A New, General Method of 3D Model Generation for Active Shape Image Segmentation

Seong-Jae Lim[1], Jayaram K. Udupa[2], Andre Souza[2],
Yong-Yeon Jeong[3], Yo-Sung Ho[1], Drew A. Torigian[4]

[1]Department of Information and Communications
Gwangju Institute of Science and Technology (GIST) – Gwangju, South Korea
[2]Medical Image Processing Group – Department of Radiology
University of Pennsylvania, Philadelphia, Pennsylvania – USA
[3]Chonnam National University Medical School – Gwangju, Korea
[4]Thoracic and Body Imaging – Department of Radiology
University of Pennsylvania - Philadelphia, Pennsylvania – USA

ABSTRACT

For 3D model-based approaches, building the 3D shape model from a training set of segmented instances of an object is a major challenge and currently remains an open problem. In this paper, we propose a novel, general method for the generation of 3D statistical shape models. Given a set of training 3D shapes, 3D model generation is achieved by 1) building the mean model from the distance transform of the training shapes, 2) utilizing a tetrahedron method for automatically selecting landmarks on the mean model, and 3) subsequently propagating these landmarks to each training shape via a distance labeling method. Previous 3D modeling efforts all had severe limitations in terms of the object shape, geometry, and topology. The proposed method is very general without such assumptions and is applicable to any data set.

Keywords: 3D active shape models, deformable models, image segmentation, image processing

1. INTRODUCTION

Motivation

Statistical models of shape variability via active shape models (*ASM*)[1] have been successfully utilized to perform segmentation and recognition tasks in 2D images. Statistical model-based methods allow modeling shape variations statistically and find in images shapes, which are consistent with the training shapes. They are very helpful when image information is missing or corrupted. 3D model-based approaches are more promising than 2D approaches since they can bring in more and realistic shape constraints for recognizing and delineating the object boundary.

Purpose

In the generation of a point distribution model (*PDM*), corresponding landmarks must be selected in all training shapes. However, manual determination of landmark correspondences is time-consuming, tedious, and error-prone. This is particularly true for 3D approaches, and perhaps impossible in 3D. Existing methods of *PDM* generation[2-4] impose serious constraints on the types of shapes that can be handled. We propose a general, automatic solution to this problem. Given a set of 3D training shapes in the form of 3D binary images, our method determines the same corresponding set of landmarks on all shape boundaries.

Correspondence to Jayaram K. Udupa: e-mail: jay@mipg.upenn.edu, Telephone: (215) 662-6780, Fax: (215) 898-9145.

2. METHODS

Input: A set of 3D binary images representing the training shapes.

Output: A 3D statistical shape model (*ASM*).

Our method consists of four steps:

Step 1: Creating a 3D mean shape $\overline{S}$ from the 3D binary images.
Step 2: Finding landmarks on $\overline{S}$.
Step 3: Propagating landmarks from $\overline{S}$ to each training shape.
Step 4: Creating the 3D *ASM*.

These steps are further described below.

Step 1: Creating Mean Shape

First, a distance transform is applied to each 3D binary image. Next the mean of all distance-transformed images so created for the training shapes is computed. This mean image is thresholded at 0. The 3D digital boundary of this 3D binary image (which represents the mean shape) is then found, as shown in Figure 1.

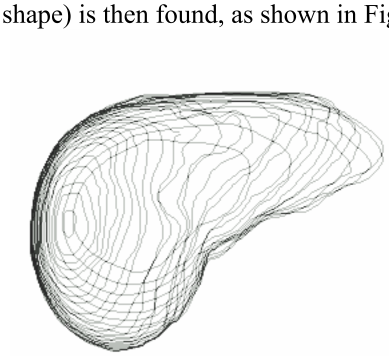

Figure 1. The mean distance image of 20 binary images representing 20 lives from 20 abdominal CT data sets.

Step 2: Finding Landmarks on the Mean Shape

Given any 3D shape as a digital surface[5] S, the method to find landmarks is as follows. For ease of description, we will present the method first for simple convex shapes (Figure 2) and subsequently for arbitrary shapes (Figure 3). For simplicity, the illustration in Figures 2 and 3 are for 2D shapes. But the method works in 3D for any shapes (and can be generalized to higher dimensions).

(1) Find the major inertia axis of S via *PCA*[6] and its points of intersection L_1, L_2, with S. Output points L_1, L_2.

(2) Find the farthest point L_3 on S from line L_1L_2. If its distance from L_1L_2 is less than a fixed threshold θ, stop.

(3) Else output L_3.

(4) Find plane P containing L_1, L_2, L_3, and for one side of P, find point L_4 farthest from P. If its distance from P is less than θ consider the other side of P and proceed recursively. Else output L_4, form new planes containing (L_1, L_2, L_4), (L_1, L_3, L_4), and (L_2, L_3, L_4) and proceed recursively.

(5) Stop when no more planes need to be considered and no points are found with "distance to plane" greater than or equal to θ.

The set of all points output by the procedure constitutes the set of landmarks defined for *S*.

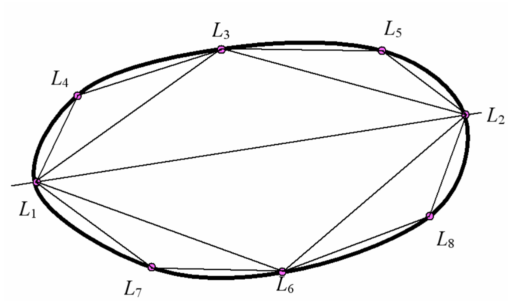

Figure 2: In 2D, there are no planes involved, and farthest distance is always found from lines. For convex shapes, the part of boundary *S* on one side of line is always connected during recursion

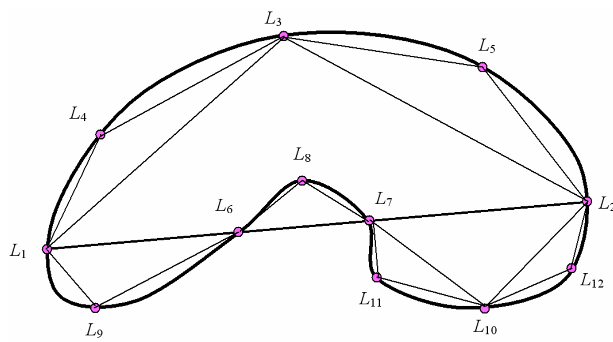

Figure 3. If boundary *S* is not convex, then, during recursion, there may be several connected components of *S* on any side of the dividing line. For example, on both sides of line L_1L_2, there are two connected components of *S*. Each component must be recursively partitions separately.

In essence, the method recursively subdivides the interior of *S* into tetrahedra. When the method stops, it will have partitioned the interior of *S* into tetrahedra of various sizes whose union approximates the interior of *S*. The vertices of the tetrahedra constitute the landmarks on *S*.

<u>Complex Shapes</u>

When *S* is not convex, during recursive subdivision, there is no guarantee that the part of *S* on one side of the current dividing plane is connected. This is illustrated in Figure 3 for the 2D cases; both sides of L_1L_2 contain two connected components of *S*. In this case, therefore, connectivity analysis becomes necessary. Further, for each disconnected part (such as the concavity in the middle in Figure 3), initial landmarks have to be found (in accordance with the first step) to start the recursive process. This is done by finding the major inertia axis of that contour resulting from intersecting *S* with the plane, which is associated with the particular connected component. Subsequently the recursive process proceeds as previously described.

When the above method is applied to the surface representing mean shape $\overline{S}$, we get a set *L* of labeled landmarks on $\overline{S}$.

Step 3: Propagation Landmarks to Each Training Shape

To avoid selecting corresponding landmarks L found on $\overline{S}$ in all training shapes, we propagate landmarks to each training shape. In this method, to propagate landmarks to the surface S of a given training shape, we use distance labeling to find the point on S that is closest to each landmark on $\overline{S}$. The distance value from each landmark of $\overline{S}$ is gradually increased. Whichever frontier (from a particular landmark) reaches a point of S first determines the landmark label to be assigned to that point. In this manner, the points on S and their labels are determined simultaneously. Figure 4 illustrates this propagation scheme. This method again is very general and can be used in any dimensional space.

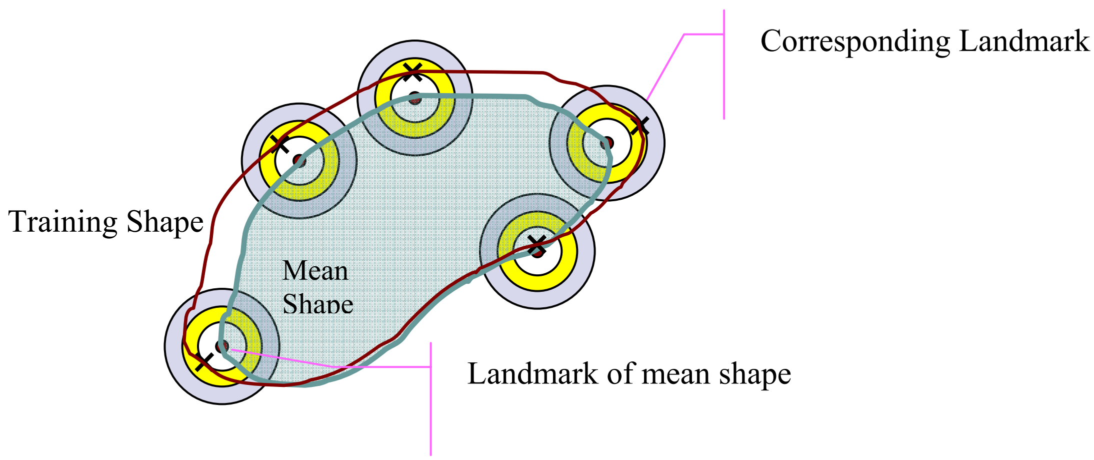

Figure 4. Distance labeling method for propagating landmarks from mean shape $\overline{S}$ to training shapes S.

Step 4: Creating 3D *ASM*

Once the landmarks are identified on all training shapes, the principles underlying the *ASM* method can be used to construct the statistical shape model. Principal component analysis is applied to the aligned shape vectors. To this end, the mean shape, the covariance matrix, and the eigensystem of the covariance matrix are computed.
The eigenvectors of the covariance matrix provide the modes of shape variation present in the training data. The eigenvectors corresponding to the largest eigenvalues account for the largest variation; a small number of modes are usually sufficient to explain most of the variation present in the training shapes.

3. RESULTS

CT liver scans of 50 patients acquired in the venous phase of enhancement were utilized for testing the new model generation method. Each 3D *CT* image data set contains up to 71 slices with an inter-slice spacing of 3.8*mm* and an in-plane pixel size of 1.00*mm* × 1.00*mm*. The 50 training shapes were generated by segmenting the liver in these data sets in a slice-by-slice manner by using live wire[7].

Qualitative Evaluation

(a)

(b)

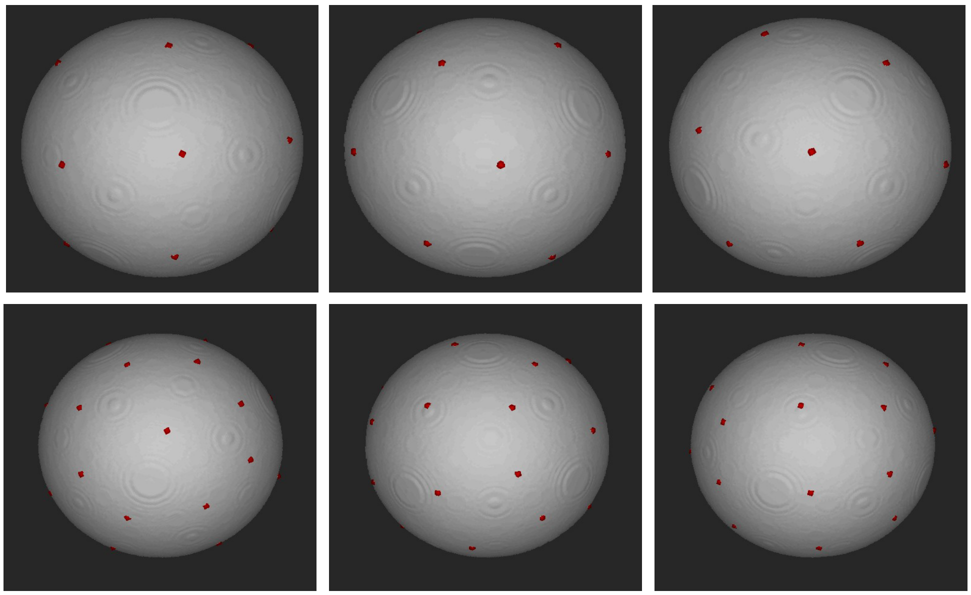

Figure 5. Landmarks on the mean shape of a simple shape (sphere), with threshold value θ: (a) 15, (b) 5 (in voxel units).

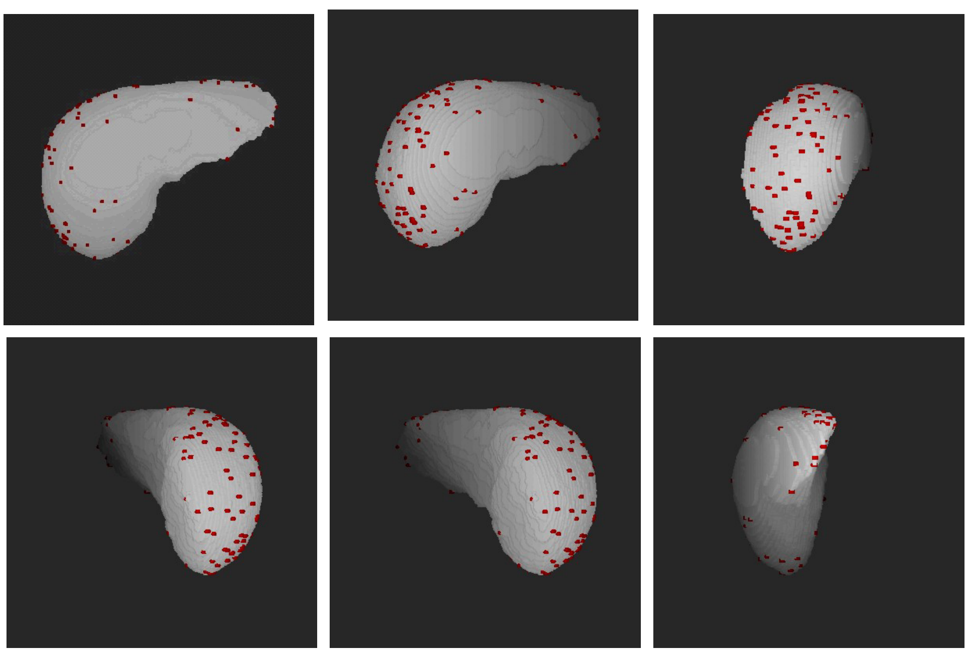

Figure 6. Landmarks on the mean shape of a complex shape (liver). Number of landmarks 120; surface size 12,757 voxels; θ = 5. The mean shape was generated from the 50 training shapes obtained from the 50 3D CT liver scan images.

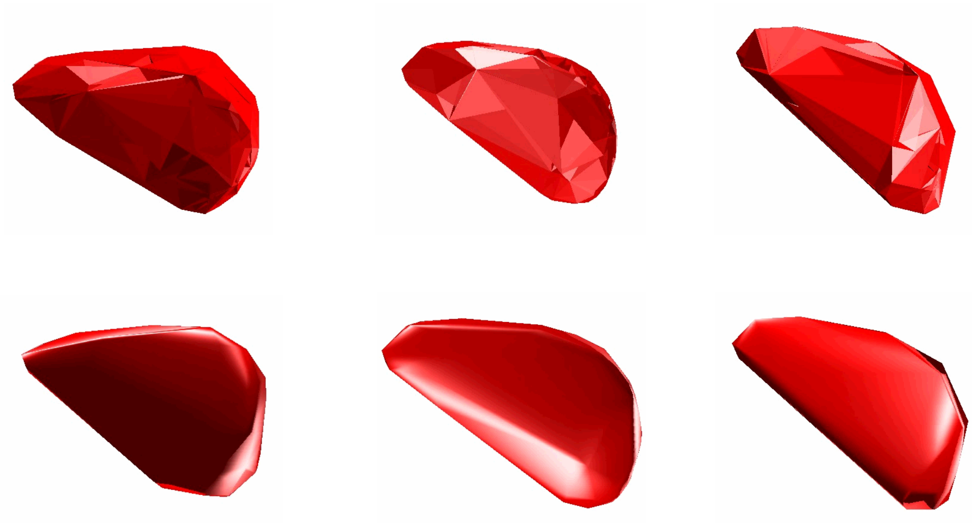

Figure 7. 3D surface representation of the largest mode of variation corresponding to $-3\sqrt{\lambda_1}$ (left), mean (middle), and mean $+3\sqrt{\lambda_1}$ (right). The surfaces displayed are created via Delaunay triangulation (top row) and a convex hull method (bottom row).

Quantitative Evaluation

The compactness of a model is its ability to describe the variability of a shape by using as few modes as possible. In Figure 8(a), we show the compactness, i.e., variance per mode, of the model generated by our method. We express compactness as % of variation per mode, given by

$$\%\text{ variation per mode} = \frac{\lambda_i}{\lambda_T} \times 100\%$$

where $\lambda_T = \sum_{i=1}^{t} \lambda_i$ is the variance, and t is the total number of eigenvalues.

In Figure 8(a), this % variation is displayed as a function of i, the eigenvalues rank. In Figure 8(b), the portion of variability captured by the first m eigenvalues is shown. It can be seen that about 94% of the shape variation is already captured by the first three eigenvectors, and this number is about 99.5% for four eigenvectors. This suggests that the model generated by the new method is very compact.

Computational Time

On a 1.7 GHz Pentium 4 PC with 768 MB RAM, the computational time required for Steps 1-4 in our method for the 50 liver 3D data sets is as follows:

Step 1: Creating mean shape : about 60 min.
Step 2: Finding landmarks on : 1 min.
Step 3: Landmark propagation : 250 min.
Step 4: Creating th 3D *ASM* : 1 min.

Steps 1 and 3 are the most time consuming. This is mainly due to the 3D distance transform involved in these steps. If the implementations are optimized and faster computers are used, we believe that the total processing time can be brought down to under 1 hour.

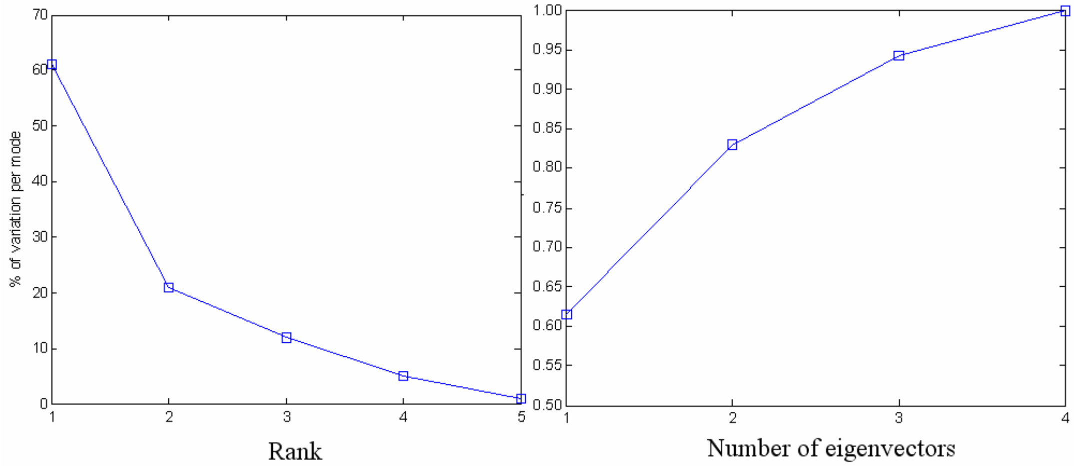

Figure 8. The % variation as a function of eigenvalues rank (a), and portion of variability captured by the first *m* eigenvectors (b).

4. CONCLUDING REMARKS

We have described a method for automatic 3D model generation for active shape model (*ASM*) based image segmentation. The tetrahedron method for automatic landmark selection guided by the mean shape is independent of the object shape, geometry, and topology. We believe that there are three groundbreaking ideas in our approach: (1) The basic method of building the mean shape. Bypassing the basic requirement of landmarks to determine the mean shape is a strength of our method. A rough mean shape is elegantly obtained via distance transforms. (2) The new method of automatically finding landmarks on the mean shape. This is the heart of our method. It does not make any assumptions on the form or shape of the objects unlike published methods most of which impose serious shape restrictions. (3) The method of propagating the landmarks to the training shapes. This guarantees that all training shapes have the same number of landmarks, and that these landmarks correspond to those on the mean shape The computation times involved in the various steps are not prohibitive and the constructed models seem to be statistically compact, which is an indication that the landmarks are selected in a sensible manner.

ACKNOWLEDGEMENT

This work was supported by DHHS grant NS37172. The first author was supported in part by the Korea Research Foundation Grant (No. M07-2004-000-10140-0), in part by the Ministry of Information and Communication (MIC) through the Realistic Broadcasting Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

REFERENCES

1. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, Active shape models – their training and application, *Comput. Vis. Image Understanding*, 61:38-59, 1995.
2. R. H. Davies, C. J. Twining, P. D. Daniel, T. F. Cootes, and C. J. Taylor, “Building optimal 2D statistical shape models,” *Image and Vision Computing*, 21:1171-1182, 2003.
3. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration, *IPMI 2001 Conference Proc*eedings, 2082:8-91, 2001.

4. Hill and C. J. Taylor, Automatic landmark generation for point distribution models, *British Machine Vision Conference*, pp. 429-438, 1994.

5. J.K. Udupa, "Multidimensional digital boundaries," *CVGIP: Graphical Models and Image Processing*, 56(4):311-323, 1994.

6. E. Jackson, *A User's Guide to Principal Components*. John Wiley & Sons, Inc., 1991 page(s): 1-25, B. Jones 3-17-94.

7. X. Falcao, J. K. Udupa, S. Samarasekera, and S. Sharma, "User-Steered Image Segmentation Paradigms: Live Wire and Live Lane," *Graphical Models and Image Processing,* 60, pp. 233-260, 1998.