

3D Medical Image Segmentation by Multiple-Surface Active Volume Models

Tian Shen and Xiaolei Huang

Department of Computer Science and Engineering,
Lehigh University, Bethlehem, PA 18015, USA

Abstract. In this paper, we propose Multiple-Surface Active Volume Models (MSAVM) to extract 3D objects from volumetric medical images. Being able to incorporate spatial constraints among multiple objects, MSAVM is more robust and accurate than the original Active Volume Models [11]. The main novelty in MSAVM is that it has 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. These two functions help MSAVM not only overcome local minima but also avoid leakage. Because of the implicit representation of AVM, the spatial information can be calculated based on the model's signed distance transform map with very low extra computational cost. The MSAVM thus has the efficiency of the original 3D AVM but produces more accurate results. 3D segmentation results, validation and comparison are presented for experiments on volumetric medical images.

1 Introduction

Object boundary extraction is an important task in medical image analysis. A large variety of 2D algorithms have been proposed over the last few decades. 3D volumetric medical images are usually analyzed as a sequence of 2D image slices [4] due to concerns over the exponential increase in computational cost in 3D.

PDE-based segmentation methods became popular after Kass et al. proposed the Snakes [2]. Because local image gradient constraints in Snakes make the model sensitive to initialization and noise, Xu et al. proposed Gradient Vector Flow (GVF) [3], which increases the attraction range of the original Snakes. Region analysis strategies [4] have also been incorporated in PDE-based models to improve robustness to noise. Another class of deformable models is level set based geometric models [5]. This approach evolves the model based on the theory of curve evolution, with speed function specifically designed to incorporate image gradient information. Because level set models are topologically free and can be easily used in any dimension, they are widely used in tubular structure and 3D cortex segmentation tasks.

In noisy medical images, statistical modeling approaches such as ASM [6] and AAM [7] are adopted by adding constraints learned offline. Integrating high-level prior knowledge, these models deform in ways constrained by the training data

and thus are often more robust in image interpretation. However, training data collection and landmark annotation are laborious in building statistical models, especially in 3D.

In many medical image applications, 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. Clearly, integrating this high-level spatial constraint into the segmentation model will further improve accuracy and robustness. A 2D method [8] segments left ventricular Epi- and Endocardial borders using coupled active contours but needs a precise manual initialization. In 3D, Zeng et al. [9] incorporated spatial constraints about gray matter and white matter into the level set framework which greatly improved cortex segmentation accuracy. In [1], a graph-theoretic approach detects multiple interacting surfaces by transforming the problem into computing a minimum s-t cut. Deformation of multiple surfaces in [10] has intersurface proximity constraints which allow each surface to guide other surfaces into place. All of the three 3D methods [9, 1, 10] require manually specifying the expected thickness between surfaces as model-based constraint.

Active volume models (AVM) were recently proposed [11] to segment 3D objects directly from volumetric medical image dataset. Compared with active contours and ASM/AAM, the AVM is a “self-contained” generative object model that does not require off-line training. In this paper, we propose the Multiple-Surface AVMs to segment coupled medical objects simultaneously. Instead of setting up a fixed distance constraint during initialization, multiple-surface AVMs dynamically update the distance constraint between the interacting surfaces based on current model surfaces’ spatial inter-relations. Integrating the distance constraint strategy with other energy terms based on image gradient and region information, 3D MSAVMs are more robust to initial positions and yield more accurate segmentation results.

2 Methodology

2.1 Review of 3D AVM and Boundary Prediction Module

An AVM is a deformable solid that minimizes internal and external energy. The internal constraint ensures the model has smooth boundary surface. The external constraints come from image data, prior, and/or user-defined features. Different from most of deformable models, one of the novel features of AVM is its unsupervised 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. Next we introduce the 3D AVM model representation and its boundary prediction module.

3D AVM [11] 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 A , which can be tetrahedron, octahedron or icosahedron. Using the finite element method (FEM),

the internal energy function can be written compactly as:

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

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

External constraints from any sources can be accounted by probabilistic integration. Suppose we have n independent constraints from image information. 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). The feature used in the k th constraint is f_k , $L(\mathbf{x})$ denotes the label of a pixel \mathbf{x} . AVM combines the multiple independent modules and applies the Bayes rule to evaluate the final confidence rate:

$$\begin{aligned} Pr(L(\mathbf{x})|f_1, f_2, \dots, f_n) &= (Pr(f_1, f_2, \dots, f_n|L(\mathbf{x}))Pr(L(\mathbf{x})))/(Pr(f_1, f_2, \dots, f_n)) \\ &\propto Pr(f_1|L(\mathbf{x}))Pr(f_2|L(\mathbf{x}))\dots Pr(f_n|L(\mathbf{x}))Pr(L(\mathbf{x})) \end{aligned} \quad (2)$$

For each independent module, the probability $Pr(f_k|L(\mathbf{x}))$ is estimated based on the AVM's current statistics about feature f_k as well as the overall feature statistics in the image [11]. Once the posterior probabilities $Pr(L(\mathbf{x})|f_1, f_2, \dots, f_n)$ are estimated, we apply the Bayesian decision rule to obtain a binary map P_B whose foreground represents the Region of Interest(ROI). That is, $P_B(\mathbf{x}) = 1$ (object pixel) if $Pr(L(\mathbf{x}) = +1|f_1, f_2, \dots, f_n) \geq Pr(L(\mathbf{x}) = -1|f_1, f_2, \dots, f_n)$, and $P_B(\mathbf{x}) = 0$ otherwise. Let signed distance transform of the ROI shape be Φ_R . Combining Φ_R and Φ_M , a region-based external energy term is defined as:

$$E_R = \int_{\Lambda} E_R(\mathbf{v}) d\Lambda = \int_{\Lambda} \Phi_M(\mathbf{v}) \Phi_R(\mathbf{v}) d\Lambda \quad (3)$$

The multiplicative term provides two-way balloon forces that deform the model toward the predicted ROI boundary.

The external energy of AVM also consists of a gradient term $E_g = -|\nabla I|^2$. Putting together internal and external energy terms, the overall energy function for AVM is defined as:

$$E = E_{int} + E_{ext} = E_{int} + (E_g + k_{reg} \cdot E_R) \quad (4)$$

where k_{reg} is a constant that balances the contributions of the gradient term and the region term.

2.2 Multiple-Surface Active Volume Models

Due to limitations in medical imaging techniques, in some regions of an image, there may not be enough information (e.g. contrast) that can be derived from the image to clearly distinguish an object boundary or surface. Therefore, a single surface based deformable model may stop at local minima or leak to converge to

an outer boundary. When spatial constraints between multiple surfaces are available, such information can help deform all interacting surfaces simultaneously with better accuracy, in a multiple surface based model framework.

The Multiple-Surface AVM we propose is initialized as several AVMs inside an outer AVM. Let i, j be the surface indices, the mean Euclidean distance value of the i th surface, M_i , from other surfaces is defined as:

$$\overline{dist}_i = \frac{\int_{\Lambda} dist_v d\Lambda}{\int_{\Lambda} d\Lambda}, \text{ where } dist_v = \{d | d = \min(|\Phi_{M_j}(\mathbf{v})|), \forall j, j \neq i\} \quad (5)$$

Φ_{M_j} is the signed distance transform of the j th surface, M_j .

To deform the multiple surfaces simultaneously with adaptive spatial constraints, we construct two distance-related Gaussian Mixtures (GM) functions to modulate the external force at each vertex. For a vertex v on the i th surface, its minimum distance value to all other surfaces $dist_v$ can be calculated based on Eq. (5). Then the GM distance functions at this vertex are defined in Eq. 6 and illustrated in Figure 1.

$$\begin{aligned} g_D(dist_v) &= (1 + \alpha) - e^{-(dist_v - \overline{dist}_i)^2 / 2\sigma_1^2} - \alpha e^{-(dist_v - \overline{dist}_i)^2 / 2\sigma_2^2} \\ g_R(dist_v) &= e^{-(dist_v - \overline{dist}_i)^2 / 2\sigma_1^2} + \alpha e^{-(dist_v - \overline{dist}_i)^2 / 2\sigma_2^2} \end{aligned} \quad (6)$$

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

Then the overall energy function of MSAVM is defined as:

$$E = E_{int} + \int_{\Lambda} (E_g(\mathbf{v}) + g_R(dist_v) \cdot E_R(\mathbf{v}) + g_D(dist_v) \cdot E_{dist_v}) \quad (7)$$

where $E_{dist_v} = dist_v$ is the energy term enforcing the distance constraint.

According to $g_R(dist_v)$ and $g_D(dist_v)$ (Fig. 1(a-b)), if $dist_v$ is close to the surface's mean distance (from other surfaces), \overline{dist}_i , then the region term $E_R(\mathbf{v})$ makes more contribution toward the surface's local deformation near v ; conversely, if $dist_v$ is far away from \overline{dist}_i , which means the local surface near the vertex may be stuck at local minima or have a leakage, the distance constraint, E_{dist_v} , is given more power to deform the surface and guide it into place.

Different from the distance constraint function in [9], 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 balance the contributions of region term and spatial constraint term. MSAVM thus has broader applications. It not only can be used for brain segmentation, but also has very good performance in extracting ventricles from heart and lungs from the thorax even though distances between these coupled ventricular surfaces vary greatly. Figure 1c and 1d show two segmentation results by distance-color (DC) mapping the spatial distance information into color space.

Instead of setting the spatial constraint manually or empirically, we update \overline{dist} based on the spatial relationship among current model surfaces. After getting \overline{dist} , we shift $g_R(dist)$ and $g_D(dist)$ functions accordingly to make sure

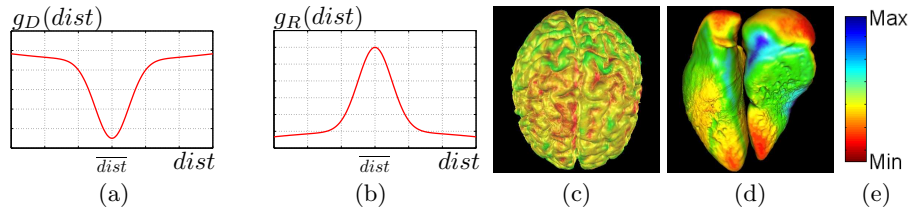


Fig. 1. (a)(b) Functions to balance the contributions of region term and spatial constraint term, (c)DC mapping of brain gray matter segmentation; the range for distance between Gray and white matters is 2~15 voxels, (d)DC mapping of heart segmentation; the distance range is 2~25 voxels, (e)the color bar used to map the spatial information.

the new \overline{dist} still corresponds to the extrema of these functions. This unsupervised learning strategy, both for multiple surface spatial constraint and for region appearance statistics, allows MSAVM to have flexible initialization and fast convergence.

2.3 The Model's Deformation

Minimization of the MSAVM's energy function can be achieved by solving several independent linear systems. For the i th surface,

$$A_{3D} \cdot V_i = L_{V_i}; \quad (8)$$

where A_{3D} is the stiffness matrix derived from Eq. 1. V_i is the vector of vertices on the i th surface. L_{V_i} is the corresponding external force vector. Using the finite differences method [12], we adopt the following steps to deform the MSAVM to match the desired object surfaces.

1. Initialize the MSAVM, stiffness matrix A_{3D} and \overline{dist} for each surface.
2. For each surface, compute Φ_M based on the current model; predict R and compute Φ_R ; and update \overline{dist} based on Eq. (5) and shift $g_R(dist_v)$ and $g_D(dist_v)$ according to \overline{dist} .
3. Deform MSAVM according to Eq. 8.
4. Adaptively decrease the degree of surface stiffness/smoothness.
5. Repeat steps 2-4 until convergence.

3 Experimental Results

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 2 shows the 3D DC mapping images during deformation. A 2D coronal projection view is also included in 2f to show the initial model and converged result.

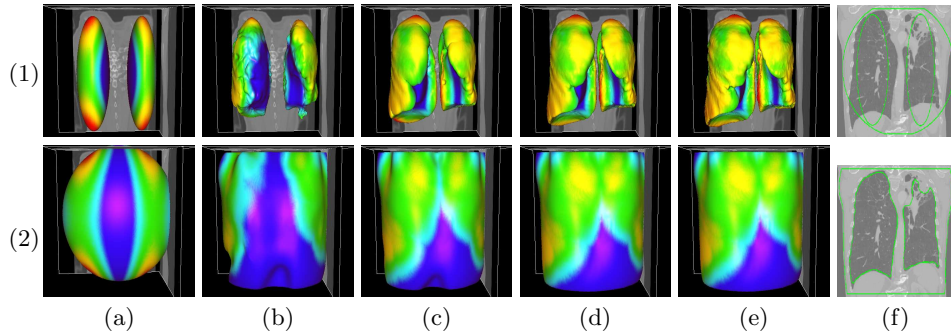


Fig. 2. DC mapping of lung surfaces segmentation using MSAVM, the distance range is 3-45 voxels. (1)(a)-(e)deformation progress of inner surfaces, (2)(a)-(e)outer surface; (a)Initial model after (b)3, (c)9, (d)21, (e)26(converged result) iterations; (1)(f)initial model in a 2D slice, (2)(f)converged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

Then we experimented with the model on segmenting heart surfaces in a cardiac CT stack. The MSAVM model 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. Figure 3 and Figure 4 show the deformation steps of the heart from two 3D viewpoints. 2D sagittal and coronal projection views are also provided in Figure 3f and Figure 4f. 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 on-line predicted object boundary with spatial constraints, MSAVM can overcome the inhomogeneity problem and extract accurately the multiple cardiac surfaces.

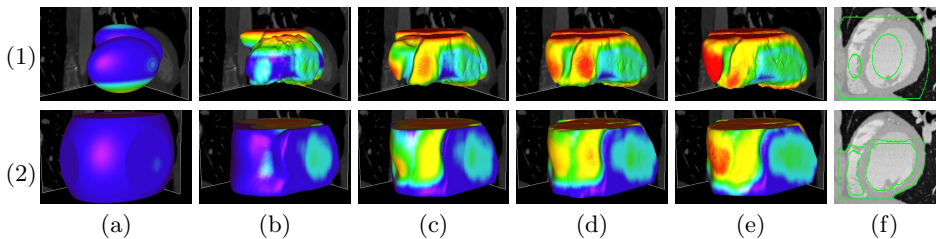


Fig. 3. DC mapping of heart segmentation using MSAVM viewed from the right, the distance range is 2-25 voxels, (1)(a)-(e)deformation progress of inner surfaces, (2)(a)-(e)DC mapping for outer surface; (a)Initial model after (b)3, (c)9, (d)21, (e)27(converged result) iterations; (1)(f)initial model in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

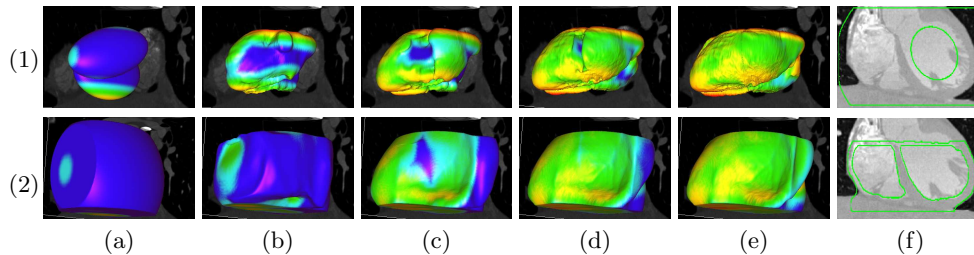


Fig. 4. DC mapping of heart segmentation using MSAVM viewed from the left, the distance range is 2-25 voxels, (1)(a)-(e)deformation progress of inner surfaces, (2)(a)-(e)outer surface; (a)Initial model after (b)3, (c)9, (d)21, (e)27(converged result) iterations; (1)(f)initial model in a 2D slice, (2)(f)converged result in a 2D slice.

Table 1 shows the running times and quantitative evaluation of *sensitivity* (P), *specificity* (Q) and *dice similarity coefficient* (DSC) on a PC workstation with Intel Duo Core 3GHz E6850 processor. Compared with the 3D 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 3D AVM.

Table 1. Quantitative evaluation and performance comparison

	MSAVM					3D AVM				
	P	Q	DSC	Iterations	Time	P	Q	DSC	Iterations	Time
Lung in Fig. 2	95.5	99.8	96.2	26	870s	92.3	99.8	94.6	33	1000s
Heart in Fig. 3	92.0	99.0	92.2	27	1535s	90.7	98.9	91.1	39	2023s

To demonstrate the MSAVM more clearly, we put a set of 2D axial projection slices from a case of 3D heart segmentation in Figure 5, and compare them with the converged result of original 3D AVM using the same initialization in Figure 5f. Due to intensity inhomogeneity inside the inner surfaces and obscure boundary of the outer surface, original 3D 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.

4 Discussions and Conclusion

In this paper, we proposed Multiple-Surface Active Volume Models to segment 3D objects in volumetric medical images. Combining high-level spatial information with predicted ROI region, MSAVM achieves better performance than the original 3D AVM. Because of the implicit representation of AVM, the spatial distance constraints can be obtained with very low extra computational cost. In fact, MSAVM reduces the running time when compared to AVM. Due to the local smoothness of simplex-mesh, it is still hard for the model to reach some

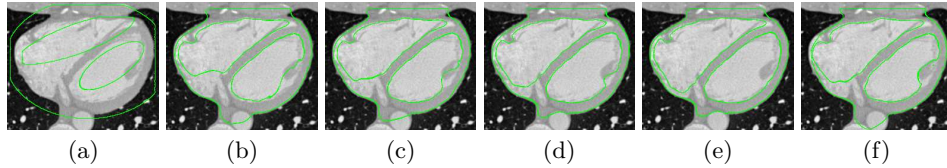


Fig. 5. Heart segmentation progress in a 2D slice projection. (a)Initial MSAVM, after (b)9, (c)15, (d)21, (e)29(converged result) iterations, (f)converged result of three separate AVMs after 36 iterations.

tip locations (e.g., top-right tip of the right ventricle in figure 5). As future work we plan to address this problem by reparameterizing the model near tips since vertices in these areas tend to be sparser than those distributed on the main body.

References

1. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images—a graph-theoretic approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **28**(1) (Jan. 2006) 119–134
2. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *Int’l Journal on Computer Vision* **1** (1987) 321–331
3. Xu, C., Prince, J.: Snakes, shapes and gradient vector flow. *IEEE Trans. on Image Processing* **7** (1998) 359–369
4. Shen, T., Zhu, Y., Huang, X., Huang, J., Metaxas, D., Axel, L.: Active volume models with probabilistic object boundary prediction module. *MICCAI* (2008) 331–341
5. Malladi, R., Sethian, J., Vemuri, B.: Shape modeling with front propagation: A level set approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **17**(2) (1995) 158–175
6. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape model - their training and application. *Computer Vision and Image Understanding* **61** (1995) 38–59
7. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. *Proc. Of European Conf. on Computer Vision* **2** (1998) 484–498
8. Spreeuwiers, L., Breeuwer, M.: Detection of left ventricular epi- and endocardial borders using coupled active contours. *Computer Assisted Radiology and Surgery* (2003) 1147–1152
9. Zeng, X., Staib, L., Schultz, R., Duncan, J.: Segmentation and measurement of the cortex from 3-d mr images using coupled-surfaces propagation. *Medical Imaging, IEEE Transactions on* **18** (1999) 927–937
10. MacDonald, D., Kabani, N., Avis, D., Evans, A.C.: Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage* **12**(3) (2000) 340–356
11. Shen, T., Li, H., Qian, Z., Huang, X.: Active Volume Models for 3D Medical Image Segmentation. *Computer Vision and Pattern Recognition, 2009. IEEE Conference on* (2009) 1–8
12. Cohen, L., Cohen, I.: 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** (1993) 1131–1147