural consequence, this potential cannot provide probability fields that can accurately drive a deformable model, since both active contours and geometric models usually assume the existence of discontinuities in the data, at the desired region boundaries, either in the form of high image gradient values (edges) or significant changes in region statistics.

To avoid the above limitation, we use a different formulation of the interaction potential, which is data dependent, i.e., penalizes the neighboring labels' dissimilarities by the distance of the sites' features. In the above definition, we replace the cost factor β with a distance function in the feature domain, namely the Bhattachayya distance,

$$\beta(i,j) = -\log \int [f_i(l)]^{1/2} [f_j(l)]^{1/2} \, \mathrm{d}l \tag{2.29}$$

which in general measures the similarity of the distributions f_i and f_j of the sites i and j, respectively. Note that if we assume pixels as sites, and intensity as the only feature, f_i and f_j can be the Gaussian distributions around the pixels intensity values; alternatively, for single pixel intensity values one can simply use as distance: $\beta(i,j) = ||f_i - f_j||$. Also, if we consider multiple features, we can generalize the cost factor β as

$$\beta(i,j) = \sqrt{\sum_{q=1}^{Q} \beta^{(q)}(i,j)}, \tag{2.30}$$

where $\beta^{(q)}$ denotes the distance of the qth features in the sets f_i and f_j, given by (2.29).

In this framework, the role of the interaction potential is also modulated by a term called *relative classification confidence*, $\gamma(K_i, K_j)$, of the neighboring sites i and j, to regulate inference in a more efficient way, as we describe below. The definition of this term is given by,

$$\gamma(K_i, K_j) = \frac{1}{1 + \exp\{\alpha(K_i - K_j)\}}, \quad \alpha > 0 \tag{2.31}$$

where K_i and K_j are the classification confidence values for the sites i and j, respectively, which we describe below, and is a constant regulating the confidence similarity.

The value of $\gamma(K_i, K_j)$ is only dependent on the relative value of K_i with respect to K_j : $K_i \gg K_j \Leftrightarrow \gamma \to 0$ and $K_i \ll K_j \Leftrightarrow \gamma \to 1$. This weighing function allows the interaction between i and j only if site j is more confidently classified than site i. This guarantees that interaction will generally "flow" from sites labeled with relative confidence to sites labeled with relative uncertainty. This constraint inherently carries the idea of the *Highest Confidence First* algorithm [7], since during the inference weakly labeled sites are encouraged to be affected by neighboring

strongly labeled sites, while the latter are discouraged to be affected by their neighbors.

Thus, taking into account the above, the interaction potential is defined as,

$$\mathbf{I}(l_i, l_j, f_i, f_j, K_i, K_j) = \frac{1}{z_{\text{int}}} \exp\left\{\frac{\delta(l_i - l_j)}{\sigma^2}\right\} \beta(i,j)\gamma(K_i, K_j) \tag{2.32}$$

Site correlations: The *correlative potential* $\mathbf{C}(l_i, l_j, f_{ij})$ is used to improve classification in instances of region ambiguities by evaluating neighboring sites that could be portraying a single region (class label),

$$\mathbf{C}(l_i, l_j, f_{ij}) = \log P(l_i = l_j | f_{ij}), \tag{2.33}$$

where f_{ij} is the joint intensity distribution of the sites i and j. To consider the joint appearance of two sites, we evaluate whether they are complimentary to each other with respect to their classification confidence: f_i and f_j are complimentary if $\{j \in N_i : K_i, K_j \leqslant K_{ij}\}$, where K_{ij} is the classification confidence for the joint distribution f_{ij}. In other words, the classifier treats neighboring sites as possible regions of the same class that have erroneously been segmented apart, and decides whether or not they belong to the same class. The formulation of (2.25) is also used or the estimation of $P(l_i = l_j | f_{ij})$, with the same values of the parameters e_1 and e_2.

Figure 2.11 illustrates the effect of the different potentials in the image likelihood field.

The Classification Confidence: According to this CoCRF framework, and along with the probabilities assigned to the image sites, one can also use the classification confidence, i.e., how confidently a probability is assigned to a site, for one main reason: we want to enforce interactions from sites classified with high confidence to neighboring sites classified with low confidence.

In a probabilistic approach, "strong" classification indicates very high probability (low feature ambiguity) or very low probability (high feature ambiguity) of a site belonging to the ROI. On the other hand, high (low) confidence does not indicate low (high) ambiguity but high confidence of assigning a probability. In other words, a site that is ambiguous, i.e., its probability of belonging to the ROI is around

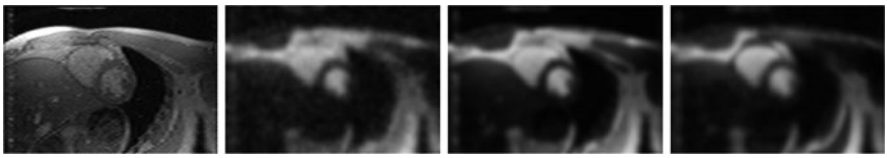

Fig. 2.11 The effect of the different potentials in the CoCRF formulation. From *left to right*: original MR cardiac image; image likelihood map using only the association potential and the SVM-based estimation; the image likelihood using the association and interaction potentials; the image likelihood after introducing the correlation potential (the complimentary sites idea) in the CRF formulation

the value 0.5, may be confidently assigned this probability. Therefore, in case of a probabilistic classification, confidence indicates the upper and lower values that a probability can take for a specific site. In such context, which is out of the scopes of this chapter, one can use the generalized notion of the probability, i.e., the belief and plausibility of the Dempster–Shafer (DS) evidence theory [31], to obtain a confidence measure for probabilistic classification.

Alternatively, and for simplicity purposes, we can simply use the probabilities obtained from (2.25) as classification confidences, i.e., $K_i = P(l_i|f_i)$ and $K_{ij} = P(l_i = l_j|f_{ij})$.

2.5 Implementation Details

Since the CRF-driven deformable model is a framework for evolving the interface with probability fields that change dynamically, based on appearance and shape, solving the probability inference in an efficient way becomes crucial for the solution stability.

Local inference: The most popular approaches for solving locally the inference have been used and compared with each other in MAP–MRF approaches. Three of the most common methods are the Iterative Conditional Modes (ICM) and the Highest Confidence First (HCF), which are deterministic, as well as Simulated Annealing (SA), which is stochastic. Li et al. [20] made an indicative comparison between these three methods, with synthetic and real examples, where SA showed overall the lowest error rates but significantly higher complexity. In synthetic images, HCF provides increased accuracy compared to ICM, with slightly higher complexity, while in real images, ICM and HCF have virtually similar accuracy. These conclusions are also justified in [6], where Chou et al. describe the main differences between ICM and HCF in solving MAP problems.

For the MAP inference in the CoCRF framework we describe above, the serial version of the HCF algorithm is adopted. This selection is justified by three important facts: (a) although ICM is faster than the serial implementation of HCF, its performance is inconsistent and its results depend to a large extent upon the initialization and the visiting order; (b) as explained in [6], HCF guarantees convergence; (c) the use of the relative classification confidence, as defined in (2.31), makes the HCF solution more straightforward.

Briefly, the HCF algorithm [6] uses the notions of "committed" and "uncommitted" sites, for strongly (confidently) and weakly classified sites, respectively. According to this algorithm, only committed sites can affect the labels of their neighbors, while uncommitted neighbors cannot affect them; in this way, inference is constrained from strongly to weakly classified sites, which, in practice, guarantees convergence. The stability of a site is determined by its classification confidence, which also plays the role of giving priority to the site, if uncommitted, to be affected by its committed neighbors, in a serial (priority queue/heap)

implementation. The local convergence is governed by a neighborhood energy functional that includes pairwise (clique) energy potentials.

For the initialization of the serial HCF algorithm [6], all sites $\mathbf{x} \in R_M$ (inside the initial seed neighborhoods) are assigned the labels $l(\mathbf{x}) = +1$. All sites $\mathbf{x}$ with classification confidence $K_\mathbf{x}$ lower than a threshold T_k are assumed uncommitted ($l(\mathbf{x}) = 0$), while the sites corresponding to the negative hypothesis (background, $\mathbf{x} \in \Omega \backslash R_M : l(\mathbf{x}) = -1$), with classification confidence higher than T_k are located. All other sites that are not in the initial seed neighborhoods but are classified as target sites are considered uncommitted. In summary, $\forall \mathbf{x} \in \Omega$:

$$\begin{cases} \mathbf{x} \in R_M \Rightarrow l(\mathbf{x}) = +1 \\ K_\mathbf{x} < T_k \Rightarrow l(\mathbf{x}) = 0 \\ K_\mathbf{x} = P(l(\mathbf{x}) = -1 | f_\mathbf{x}) \geqslant T_k \Rightarrow l(\mathbf{x}) = -1 \\ \mathbf{x} \in \Omega \backslash R_M : K_\mathbf{x} = P(l(\mathbf{x}) = +1 | f_\mathbf{x}) \geqslant T_k \Rightarrow l(\mathbf{x}) = 0 \end{cases}$$

Figure 2.12 illustrates the effect of choosing different values for the confidence T_k, on (a) the segmentation accuracy and (b) the HCF iterations. These plots were obtained using 200 different images, including 60 different SDOCT images, similar to the examples in Figs. 2.2 and 2.4 (mouse retinal tumor and geometric atrophy of the human retinal RPE, respectively). From these plots one can see that a reasonable choice is $T_k = 0.5$, which indicates equal probability of a site belonging to the ROI and the background. The rates in this plot can be directly compared with error rates reported in the literature. Also, one may notice that the error rate decreases for higher values of T_k; this is because higher values of T_k make the choice of stable sites more strict. However, fewer stable sites correspond to higher number of HCF iterations locally, which can be seen in the right plot.

Implementation in steps: The overall algorithm for the CRF-driven deformable model evolution consists of the following steps.

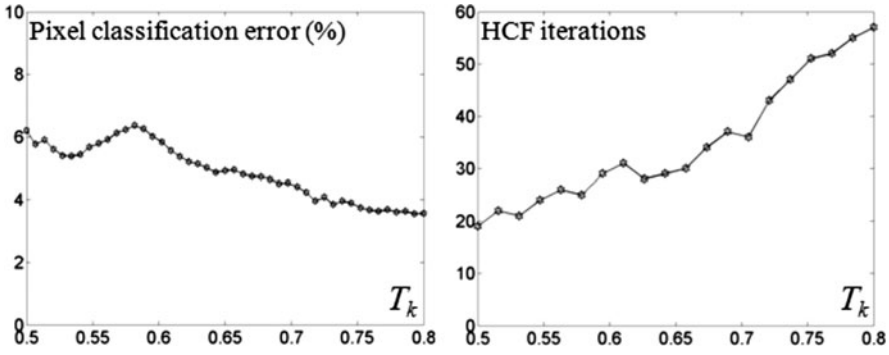

Fig. 2.12 Influence of the thresholding parameter T_k on the average segmentation error and the computational cost. *Left*: x-axis: T_k that determines strong site classification. *Right*: average HCF iterations for different T_k values

1. Model initialization inside the ROI and learning the model interior statistics. In this first step, we estimate the feature set that a site – a pixel and an image patch – must have to be part of the desired object; our goal here is to obtain an initial knowledge for the ROI, which will be updated during the model evolution. In case we have samples for the target ROI available, we also use them in this initial training. Alternatively, one can use the idea of line drawing in a user-guided segmentation manner, similar to some methods in the Computer Vision literature [1, 9].
2. Calculation of the internal energy of the deformable model $E_{\text{int}}(\Phi_M)$, the priors $P(\Phi_M)$, $P(F)$, and the likelihood $P(L|\Phi_M)$, described in Sect. 2.4.2.
3. Estimate the image probability distribution $P(L|F)$, using the CoCRF definitions in Sect. 2.4.2, and integrating intensity distributions and high-scale texture in fixed-size image patches, if needed. For the latter, one can use the Gabor filter decomposition, assuming conditional independence between the responses in different scales and orientations [37]. Obviously, the spatial scale of the image patches must be greater than the size of the texton, i.e., the size of the structure element of the object texture; for intensity-wise homogeneous regions, the texton size is equal to one. The optimal texton size can be automatically estimated using the information-theoretic Kullback–Leibler (K–L) Divergence criterion as in [13]. For computational efficiency, one can calculate only the probability field around the position of the evolving front, instead of the entire image.
4. Evolve the model within a narrow band around it, using as external (data-driven) energy the probability field from step (iii), by solving the MAP problem in (2.19). Obviously, constraining the solution inside a narrow band around the evolving front can yield much lower computational cost (see above).
5. For the new position of the model, update the interior statistics and repeat the steps (2)–(4).

Example: Segmentation of the Geographic Atrophy from OCT data: Here we give a brief description of how the learning-driven deformable model described above can be applied to a clinical application, namely the segmentation of the Geographic Atrophy in the human retina, from SDOCT data [34].

Problem Statement: AMD has become the most common cause of severe irreversible vision loss in developed countries. In patients with advanced dry AMD, most of the severe vision loss results from atrophy of the RPE. Confluent areas of RPE atrophy are clinically referred to as "Geographic Atrophy" (GA), which can cause legal blindness if it affects the central macula. There is currently no effective treatment for GA and there is only a rudimentary understanding of its pathophysiology. Furthermore, its visibility by standard photography depends on the degree of pigmentation present in the surrounding intact RPE.

Imaging GA with SDOCT: A relatively new imaging modality, SDOCT, demarcates areas of GA precisely even when it cannot be identified by photography. It utilizes the principles of reflectometry and interferometry to obtain structural information from the retina and layers under the retina at different depths along each axial scan (A-scan). SDOCT can be reconstructed to generate an *en face*

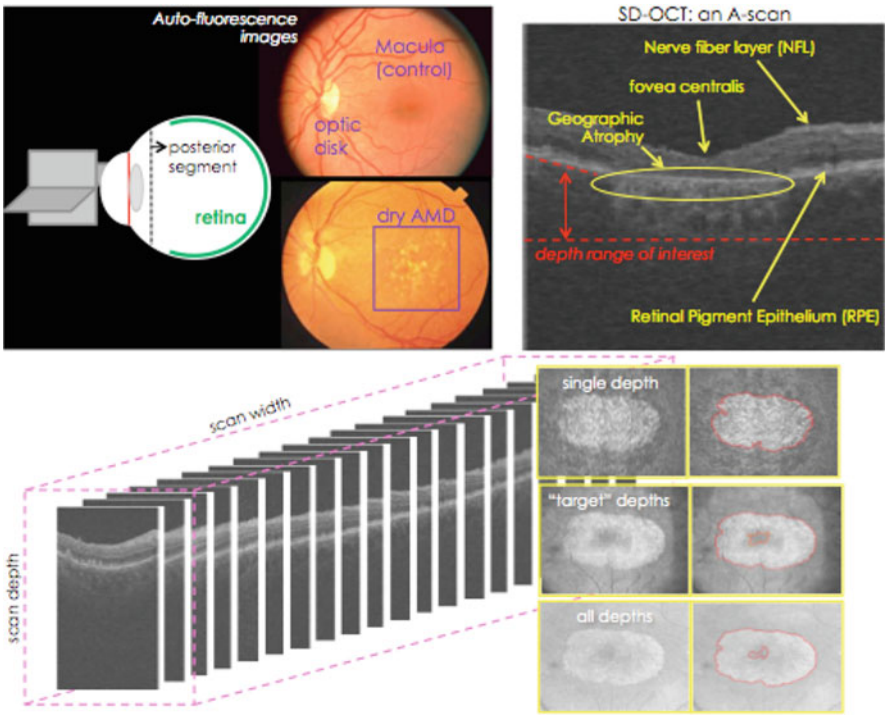

Fig. 2.13 SDOCT for imaging the Geographic Atrophy in the RPE. *Upper-left panel*: the GA appearance in Autofluorescence images. *Upper-right panel*: the morphology of the retina and how we can visualize the Geographic Atrophy; highlighted in the *yellow circle* is the part of the RPE that is thinner (GA), and therefore the light (OCT) passes through that layer (scattered intensity below the RPE). *Bottom row*: the construction of *en face* images from the A-scans; choosing a proper depth range in the A-scans for averaging the intensity can yield 2D image data with appropriate region contrast between the GA (brighter region) and the background. Note that the *x*–*y* axes in the *en face* images correspond to the *x*–*z* axes in the A-scan stack

SDOCT image, allowing for precise topographic localization of GA [15], as described in Fig. 2.13.

Segmentation of the Geographic Atrophy in the Retinal Pigment Epithelium: The CRF-driven geometric model was used to automatically segment the GA from *en face* SDOCT images [34]. In the example of Fig. 2.14, a circle was used as target shape (for the shape-based energy term). The image resolution is 200 $\times$ 200 pixels, and 5 $\times$ 5 patches were used as sites in the CRF. In panel (a) the original image (left) and the segmentation result (right) are shown. Panel (b) shows the model evolution using five markers for initialization. During the evolution, new regions are detected by the CoCRF, due to the dynamic updating of the model interior statistics, and the confidence-driven classification. In this case, the calculated image likelihood field overpowers the effect of the partitioning energy term of

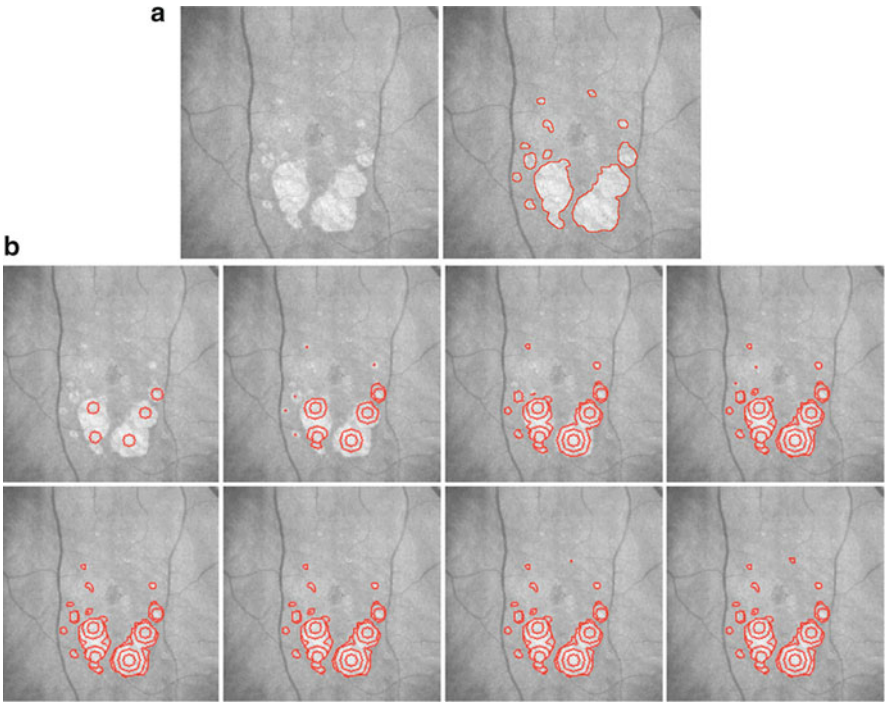

Fig. 2.14 Segmentation of the GA in en face SDOCT images: the dynamically updated CoCRF detects new GA regions during the model evolution. The model boundaries are shown in *red*. (**a**) *Left*: original image; *right*: final result. (**b**) Model evolution

(2.17), which forces the zero-level of the model distance function toward a connected form.

In [34], a numerical performance validation of this CRF-driven model is shown, using 15 subjects (A-scan sets) and manual segmentation as ground-truth (control). The validation measurement is the ratio,

$$r_{\mathrm{ga}} = \frac{A(R_{\mathrm{m}} \cap R_{\mathrm{a}})}{A(R_{\mathrm{m}})}, \tag{2.34}$$

where R_m and R_a denote the manually and automatically estimated regions, respectively, and $A(.)$ denotes the area. In this experiment, the results are (1) single A-scan depth: $r_{\mathrm{ga}} = 98.6 - 99.3\%$, (2) entire A-scan depth range: $r_{\mathrm{ga}} = 98.9 - 99.7\%$, and (3) anatomic contour-determined A-scan depth range: $r_{\mathrm{ga}} = 99.7 - 99.8\%$.

2.6 Summary

In this chapter, we described a major class of medical image segmentation methods, namely the deformable model based methods, including approaches that

successfully integrate such modeling with machine learning-based classification, for increased robustness.

Understanding image content and extracting useful image features are critical to medical image interpretation, representation, search, and mining. The automated methods that we described here can be run online, for image evaluation, or off-line on medical image archives to partition images into regions and to extract image features. In the latter case, image features can then serve indexing mechanisms for efficient search.

A challenge is to partition the images in a way that is close to the "ideal" manual segmentation from an expert, using domain prior knowledge and intuition. Along this line, the "traditional" active contours and the geometric models can provide efficient solutions, but their robustness is limited to specific types of data, namely images that provide sufficient edge and/or region-based information, for explicit modeling in a deterministic or probabilistic manner with parametric assumptions.

In an effort to provide a generalized solution to the medical image segmentation problem, we focused on a specific type of deformable models, the CRF-driven model, an indicative approach among the learning-driven models, which integrate shape modeling, learned prior information, and raw image observations. This kind of methods has shown promising results in terms of handling ambiguities in the data, and most importantly, they can be scalable to different image modalities and applications.

Specifically, in the CRF-driven model, there are smoothness constraints that smooth out or discard pixels with incoherent appearance statistics so that the segmentation is more robust in the presence of noise and intensity inhomogeneity. However, there are some issues that remain open and are to be tackled in future work to make this methodology as generalized as possible. First, it is topology independent; however, due to the integration of the CRF, a discriminative learning framework, the model initialization has to be inside the target regions, due to the region growing like evolution and for training. Second, although this model handles local feature ambiguities, by updating the region statistics dynamically, its initialization for the target and the background (positive and negative hypotheses for the classification) must include indicative samples of both regions; again, this is because of the discriminative nature of the integrated classifier. Third, the computational cost is in average higher than the cost of most recent deformable models and learning-based segmentation approaches; this is mainly due to the inference process and the updating of the region statistics in every iteration of the evolution. In this sense, this method focuses on the robustness w.r.t. the local feature ambiguities and the complex boundaries, rather than the computational efficiency.

Finally, several types of priors, such as data priors and shape priors are introduced in most model-based segmentation frameworks. Important questions remain however, regarding whether or not to use prior or what types of priors are appropriate, given specific image segmentation and search tasks.

Acknowledgment We thank Xiaolei Huang for her contribution related to the Metamorphs model [11].

References

1. Blake A, Rother C, Brown M, Perez P, Torr P (2004) Interactive image segmentation using an adaptive GMMRF model. European Conference on Computer Vision, 2004
2. Caselles V, Kimmel R, Sapiro G (1995) Geodesic active contours. IEEE International Conference on Computer Vision, 1995
3. Chan T, Vese L (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277
4. Chan T, Zhu W (2005) Level set based shape prior segmentation. IEEE Conference on Computer Vision and Pattern Recognition, 2005
5. Chen T, Metaxas D (2000) Image segmentation based on the integration of markov random fields and deformable models. International Conference on Medical Imaging Computing and Computer-Assisted Intervention, 2000
6. Chou PB, Cooper PR, Swain MJ, Brown CM, Wixson LE (1993) Probabilistic network inference for cooperative high and low level vision. In: Chellappa R, Jain A (eds) Markov random fields: theory and applications. Academic Press, pp 211–243
7. Chou P, Brown C (1990) The theory and practice of bayesian image labeling. Int J Comput Vis 4:185–210
8. Cohen LD, Cohen I (1993) Finite-element methods for active contour models and balloons for 2-D and 3-D images. IEEE Trans Pattern Anal Mach Intell 15:1131–1147
9. Duchenne O, Audibert J, Keriven R, Ponce J, Segonne F (2008) Segmentation by transduction. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2008
10. He X, Zemel R, Carreira-Perpinan M (2004) Multiscale conditional random fields for image labeling. IEEE Conference on Computer Vision and Pattern Recognition, 2004
11. Huang X, Metaxas D, Chen T (2008) Metamorphs: deformable shape and appearance models. IEEE Trans Pattern Anal Mach Intell 30(8):1444–1459
12. Huang R, Pavlovic V, Metaxas D (2006) A tightly coupled region-shape framework for 3D medical image segmentation. IEEE International Symposium on Biomedical Imaging, 2006
13. Huang X, Qian Z, Huang R, Metaxas D (2005) Deformable model based textured object segmentation. International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, 2005
14. Huang R, Pavlovic V, Metaxas D (2004) A Graphical Model Framework for Coupling MRFs and Deformable Models, IEEE Conference on Computer Vision and Pattern Recognition, 2004
15. Jiao S, Knighton R, Huang X, Gregori G, Puliafito C (2005) Simultaneous acquisition of sectional and fundus ophthalmic images with spectral-domain optical coherence tomography. Opt Express 13(2):444–452
16. Jones T, Metaxas D (1997) Automated 3D Segmentation using deformable models and fuzzy affinity, Information Processing in Medical Imaging, 1997.
17. Kass M, Witkin A, Terzopoulos D (1988) Snakes: Active contour models. Int J Comput Vis 1:321–331
18. Kumar S, Hebert M (2004) Discriminative fields for modeling spatial dependencies in natural images, Advances in Neural Information Processing Systems, 2004
19. Lafferty J, McCallum A, Pereira F (2001) Conditional random fields: probabilistic models for segmenting and labeling sequence data, International Conference on Machine Learning, 2001
20. Li SZ, Chan KL, Wang H (1996) Bayesian image restoration and segmentation by constrained optimization, In: Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, 1996

21. Malladi R, Sethian J, Vemuri B (1995) Shape modeling with front propagation: a level set approach. IEEE Trans Pattern Anal Mach Intell 17(2):158–175
22. McCullagh P, Nelder JA (1987) Generalized linear models. Chapman and Hall, pp 21–42
23. McInerney T, Terzopoulos D (1996) Deformable models in medical image analysis: a survey. Med Image Anal, 1(2):91–108
24. Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pure Appl Math 42(5):577–685
25. Paragios N, Deriche R (2002) Geodesic active regions and level set methods for supervised texture segmentation. Int J Comput Vis 46(3):223–247
26. Paragios N, Rousson M, Ramesh V (2002) Matching distance functions: a shape-to-area variational approach for globalto-local registration, European Conference on Computer Vision, 2002
27. Platt JC (1999) Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Smola A, Bartlett P, Schlkopf B, Schuurmans D (eds) Advances in Large Margin Classifiers, MIT Press, pp 61–74
28. Ronfard R (1994) Region-based strategies for active contour models. Int J Comput Vis 13(2):229–251
29. Ruggeri M, Tsechpenakis G, Jiao S, Jockovich ME, Cebulla C, Hernandez E, Murray TG, Puliafito CA (2009) Retinal tumor imaging and volume quantification in mouse model using spectral-domain optical coherence tomography. Opt Express 17(5):4074–4083
30. Samson C, Blanc-Feraud L, Aubert G, Zerubia J (2000) A level set model for image classification. Int J Comput Vis 40(3):187–198
31. Shafer G (1990) Perspectives on the theory and practice of belief functions. Int J Approx Reason 4(5):323–362
32. Smith W, Assink J, Klein R, Mitchell P, Klaver CC, Klein BE, Hofman A, Jensen S, Wang JJ, de Jong PT (2001) Risk factors for age-related macular degeneration: pooled findings from three continents. Ophthalmology 108(4):697–704
33. Tsechpenakis G, Metaxas D (2009) CoCRF deformable model: a geometric model driven by collaborative conditional random fields. IEEE Trans Image Process 18(10):2316–2329
34. Tsechpenakis G, Lujan B, Martinez O, Gregori G, Rosenfeld PJ (2008) Geometric Deformable Model Driven by CoCRFs: Application to Optical Coherence Tomography, International Conference on Medical Image Computing and Computer Assisted Intervention, 2008
35. Tsechpenakis G, Wang J (2007) CRF-based segmentation of human tear meniscus obtained with optical coherence tomography, IEEE International Conference on Image Processing, 2007
36. Tsechpenakis G, Wang J, Mayer B, Metaxas D (2007) Coupling CRFs and deformable models for 3D medical image segmentation, IEEE Mathematical Methods in Biomedical Image Analysis, 2007
37. Tsechpenakis G, Metaxas D (2007) CRF-driven implicit deformable model, IEEE Conference on Computer Vision and Pattern Recognition, 2007
38. Vese LA, Chan TF (2002) A multiphase level set framework for image segmentation using the Mumford and Shah model. Int J Comput Vis 50(3):271–293
39. Xu C, Prince JL (1998) Snakes, shapes and gradient vector flow. IEEE Trans Image Process 7(3):359–369
40. Zhu S, Yuille A (1996) Region competition: unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. IEEE Trans Pattern Anal Mach Intell 18(9):884–900
41. Yezzi AJ, Tsai A, Willsky A (1999) A statistical approach to snakes for bimodal and trimodal imagery. Proc IEEE Int Conf Comput Vis 2:898–903
42. Yezzi A, Kichenssamy S, Kumar A, Olver P, Tannebaum A (1997) A geometric snake model for segmentation of medical imagery. IEEE Trans Med Imaging 16(2):199–209

Biography

Gavriil Tsechpenakis is an assistant professor at the Computer and Information Science Dept. of Indiana University-Purdue University Indianapolis (IUPUI). He received his Ph.D. in Electrical and Computer Eng. in 2003, at the National Technical University of Athens, Greece. In 2004–2006, he was a postdoctoral fellow at Rutgers University, NJ; in 2007 he was a visiting assistant professor at University of Miami (UM); in 2008–2009 he was a senior researcher at the UM Center for Computational Science. His research focus is on computational biomedicine, computer vision, and machine learning. He is an ACM and IEEE member.

http://www.springer.com/978-1-4419-8194-3

Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies
Volume 1
El-Baz, A.S.; Acharya U, R.; Mirmehdi, M.; Suri, J.S. (Eds.)
2011, XII, 410 p., Hardcover
ISBN: 978-1-4419-8194-3