Chapter 1
Integrating Shape and Texture in 3D Deformable Models: From Metamorphs to Active Volume Models

Tian Shen, Shaoting Zhang, Junzhou Huang, Xiaolei Huang, and Dimitris N. Metaxas

Abstract For the purpose of object boundary extraction, traditional shape-based deformable models rely on external image forces that come primarily from edge or image gradient information. Such reliance on local edge information makes the models prone to get stuck in local minima due to image noise and various other image artifacts. In this chapter, we review a 2D deformable model – Metamorphs, which integrates region texture constraints so as to achieve more robust segmentation. Compared with traditional shape-based models, Metamorphs segmentation result is less dependent on model initialization and not sensitive to noise and spurious edges inside the object of interest. Then, we review Active Volume Model (AVM), a similar and improved approach for 3D segmentation. The shape of this 3D model is considered as an elastic solid, with a simplex-mesh surface made of thousands of vertices. Deformations of the model are derived from a linear system that encodes external forces from the boundary of a Region of Interest (ROI), which is a binary mask representing the object region predicted by the current model. Efficient optimization and fast convergence of the model are achieved using the Finite Element Method (FEM). To further improve segmentation performance, a multiple-surface constraint is also employed to incorporate spatial constraints among multiple objects. It uses two surface distance-based functions to adaptively adjust the weights of contribution from the image-based region information and from spatial constraints among multiple interacting surfaces. Several applications are shown to demonstrate the benefits of these segmentation algorithms based on deformable models that integrate multiple sources of constraints.

Keywords Metamorphs · Active volume models · Deformable models · Implicit representation · Texture · Distance transform · Nonparametric region statistics · Multiple-surface constraint · Finite element method

S. Zhang (✉)
Department of Computer and Information Science, Rutgers,
The State University of New Jersey, Piscataway, NJ 08854, USA
e-mail: shaoting@cs.rutgers.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_1,

1.1 Introduction

Object boundary extraction is an important task in image analysis. The main challenge is to retrieve high-level information from low-level image signals while minimizing the effect of noise, intensity inhomogeneity, and other factors. Model-based methods have been widely used with considerable success. Most noticeable are two types of models: deformable models [1, 2], and statistical shape and appearance model [3, 4].

Kass et al. proposed Snakes [1], which are energy-minimizing splines with smoothness constraints and influenced by image forces. Other parametric deformable models were proposed to incorporate overall shape and motion constraints [5, 6] and to increase the attraction range of the original Snakes by Gradient Vector Flow (GVF) [7]. Depending solely on image gradient information, however, these methods may be trapped by noise and spurious edges. Region analysis strategies [8, 9] have been incorporated in Snake-like models to improve their robustness to noise.

Another class of deformable models is level set-based geometric models [2, 9]. This approach represents curves and surfaces implicitly as the level set of a higher dimensional scalar function, and the evolution of these implicit models is based on the theory of curve evolution, with speed function specifically designed to incorporate image gradient information. The integration of region information in geometric models has been mostly based on solving the frame partition problem as in *Geodesic Active Region* [10] and *Active contours without edges* [11]. The level set models are more computationally expensive and often require knowing the number of regions and appearance statistics of each region a priori, but they are free in topology and do not need explicit parameterization. So the level set approach is commonly used in segmenting multiple objects [12] and achieves good result in tubular structure segmentation [13]. Coupled surface constraints and dual-front implementation of level set active contours [14] also provide the flexibility of capturing variable degrees of localness in optimization.

In noisy images, statistical modeling approaches can add constraints from offline learning. Cootes et al. proposed methods for building *active shape models* (ASM) [4] and *active appearance models* [3], by learning patterns of variability from a training set of annotated images. Integrating high-level knowledge, these models deform in ways constrained by the training data and are often more robust in image interpretation. Image interpretation by shape-appearance joint prior models can be based on image search [4], or by maximizing posterior likelihood of the model given image information, in a Bayesian framework [15]. In medical imaging, shape priors particularly have been introduced to cardiac segmentation [16–18], and to deformable models for constrained segmentation of bladder and prostate [19]. One limitation of the statistical models is in the laborious training data collection and annotation process.

Recently, Huang et al. present a new deformable modeling strategy aimed at integrating shape and appearance in a unified space, which is named as Metamorphs

[20, 21]. The model has not only boundary shape, but also interior appearance. The model shape is implicitly embedded in a higher dimensional space of distance transforms, thus represented by a distance map "image." In this way, both shape and appearance of the model are defined in the pixel space. A common deformation scheme, the Free Form Deformations (FFD), parameterizes warping deformations of the volumetric space in which the model is embedded in, hence deforming both model boundary and interior simultaneously. As the model deforms, the model's interior appearance statistics are updated adaptively. Both shape and appearance constraints are taken into account for determining deformation of the model toward object boundary. The robustness of Metamorphs is demonstrated using both natural and medical 2D images that have high noise levels, intensity inhomogeneity, and complex texture. The details are discussed in Sect. 1.2.

Compared with shape-based deformable models, Metamorphs interior appearance provides additional region-based constraints, making it more robust and efficient in segmentation. Efforts have been made on the integration of interior appearance into 3D models. Huang et al. introduces an adaptive model-based segmentation framework [22], in which edge and region information are integrated and used adaptively while a solid model deforms toward the object boundary. Instead of performing segmentation in an entire 3D volume, they propose model-based segmentation in an adaptively changing subvolume of interest. The subvolume is determined based on appearance statistics of the evolving object model, and within the subvolume, more accurate and object-specific edge and region information can be obtained. After calculating ROI and distance function, marching cubes [23] is applied to reconstruct the 3D surface. Shen et al. [24] and Zhang et al. [25] propose two pseudo-3D segmentation methods to reconstruct 3D object surface from 2D Metamorphs results; the details of these methods are discussed in Sects. 1.2.4 and 1.2.5, respectively. However, these methods have not incorporated real 3D deformable models. Texture information from 2D cannot reflect the 3D object appearance well. The efficiency of these pseudo-3D models and their sensitivity to the tuning of parameters are also problems. Further, the reconstruction of 3D surface from 2D contours is often subject to difficulties caused by discontinuity between contours obtained from adjacent image slices.

To address these limitations, a novel volumetric deformable model is proposed recently, which is called the Active Volume Model (AVM) Table 1.1 [26]. The AVM's shape is represented by a simplex mesh (or finite-element triangulation) and its volumetric interior carries the various visual appearance feature statistics. An advantage of the AVM formulation is that it allows the predicted object-region information to naturally become part of a linear system, the solution of which gives the deformation of the model to minimize an energy function designed to deform the model toward object boundary. Being a parametric model fitting approach, the AVM convergence is fast, typically taking no more than 40 iterations. Several factors contribute to this efficiency: (1) AVM focuses on modeling the foreground object and then reasons about the background, instead of modeling the background explicitly, (2) the model's deformations can be solved in a linear system, and

(3) multiple external constraints are combined in a probabilistic framework and together contribute to long-range forces coming directly from the predicted object region boundary. The converged AVM is directly a smooth mesh representing the segmented object surface. Therefore, it does not need any postprocessing step, such as surface reconstruction, as required by other methods, including region growing, level set, graph cut, and MRF.

To further enable simultaneous segmentation of multiple objects or multiple object parts, AVM is augmented by integrating multiple-surface spatial constraints. The improved model is named as multiple-surface AVM (MSAVM) [24]. MSAVM is especially in many medical image applications, where we are interested in extracting boundaries of several surfaces that are coupled in such a way that their relative positions are known and the distances between them are within a specific range. Integrating this high-level spatial constraint into the segmentation model can improve accuracy and robustness. In the literature, several methods have been proposed to segment coupled medical objects simultaneously. A 2D method [27] segments left ventricular Epi-and Endocardial borders using coupled active contours but needs a precise manual initialization. In 3D, Zeng et al. [28] incorporated spatial constraints about gray matter and white matter into a level set framework which greatly improved cortex segmentation accuracy. In [29], a graph-theoretic approach detects multiple interacting surfaces by transforming the problem into computing a minimum s-t cut. Deformation of multiple surfaces in [30] has intersurface proximity constraints which allow each surface to guide other surfaces into place. However, all of the above three 3D methods [28–30] require manually specifying the expected thickness between surfaces as model-based constraint. In contrast, this requirement is removed in MSAVM: Instead of setting up a fixed distance constraint during initialization, MSAVM dynamically updates the distance constraint between the interacting surfaces based on current model surfaces' spatial interrelations. Integrating the distance constraint strategy with other energy terms based on image gradient and region information, MSAVM is less sensitive to the initialization of model positions and yields more accurate segmentation results than AVM.

The remainder of the chapter is organized as follows. In Sect. 1.2, we review Metamorphs – a deformable shape and texture model, and its pseudo-3D versions. In Sect. 1.3, we review AVM's properties, including the 3D deformation, the representation, the boundary prediction module, as well as MSAVM's multiple-surface distance constraints. Section 1.4 demonstrates several applications of the above methods. We conclude this chapter and discuss future work in Sect. 1.5.

1.2 Metamorphs

The Metamorphs deformable model [20, 21] was proposed with the aim to naturally integrate shape and texture information in model-based segmentation. A limitation of many previous segmentation methods is that, the region-based module and the boundary-based module are used separately, thus information from both sources are

not integrated during the evolution of a deformable model. To address this problem, Metamorphs is modeled not a shape model but an integrated model with both boundary shape and interior texture. In 2D, the model's representation therefore is a "disk" rather than a spline curve. The dynamics of a Metamorphs model are derived coherently from both boundary and region information during the whole course of model evolution in a common variational framework. In this section, we briefly review Metamorphs and its variations.

1.2.1 2D Shape Representation and Deformations

The Model's Shape Representation: The model's shape is embedded implicitly in a higher dimensional space of distance transforms. The Euclidean distance transform is used to embed the boundary of an evolving model as the zero level set of a higher dimensional distance function. Let $\Phi : \Omega \rightarrow R^+$ be a Lipschitz function that refers to the distance transform for the model shape $\mathcal{M}$. By definition, Ω is bounded since it refers to the image domain. The shape defines a partition of the domain: the region that is enclosed by $\mathcal{M}$, $[\mathcal{R}_{\mathcal{M}}]$, the background $[\Omega - \mathcal{R}_{\mathcal{M}}]$, and on the model, $[\partial\mathcal{R}_{\mathcal{M}}]$. Given these definitions, the following implicit shape representation for $\mathcal{M}$ is considered:

$$\Phi_{\mathcal{M}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \partial\mathcal{R}_{\mathcal{M}} \\ +\mathcal{D}(\mathbf{x}, \mathcal{M}), & \mathbf{x} \in \mathcal{R}_{\mathcal{M}} \\ -\mathcal{D}(\mathbf{x}, \mathcal{M}), & \mathbf{x} \in [\Omega - \mathcal{R}_{\mathcal{M}}] \end{cases} \tag{1.1}$$

where $\mathcal{D}(\mathbf{x}, \mathcal{M})$ refers to the minimum Euclidean distance between the image pixel location $\mathbf{x} = (x, y)$ and the model $\mathcal{M}$.

Such implicit embedding makes the model shape representation a distance map "image," which greatly facilitates the integration of shape and appearance information. It also provides a feature space in which objective functions that are optimized using a gradient descent method can be conveniently used.

The Model's Deformations: The deformations that Metamorphs models can undergo are defined using a space warping technique, the FFD. The essence of FFD is to deform an object by manipulating a regular control lattice F overlaid on its volumetric embedding space. In Metamorphs, we consider an Incremental Free Form Deformations (IFFD) formulation using the cubic B-spline basis [31].

Let us consider a lattice of control points

$$F = \{F_{m,n}\} = \{(F^x_{m,n}, F^y_{m,n})\}; \quad m = 1, \ldots, M, \quad n = 1, \ldots, N \tag{1.2}$$

overlaid on a region $\Gamma = \{\mathbf{x}\} = \{(x, y) | l_x \leq x \leq h_x, l_y \leq y \leq h_y\}$ in the embedding space that encloses the model. Let us denote its initial regular configuration with no deformation as F^0, and the deforming configuration as $F = F^0 + \delta F$. Then, the IFFD

parameters $\mathbf{q}$ are the deformation improvements of the control points in both x and y directions:

$$\mathbf{q} = \delta F = \{(\delta F^x_{m,n}, \delta F^y_{m,n})\}; (m,n) \in [1,M] \times [1,N] \tag{1.3}$$

$\mathbf{q}$ is the deformation parameter. The deformed position of a pixel $\mathbf{x} = (x, y)$ given the deformation of the control lattice from F^0 to F, is defined in terms of a tensor product of Cubic B-spline polynomials:

$$D(\mathbf{x}) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u) B_l(v) F_{i+k,j+l} \tag{1.4}$$

where $i = \left\lfloor \frac{x-l_x}{h_x-l_x} \cdot (M-1) \right\rfloor$, $j = \left\lfloor \frac{y-l_y}{h_y-l_y} \cdot (N-1) \right\rfloor$. This is the familiar definition for cubic B-spline based interpolation. As a space warping technique, IFFD also integrates naturally with the implicit shape representation which embeds the model shape in a higher dimensional space. More details are available at [31].

1.2.2 Model Dynamics and Evolution

The motion of the model is driven by two types of energy terms derived from the image: the gradient data terms E_{g}, and the region data terms E_{R}. So the overall external energy functional E is defined by:

$$E_{\mathrm{ext}} = E_{\mathrm{g}} + kE_{\mathrm{R}} \tag{1.5}$$

where k is a constant balancing the contributions from the two types of terms. In the Metamorphs formulation, we are able to omit the model smoothness term, since this smoothness is implicit by using FFD. The details of the E_{g} and E_{R} energy terms are revisited in Sect. 1.3. In this section, we introduce the representation of the region, or the model's texture, used in the region data terms, as well as the algorithmic steps in model evolution.

The Model's Texture: Rather than using traditional statistical parameters (such as mean and variance) to approximate the intensity distribution of the model interior, we model the distribution using a nonparametric kernel-based method. The nonparametric approximation is differentiable, more generic and can represent complex multimodal intensity distributions.

Suppose the model is placed on an image I, the image region bounded by current model $\Phi_{\mathcal{M}}$ is $\mathcal{R}_{\mathcal{M}}$, then the intensity p.d.f. of the model interior region can be represented using a Gaussian kernel-based density estimation:

$$\mathbf{P}(i|\Phi_{\mathcal{M}}) = \frac{1}{V(\mathcal{R}_{\mathcal{M}})} \int\int_{\mathcal{R}_{\mathcal{M}}} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{\frac{-(i-I(\mathbf{y}))^2}{2\sigma^2}} \mathrm{d}\mathbf{y} \tag{1.6}$$

where $i = 0, \ldots, 255$ denotes the pixel intensity values, $V(\mathcal{R}_{\mathcal{M}})$ denotes the volume of $\mathcal{R}_{\mathcal{M}}$, $\mathbf{y}$ represents pixels in the region $\mathcal{R}_{\mathcal{M}}$, and σ is a constant specifying the width of the Gaussian kernel [32]. One example of this nonparametric density estimation can be seen in Fig. 1.1. The zero level set of the evolving models $\Phi_{\mathcal{M}}$ are drawn on top of the original image in Fig. 1.1a. The model interior regions $\mathcal{R}_{\mathcal{M}}$ are cropped and shown in Fig. 1.1b. Given the model interiors, their nonparametric intensity p.d.f.s $\mathbf{P}(i|\Phi_{\mathcal{M}}))$ are shown in Fig. 1.1c, where the horizontal axis denotes the intensity values $i = 0, \ldots, 255$, and the vertical axis denotes the probability values $P \in [0, 1]$. Finally, over the entire image I, we evaluate the probability of every pixel's intensity according to the model interior intensity p.d.f., and the resulting probability (or likelihood) map is shown in Fig. 1.1d.

Using this nonparametric estimation, the intensity distribution of the model interior gets updated automatically while the model deforms to cover a new set of interior pixels; and it avoids having to estimate and keep a separate set of intensity parameters, such as the mean and variance if a Gaussian or Mixture-of-Gaussian model was used. Moreover, this kernel-based estimation in (1.6) is a continuous function, which facilitates the computation of derivatives in a gradient descent-based optimization framework.

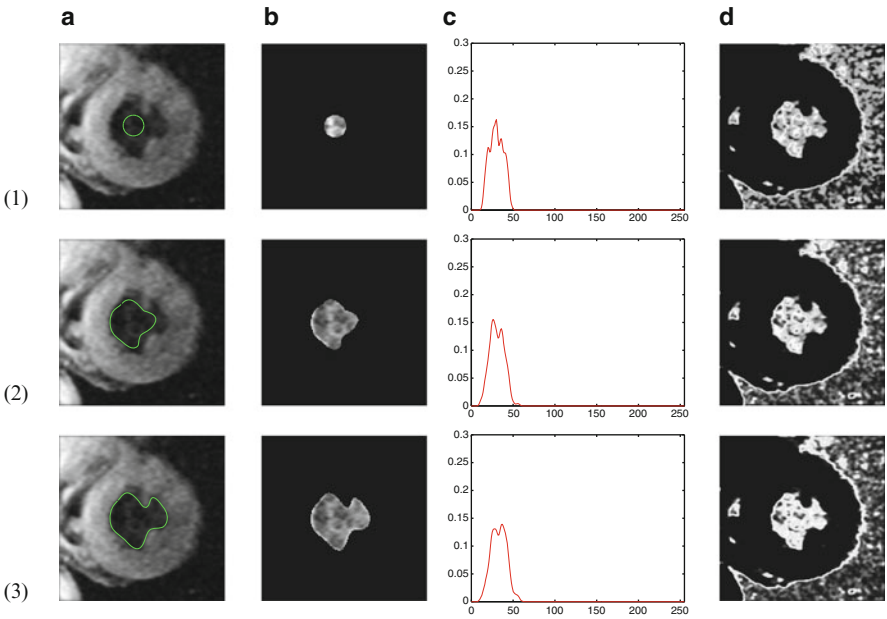


Fig. 1.1 Left ventricle endocardium segmentation, demonstrating Metamorphs appearance representation. (**1**) Initial model. (**2**) Intermediate result after four iterations. (**3**) Final converged result after ten iterations. (**a**) The evolving model drawn on original image. (**b**) Interior region of the evolving model. (**c**) The intensity p.d.f. of the model interior. (**d**) The image intensity probability map according to the p.d.f. of the model interior

Model Evolution: Since the energy terms in Metamorphs model-based segmentation framework are all differentiable with respect to the model deformation parameters $\mathbf{q}$ (1.3), a unified gradient descent-based parameter updating scheme can be derived using both edge and region information. Based on the energy term definitions, one can derive an evolution equation for each element $\mathbf{q}_i$ in the deformation parameters $\mathbf{q}$. The detailed derivations for each term can be found in [20].

The overall model fitting algorithm consists of the following steps:

1. Initialize the deformation parameters $\mathbf{q}$ to be $\mathbf{q}^0$, which indicates no deformation.
2. Compute $\partial E/\partial \mathbf{q}_i$ for each element $\mathbf{q}_i$ in the deformation parameters $\mathbf{q}$.
3. Update the parameters $\mathbf{q}'_i = \mathbf{q}_i - \lambda \cdot (\partial E/\partial \mathbf{q}_i)$. λ is the gradient descent step size.
4. Using the new parameters, compute the new model $\mathcal{M}' = D(\mathbf{q}'; \mathcal{M})$
5. Update the model. Let $\mathcal{M} = \mathcal{M}'$ recompute the implicit shape representation $\Phi_{\mathcal{M}}$, and the new partition of the image domain by the new model: $[\mathcal{R}_{\mathcal{M}}]$, $[\Omega - \mathcal{R}_{\mathcal{M}}]$ and $[\partial \mathcal{R}_{\mathcal{M}}]$. Also reinitialize a regular FFD control lattice to cover the new model, update the predicted object ROI "shape image" ϕ_{r} based on the new model interior, and recompute the energy terms.
6. Repeat steps 1–5 until convergence.

In the algorithm, after each iteration, both model shape and model-interior intensity statistics get updated, and deformation parameters get reinitialized for the new model. This allows continuous, both large-scale and small-scale deformations for the model to converge to the energy minimum. Figure 1.2 compares the results of GVF Snakes and Metamorphs. Combining both image gradient and model appearance constraints, Metamorphs performs more robustly and accurately than GVF.

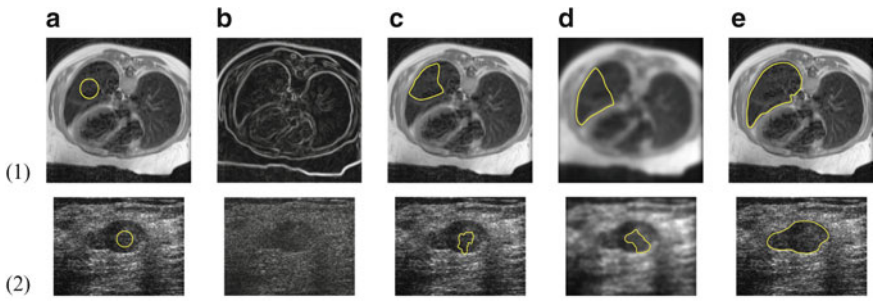


Fig. 1.2 Comparison between GVF snakes and Metamorphs. (**a**) Original image with initial model drawn on *top*. (**b**) Gray-level edge map. (**c**) Result using GVF snake on original image. (**d**) Result using GVF snake on smoothed image. (**e**) Result using Metamorphs without image smoothing

1.2.3 Adaptive 3D Metamorphs

Efforts have been put into extending 2D Metamorphs to 3D. Huang et al. proposed adaptive Metamorphs for 3D segmentation [22]. Instead of performing segmentation in an entire 3D volume, they apply model-based segmentation in an adaptively changing subvolume of interest. The subvolume is determined based on appearance statistics of the evolving object model, and within the subvolume, more accurate and object-specific edge and region information can be obtained. Still using the implicit shape representation in 3D, this method requires applying Marching Cubes algorithm [23] to reconstruct the 3D surface, after the distance function is obtained. However, the reconstructed 3D surface may not be smooth. And topologies of the meshes reconstructed at different stages of deformation may be neither consistent with each other nor consistent with the true object topology. Figure 1.3 shows some tumor segmentation results using this adaptive 3D Metamorphs model. Zooming in on these figures, one can observe discontinuity on some of the reconstructed surface meshes.

1.2.4 Pseudo-3D Segmentation

One can also perform 2D segmentation in each slice, then find correspondence between sample points on each pair of adjacent contours and reconstruct the 3D surface mesh [24]. Since most 3D volumetric medical images consist of stacks of 2D slices, the pseudo-3D method is applicable in a variety of 3D segmentation problems. The basic idea is to perform 2D segmentation using a deformable

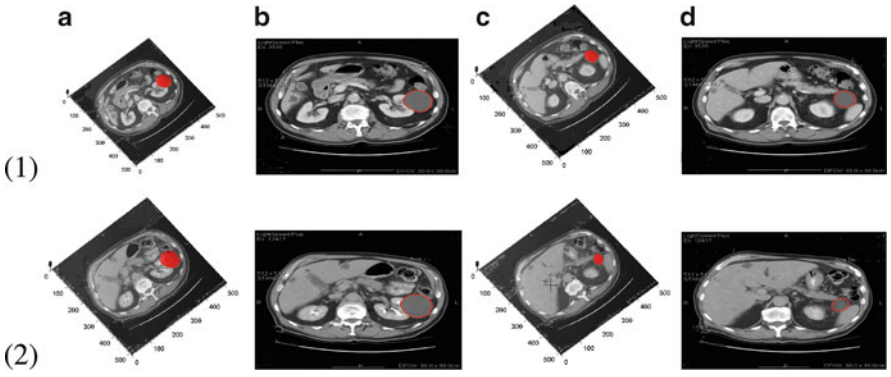


Fig. 1.3 Tumor segmentation examples. (**1.a**) The converged model representing the segmented tumor, plus one slice in a 3D CT image; (**1.b**) 2D view of the segmented contour in the same image slice; (**1.c**) The model plus another slice of the same 3D image; (**1.d**) The segmented contour in the same slice as in (**1.c**); (**2**) similar to (**1**) but for a different 3D image

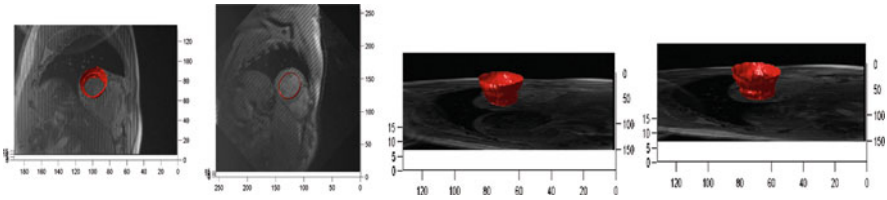


Fig. 1.4 Left ventricle endocardium segmentation example. Converged 3D model plus one slice in the 3D image, shown from several different views

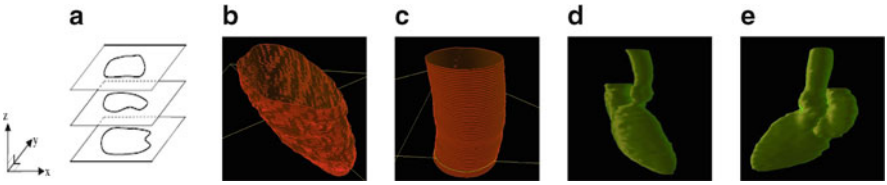


Fig. 1.5 Pseudo-3D segmentation and reconstruction. (**a**) Illustrating the "stack of contours" concept. (**b**, **c**) left ventricle (LV) and aorta showing segmentation on individual slices; LV is based on 82 slices and aorta 50 slices. (**d**, **e**) Complete reconstruction result with aorta, left atrium (LA) and LV. The aorta consists of 136 slices, LA consists of 101 and LV of 146 slices

model contour in one slice, and then propagate the contour to initialize models in neighboring slices (e.g., above and below). The initialized model on a new slice is then allowed to deform till convergence. To construct a 3D mesh model from the stack of converged 2D contours, we apply a shape registration algorithm [33] on pair-wise contours. Fifty sample points are taken from the first contour model, and correspondences for these points are computed sequentially on all other contours by shape registration. Once correspondences between points on the stack of contours are established, a 3D triangular mesh representing the segmented object boundary is reconstructed using contour points as surface vertices (Fig. 1.4).

Once the segmentation is complete in the 3D volume and correspondences between the stack of 2D contours are established, the segmented 3D volume is rendered as a triangular mesh. Interactive editing of the segmentation can be performed on individual 2D slices, and after editing, correspondences need to be recomputed only for the slices immediately adjacent to the edited slice. Figure 1.5 shows an example pseudo-3D reconstruction result of the segmented left ventricle using a heart CT volume.

1.2.5 Variational Methods for Surface Reconstruction

Another approach for pseudo-3D segmentation is to iteratively employ variational methods for surface reconstruction [25]. 2D Metamorphs models are initialized in

several slices of a 3D medical image. These 2D contours can be viewed as scattered data. Variational interpolation techniques, such as thin-plate spline (TPS) can be applied on the scattered data to obtain 3D surface shape [34]. In 3D, the interpolation function to be reconstructed, $f\mathbf{x}$, can be represented using a weighted sum of appropriate radial basis functions $\phi = |\mathbf{x}|^2 \log(|\mathbf{x}|)$.

$$f(\mathbf{x}) = \sum_{j=1}^{n} d_j \phi(\mathbf{x} - \mathbf{c}_j) + P(\mathbf{x}) \tag{1.7}$$

Where $\mathbf{c}_j$ are contour point coordinates in 3D, $\mathbf{d}_j$ are the weights, and $P(\mathbf{x})$ is a degree one polynomial that accounts for the linear and constant portions of f.

To solve for the weight parameters, $\mathbf{d}_j$, and the polynomial coefficients in P, we can use known constraints $h_i = f(\mathbf{c}_i)$.

$$h_i = \sum_{j=1}^{k} d_j \phi(\mathbf{c}_i - \mathbf{c}_j) + P(\mathbf{c}_i) \tag{1.8}$$

Let $\mathbf{c}_i = (c_i^x, c_i^y, c_i^z)$, $\phi_{ij} = \phi(\mathbf{c}_i - \mathbf{c}_j)$, we can write down the following linear system [34]:

$$\begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1k} & 1 & c_1^x & c_1^y & c_1^z \\ \phi_{21} & \phi_{22} & \dots & \phi_{2k} & 1 & c_2^x & c_2^y & c_2^z \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_{k1} & \phi_{k2} & \dots & \phi_{kk} & 1 & c_k^x & c_k^y & c_k^z \\ 1 & 1 & \dots & 1 & 0 & 0 & 0 & 0 \\ c_1^x & c_2^x & \dots & c_k^x & 0 & 0 & 0 & 0 \\ c_1^y & c_2^y & \dots & c_k^y & 0 & 0 & 0 & 0 \\ c_1^z & c_2^z & \dots & c_k^z & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \\ p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_k \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1.9}$$

In practice, we use sampled points on the initialized models as well as some nearby points to constrain the linear system. Employing the isosurface method, the f function value for all model points is set to be zero, and the function value of any nearby point is set to be its Euclidean distance to the closest model point. Solving this linear system gives us the weights d_j and the linear coefficients in P. Substituting these values into the function in (1.7), the function value at any voxel can be evaluated. Therefore, we can obtain the intersection of the zero-level isosurface with the 3D volume image, and reconstruct 2D isocontours on those slices without model initialization; then, these isocontours are used to initialize 2D Metamorphs models on those slices. The model contours are allowed to deform in 2D till convergence. During the deformation process, new isosurfaces can be generated by calculating new weights and substituting them to the interpolation function (1.7). The final 3D segmentation result is obtained from the stack of converged

2D contours. Note that using the variational surface reconstruction technique avoids the registration and correspondence finding step in Sect. 1.2.4.

1.3 Active Volume Models

One limitation of these pseudo-3D segmentation methods is that the spatial continuity between slices is not considered and thus the reconstructed 3D boundary surface is often not smooth due to discontinuity between 2D segmentations on neighboring slices. The adaptive Metamorphs [22] method, although in 3D, can generate topologically incoherent boundary. To address these limitations, AVM was proposed [26] as a true 3D deformable model that incorporates the merits of Metamorphs [20] while being computationally efficient in 3D and generating smooth, topologically coherent boundary surfaces. Instead of FFD-based deformation, AVM employs the more efficient Finite Element Method (FEM) to represent and solve for model deformations (Sect. 1.3.1). In Metamorphs, another problem is the sensitivity of result to the tuning of parameters, especially the weight parameters that balance the contributions of different energy terms and thresholds on the probability map for object vs. background classification. In AVM, the number of parameters is reduced, and a probabilistic ROI boundary-prediction module provides a meaningful classification (Sect. 1.3.2) without any ad hoc thresholding. Another advantage of AVM is that the spatial information between multiple surfaces can be incorporated to further improve accuracy and robustness (Sect. 1.3.3).

1.3.1 3D Shape Representation and Deformation

Explicit 3D Shape Representation: The 3D surface of AVM is defined by a mapping

$$\begin{aligned} \Lambda = [0,1] \times [0,1] &\rightarrow R^3 \\ (s,r) \rightarrow \mathbf{v}(s,r) &= (x(s,r), y(s,r), z(s,r)) \end{aligned} \tag{1.10}$$

Cohen et al. used tessellation to build a 3D FEM surface [35]. The tessellation can be either a cylinder or an ellipsoid. The model works well in the cylinder case, but fails using an ellipsoid. The problem is that the distribution of vertices on the ellipsoid is in quite an irregular manner. The vertices near the two poles exert strong internal forces to drag other vertices, which causes incorrect results. In order to solve the above problem in tessellation FEM mesh and enable the model to match closely object boundary, AVM adopts a polyhedron mesh as the model representation which places vertices regularly on the model. More specifically, an AVM is considered as an elastic solid and defined as a finite element triangulation Λ, which

can be tetrahedron, octahedron, or icosahedron. Using the FEM, the internal energy function can be written compactly as:

$$E_{\text{int}} = \frac{1}{2}\int_{\Lambda} (\mathbf{Bv})^T D(\mathbf{Bv}) \mathrm{d}\Lambda \tag{1.11}$$

where $\mathbf{B}$ is the differential operator for the model vertices $\mathbf{v}$ and D is the stress matrix (or constitutive matrix). A more detailed explanation of the matrices can be found in [36].

The Model's Deformation: Given the above shape representation and internal energy term, and external energy terms in the segmentation framework (Sect. 1.3.2), minimization of the AVM energy function can be achieved by solving the following linear system:

$$A_{3\text{D}} \cdot V = L_{\text{V}} \tag{1.12}$$

where $A_{3\text{D}}$ is the stiffness matrix derived from (1.11) by using the basis function in (1.13). $A_{3\text{D}}$ is symmetric and positive definite. V is the vector of vertices on the surface of AVM. L_{V} is the external force vector corresponding to the vertex vector and is obtained from the external energy terms (Sect. 1.3.2). To facilitate the computation, AVM adopts a continuous piecewise linear basis function,

$$\phi_j(v_i) = \delta_{ij} \equiv \begin{cases} \gamma & i = j \\ 0 & i \neq j \end{cases} \tag{1.13}$$

where v_i is the ith vertex on the finite element triangulation and γ is a positive value to control the smoothness of the model.

Equation (1.12) can be solved by using finite differences [35]. After initializing the AVM, the final converged result can be obtained iteratively based on equation:

$$\frac{V^t - V^{t-1}}{\tau} + A_{3\text{D}} \cdot V^t = L_{V^{t-1}} \tag{1.14}$$

where V^0 is the initial AVM vertex vector and τ is the time step size. Equation (1.14) can be written in a finite differences formulation, which yields

$$\begin{aligned} M \cdot V^t &= V^{t-1} + \tau L_{V^{t-1}} \\ M &= (I + \tau A_{3\text{D}}) \end{aligned} \tag{1.15}$$

This matrix M is sparse, so the linear system can be solved efficiently using numerical tools.

1.3.2 Model Dynamics and Evolution

In order to fit to the boundary of an object, the AVM is driven by both a gradient based data term and a region data term which are derived from image information.

The overall external energy function consists of two terms: the gradient term E_g and the region term E_R. So the overall energy function is:

$$E = E_{int} + E_{ext} = E_{int} + (E_g + k_{reg} \cdot E_R) \tag{1.16}$$

where k_{reg} is a constant that balances the contributions of the two external energy terms. The weight factor between E_{int} and E_{ext} is implicitly embedded in the FEM basis function (γ in (1.13)).

The Image Gradient Data Term: The gradient data term can be defined using the gradient map, edge distance map, or a combination of both. Denote a gradient magnitude map or the distance transform of an edge map as F_g, the gradient data term is defined as:

$$E_g = \int_\Lambda F_g(\mathbf{x}) d\Lambda \tag{1.17}$$

$$F_g = \begin{cases} D_{edge}^2, & \text{edge distance map} \\ -|\nabla I|^2, & \text{gradient magnitude map} \end{cases} \tag{1.18}$$

where D_{edge} refers to the unsigned distance transform of the edge map, and ∇I represents the image gradient.

The Object Region Data Term: A novel aspect of the AVM is that it learns the appearance statistics of the object of interest dynamically and the model's deformation is driven by the predicted object-region boundary. External constraints from various sources can be accounted in the region data term by probabilistic integration. Let us consider that each constraint corresponds to a probabilistic boundary prediction module, and it generates a confidence-rated probability map to indicate the likelihood of a pixel being: +1 (object class), or −1 (nonobject class). Suppose we have n independent external constraints, the feature used in the kth constraint is f_k, and $L(\mathbf{x})$ denotes the label of a pixel $\mathbf{x}$, our approach to combining the multiple independent modules is applying the Bayes rule in order to evaluate the final confidence rate:

$$\begin{aligned} \Pr(L(\mathbf{x})|f_1, f_2, \ldots, f_n) &= \frac{\Pr(f_1, f_2, \ldots, f_n|L(\mathbf{x}))\Pr(L(\mathbf{x}))}{\Pr(f_1, f_2, \ldots, f_n)} \\ &\propto \Pr(f_1|L(\mathbf{x}))\Pr(f_2|L(\mathbf{x})) \ldots \Pr(f_n|L(\mathbf{x}))\Pr(L(\mathbf{x})) \end{aligned} \tag{1.19}$$

For each independent module, the probability $\Pr(f_k|L(\mathbf{x}))$ is estimated based on the AVM's interior statistics. Considering a module using intensity statistics, the object region can be predicted according to the current model-interior intensity distribution. For instance, for a pixel $\mathbf{x}$ with intensity feature value $I(\mathbf{x}) = i$, where i ranges from 0 to 255, we have:

$$\begin{aligned} \Pr(i|I) &= \Pr(i, \text{object}|I) + Pr(i, \text{non-object}|I) \\ &= \Pr(i|\text{object}, I)\Pr(\text{object}|I) + \Pr(i|\text{non-object}, I)\Pr(\text{non-object}|I) \end{aligned} \tag{1.20}$$

In the equation,

- the intensity distribution over the entire image I, $\Pr(i|I)$ is known,
- we estimate the object-interior distribution $\Pr(i|\text{object}, I)$ by the current model-interior intensity distribution.
- Therefore, we can compute the background intensity distribution as:

$$\Pr(i|\text{non-object}, I) = \frac{\Pr(i|I) - \Pr(i|\text{object}, I)\Pr(\text{object}|I)}{\Pr(\text{non-object}|I)} \tag{1.21}$$

where we assume a uniform prior, $\Pr(\text{object}|I) = \Pr(L(\mathbf{x}) = \text{object}) = 0.5$ and $\Pr(\text{non-object}|I) = \Pr(L(\mathbf{x}) = \text{non-object}) = 0.5$.

Having both foreground object and background probabilities, we can obtain a binary map P_{B} that represents the predicted object region by applying the Bayesian Decision rule. That is, $P_{\mathrm{B}}(\mathbf{x}) = 1$ if $\Pr(i|\text{object}, I) \geq \Pr(i|\text{non-object}, I)$, and $P_{\mathrm{B}}(\mathbf{x}) = 0$ otherwise. We then apply a connected component analysis algorithm on P_{B} to retrieve the connected component that overlaps the current model. This connected region is considered as the current ROI. Due to noise, there might be small holes that need to be filled before extracting the shape of the ROI, R. Let us denote the signed distance transform of the current model's surface shape as Φ_Λ, and the signed distance transform of the ROI boundary shape as Φ_{R}, the region-based external energy term is defined using voxels within a narrow band around the model surface as:

$$E_{\mathrm{R}} = \int_{\Lambda} \Phi_\Lambda(\mathbf{v})\Phi_{\mathrm{R}}(\mathbf{v})\mathrm{d}\Lambda \tag{1.22}$$

The multiplicative term provides two-way balloon forces that deform the model toward the predicted ROI boundary. This allows flexible model initializations either overlapping the object or inside the object.

As one can see in Fig. 1.6c, the ROI evolves according to the changing object appearance statistics (estimated by model-interior statistics). And the image forces generated by the region term deform the model to converge to the object boundary. The Bayesian Decision-based ROI boundary prediction method outperforms other simple thresholding-on-the-probability-map techniques. For instance, we show the binary map P_{B} generated by applying a threshold of the mean of the model-interior probability in Fig. 1.6 (5) for comparison purposes; the ROIs and the converged model result significantly underestimate the true object volume.

Model Evolution: Using (1.15), we adopt the following steps to deform the AVM toward matching the desired object boundary.

1. Initialize the AVM, stiffness matrix $A_{3\mathrm{D}}$, step size τ, and calculate the gradient magnitude or edge map.
2. Compute Φ_Λ based on the current model; predict object ROI R by applying the Bayesian Decision rule to binarizing the estimated *object* probability map, and

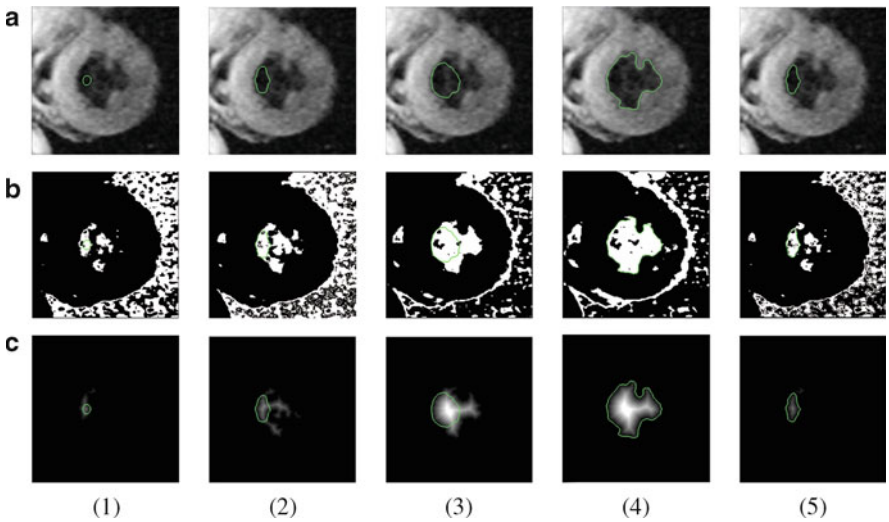


Fig. 1.6 Left ventricle endocardium segmentation using active volume model. (**a**) The model drawn on the original cardiac image. (**b**) The binary map estimated by intensity-based likelihood maps and applying the Bayesian Decision rule. (**c**) Distance transform of the ROI boundary. (**1**) Initial model. (**2**) The model after eight iterations. (**3**) The model after 18 iterations. (**4**) Final converged result after 26 iterations. (**5**) The converged result using the mean model-interior intensity probability as the threshold

compute Φ_R. Calculate the external force vector L_V which is a vector consisting of external forces at every vertex $\mathbf{v}$: $L_v = \partial E_{ext}/\partial \mathbf{v}$.
3. Deform the model according to (1.15).
4. Adaptively increase the external force factor in (1.16), decrease the step size τ in (1.15), and reduce γ in (1.13).
5. Repeat steps 2–4 until convergence.

In Step 4, adaptively changing the weight factors guarantees the model can not only reach the desired object boundary, but also capture a lot of details on the boundary.

1.3.3 Multiple-Surface Constraints

In some medical images, there may not be enough information (e.g., contrast) that can be derived from the images to clearly distinguish the object boundaries of interest. This could be due to neighboring objects having very similar tissue types or due to limitations in medical imaging technology. Therefore, a single surface-based deformable model may stop at local minima or leak out to incorrectly converge at a nearby object's boundary. Often such mistakes can be avoided by considering spatial constraints between multiple objects – for instance, by integrating the spatial

constraints in a multiple-surface-based deformable model framework [19, 24, 28] and deforming all interacting surfaces simultaneously to extract the object boundaries with better accuracy.

The Multiple-Surface AVM is initialized as several AVMs inside an outer AVM [1]. And each AVM has its own predicted ROI. To deform the multiple surfaces simultaneously with adaptive spatial constraints, the constraints are integrated into the model's energy function. We construct two distance-related Gaussian Mixtures functions, $g_{\mathrm{R}}(\mathrm{dist})$ and $g_{\mathrm{D}}(\mathrm{dist})$, which are defined based on the distance value between surfaces. Let i, j be surface indices, the mean distance value of the ith surface to other surfaces is defined as:

$$\overline{\mathrm{dist}}_i = \frac{\int_{\Lambda_i} \mathrm{dist}(\mathbf{v})\mathrm{d}\Lambda_i}{\int_{\Lambda_i} \mathrm{d}\Lambda_i} \tag{1.23}$$

$$\mathrm{dist}(\mathbf{v}) = \min_{\forall j,\, j \neq i} \left(|\Phi_{\Lambda_j}(\mathbf{v})|\right) \tag{1.24}$$

where $\mathbf{v}$ is a vertex on the ith surface Λ_i and Φ_{Λ_j} is the implicit representation (i.e., signed distance transform) of the jth surface Λ_j.

The two distance-related Gaussian Mixtures functions of the ith surface are defined in (1.25) and illustrated in Fig. 1.7.

$$\begin{aligned} g_{\mathrm{D}}(\mathrm{dist}) &= (1+\alpha) - \mathrm{e}^{-(\mathrm{dist}-\overline{\mathrm{dist}}_i)^2/2\sigma_1^2} - \alpha \mathrm{e}^{-(\mathrm{dist}-\overline{\mathrm{dist}}_i)^2/2\sigma_2^2} \\ g_{\mathrm{R}}(\mathrm{dist}) &= \mathrm{e}^{-(\mathrm{dist}-\overline{\mathrm{dist}}_i)^2/2\sigma_1^2} + \alpha \mathrm{e}^{-(\mathrm{dist}-\overline{\mathrm{dist}}_i)^2/2\sigma_2^2} \end{aligned} \tag{1.25}$$

where $\alpha \in (0, 1)$ is a weighting parameter, σ_1 and σ_2 ($\sigma_1 < \sigma_2$) are the standard deviations of two Gaussians, respectively.

Then, the energy function for the ith surface of MSAVM is defined as:

$$E = E_{\mathrm{int}} + E_{\mathrm{R}} + E_{\mathrm{dist}} \tag{1.26}$$

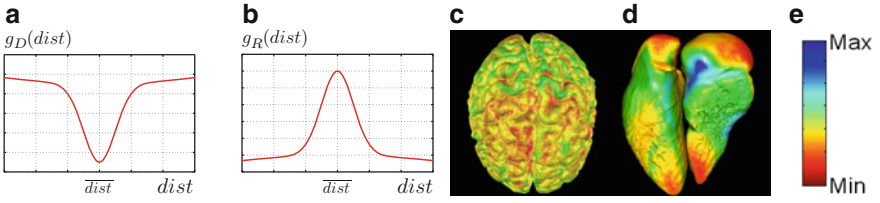


Fig. 1.7 (**a**, **b**) Functions to balance the contributions of region term and spatial constraint term, (**c**) DC mapping of brain *gray* matter segmentation; the range for distance between *gray* and *white* matters is 2–15 voxels, (**d**) DC mapping of heart segmentation; the distance range is 2–25 voxels, (**e**) the *color bar* used to map the spatial information

[1] An MSAVM can also be without an outer AVM.

where E_{int} is the same as the internal energy in (1.16). E_{R} is the external energy term derived from the predicted object ROI, defined as:

$$E_{\text{R}} = \int_{\Lambda} g_{\text{R}}(\text{dist}(\mathbf{v}))\Phi_{\text{M}}(\mathbf{v})\Phi_{\text{R}}(\mathbf{v})\text{d}\Lambda \tag{1.27}$$

and E_{dist} is the external energy term derived from the spatial distance constraint, defined as:

$$E_{\text{dist}} = \int_{\Lambda} g_{\text{D}}(\text{dist}(\mathbf{v}))(\text{dist}(\mathbf{v}) - \overline{\text{dist}}_i)^2\text{d}\Lambda \tag{1.28}$$

Given a vertex $\mathbf{v}$ on the ith surface, its minimum distance value to all the other surfaces dist($\mathbf{v}$) can be calculated based on (1.23). According to g_{R}(dist($\mathbf{v}$)) and g_{D}(dist($\mathbf{v}$)) (Fig. 1.7a, b), if dist($\mathbf{v}$) is close to the ith surface's mean distance (to other surfaces), $\overline{\text{dist}}_i$, then g_{R}(dist($\mathbf{v}$)) is large and the region term E_{R} makes more contribution toward the surface's local deformation near $\mathbf{v}$; conversely, if dist($\mathbf{v}$) is far away from $\overline{\text{dist}}_i$, which means the local surface near the vertex may be stuck at local minima or have a leakage, the energy term for distance constraint E_{dist} is given more power to deform the local surface to satisfy the distance constraint and guide it into place.

Comparing with the distance constraint function in [28], which only works well in the case of brain segmentation since the cortical layer has a nearly constant thickness, MSAVM adopts the above g_{R}(dist) and g_{D}(dist) functions to adaptively control the spatial constraints. MSAVM thus has broader applications. It can not only be used to segment brain gray matter and white matter, but also has very good performance in extracting ventricles from heart and lungs in the thorax even though distances between these coupled ventricular surfaces vary greatly. Figure 1.7c, d show two segmentation results by distance-color (DC) mapping the spatial distance information into color space.

MSAVM maintains the fast convergence and flexible initialization properties of AVM. Instead of setting a static spatial constraint manually or empirically, after each iteration, MSAVM updates each surface's mean distance value $\overline{\text{dist}}$ based on the spatial relationship among its current model surfaces. The g_{R}(dist) and g_{D}(dist) functions for each surface are then shifted accordingly to make sure the new $\overline{\text{dist}}$ still corresponds to the centerline of these functions. This unsupervised strategy for online learning of spatial distance constraints between MSAVM's multiple surfaces, coupled with its AVM's online learning of region appearance statistics, make MSAVM possess both adaptive spatial constraints and adaptive region-based constraints. These properties allow MSAVM to often have even more flexible initialization and faster convergence than the original AVM.

Model Evolution: For MSAVM, each surface is treated with an independent linear system. For the ith surface,

$$A_i \cdot V_i = L_{V_i} \tag{1.29}$$

where A_i is the stiffness matrix defined the same way as 3D AVM, V_i is the vector of vertices of the ith surface and L_{V_i} is the corresponding external force vector. The difference from AVM is that the spatial constraint is a part of the external force vector (1.28) in MSAVM. Thus, deforming MSAVM can be achieved by solving several independent linear systems.

The algorithmic steps for MSAVM evolution are as follows:

1. Initialize the MSAVM, stiffness matrix A_{3D} and $\overline{\text{dist}}$ for each surface.
2. For each surface, compute Φ_Λ based on the current model; predict R and compute Φ_R; and update $\overline{\text{dist}}$ based on (1.24) and shift $g_R(\text{dist})$ and $g_D(\text{dist})$ according to $\overline{\text{dist}}$; calculate the external force vector.
3. Deform each surface of MSAVM according to (1.29).
4. Adaptively decrease the degree of surface stiffness/smoothness.
5. Repeat steps 2–4 until convergence.

1.4 Applications

In this section, we apply these methods described above to different applications, and compare them with other 3D segmentation methods, including Geodesic Active Contours (GAC) [37] and Level Set Evolution Without Reinitialization (LSEWR) [38] by measuring the running times and validating the segmentation results using expert ground truth markings.

1.4.1 3D Segmentation Using AVM

On a PC workstation with an Intel Duo Core 3 GHz E6850 processor, we compared AVM with other 3D segmentation methods. Figure 1.8 shows the final triangulation surface of AVM, compared with the final converged results of GAC and LSEWR after surface reconstruction. Table 1.1 presents the *sensitivity* (P), *specificity* (Q), and Dice Similarity Coefficient (DSC) [39] values and running times for various experiments.

Since AVM is represented by FE triangulation, comparing with GAC and LSEWR, smooth surfaces can be obtained directly by AVM without any postprocessing, such as morphological operations and surface reconstruction. AVM also preserves topology during deformation. As shown for a 2D slice projection example in Fig. 1.10, level set segmentation results (Fig. 1.10 (2) and (3)) have complex topology, while AVM result is one smooth surface delineating lung boundary despite tumor presence. Thus, AVM is very suitable for extracting organ boundaries from volumetric medical images.

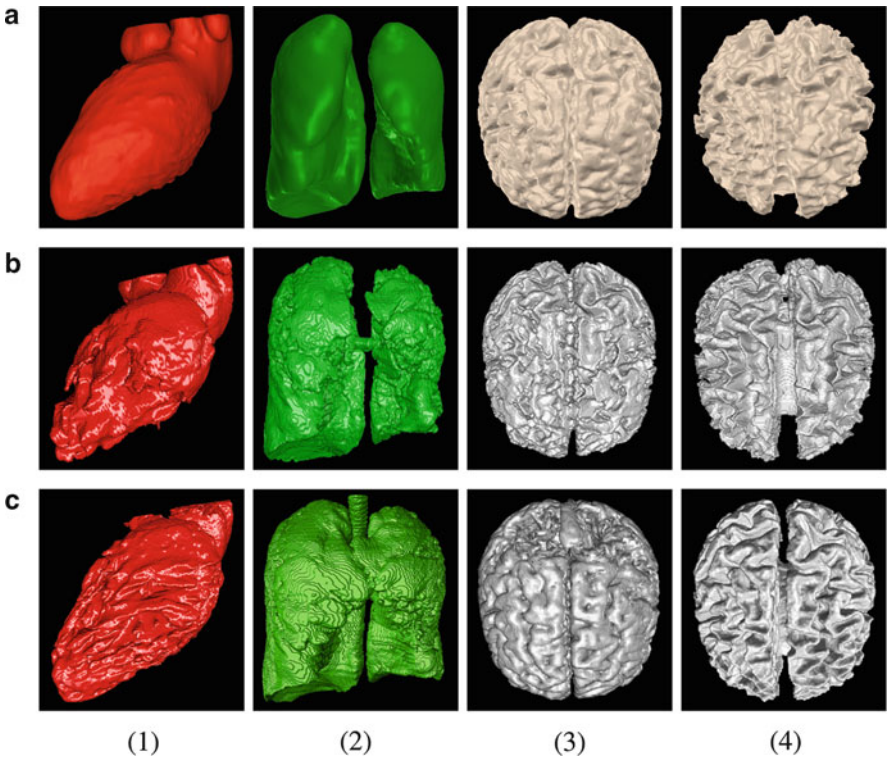


Fig. 1.8 Comparing AVM with geodesic active contours (GAC) and level set evolution without reinitialization (LSEWR). (**a**) AVM, (**b**) GAC, (**c**) LSEWR. (**1**) heart LV segmentation, (**2**) lung segmentation, (**3**) brain GM segmentation, (**4**) brain WM segmentation

Table 1.1 Quality evaluation and performance comparison

Organ	Methods	P	Q	DSC	Time
Lungs	AVM	93.6	99.8	95.2	1000s
	GAC	75.7	99.9	85.8	2149s
	LSEWR	91.4	99.7	94.6	1840s
Heart	AVM	91.8	99.6	94.3	1044s
	GAC	78.0	99.8	87.6	1752s
	LSEWR	80.1	99.9	88.5	1452s
GM	AVM	87.6	98.3	91.5	1620s
	GAC	75.7	99.0	85.0	2332s
	LSEWR	86.4	99.9	89.4	621s
WM	AVM	76.8	96.2	78.3	1833s
	GAC	66.9	95.5	72.5	2453s
	LSEWR	81.1	99.8	89.0	643s

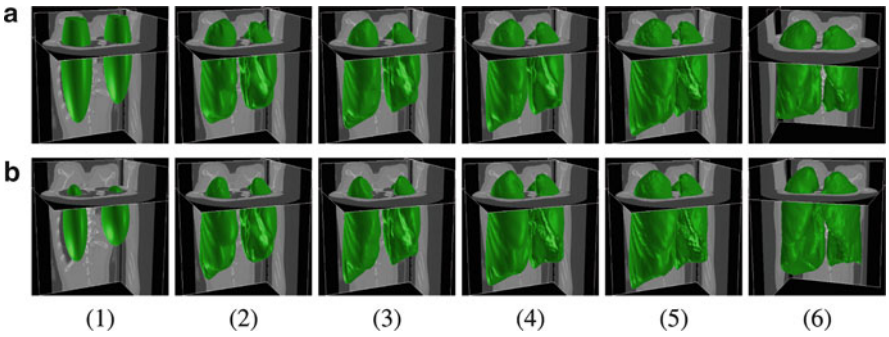


Fig. 1.9 AVM has less dependence on model initialization and parameter settings. (**a**) (**1**) Initial AVM, (**2**) after three iterations, (**3**) after 12 iterations, (**4**) after 21 iterations, (**5**) final converged result after 33 iterations, (**6**) final result viewed from a different viewpoint. (**b**, **a**) (**1**) A different AVM initialization on the same image stack, (**2**) after three iterations, (**3**) after 12 iterations, (**4**) after 24 iterations, (**5**) final converged result after 36 iterations, (**6**) final result viewed from a different viewpoint

In our experiments, AVM demonstrates several other advantages.

1. The model is very robust to noise. All the image data used in the testing are the original data without any preprocessing, e.g., smoothing and morphological operations.
2. The AVM is less dependent on the setting of parameters and on the initialization of model shape and position. We can thus use the same one parameter setting for all the lung segmentation examples, another parameter setting for all the heart segmentation examples, and so on. Figure 1.9 shows in a left and right lungs segmentation example that the AVM was initialized as ellipsoids with different radii and initial positions. The model always converged to the same result.
3. The model has good performance in avoiding leakage and overcoming local minima. Figure 1.10 shows a slice from the thorax CT DICOM stack. Note that the patient has lung disease which causes part of the lung interior region to have similar texture as the lung exterior. Due to the abnormality, our previous 2D method [24] failed to reach the object boundary but stopped at a local minima. And the abnormal interior region had negative effect on GAC and LSEWR results, which have complex topology and contain small holes and islands inside the desired object boundary. In contrast, AVM can avoid these problems and reach the true boundary of the lung despite the abnormality.
4. Since there are thousands of vertices on the FE triangulation surface, the model is able to extract very detailed information on object surfaces. This advantage can be seen from the human brain gray matter (GM) and white matter (WM) segmentation example in Fig. 1.11. The 3D simulated MRI brain images are provided by BrainWeb (http://www.bic.mni.mcgill.ca/brainweb/). The MRI stack, which is of T_1 modality, has 1 mm slice thickness, 3% noise level, and 20% INU. Both models (one for GM and one for WM) are initialized as

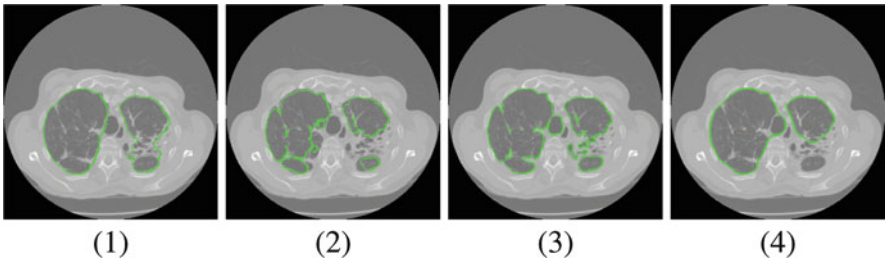


Fig. 1.10 AVM preserves topology and is good at avoiding leakage and overcoming local minima. (**1**) 2D projection of AVM, (**2**) 2D projection of geodesic active contours' result, (**3**) 2D projection of level set evolution without reinitialization's result, (**4**) 2D AVM result [24]

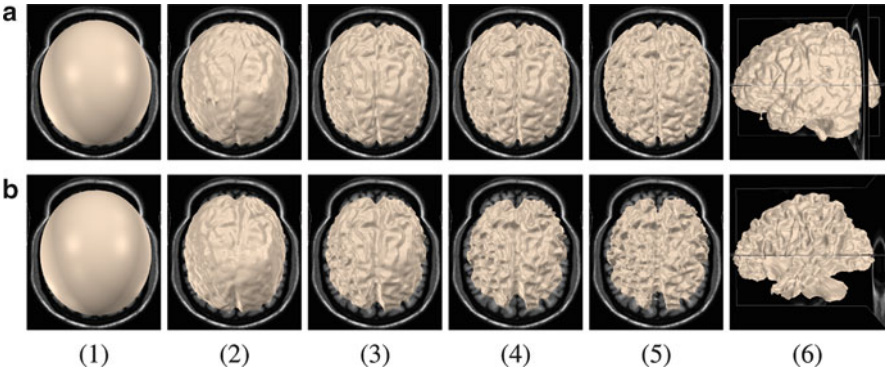


Fig. 1.11 GM and WM segmentation using AVM. The GM and WM model surfaces each has 131,074 control vertices. (**a**) (**1**) Initial model of GM, (**2**) after three iterations, (**3**) after 12 iterations, (**4**) after 24 iterations, (**5**), (**6**) final converged result after 36 iterations. (**b**) (**1**) Initial model of WM, (**2**) after three iterations, (**3**) after 12 iterations, (**4**) after 24 iterations, (**5**), (**6**) final converged result after 39 iterations

ellipsoids with strong initial smoothness constraint. As the models are getting closer to the approximated object boundary, the models decrease the smoothness constraint automatically based on the deformation strategy. Then, a lot of details on the object surfaces appear on the models.

1.4.2 Multiple-Surface Segmentation Using MSAVM

We applied MSAVM to segmenting various organ surfaces in volumetric medical images. First, we put the model into a thorax CT stack to segment the lungs. The model was initialized as one outer ellipsoid around the thorax and two inside ellipsoids whose long axes are perpendicular to the axial image plane. Figure 1.12 shows the 3D DC mapping images during deformation. A 2D coronal projection view is also included in Fig. 1.12f to show the initial model and converged result.

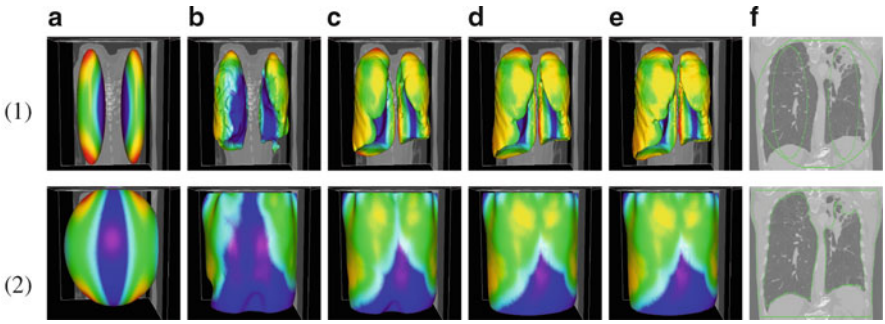


Fig. 1.12 DC mapping of lung surfaces segmentation using MSAVM, the distance range is 3–45 voxels. (**1**) (**a**)–(**e**) deformation progress of inner surfaces, (**2**) (**a**)–(**e**) outer surface; (**a**) Initial model after (**b**) 3, (**c**) 9, (**d**) 21, (**e**) 26 (converged result) iterations; (**1**) (**f**) initial model in a 2D slice, (**2**) (**f**) converged result in a 2D slice

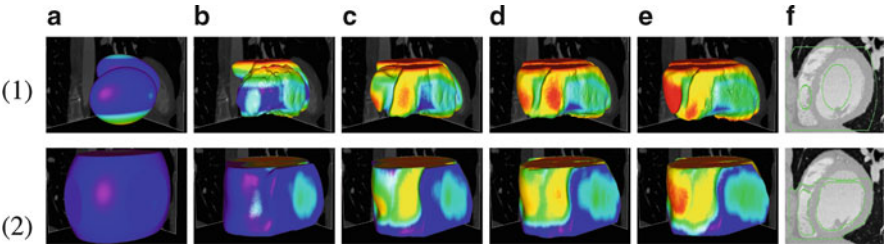


Fig. 1.13 DC mapping of heart segmentation using MSAVM viewed from the right, the distance range is 2–25 voxels, (**1**) (**a**)–(**e**) deformation progress of inner surfaces, (**2**) (**a**)–(**e**) DC mapping for outer surface; (**a**) Initial model after (**b**) 3, (**c**) 9, (**d**) 21, (**e**) 27 (converged result) iterations; (**1**) (**f**) initial model in a 2D slice, (**2**) (**f**) converged result in a 2D slice

Then, we experimented with the model on segmenting heart surfaces in a cardiac CT stack. The MSAVM is initialized as three ellipsoids: one for epicardial surface of the myocardium, one for endocardial surface of the left ventricle, and a third one for endocardial surface of the right ventricle. Some boundary condition is also specified so that the model does not deform beyond the top and bottom slices. Figures 1.13 and 1.14 show the deformation steps of the heart from two 3D viewpoints. 2D sagittal and coronal projection views are also provided in Figs. 1.13f and 1.14f. Due to intensity inhomogeneity caused by papillary muscles inside the left ventricle, it would be difficult for a single surface deformable model to reach the desired boundary without supervised learning priors. However, deforming according to the online predicted object boundary with spatial constraints, MSAVM can overcome the inhomogeneity problem and extract accurately the multiple cardiac surfaces.

To demonstrate the MSAVM more clearly, we put a set of 2D axial projection slices from a case of 3D heart segmentation in Fig. 1.15, and compare them with the converged result of original AVM using the same initialization in Fig. 1.15f. Due to intensity inhomogeneity inside the inner surfaces and obscure boundary of the outer

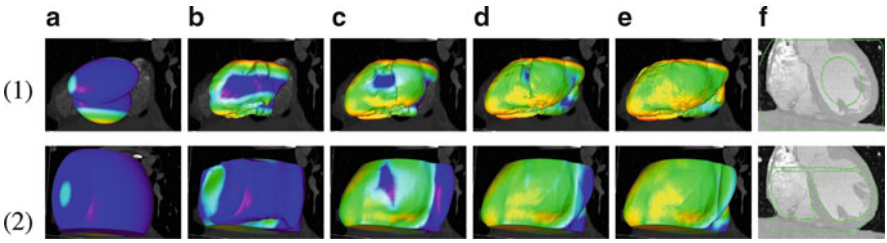


Fig. 1.14 DC mapping of heart segmentation using MSAVM viewed from the *left*, the distance range is 2–25 voxels, (**1**) (**a**)–(**e**) deformation progress of inner surfaces, (**2**) (**a**)–(**e**) outer surface; (**a**) Initial model after (**b**) 3, (**c**) 9, (**d**) 21, (**e**) 27(converged result) iterations; (**1**) (**f**) initial model in a 2D slice, (**2**) (**f**) converged result in a 2D slice

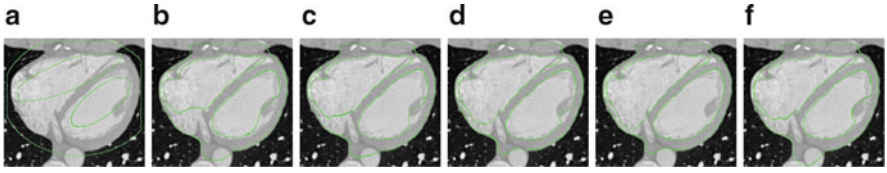


Fig. 1.15 Heart segmentation progress in a 2D slice projection. (**a**) Initial MSAVM, after (**b**) 9, (**c**) 15, (**d**) 21, (**e**) 29(converged result) iterations, (**f**) converged result of three separate AVMs after 36 iterations

Table 1.2 Quantitative evaluation and performance comparison

Data	Methods	P	Q	DSC	Iterations	Time
Lung in Fig. 1.12	MSAVM	95.5	99.8	96.2	26	870s
	AVM	92.3	99.8	94.6	33	1000s
Heart in Fig. 1.13	MSAVM	92.0	99.0	92.2	27	1535s
	AVM	90.7	98.9	91.1	39	2023s

surface, original AVM either leaks to the outer-most (e.g., outer surface) or stops at local minima (e.g., left ventricle). However, deforming under the spatial constraints, MSAVM can avoid such leakage and overcome the local minima to find the desired object boundary.

Table 1.2 summarizes the MSAVM running times and quantitative evaluation of *sensitivity* (P), *specificity* (Q), and DSC on a PC workstation with Intel Duo Core 3 GHz E6850 processor. Compared with the AVM without spatial constraint, MSAVM improved segmentation results in all the cases. Even though MSAVM needs extra time to calculate the spatial distances among surfaces, it has faster convergence so MSAVM is actually faster than AVM.

1.4.3 Prediction of Missing Structures

Being data-driven segmentation approaches, Metamorphs and AVM find object boundary based on image observations. On the other hand, statistical shape and appearance models, such as the Active Shape Models (ASM) [4], perform segmentation based on both priors learned offline and image observations extracted online. We investigated one application of combining Metamorphs and AVM, in predicting missing structures in rat brains due to abnormality. The basic idea is to generate 3D segmentation results using both ASM and Metamorphs. Comparing the volume ratio of the two results, topology changes (e.g., because of some missing brain structure) can be detected. If the structure being segmented is present, the ratio of the two volumes will be close to 1. If the structure is missing, ASM will deform little during segmentation thus remain close to the mean shape of the structure; on the contrary, the data-driven Metamorphs model shrinks to a much smaller volume since there is no obvious contrast boundary to attract the model. Therefore, in the case of the structure being missing, the ratio between the volume of the Metamorphs result and that of that ASM result is much less than 1.

In our experiment, four 4-month-old female rats were anesthetized with chloral hydrate and transcardially perfused with PBS (0.1 M, pH 7.4), followed by 4% paraformaldehyde in PBS, using a Perfusion One apparatus. The brains were left in the cranium to prevent distortions and damage and the heads were stored in 4% paraformaldehyde in PBS. The specimens were scanned on a 21T, 900 MHz, Bruker Biospin MRI, with TE = 5 ms, TR = 100 ms, f.o.v. $3 \times 2.16 \times 2.58$ and voxel size 0.06 mm, isotropic. The MR images were segmented with the use of the modeling and visualization package Amira 4.1 by experts to establish the ground truth. For the evaluation of the method, we selected 15 test datasets from the four rats' brain MRI datasets which contains 8 datasets with the cerebellum and 7 datasets without the cerebellum. We then segmented the testing brain datasets with and without the cerebellum using the ASM method slice by slice and pseudo-3D Metamorphs models simultaneously.

When cerebellum exists, the volume segmented by ASM slice by slice and the volume segmented by 3D Metamorphs model are approximately the same and the ratio of the volumes is close to 1. On the other hand, without cerebellum superimposed, the volume segmented by ASM is much larger than the volume segmented by Metamorphs model due to the absence of structure boundary and Metamorphs insufficient balloon force. The ratio between the volumes segmented by ASM slice by slice and the volumes segmented by Metamorphs model is much large than 1. The Metamorphs results on 4 datasets (out of 15) are shown in Fig. 1.16. Two are normal cases and the others are abnormal. For better comparison and visualization, normal and abnormal cases are displayed together, giving us two pairs of segmentation results. Each pair is shown in two columns (i.e. two images), which are the segmented volumes viewed from two different viewpoints. The top and bottom parts of the volumes are clipped. The outer shapes are results of segmented normal structures, while the inner shapes come from abnormal cases

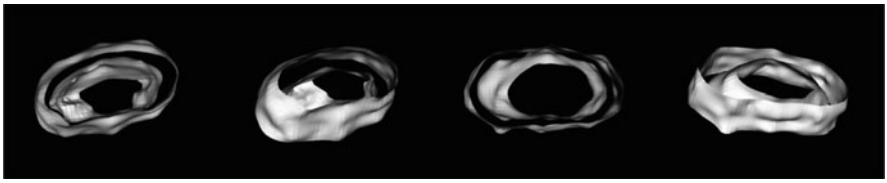


Fig. 1.16 Metamorphs' 3D segmentation results on four datasets (two normal cases and two abnormal cases). For better comparison and visualization, normal and abnormal cases are displayed together, and top as well as bottom parts of the volumes are clipped. The outer shapes are segmented normal structures, while the inner shapes come from abnormal cases with the structure missing. The volumes of inner shapes are much smaller than those of outer ones (less than one half). ASM's volumes in all cases are similar to Metamorphs results in normal cases, which are not shown here

with structure missing. One can see the inner-shape volumes are much smaller than the outer ones (less than one half). ASM's segmented volumes are similar to Metamorphs results on normal cases, which are not shown here. Out of 15 test datasets, there were 7 correct predictions of cerebellum (out of 8) and 6 correct predictions of noncerebellum (out of 7). Thus, the success rate for correctly detecting the presence of cerebellum is 86.6%. With the proposed method, we can combine statistical prior models and data-driven models to address the issue of detecting missing or emerging structures of interest.

1.5 Conclusions and Future Works

In this chapter, we have reviewed Metamorphs – a class of deformable models that integrate region texture constraints in 2D segmentation, as well as some pseudo-3D segmentation methods based on Metamorphs. We also reviewed a recently proposed AVM, which is a natural extension of parametric deformable models to integrate object appearance and region information in 3D. Several applications were introduced to demonstrate the benefits of these algorithms.

The results are promising. However, due to the local smoothness of simplex-mesh and parametric models, it is still hard for the model to reach some tip locations (e.g., top-right tip of the right ventricle in Fig. 1.15). As future work we plan to address this problem by remeshing the model near high-curvature areas such as tips since vertices in these areas tend to be sparser than those distributed on the main body. We are also interested in the incorporation offline-learned prior models into AVM. 4D segmentation or motion reconstruction is another research direction.

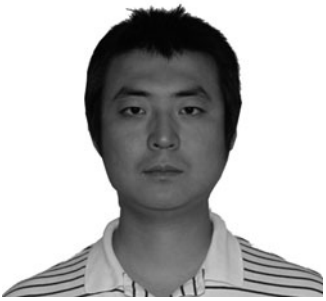
Acknowledgments The authors would like to thank Prof. Leon Axel (NYU) for providing the heart CT volume data, and Prof. Panayotis K. Thanos (Brookhaven National Lab) for providing the rat brain data.

References

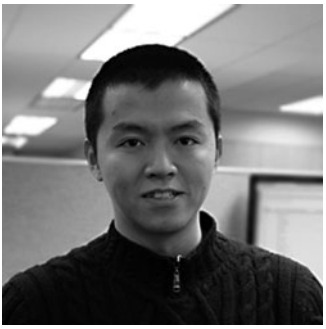
1. Kass M, Witkin A, Terzopoulos D (1987) Snakes: active contour models. Int J Comput Vis 1:321–331
2. Malladi R, Sethian J, Vemuri B (1995) Shape modeling with front propagation: a level set approach. IEEE Trans Pattern Anal Mach Intell 17(2):158–175
3. Cootes T, Edwards G, Taylar C (1998) Active appearance models. Proc Eur Conf Comput Vis 2:484–498
4. Cootes T, Taylor C, Cooper D, Graham J (1995) Active shape model their training and application. Comput Vis Image Underst 61:38–59
5. Staib L, Duncan J (1992) Boundary finding with parametrically deformable models. IEEE Trans Pattern Anal Mach Intell 14(11):1061–1075
6. Metaxas DN, Terzopoulos D (1993) Shape and nonrigid motion estimation through physics-based synthesis. IEEE Trans Pattern Anal Mach Intell 15(6):580–591
7. Xu C, Prince J (1998) Snakes, shapes and gradient vector flow. IEEE Trans Image Process 7:359–369
8. 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
9. Cremers D, Rousson M, Deriche R (2007) A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. Int J Comput Vis 72(2):195–215
10. Paragios N, Deriche R (2002) Geodesic active regions and level set methods for supervised texture segmentation. Int J Comput Vis 46(3):223–247
11. Chan T, Vese L (2001) Active contours without edges. IEEE Trans Image Process 10:266–277
12. Fan X, Bazin P-L, Prince J (2008) A multi-compartment segmentation framework with homeomorphic level sets. In: CVPR, June 2008, pp 1–6
13. Subakan O, Vemuri B (2008) Image segmentation via convolution of a level-set function with a Rigaut kernel. In: CVPR, June 2008, pp 1–6
14. Li H, Yezzi A (2007) Local or global minima: flexible dual-front active contours. IEEE Trans Pattern Anal Mach Intell 29(1):1–14
15. Yang J, Duncan J (2004) 3D image segmentation of deformable objects with joint shape-intensity prior models using level sets. Med Image Anal 8(3):285–294
16. Kohlberger T, Cremers D, Rousson M, Ramaraj R, Funka-Lea G (2006) 4D shape priors for a level set segmentation of the left myocardium in SPECT sequences. In: MICCAI (1), pp 92–100
17. Zhu Y, Papademetris X, Sinusas A, Duncan J (2008) Segmentation of left ventricle from 3D cardiac MR image sequences using a subject-specific dynamical model. In: CVPR, June 2008, pp 1–8
18. Zheng Y, Barbu A, Georgescu B, Scheuering M, Comaniciu D (2008) Four-chamber heart modeling and automatic segmentation for 3-D cardiac ct volumes using marginal space learning and steerable features. IEEE Trans Med Imaging 27:1668–1681
19. Costa M, Delingette H, Novellas S, Ayache N (2007) Automatic segmentation of bladder and prostate using coupled 3D deformable models. In: MICCAI (1), pp 252–260
20. Huang X, Metaxas D, Chen T (2004) Metamorphs: deformable shape and texture models. In: CVPR, 2004, pp 496–503
21. Huang X, Metaxas DN (2008) Metamorphs: deformable shape and appearance models. IEEE Trans Pattern Anal Mach Intell 30(8):1444–1459
22. Huang J, Huang X, Metaxas DN, Axel L (2007) Adaptive metamorphs model for 3D medical image segmentation. In: MICCAI, pp 302–310
23. Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. Comput Graph 21(4):163–169
24. Shen T, Zhu Y, Huang X, Huang J, Metaxas D, Axel L (2008) Active volume models with probabilistic object boundary prediction module. In: MICCAI, pp 331–341

25. Zhang S, Zhou J, Wang X, Chang S, Metaxas D, Pappas G, Delis F, Volkow N, Wang G, Thanos P, Kambhamettu C (2009) 3D segmentation of rodent brains using deformable models and variational methods. In: MMBIA'09
26. Shen T, Li H, Qian Z, Huang X (2009) Active volume models for 3D medical image segmentation. In: CVPR
27. Spreeuwers L, Breeuwer M (2003) Detection of left ventricular epi-and endocardial borders using coupled active contours. In: Computer assisted radiology and surgery, pp 1147–1152
28. Zeng X, Staib L, Schultz R, Duncan J (1999) Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation. IEEE Trans Med Imaging 18:927–937
29. Li K, Wu X, Chen D, Sonka M (2006) Optimal surface segmentation in volumetric images – a graph-theoretic approach. IEEE Trans Pattern Anal Mach Intell 28(1):119–134
30. MacDonald D, Kabani N, Avis D, Evans AC (2000) Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage 12(3):340–356
31. Huang X, Paragios N, Metaxas D (2003) Establishing local correspondences towards compact representations of anatomical structures. In: MICCAI, pp 926–934
32. Elgammal AM, Harwood D, Davis LS (2000) Non-parametric model for background subtraction. In: ECCV, pp 751–767
33. Huang X, Paragios N, Metaxas D (2006) Shape registration in implicit spaces using information theory and free form deformations. IEEE Trans Pattern Anal Mach Intell 28 (8):1303–1318
34. Turk G, O'Brien JF (1999) Shape transformation using variational implicit functions. In: The Proceedings of ACM SIGGRAPH 99, pp 335–342
35. Cohen L, 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
36. Logan DL (2006) A first course in the finite element method, 4th edn. CL Engineering, India
37. Caselles V, Kimmel R, Sapiro G (1997) Geodesic active contours. Int J Comput Vis 22:61–79
38. Li C, Xu C, Gui C, Fox MD (2005) Level set evolution without re-initialization: a new variational formulation. In: CVPR, vol 1. pp 430–436
39. Popovic A, de la Fuente M, Engelhardt M, Radermacher K (2007) Statistical validation metric for accuracy assessment in medical image segmentation. Int J Comput Assist Radiol Surg 2:169–181

Biography



Tian Shen received the B.S. and M.S. degrees from Xidian University, China, in 2004 and 2007, respectively. Currently, he is working toward the Ph.D. degree in the Department of Computer Science and Engineering in Lehigh University. His research interests are in the fields of image processing and computer vision, in particular PDEs, active contour models for 3D image segmentation and analysis with applications to medical imaging. He is a student member of IEEE.



Shaoting Zhang received the B.E. degree from Zhejiang University in 2005, and M.S. degree from Shanghai Jiao Tong University in 2007. Currently, he is working toward the Ph.D. in the Computer Science Department at Rutgers University. His research interests are in the fields of deformable models, geometry processing, and sparse learning techniques.

Junzhou Huang received the Bachelor's degree in Control Science and Engineer from Huazhong University of Science and Technology, Wuhan, China in 1996, and received the Master's degree in Pattern Recognition and Intelligence Systems from the Institute of Automation, Chinese Academy of Sciences, Beijing, China in 2003, and is currently working toward the Ph.D. in Computer Science at Rutgers, The State University of New Jersey. His current research interests include sparse learning, compressive sensing, video/image processing, large-scale data optimization, etc. He is a student member of IEEE and SIAM.

Xiaolei Huang is an assistant professor in the Computer Science and Engineering Department at Lehigh University since Aug. 2006. She is directing the Image Data emulation and Analysis (idea) lab. From August 2005 to August 2006, she was a research scientist at Siemens Medical Solutions USA, Inc. at Malvern, PA. She received the Ph.D. degree in Computer Science from Rutgers, The State University of New Jersey at New Brunswick, NJ, in 2006, the M.S. in Computer Science also from Rutgers University in 2001, and the B.E. degree in Computer Science from Tsinghua University, China in 1999.

Dr. Dimitris N. Metaxas is a professor in the Computer Science Department at Rutgers University. He is directing the Computational Biomedicine Imaging and Modeling Center (CBIM). He received the B.E. degree from the National Technical University of Athens Greece in 1986, M.S. degree from the University of Maryland in 1988, and Ph.D. from the University of Toronto in 1992. He has been conducting research toward the development of formal methods upon which computer vision, computer graphics, and medical imaging can advance synergistically.