

Diaphragm Dome Surface Segmentation in CT Data Sets: A 3D Active Appearance Model Approach

Reinhard Beichel^a, Georg Gotschuli^b, Erich Sorantin^b, Franz Leberl^a and Milan Sonka^c

^aInstitute for Computer Graphics and Vision, Graz University of Technology,
Inffeldgasse 16/2, A-8010 Graz, Austria

^b Department of Radiology, Graz University Hospital,
Auenbruggerplatz 9, A-8036 Graz, Austria

^cDepartment of Electrical and Computer Engineering, The University of Iowa,
Iowa City, IA 52242, USA

ABSTRACT

Knowledge about the location of the diaphragm dome surface, which separates the lungs and the heart from the abdominal cavity, is of vital importance for applications like automated segmentation of adjacent organs (e.g., liver) or functional analysis of the respiratory cycle. We present a new 3D Active Appearance Model (AAM) approach to segmentation of the top layer of the diaphragm dome. The 3D AAM consists of three parts: a 2D closed curve (reference curve), an elevation image and texture layers. The first two parts combined represent 3D shape information and the third part image intensity of the diaphragm dome and the surrounding layers. Differences in height between dome voxels and a reference plane are stored in the elevation image. The reference curve is generated by a parallel projection of the diaphragm dome outline in the axial direction. Landmark point placement is only done on the (2D) reference curve, which can be seen as the bounding curve of the elevation image. Matching is based on a gradient-descent optimization process and uses image intensity appearance around the actual dome shape. Results achieved in 60 computer generated phantom data sets show a high degree of accuracy (positioning error -0.07 ± 1.29 mm). Validation using real CT data sets yielded a positioning error of -0.16 ± 2.95 mm. Additional training and testing on in-vivo CT image data is ongoing.

Keywords: Active Appearance Models, diaphragm dome surface segmentation, 3D border detection

1. INTRODUCTION

Computer-aided and highly automated segmentation algorithms are needed to fully utilize the structural and functional analysis possibilities of modern three-dimensional (3D) image modalities like X-ray computed tomography (CT) or magnetic resonance imaging (MRI). This is especially true for large anatomical structures which would require a considerable amount of time if segmented manually. Segmentation of the diaphragm dome surface* (Fig. 1) is such an application where a highly automated segmentation is of benefit for applications like diaphragm motion analysis/modeling, mechanical analysis, functional analysis of the respiratory cycle and planning of lung volume reduction surgery.

Another not so obvious application is to support automated segmentation of neighboring organs like the liver, where leakage into the heart region is one of the main problems.¹ Obstacles like this have been mainly solved by the development of semiautomatic approaches for liver segmentation.^{2,3} A solution for separating heart and liver has been presented in.⁴ Points from a lung segmentation which are adjacent to the diaphragm are used to interpolate the gap in the region of the heart-liver transition. The method is prone to error and additional interactively-defined points may be needed for the interpolation.⁴ Another shortcoming of this approach is that no gray value based evidence is used to support the segmentation process.

Further author information: (Send correspondence to R.B.)

R.B.: E-mail: beichel@icg.tu-graz.ac.at, phone: +43 316 873 5022

*We distinguish between two parts of the diaphragm: the dome and the area of apposition to the rib cage and abdominal side walls.

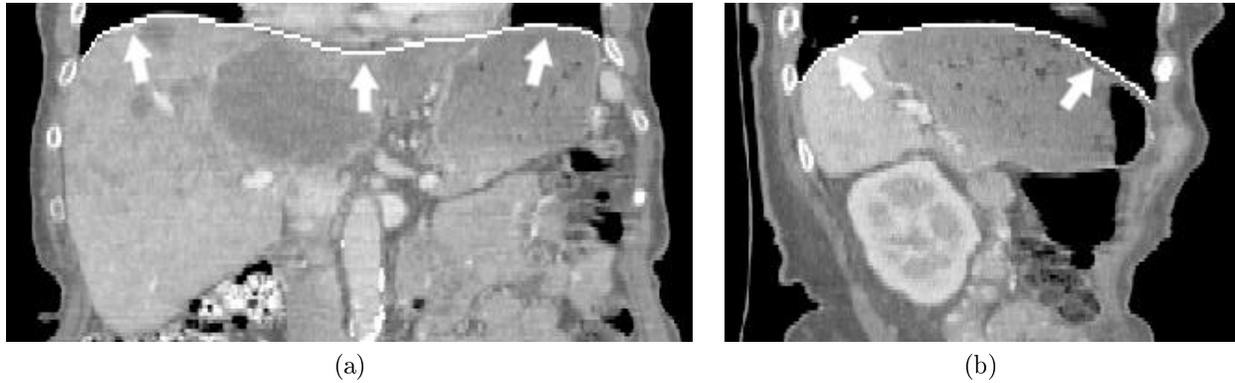


Figure 1. Coronal (a) and sagittal (b) views of the diaphragm dome surface marked with white color. The arrows are pointing towards the location of the dome surface.

Active Appearance Models (AAMs) have been successfully used for different medical applications.⁵⁻⁸ We propose a new three-dimensional (3D) Active Appearance Model (AAM) approach for segmenting the upper top layer of the diaphragm dome. In this way a priori knowledge of the diaphragm dome shape and appearance are used in combination with gray value based evidence found in the application data set, avoiding the problems discovered in.⁴ By using a two component representation of the 3D dome surfaces, the process of generating learning samples for the model building step is simplified. Our approach mainly aims to facilitate automated liver segmentation, but is also suitable for other applications mentioned above.

2. ACTIVE APPEARANCE MODEL BASED 3D SURFACE SEGMENTATION

Robust segmentation of medical images can be challenging due to low contrast, partial volume effect or noise. Active Appearance Models (AAMs), a top-down segmentation approach developed by Cootes and Taylor,^{9,10} are capable to solve this problems by utilizing a priori knowledge of anatomy. Using a statistical model, the AAM describes appearance and shape of an object class previously learned from a set of samples. Image segmentation is achieved by minimizing the difference between the model and an image considering statistically plausible shape and intensity variations. The diaphragm dome surface typically shows a consistent shape pattern and is therefore well suited for an AAM based segmentation approach. We have extended the 2D AAM to segment 3D surfaces which can be represented by an elevation image.

2.1. Extension to 3D

One of the main problems in building 3D AAMs is the representation of 3D shapes, due to the need of finding key landmark points on the 3D dome surface. A manual landmark point placement in 3D is time consuming, error prone, and also presents visualization challenges. Our approach transforms the 3D landmark point placement problem to a 2D problem and therefore substantially reduces the amount of necessary user interaction .

2.2. 3D shape representation

The 3D diaphragm shape representation can be divided in two parts: a 2D closed curve (reference curve) and an elevation image in which elevation is coded using gray values. Both parts refer to an axial image plane, which will be denoted as reference plane. The reference plane is placed in each sample data set on the top of the diaphragm dome. The reference curve represents the outline of the axial projection of the diaphragm dome (Fig. 2). Differences in height between dome voxels and the reference plane are stored in the elevation image. Combined, the chosen 3D shape representation can be viewed as a 2D AAM in which the gray levels represent shape. Landmark points are only needed for the 2D reference curve.

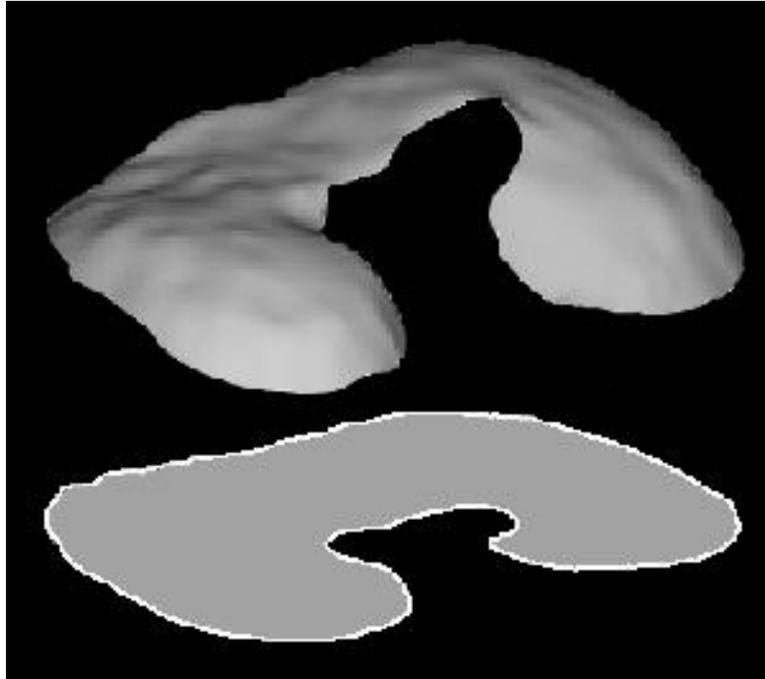


Figure 2: Axial projection of the 3D diaphragm dome surface. The reference curve is shown in white.

Using the landmark points of each data set, a statistical model of reference curve shape variations can be generated by means of a Principal Component Analysis (PCA). The linear model

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_x \mathbf{b}_x \quad (1)$$

can be used to approximate examples of the learned class of reference curves, where $\bar{\mathbf{x}}$ denotes the mean reference curve, \mathbf{P}_x the reference curve eigenvector matrix and \mathbf{b}_x the reference curve parameters. By warping the elevation images to the mean reference curve shape and sampling them into elevation vectors, a linear model can be build

$$\mathbf{h} = \bar{\mathbf{h}} + \mathbf{P}_h \mathbf{b}_h \quad , \quad (2)$$

where $\bar{\mathbf{h}}$ is the mean elevation vector, \mathbf{P}_h the elevation eigenvector matrix and \mathbf{b}_h the elevation parameters. Thus, the diaphragm dome shape can be represented by the parameter vectors \mathbf{b}_x and \mathbf{b}_h . The two parts of the shape model are linked together via the warping function.

2.3. 3D appearance representation

To describe the 3D appearance, gray values are collected from layers parallel to the diaphragm dome voxel surface. Each texture layer is represented by an image. All layers are then warped to the average reference curve. The warped layers are expressed as an intensity vector \mathbf{g} by concatenating the intensity vectors from each layer. Then, an intensity-normalization to the average intensity of 0 and a variance of 1 is carried out. Applying PCA to the normalized data, a linear model

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g \quad (3)$$

can be obtained for the intensity vector \mathbf{g} , where $\bar{\mathbf{g}}$ denotes the mean intensity, \mathbf{P}_g the intensity eigenvector matrix and \mathbf{b}_g the intensity parameters.

It is important to provide additional appearance information of the surrounding non-diaphragm voxels to assist the matching process. This can be done by generating a fringe around the diaphragm dome using

additional landmark points. For the extension of the elevation image in the fringe area, an extrapolation or a nearest neighbor approach can be used.

2.4. Building a 3D AAM

To build a 3D Model, the two shape coefficient vectors (\mathbf{b}_x , \mathbf{b}_h) and the gray-level intensity coefficient vector (\mathbf{b}_g) are concatenated in the following manner

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_x \mathbf{b}_x \\ \mathbf{W}_h \mathbf{b}_h \\ \mathbf{b}_g \end{bmatrix} = \begin{bmatrix} \mathbf{W}_x \mathbf{P}_x^T (\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{W}_h \mathbf{P}_h^T (\mathbf{h} - \bar{\mathbf{h}}) \\ \mathbf{P}_g^T (\mathbf{g} - \bar{\mathbf{g}}) \end{bmatrix}, \quad (4)$$

where \mathbf{W}_x and \mathbf{W}_h are diagonal matrices relating to different units of shape, elevation information and intensity. A PCA is applied to the sample set of all \mathbf{b} vectors, yielding the model

$$\mathbf{b} = \mathbf{P}_c \mathbf{c}, \quad (5)$$

where \mathbf{P}_c is a matrix consisting of eigenvectors and \mathbf{c} are the resulting appearance model coefficients. Applying this procedure to a set of 3D training objects results in an ‘average object’ and its characteristic variations in 3D shape and surrounding intensity layers. The three basic components of the 3D AAM can be expressed as functions of model coefficients \mathbf{c} :

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_x \mathbf{W}_x^{-1} \mathbf{P}_{cx} \mathbf{c}, \quad (6)$$

$$\mathbf{h} = \bar{\mathbf{h}} + \mathbf{P}_h \mathbf{W}_h^{-1} \mathbf{P}_{ch} \mathbf{c}, \quad (7)$$

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{P}_{cg} \mathbf{c}, \quad (8)$$

and

$$\mathbf{P}_c = \begin{bmatrix} \mathbf{P}_{cx} \\ \mathbf{P}_{ch} \\ \mathbf{P}_{cg} \end{bmatrix}. \quad (9)$$

Given a model coefficient vector \mathbf{c} and a reference plane position, a corresponding diaphragm dome shape and surrounding gray value layers in the volume data frame can be generated as follows:

1. Generate a new reference curve shape \mathbf{x} and elevation vector \mathbf{h} by using Equations (6) and (7), respectively.
2. Transform the reference curve shape \mathbf{x} to the x, y -coordinate system of the volume data set by applying the similarity transformation

$$\hat{\mathbf{x}} = A_{\mathbf{t}}(\mathbf{x}) = \begin{pmatrix} 1+a & -b \\ b & 1+a \end{pmatrix} \mathbf{x} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (10)$$

using the pose parameter vector $\mathbf{t} = (a, b, t_x, t_y)^T$ where $a = s \cos \theta - 1$ and $b = s \sin \theta$.

3. Calculate a new diaphragm dome shape by converting \mathbf{h} to an image, multiply it by the height scaling factor ξ , which corresponds to the maximal difference in height between the diaphragm dome shape and reference plane, and warp it according to $\hat{\mathbf{x}}$.
4. Generate a new intensity vector \mathbf{g} by using Equation (8), transform the intensity values (texture) to the volume frame by $\mathbf{g}_v = B_{\mathbf{u}}(\mathbf{g}) = (1 + u_1)\mathbf{g} + u_2 \mathbf{i}^\dagger$ using the intensity parameter vector $\mathbf{u} = (u_1, u_2)^T$ and split it up into layer intensity vectors.
5. Convert the layer intensity vectors to images, warp them according to $\hat{\mathbf{x}}$ and translate the gray values in elevation according to the elevation image and relative layer position.

[†] \mathbf{i} denotes the unit vector.

2.5. Matching the 3D AAM to volumetric images

The AAM of the diaphragm dome can be used for segmentation by minimizing the difference between the 3D model appearance and a target volume data set by using a gradient descent minimization. The actual dome shape in the volume frame is defined by the model parameters \mathbf{c} , pose parameters \mathbf{t} and the height scaling factor ξ . The surrounding gray value layers are sampled into images, converted to an intensity vector and transformed by $B_{\mathbf{u}}^{-1}$ into an intensity vector \mathbf{g}_{target} in the model frame. The actual gray value appearance of the model \mathbf{g}_{model} can be calculated from \mathbf{c} by using Equation (8). During the matching process the error $|\delta\mathbf{g}|^2 = |\mathbf{g}_{target} - \mathbf{g}_{model}|^2$ is minimized by varying the parameter vector $\mathbf{p}^T = (\mathbf{c}^T | \mathbf{t}^T | \mathbf{u}^T)$ consisting of model parameters \mathbf{c} , pose parameters \mathbf{t} and global intensity parameters \mathbf{u} .

For an effective update of the parameter vector \mathbf{p} , the linear relations $\delta\mathbf{c} = \mathbf{R}_c\delta\mathbf{g}$, $\delta\mathbf{t} = \mathbf{R}_t\delta\mathbf{g}$ and $\delta\mathbf{u} = \mathbf{R}_u\delta\mathbf{g}$, predicting the necessary change of parameter values given an observed gray value difference vector $\delta\mathbf{g}$, have to be found in a training process before the matching process is started. This can be done by a method based on a first order Taylor expansion as described by Cootes in.^{10,11} For matching, the reference plane is placed at the top of the diaphragm dome, the 3D AAM is roughly moved to a good starting position and the height scaling factor ξ is determined. Then, starting from an initial estimate of the current model parameters (e.g., 'mean model'), the following steps are repeated as long as the error $|\delta\mathbf{g}|^2$ improves:

1. Evaluate the error vector $\delta\mathbf{g} = \mathbf{g}_{target} - \mathbf{g}_{model}$
2. Compute the current error $E = |\delta\mathbf{g}|^2$
3. Set $k = 1$
4. Update the model parameters: $\tilde{\mathbf{c}} = \mathbf{c} - k\delta\mathbf{c}$, $\tilde{\mathbf{t}} = \mathbf{t} - k\delta\mathbf{t}$ and $\tilde{\mathbf{u}} = \mathbf{u} - k\delta\mathbf{u}$
5. Calculate a new error vector $\delta\tilde{\mathbf{g}}$ using the updated parameters $\tilde{\mathbf{c}}$, $\tilde{\mathbf{t}}$ and $\tilde{\mathbf{u}}$
6. If $|\delta\tilde{\mathbf{g}}|^2 < E$ then accept the new estimates: $\mathbf{c} = \tilde{\mathbf{c}}$, $\mathbf{t} = \tilde{\mathbf{t}}$ and $\mathbf{u} = \tilde{\mathbf{u}}$
7. Else try at $k = 1.5$, $k = 0.5$, $k = 0.25$ etc.

Note that compared to a conventional 2D AAM matching procedure, the error $|\delta\mathbf{g}|^2$ is calculated using the gray value layers in 3D space.

3. EVALUATION ON PHANTOM DATA

To fully validate the basic concept of this new AAM approach, 100 computer generated diaphragm phantom data sets were generated. Figure 3 depicts one of them. The phantom simulated the following structures by randomly selected 8 bit gray values out of the range specified within brackets: lungs (30-60), spine (200-255), side walls (70-85), abdomen (85-130), heart (130-200) and a region of transition between heart and abdomen (100-160). The diaphragm dome surface was represented by the top layer of the abdominal region. The number of slices ranged between 36 and 49. Each slice consisted of 256×256 voxels. All phantom data sets had the same voxel size of $0.65 \times 0.65 \times 2.0$ mm³. The phantom body was placed around the axial image center with a randomly selected displacement in the axial plane of up to 10% of the mean object size in x -direction. Ground truth data were provided directly by the generation process.

The data sets were divided in a training set consisting of 40 cases and a test data set of 60 cases. When building the model, five texture layers were considered, two located above and two below the actual diaphragm dome surface. An extrapolation was used to generate height values in the fringe area. For matching the 3D AAM was placed in the axial image center. The corresponding reference plane positions and the height scaling factors ξ were provided.

For border positioning error calculation the signed error between all voxels of the reference surface (ground truth) and the closest voxels of the model dome surface were determined. A negative error value indicates that the voxel was below the reference surface. The signed error is given as mean \pm standard deviation. On the

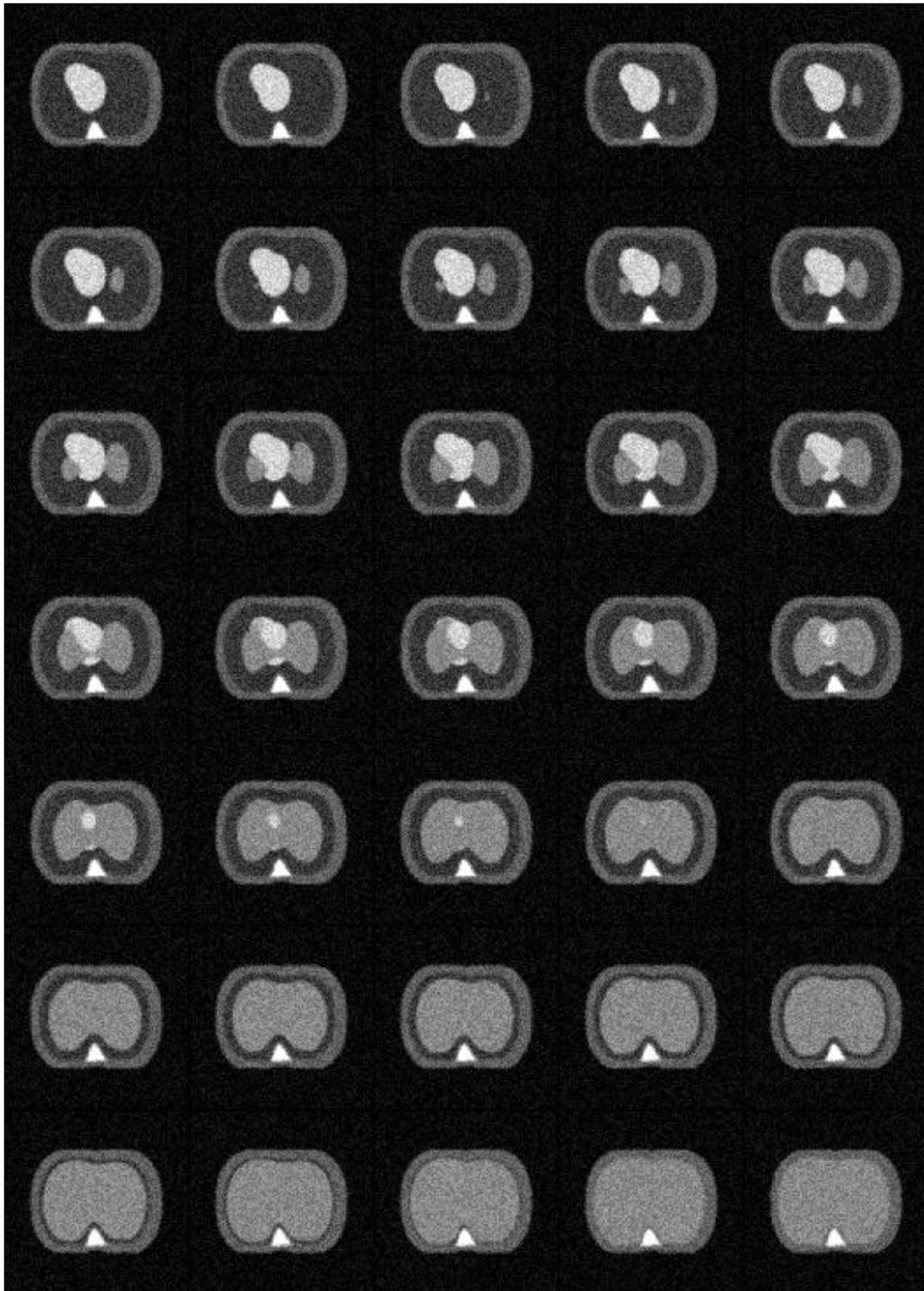


Figure 3. Example for a computer generated diaphragm phantom data set. The top slice is shown in the upper left corner. The slices are ordered from the left to the right.

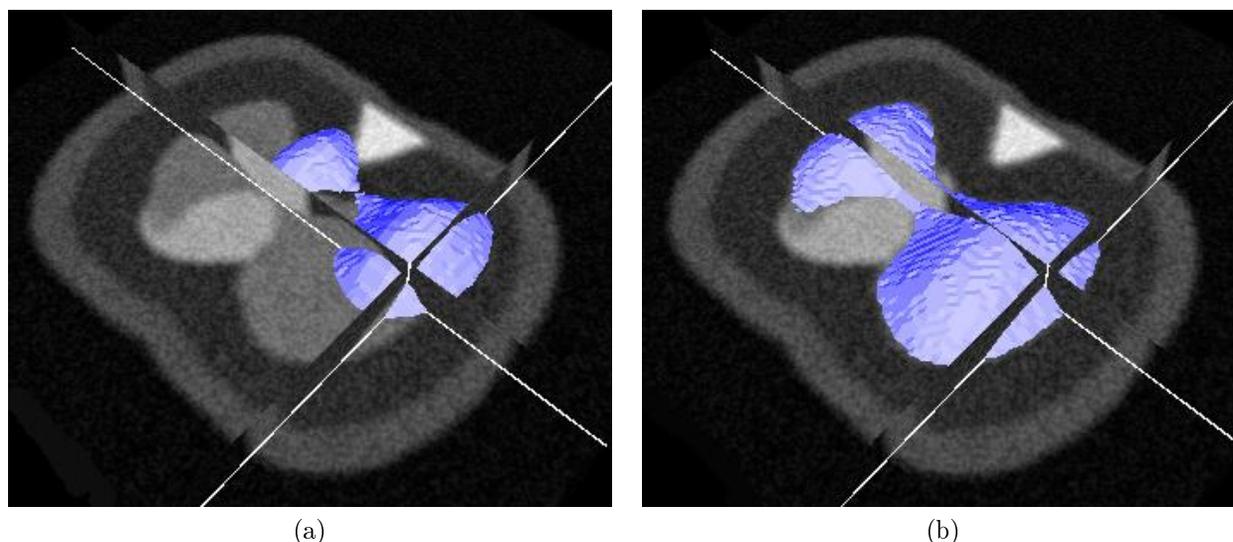


Figure 4. Visualization of the matching progress of the 3D Active Appearance Model applied to a phantom volume data set. (a) Start position of the 3D surface. (b) Final match.

40 test cases, the overall border positioning error was -0.07 ± 1.29 mm, showing almost no bias. Figure 4 shows the initial and end stages of the 3D model applied to a phantom data set. The region of transition between heart and abdominal region (diaphragm) was correctly separated by the model (Fig. 4.b).

4. EXPERIMENTS WITH CT DATA SETS

For experiments, 14 routinely acquired contrast-enhanced spiral CT liver scans were available. Images were acquired using a standard protocol for liver tumor screening. Each slice consisted of 512×512 voxels and between 47 and 109 slices. In-plane voxel dimension ranged between 0.55 to 0.78 mm with most of the voxel sizes falling between 0.6 and 0.7 mm. Therefore no resampling of the data sets was performed. The dimension along the z -direction was 2, 2.5, 4, and 5 mm, respectively. Different z -dimension values were taken into account during the calculation of the differences in height between the dome voxels and reference plane. The CT scans were acquired during end expiration, no volume controller was used. In 2 cases, a small part of the top of the diaphragm dome was not completely imaged. Only the available portion of the captured diaphragm was used for building the AAM in those cases and a value of -700 Hounsfield units (HU) was assumed for missing lung tissue above the diaphragm dome.

Ground truth data were generated manually by an expert. Since an average diaphragm dome surface consists of approximately 70,000 voxels, an interpolation scheme was used. Diaphragm dome surface voxels were only identified at a 16×16 sampling grid. Near the border to the side walls, additional voxels were added to preserve a precise shape representation. The dome surface was generated by thin-plate spline interpolation.

To increase the number of samples available for the pilot work reported here, 14 additional data sets were generated by applying a nonlinear geometric transformation to the axial slices of the existing CT scans. The transformation was chosen so that the new 3D dome shapes generated, were plausible enough to descend from the original sample population (Fig. 5). The relation between coordinates in axial plane (x, y) of the newly generated data sets and the coordinates (\tilde{x}, \tilde{y}) of the source data sets can be described by

$$(\tilde{x}, \tilde{y}) = (f_\gamma(x), f_\gamma(y)) \quad (11)$$

using $\gamma = 256$ and

$$f_\gamma(a) = \begin{cases} \gamma(1 + (a/\gamma - 1)^{1.2}) & : a \geq 0 \\ -\gamma(1 + (-a/\gamma - 1)^{1.2}) & : a < 0 \end{cases} \quad (12)$$

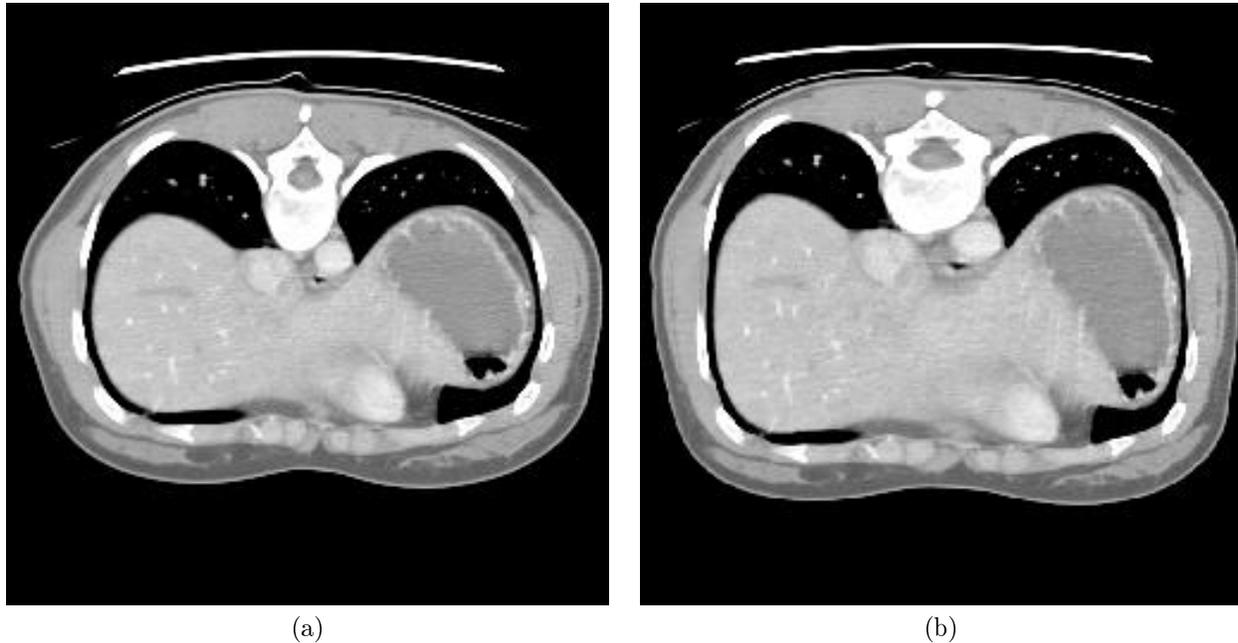


Figure 5: Slice of a volume data set before (a) and after (b) application of the nonlinear geometric transformation.

For gray-level interpolation a nearest neighbor interpolation was used. Ground truth data for the newly generated data sets were also transformed according to Equation 11. Gray-level appearance was not changed.

Texture was captured in 9 layers where 4 of them were placed above and 4 placed below the model shape. Height values in the fringe area were set to the value of the nearest voxel of the dome shape. For testing, 8 cases, consisting of 6 original and 2 newly generated data sets, were selected. The remaining 20 cases were used for training. Selection of the reference plane position, the height scaling factor ξ and the starting x - and y -positions were defined interactively. In all cases the mean model was scaled by $s = 1.15$.

Errors were measured and reported as outlined in Section 3. The border positioning error on the complete test data set was -0.16 ± 2.95 mm. Figure 6 depicts two examples out of the test data set where the matched model is shown in coronal and sagittal slices of the CT volumes. The first example (Figs. 6.a and 6.b) shows a good match between the AAM-detected dome surface and reality with the border positioning error of 0.048 ± 1.29 mm. Matching errors in the region of apposition led to a somewhat higher border positioning error of 0.50 ± 3.50 mm on the second example (Figs. 6.c and 6.d). In this case the appearance pattern of the abdominal region was influenced by a fluid accumulation in the abdomen (ascites) compared to other data sets.

5. DISCUSSION AND CONCLUSIONS

A new 3D AAM for the segmentation of the diaphragm dome surface was presented. On the 60 phantom test cases the method showed good matching results with low bias and variance. Results on CT data sets were promising and full-scale evaluation on CT data is underway. Problems of the model to fully match the dome surface in local regions show the need for further refinements of the developed method. Adding a semi- or fully automatic initialization method that places the model close to the dome surface before the matching/minimization process is started, would be one example. A good initialization would also reduce the risk that the model converges to an undesired local minima.

A very important issue for AAMs in general is to have a representative learning set of the target population, since the ability of the AAM to match to new data is closely linked to the samples used for model building/training. Here, clearly a larger learning set is needed where frequent pathological findings are also included. Usage of a lung volume controller would help to increase the quality of the training sets.

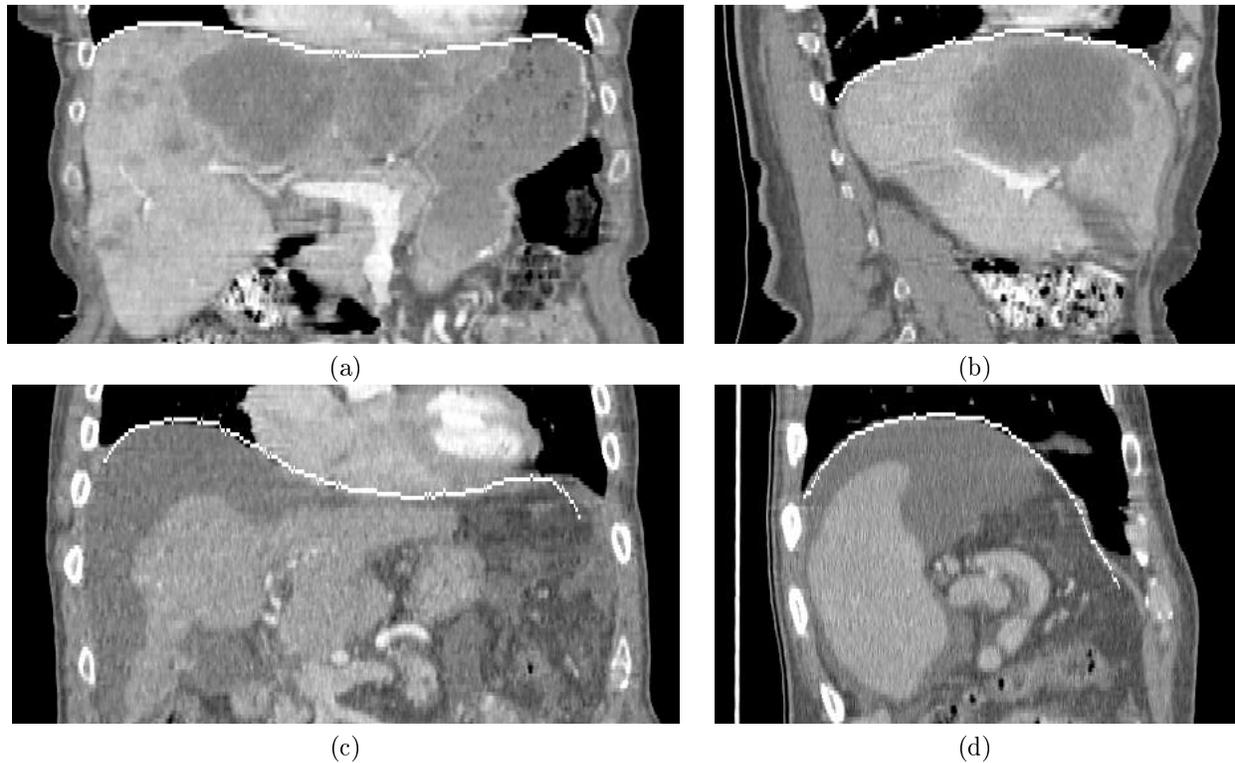


Figure 6. Examples for matching results of the 3D AAM on CT data (see text). (a) Coronal and (b) sagittal view of the first example. (c) Coronal and (d) sagittal view of the second example. The model surface is shown in white.

ACKNOWLEDGMENTS

This work was partly supported by the Austrian Science Foundation (FWF) under grant P14897. Thanks to Mark E. Olszewski and Steven C. Mitchell for their support.

REFERENCES

1. S. Pan and M. Dawant, "Automatic 3D segmentation of the liver from abdominal CT images: A level-set approach," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, eds., *Proc. SPIE* **4322**, pp. 128–138, 2001.
2. A. Schenk, G. Prause, and H.-O. Peitgen, "Local cost computation for efficient segmentation of 3D objects with live wire," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, eds., *Proc. SPIE* **4322**, pp. 1357–1364, 2001.
3. G. Glombitza, W. Lamadé, A. M. Demiris, M. R. Göpfert, A. Mayer, M. L. Bahner, H.-P. Meinzer, G. Richter, T. Lehnert, and C. Herfarth, "Virtual planning of liver resections: Image processing, visualization and volumetric evaluation," *Int. Journal of Medical Informatics* **53**, pp. 225–237, 1999.
4. P. Hassenpflug, G. Glombitza, C. Cárdenas, I. Wolf, A. Benner, and H.-P. Meinzer, "Komponenten zur Unterstützung der automatischen Bildsegmentierung von CT-Daten der Leber," in *Bildverarbeitung für die Medizin 2000*, A. Horsch and T. Lehmann, eds., pp. 101–105, Springer-Verlag, 2000.
5. S. C. Mitchell, B. P. F. Lelieveldt, R. van der Geest, J. Schaap, J. H. C. Reiber, and M. Sonka, "Segmentation of cardiac MR images: An active appearance model approach," in *Medical Imaging: Image Processing*, *Proc. SPIE* **3979**, pp. 224–234, 2000.

6. S. C. Mitchell, B. P. F. Lelieveldt, R. van der Geest, J. H. C. Reiber, and M. Sonka, "Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac MR images," *IEEE Transactions on Medical Imaging* **20**, pp. 415–423, 2001.
7. S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka, "Time-continuous segmentation of cardiac MR image sequences using active appearance motion models," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, eds., *Proc. SPIE* **4322**, pp. 249–256, 2001.
8. H. Bosch, S. Mitchell, B. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. Reiber, "Active appearance motion models for endocardial contour detection in time sequences of echocardiograms," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, eds., *Proc. SPIE* **4322**, pp. 257–268, 2001.
9. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. European Conference on Computer Vision*, H. Burkhardt and B. Neumann, eds., **2**, pp. 484–498, Springer, 1998.
10. T. F. Cootes, "Statistical models of appearance for computer vision," *Technical report, available at http://www.isbe.man.ac.uk/~bim/Models/app_model.ps.gz*, 2001.
11. T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Medical Imaging: Image Processing*, M. Sonka and K. M. Hanson, eds., *Proc. SPIE* **4322**, pp. 236–248, 2001.