Active Volume Models for 3D Medical Image Segmentation

Tian Shen[1] Hongsheng Li[1] Zhen Qian[2] Xiaolei Huang[1]
[1] Department of Computer Science & Engineering, Lehigh University, Bethlehem, PA 18015
[2] Piedmont Heart Institute, Piedmont Healthcare, Atlanta, GA 30309
{tis207, h.li}@lehigh.edu, zhen.qian@piedmont.org, xih206@lehigh.edu

Abstract

In this paper, we propose a novel predictive model for object boundary, which can integrate information from any sources. The model is a dynamic "object" model whose manifestation includes a deformable surface representing shape, a volumetric interior carrying appearance statistics, and an embedded classifier that separates object from background based on current feature information. Unlike Snakes, Level Set, Graph Cut, MRF and CRF approaches, the model is "self-contained" in that it does not model the background, but rather focuses on an accurate representation of the foreground object's attributes. As we will show, however, the model is capable of reasoning about the background statistics thus can detect when is change sufficient to invoke a boundary decision. The shape of the 3D model is considered as an elastic solid, with a simplex-mesh (i.e. finite element triangulation) 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). Other advantages of the model include the ease of dealing with topology changes and its ability to incorporate human interactions. Segmentation and validation results are presented for experiments on noisy 3D medical images.

1. Introduction

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 [10, 14], and statistical shape and appearance model [4, 5].

Kass *et al*. proposed Snakes [10], 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 [20, 16] and to increase the attraction range of the original Snakes by *Gradient Vector Flow* (GVF) [22]. Depending solely on image gradient information, however, these methods may be trapped by noise and spurious edges. Region analysis strategies [24, 9, 7] have been incorporated in Snake-like models to improve their robustness to noise.

Another class of deformable models is level set based geometric models [14, 7]. 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* [17] and *Active contours without edges* [2]. 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 [8] and achieves good result in tubular structure segmentation [21]. Coupled surface constraints and dual-front implementation of level set active contours [13] also provide the flexibility of capturing variable degrees of localness in optimization.

In noisy images, statistical modeling approaches can add constraints from prior off-line learning. Cootes *et al*. proposed methods for building *active shape models* [5] and *active appearance models* [4], 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 [5], or by maximizing posterior likelihood of the model given image information, in a Bayesian framework [23]. In medical imaging, shape priors particularly have been introduced to cardiac segmentation [11, 25], and to de-

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

formable models for constrained segmentation of bladder and prostate [6]. One limitation of the statistical models is in the laborious training data collection and annotation process.

In this paper, we propose a novel volumetric deformable model which we term the Active Volume Models (AVM). Compared with active contour and active shape/appearance models, the AVM is a "self-contained" generative object model that does not require off-line training but generates useful priors about the object. It focuses on an accurate modeling of the foreground object's attributes and embodies object feature statistics learned adaptively as the model deforms toward object boundary. With additional information about the overall image feature statistics, the model also estimates the background statistics so that a Bayes classifier can be applied to predicting dynamically the object region (Section 2.3). In its object-region prediction framework, the model is capable of probabilistically integrating constraints from multiple visual cues.

The AVM model'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 (Section 3.3), 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 model 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 reason 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. In our experimental evaluation on various 3D noisy medical images, AVM deforms much faster than level set methods such as *Active contours without edges*(ACWE) [2]. Even though there are thousands of vertices (or nodes) on the surface of 3D AVM, the model's deformation speed is still less than 20 seconds per iteration. The converged 3D AVM model is directly a smooth mesh representing the segmented object surface. Therefore it does not need any post-processing step such as surface reconstruction, as required by other methods including region growing, level set, graph cut and MRF.

The remainder of the paper is organized as follows. In Section 2, we introduce the representation and the boundary prediction module of 3D AVM. In Section 3, we present the deformation of AVM based on gradient and object region information. In Section 4, experimental results are presented. We conclude the paper and discuss future work in Section 5.

2. The Active Volume Model

An AVM is a deforming solid that minimizes internal and external energy [19]. The internal constraint ensures the model has smooth boundary or surface. The external constraints come from image data, prior, and user defined features. In this section, we introduce the representations of 3D AVM, and introduce a novel object boundary prediction module by AVM.

2.1. Explicit Shape Representation of 3D AVM

The shape of a 2D AVM is defined similarly to active contours [10], as a spline curve that has associated elasticity and rigidity. In 3D, the surface of a model can be defined by a mapping

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

and the model's internal energy in 3D is defined as a Sobolev smoothing term

$$\begin{array}{c} E_{int} = \int_\Lambda (\alpha(s)|\mathbf{v}_s|^2 + \alpha(r)|\mathbf{v}_s|^2 + \beta(s)|\mathbf{v}_{ss}|^2 \\ +\beta(r)|\mathbf{v}_{rr}|^2 + \beta(sr)|\mathbf{v}_{sr}|^2) ds dr \end{array} \tag{2}$$

In [3], Cohen *et al*. used tessellation to build a 3D finite element method (FEM) surface. 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. Figure 1 shows the disadvantage of using a tessellation ellipsoid—strong internal forces generated near the poles adversely affect the model's deformation.

In order to solve the above problem in tessellation FEM mesh and enable the model to match closely object boundary, 3D AVM adopts a polyhedron mesh as the model representation which places vertices regularly on the model. More specifically, a 3D AVM is considered as an elastic solid and defined as a finite element triangulation Λ, which can be tetrahedron, octahedron or icosahedron. Using the finite element method, the internal energy function can be written compactly as:

$$E_{int} = \frac{1}{2} \int_\Lambda (\mathbf{B}\mathbf{v})^T D (\mathbf{B}\mathbf{v}) d\Lambda \tag{3}$$

where $\mathbf{B}$ is the differential operator for the model vertices $\mathbf{v}$ and D is the stress matrix (or constitutive matrix).

2.2. Implicit Shape Representation of 3D AVM

The implicit shape representation using level set [14] makes a model's shape an "image", which greatly facilitates

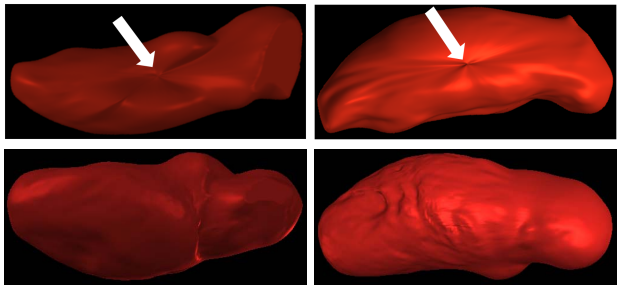

Figure 1. Comparing heart Left Ventricle segmentation by tessellation and by triangulation models. (A) Tessellation model with 400×120 vertices. Notice that the two poles exert very strong force to drag other vertices which causes the incorrect result. (B) Finite Element Triangulation model with 40,962 vertices. (1)&(2) Two different views.

the integration of boundary and region information [9]. For 3D AVM, we compute the implicit representation of model shape to be used in region-based external energy terms. The Euclidean distance transform is applied to embed implicitly an evolving model's surface in 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 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 definition the following equation is considered:

$$\Phi_{\mathcal{M}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \partial\mathcal{R}_{\mathcal{M}} \\ +ED(\mathbf{x},\mathcal{M}) > 0, & \mathbf{x} \in \mathcal{R}_{\mathcal{M}} \\ -ED(\mathbf{x},\mathcal{M}) < 0, & \mathbf{x} \in [\Omega - \mathcal{R}_{\mathcal{M}}] \end{cases} \quad (4)$$

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

2.3. AVM Model's Boundary Prediction Module

Different from most of deformable models, one of the novel features of AVM is its adaptive object boundary prediction scheme. The model alternates between two operations: deform according to the current object boundary prediction, and predict according to current appearance statistics of the model. Using this on-line prediction mechanism, the expected object information updates automatically while the model deforms. And long-range external forces are generated from the predicted object boundary to effectively deform the model.

External constraints from any sources can be accounted 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 (*non_object* class). Suppose we have n independent external constraints derived from image information, the feature used in the kth constraint is f_k, $L(\mathbf{x})$ 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:

$$\begin{aligned} &Pr(L(\mathbf{x})|f_1, f_2, ..., f_n) \\ &= (Pr(f_1, f_2, ..., f_n|L(\mathbf{x}))Pr(L(\mathbf{x}))/(Pr(f_1, f_2, ..., f_n)) \\ &\propto Pr(f_1|L(\mathbf{x}))Pr(f_2|L(\mathbf{x}))...Pr(f_n|L(\mathbf{x}))Pr(L(\mathbf{x})) \end{aligned} \quad (5)$$

For each independent module, the probability $Pr(f_k|L(\mathbf{x}))$ is estimated based on the AVM model's current statistics about feature f_k as well as the overall feature statistics in the image. The derivation is as follows.

$$\begin{aligned} Pr(f_k) &= Pr(f_k, L(\mathbf{x}) = +1) + Pr(f_k, L(\mathbf{x}) = -1) \\ &= Pr(f_k|L(\mathbf{x}) = +1)Pr(L(\mathbf{x}) = +1) \\ &\quad + Pr(f_k|L(\mathbf{x}) = -1)Pr(L(\mathbf{x}) = -1) \end{aligned} \quad (6)$$

Assuming the current AVM model embodies priors learned dynamically about the foreground object, we approximate the probabilistic distribution of feature f_k in the object, $Pr(f_k|L(\mathbf{x}) = +1)$, by the feature's distribution in the current AVM model. The overall distribution of f_k in the image, $Pr(f_k)$, is also known. Both probability density functions, $Pr(f_k|L(\mathbf{x}) = +1)$ and $Pr(f_k)$, are estimated using a nonparametric kernel-based density estimation method [9]. The p.d.fs are differentiable and can represent complex multi-modal distributions. Therefore, we can now reason about the feature distribution in the background

$$Pr(f_k|L(\mathbf{x}) = -1) = \frac{Pr(f_k) - Pr(f_k|L(\mathbf{x}) = +1)Pr(L(\mathbf{x}) = +1)}{Pr(L(\mathbf{x}) = -1)}. \quad (7)$$

The prior independent of image features, $Pr(L(\mathbf{x}))$, in Equations 5 and 7 can be assumed uniform: $Pr(L(\mathbf{x}) = +1) = 0.5$ and $Pr(L(\mathbf{x}) = -1) = 0.5$. Spatially-varying prior is another choice. For instance, a Gaussian distance model can be adopted so that pixels close to the AVM model have higher prior probability being part of the object.

Once the posterior probabilities $Pr(L(\mathbf{x})|f_1, f_2, ..., f_n)$ are estimated, we apply the Bayesian decision rule to obtain a binary map P_B whose foreground represents the predicted object region. That is, $P_B(\mathbf{x}) = 1$ (object pixel) if $Pr(L(\mathbf{x}) = +1|f_1, f_2, ..., f_n) \geq Pr(L(\mathbf{x}) = -1|f_1, f_2, ..., f_n)$, and $P_B(\mathbf{x}) = 0$ otherwise. The probability of error for the decision at pixel $\mathbf{x}$ is $\min(Pr(L(\mathbf{x}) = +1|f_1, f_2, ..., f_n), Pr(L(\mathbf{x}) = -1|f_1, f_2, ..., f_n))$.

In this paper, we show that by considering two types of features—pixel intensity $i(\mathbf{x})$ and pixel gradient magnitude $g(\mathbf{x})$, and assuming a uniform prior for $Pr(L(\mathbf{x}))$, the above framework generates reasonable estimates of background feature statistics (Equation 7) and consistently gives good predictions of the object region on a variety of medical images.

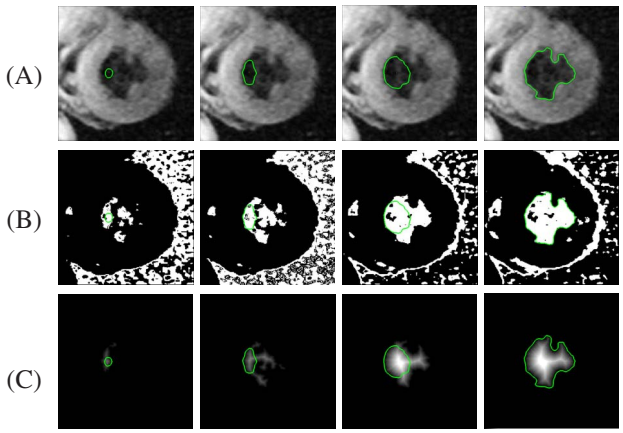

Figure 2. Left Ventricle endocardium segmentation using a 2D active volume model. (A) The model drawn on the original heart image. (B) The binary map P_B estimated by the boundary prediction module. (C) Distance transform of the ROI boundary. (1) Initial model. (2) The model after 8 iterations. (3) The model after 18 iterations. (4) Final converged result after 26 iterations.

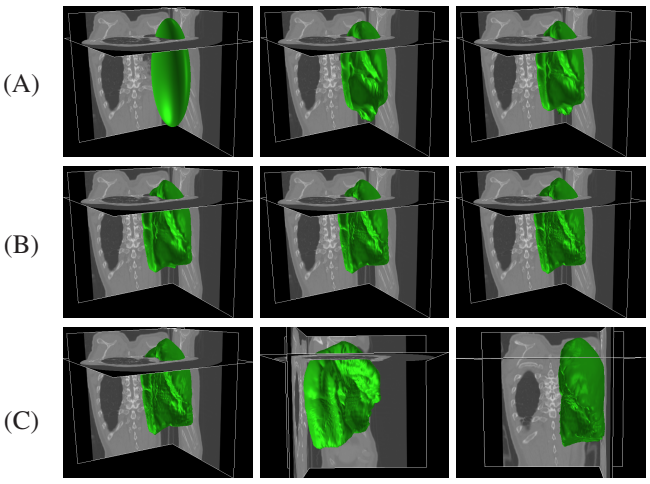

Figure 3. Left Lung segmentation using 3D active volume model. (A)(1) Initial model, (2) after 3 iterations, (3) after 6 iterations. (B)(1) after 12 iterations, (2) after 18 iterations, (3) after 24 iterations. (C) Final converged result after 30 iterations.

Given the binary map P_B, we 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 Region Of Interest (ROI) and its boundary represents the predicted object boundary. Due to noise, there might be small holes that need to be filled before extracting the boundary of the ROI, R. The progressive ROI updating can be clearly seen from a 2D AVM example in Figure 2. In the figure, the ROI evolves according to the changing object appearance statistics (estimated by current volumetric model's statistics). And the image forces (Section 3) generated by the ROI region energy term deform the model to converge to the object boundary.

The initialization of a 3D AVM model is very flexible. We can either initialize the model with a simple shape, for instance an ellipsoid, or initialize it with a standard atlas that has similar structure as the object to be segmented. One example of the 3D AVM can be seen in Figure 3 by using the chest CT data provided by [15]. The initialization is a simple ellipsoid near the object. Notice that the model only partially overlaps the predicted ROI. External forces from the adaptively changing ROI deform the model to finally converge on the object boundary.

3. Model Dynamic Deformation

In order to fit to the boundary of an object, the AVM model 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:

$$E = E_{int} + E_{ext} = E_{int} + (E_g + k_{reg} \cdot E_R) \quad (8)$$

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 Eqn. 13).

3.1. 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:

$$E_g = \int_\Lambda F_g(\mathbf{x}) d\Lambda \quad (9)$$

$$F_g = \begin{cases} D^2_{edge}, & edge\ distance\ map \\ -|\nabla I|^2, & gradient\ magnitude\ map \end{cases} \quad (10)$$

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

3.2. The Object Region Data Term

Given the current model, a Region of Interest (ROI) R representing the predicted object region can be computed by the prediction module introduced in Section 2.3. Let us denote the signed distance transform of the ROI shape as Φ_R. Combining Φ_R and the current model's implicit shape representation Φ_M (Section 2.2), the region-based external energy term is defined as:

$$E_R = \int_\Lambda \Phi_M(\mathbf{x}) \Phi_R(\mathbf{x}) d\Lambda \quad (11)$$

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.

3.3. The Model's Deformation

Minimization of the 3D AVM energy function can be achieved by solving the following linear system

$$A_{3D} \cdot V = L_V; \tag{12}$$

where A_{3D} is the stiffness matrix derived from Equation 3 by using the basis function in Equation 13. A_{3D} 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 gradient data term and region data term. To facilitate the computation, 3D AVM adopts a continuous piecewise linear basis function,

$$\phi_j(v_i) = \delta_{ij} \equiv \begin{cases} \gamma & i = j \\ 0 & i \neq j \end{cases} \tag{13}$$

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 12 can be solved by using finite differences [3]. After initializing the 3D AVM, the final converged result can be obtained iteratively based on equation:

$$(V^t - V^{t-1})/\tau + A_{3D} \cdot V^t = L_{V^{t-1}} \tag{14}$$

where V^0 is the initial AVM vertex vector and τ is the time step size. Equation 14 can be written in a finite differences formulation, which yields

$$\begin{array}{c} M \cdot V^t = V^{t-1} + \tau L_{V^{t-1}} \\ M = (I + \tau A_{3D}) \end{array} \tag{15}$$

Using Equation 15, we adopt the following steps to deform the 3D AVM toward matching the desired object boundary.

1. Initialize the AVM, stiffness matrix A_{3D}, step size τ, and calculate the gradient magnitude or edge map.
2. Compute Φ_M based on the current model; predict R by applying the Bayesian Decision rule to binarizing the estimated *object* probability map, and compute Φ_R.
3. Deform the model according to Equation 15.
4. Adaptively increase the external force factor in Equation 8, decrease the step size τ in Equation 15 and reduce γ in Equation 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.

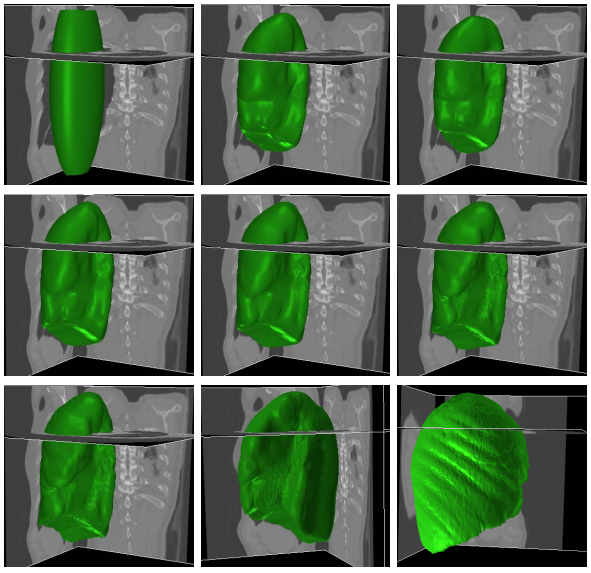

Figure 4. Right Lung segmentation using 3D AVM with a FE triangulation mesh consisting of 32,770 control vertices. (A)(1) Initial model, (2) after 3 iterations, (3) after 9 iterations. (B)(1) after 15 iterations, (2) after 21 iterations, (3) after 27 iterations. (C) Final converged result after 33 iterations.

4. Experimental Results

In this section, 3D AVM is used to segment human organs from volumetric medical images.

First, we put the model into a thorax CT stack to segment the right lung. The model was initialized as an ellipsoid whose long axis is perpendicular to the axial image plane. Figure 4 shows the initial model, segmentation results in several steps and the final converged result.

Then we tested the model by segmenting the left ventricle (LV) in a CT stack. The model is again initialized as an ellipsoid and is placed in the stack after a set of rotation operations so that the ellipsoid's major axis roughly aligns with the LV's long axis. Some boundary condition is also specified so that the model does not deform beyond the z-range defined by the top and bottom slices of the stack. Since the shape of the human heart is more complex than that of the lung, we use a 3D AVM with more vertices than the one used for lung segmentation. We also adopt stricter smoothness constraints. Figure 5 shows the initial model, segmentation results in several steps and the final converged result.

On a PC workstation with an Intel Duo Core 3GHz E6850 processor, we compared 3D AVM with *Geodesic Active Contours* (GAC) [1] and *Level Set Evolution Without Reinitialization* (LSEWR) [12] by measuring the running times and validating the segmentation results using expert ground truth markings. Figure 6 shows the final triangulation surface of 3D AVM, compared with the final converged results of GAC and LSEWR after surface reconstruction.

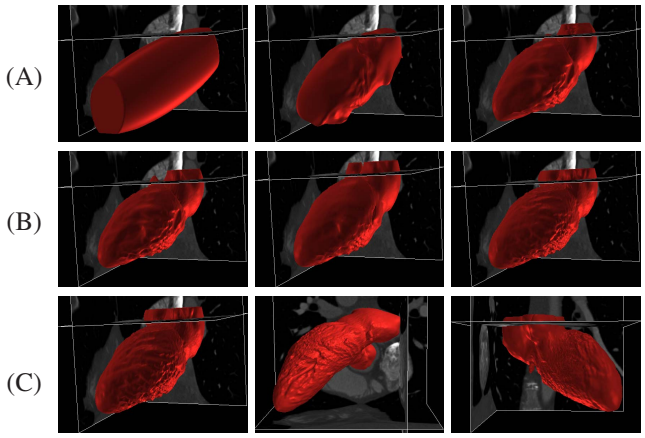

Figure 5. Left Ventricle segmentation using 3D AVM with 65,538 control vertices. (A)(1) Initial model, (2) after 3 iterations, (3) after 9 iterations. (B)(1) after 15 iterations, (2) after 21 iterations (3) after 30 iterations. (C) Final converged result after 39 iterations.

Table 1 presents the *sensitivity* (P), *specificity* (Q) and *Dice Similarity Coefficient* (DSC) [18] values and running times for various experiments.

Since 3D AVM model is represented by FE triangulation, comparing with GAC and LSEWR, smooth surfaces can be obtained directly by AVM without any post-processing, such as morphological operations and surface reconstruction. 3D AVM also preserves topology during deformation. As shown for a 2D slice projection example in Figure 8, level set segmentation results (Fig. 8B and C) have complex topology, while 3D AVM result is one smooth surface delineating lung boundary despite tumor presence. Thus 3D AVM is very suitable for extracting organ boundaries from volumetric medical images.

Furthermore, in our experiments, 3D 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 3D 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 7 shows in a left and right lungs segmentation example that the 3D 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 8 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, 2D AVM 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, 3D AVM model 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 Figure 9. The 3D simulated MRI brain images are provided by BrainWeb (http://www.bic.mni.mcgill.ca/brainweb/). The MRI stack is of T_1 modality, has $1mm$ slice thickness, 3% noise level and 20% INU. Both models (one for GM and one for WM) are initialized as ellipsoids with strong initial smoothness constraint. As the

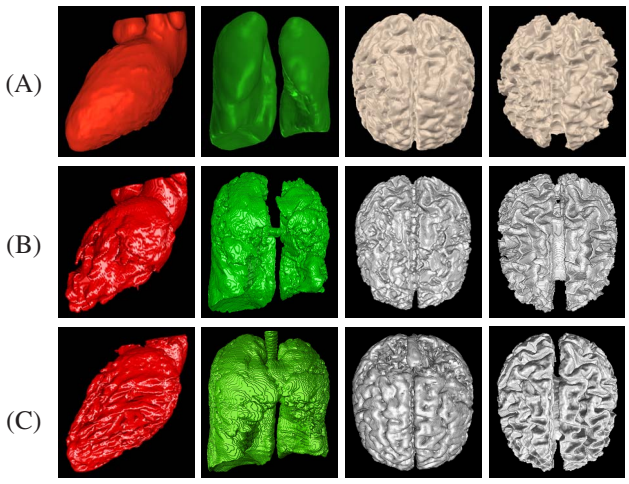

Figure 6. Comparing 3D AVM with Geodesic Active Contours (GAC) and Level Set Evolution Without Reinitialization (LSEWR). (A) 3D AVM, (B) GAC, (C) LSEWR. (1) heart LV segmentation, (2) lung segmentation, (3) brain GM segmentation, (4) brain WM segmentation.

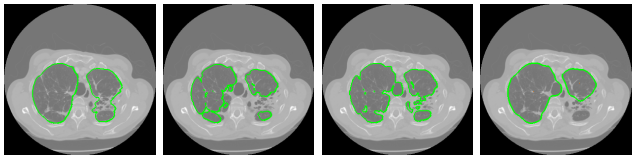

Figure 8. 3D AVM preserves topology and is good at avoiding leakage and overcoming local minima. (1) 2D projection of 3D AVM model, (2) 2D projection of Geodesic Active Contours' result, (3) 2D projection of Level Set Evolution Without Reinitialization's result, (4) 2D AVM result.

Table 1. Quality evaluation and performance comparison

		3D AVM				GAC				LSEWR			
Organ	size (voxels)	P	Q	DSC	Time	P	Q	DSC	Time	P	Q	DSC	Time
Lungs	$(307 \times 307 \times 187)$	93.6	99.8	95.2	1000s	75.7	99.9	85.8	2149s	91.4	99.7	94.6	1840s
Heart	$(256 \times 256 \times 153)$	91.8	99.6	94.3	1044s	78.0	99.8	87.6	1752s	80.1	99.9	88.5	1452s
GM	$(181 \times 217 \times 180$	87.6	98.3	91.5	1620s	75.7	99.0	85.0	2332s	86.4	99.9	89.4	621s
WM	$(181 \times 217 \times 180)$	76.8	96.2	78.3	1833s	66.9	95.5	72.5	2453s	81.1	99.8	89.0	643s

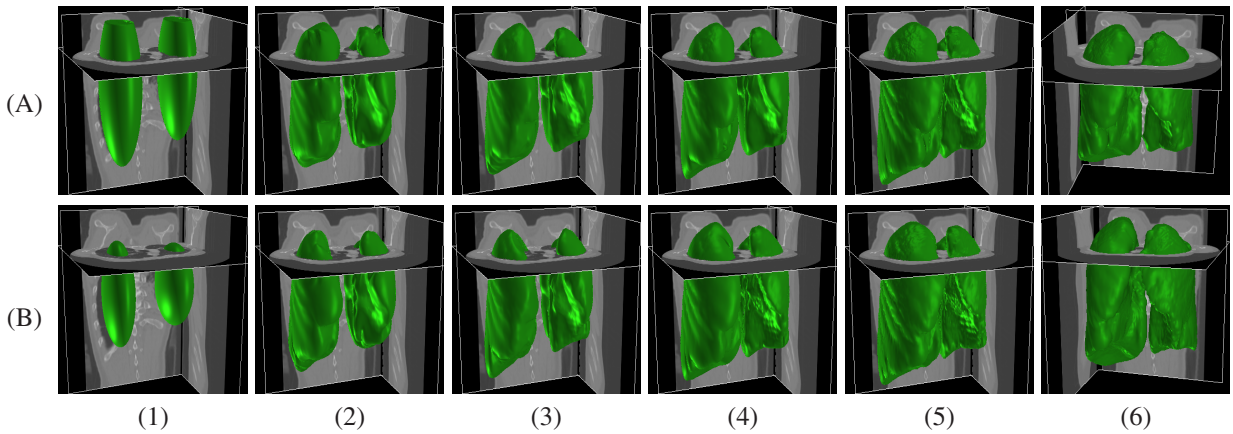

Figure 7. 3D AVM has less dependence on model initialization and parameter settings. (A)(1) Initial 3D AVM model, (2) after 3 iterations, (3) after 12 iterations, (4) after 21 iterations, (5) final converged result after 33 iterations, (6) final result viewed from a different viewpoint. (B)(A)(1) A different 3D AVM model initialization on the same image stack, (2) after 3 iterations, (3) after 12 iterations, (4) after 24 iterations, (5) final converged result after 36 iterations, (6) final result viewed from a different viewpoint.

models are getting closer to the approximated object boundary, the models decrease the smoothness constraint automatically based on the deformation strategy described in Section 3.3. Then a lot of details on the object surfaces appear on the models.

5. Conclusion and Future Work

In this paper, we proposed a novel active volume model, which is a natural extension of parametric deformable models to integrate object appearance and region information. The main contributions include: (1) a clean formulation to integrate online learning and region statistics into active contours and surfaces, which provides flexible initialization and rapid convergence, (2) the finite differences optimization framework that enables very fast gradient- and appearance-statistics based model deformations, (3) the combination of multiple sources of information in a unified framework for predicting object region and boundary. Using various experiments on 3D medical images, we demonstrate that the AVM model can perform segmentation efficiently and reliably on CT and MRI images. However, due to the local smoothness of simplex-mesh, it is still hard for the model to reach details on branch structures (e.g. WM in Figure 9B). In the future, we plan to address this problem by reparameterizing the model near branches since vertices in such areas are sparser than those on the main body.

6. Acknowledgements

The authors would like to thank Prof. Leon Axel (NYU) for providing the heart CT volume data and also acknowledge stimulating discussions with Prof. Dimitris Metaxas (Rutgers), Junzhou Huang (Rutgers) and Yaoyao Zhu (Lehigh). This work was supported by NSF grant IIS-0812120.

References

[1] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Internal Journal of Computer Vision*, 22, 1997.

[2] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. on Image Processing*, 10:266–277, 2001.

[3] L. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15:1131–1147, 1993.

[4] T. Cootes, G. Edwards, and C. Taylar. Active appearance models. *Proc. Of European Conf. on Computer Vision*, 2:484–498, 1998.

[5] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape model - their training and application. *Computer Vision and Image Understanding*, 61:38–59, 1995.

[6] M. Costa, H. Delingette, S. Novellas, and N. Ayache. Automatic segmentation of bladder and prostate using coupled 3D deformable models. In *MICCAI (1)*, pages 252–260, 2007.

[7] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color,

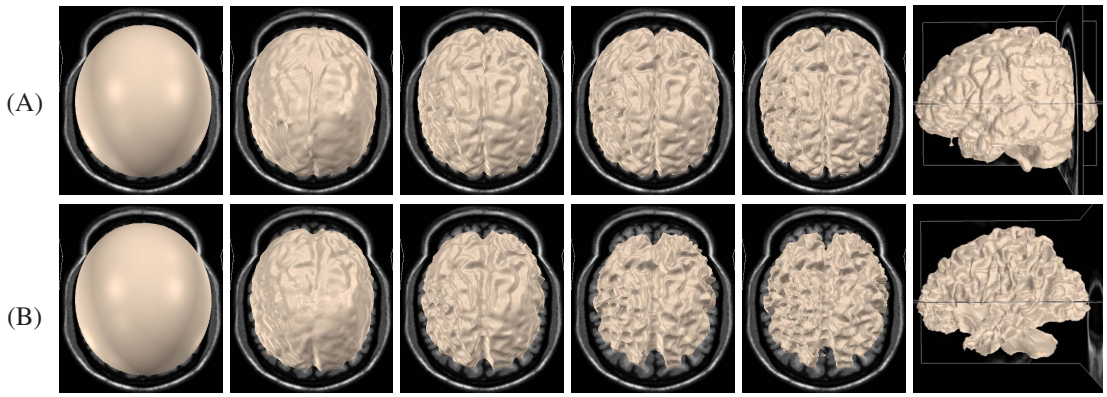

Figure 9. GM and WM segmentation using 3D AVM. The GM and WM model surfaces each has 131,074 control vertices. (A)(1) Initial model of GM, (2) after 3 iterations, (3) after 12 iterations, (4) after 24 iterations, (5)&(6) final converged result after 36 iterations. (B)(1) Initial model of WM, (2) after 3 iterations, (3) after 12 iterations, (4) after 24 iterations, (5)&(6) final converged result after 39 iterations.

texture, motion and shape. *Int. J. Comput. Vision*, 72(2):195–215, 2007.

[8] X. Fan, P.-L. Bazin, and J. Prince. A multi-compartment segmentation framework with homeomorphic level sets. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–6, June 2008.

[9] X. Huang, D. Metaxas, and T. Chen. Metamorphs: Deformable shape and texture models. In *CVPR*, pages 496–503, 2004.

[10] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int'l Journal on Computer Vision*, 1:321–331, 1987.

[11] T. Kohlberger, D. Cremers, M. Rousson, R. Ramaraj, and G. Funka-Lea. 4D shape priors for a level set segmentation of the left myocardium in spect sequences. In *MICCAI (1)*, pages 92–100, 2006.

[12] C. Li, C. Xu, C. Gui, and M. D. Fox. Level set evolution without re-initialization: A new variational formulation. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Conference on*, 1:430–436, 2005.

[13] H. Li and A. Yezzi. Local or global minima : Flexible dual-front active contours. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(1):1–14, 2007.

[14] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.

[15] R. A. McLaughlin, J. P. Williamson, M. J. Phillips, J. J. Armstrong, S. Becker, D. R. Hillman, P. R. Eastwood, and D. D. Sampson. Applying anatomical optical coherence tomography to quantitative 3d imaging of the lower airway. *Optics Express*, 16:17521–17529, 2008.

[16] D. N. Metaxas and D. Terzopoulos. Shape and non-rigid motion estimation through physics-based synthesis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(6):580–591, 1993.

[17] N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *The International Journal of Computer Vision*, 46(3):223–247, 2002.

[18] A. Popovic, M. de la Fuente, M. Engelhardt, and K. Radermacher. Statistical validation metric for accuracy assessment in medical image segmentation. *International Journal of Computer Assisted Radiology and Surgery*, 2:169–181, 2007.

[19] T. Shen, Y. Zhu, X. Huang, J. Huang, D. Metaxas, and L. Axel. Active volume models with probabilistic object boundary prediction module. *MICCAI '08*, pages 331–341, 2008.

[20] L. Staib and J. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, 1992.

[21] O. Subakan and B. Vemuri. Image segmentation via convolution of a level-set function with a Rigaut kernel. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–6, June 2008.

[22] C. Xu and J. Prince. Snakes, shapes and gradient vector flow. *IEEE Trans. on Image Processing*, 7:359–369, 1998.

[23] J. Yang and J. Duncan. 3D image segmentation of deformable objects with joint shape-intensity prior models using level sets. *Medical Image Analysis*, 8(3):285–294, September 2004.

[24] S. Zhu and A. Yuille. Region Competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.

[25] Y. Zhu, X. Papademetris, A. Sinusas, and J. Duncan. Segmentation of left ventricle from 3D cardiac MR image sequences using a subject-specific dynamical model. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.