3-D Active Appearance Models: Segmentation of Cardiac MR and Ultrasound Images

Steven C. Mitchell, Johan G. Bosch, Boudewijn P. F. Lelieveldt, Rob J. van der Geest, Johan H. C. Reiber, and Milan Sonka*, *Fellow, IEEE*

Abstract—A model-based method for three-dimensional image segmentation was developed and its performance assessed in segmentation of volumetric cardiac magnetic resonance (MR) images and echocardiographic temporal image sequences. Comprehensive design of a three-dimensional (3-D) active appearance model (AAM) is reported for the first time as an involved extension of the AAM framework introduced by Cootes *et al.* The model's behavior is learned from manually traced segmentation examples during an automated training stage. Information about shape and image appearance of the cardiac structures is contained in a single model. This ensures a spatially and/or temporally consistent segmentation of three-dimensional cardiac images.

The clinical potential of the 3-D AAM is demonstrated in short-axis cardiac MR images and four-chamber echocardiographic sequences. The method's performance was assessed by comparison with manually identified independent standards in 56 clinical MR and 64 clinical echo image sequences. The AAM method showed good agreement with the independent standard using quantitative indexes of border positioning errors, endo- and epicardial volumes, and left ventricular mass. In MR, the endocardial volumes, epicardial volumes, and left ventricular wall mass correlation coefficients between manual and AAM were $R^2 = 0.94, 0.97, 0.82$, respectively. For echocardiographic analysis, the area correlation was $R^2 = 0.79$. The AAM method shows high promise for successful application to MR and echocardiographic image analysis in a clinical setting.

Index Terms—Active appearance model, active shape model, cardiac segmentation, echocardiographic image analysis, magnetic resonance image analysis.

I. INTRODUCTION

CARDIOVASCULAR disease is the number one cause of death in the Western world. Cardiac imaging is an established approach to diagnosing cardiovascular disease and plays an important role in its interventional treatment. Three-dimensional (3-D) imaging of the heart and the cardiovascular system is now possible with X-ray computed tomography, magnetic resonance (MR), positron emission tomography, single photon emission tomography, and ultrasound, to name just the main imaging modalities. While cardiac imaging capabilities are developing rapidly, the images are mostly analyzed visually, and

therefore qualitatively. The ability to quantitatively analyze the acquired image data is still not sufficiently available in routine clinical care. Large amounts of acquired data are not fully utilized because of the tedious and time-consuming character of manual analyses. This is even more so when three-dimensional image data need to be processed and analyzed. Image segmentation is a prerequisite to quantitative analysis, and thus developing methods for highly automated three-dimensional cardiac image segmentation is of primary importance.

There are three main reasons why existing methods frequently exhibit lower the success rate in comparison with human expert observers, especially when applied to clinical-quality images: existing methods do not incorporate a sufficient amount of a priori knowledge about the segmentation problem; do not consider three-dimensional or temporal context as an integral part of their functionality; and position the segmentation boundaries at locations of the strongest local image features not considering true anatomical boundary locations.

A number of 3-D medical image analysis approaches have occurred recently, many of them addressing one or more of the above-mentioned shortcomings of available segmentation techniques. A detailed review of existing 3-D cardiac modeling approaches is provided in [1]. In the context of our work and considering the goal of segmenting three-dimensional volumetric and temporal cardiac images and image sequences, statistical modeling of 3-D shape and 3-D image properties is crucial. Vemuri *et al.* concentrated on a 3-D model that combines deformed superquadric primitives with a local displacement field expressed on an orthonormal wavelet basis [2]. As a result of this orthonormal basis, the shape parameters become physically meaningful, and thus a preferred shape can be imposed based on parameter distributions in a set of training samples. Similarly, Staib *et al.* developed a three-dimensional balloon model [3]. The model is parameterized on an orthonormal Fourier basis such that the statistics of the Fourier coefficients in a training set allow a constrained image search. Model fitting in these two methods is performed by balancing an internal energy term with an external, gradient-derived scalar field.

Cootes and Taylor developed a statistical point distribution model (PDM) and demonstrated its utility for two-dimensional (2-D) image segmentation [4], [5]. One of the primary contributions was an ease of automated learning of the model parameters from sets of corresponding points as well as the PDMs ability to incorporate shape and boundary gray-level properties and their allowed variations. Applications to segmentation of echocardiographic data [6] and deep neuroanatomical structures from MR images of the brain may serve as examples [7]. Following

the point distribution model ideas, Kelemen *et al.* built a statistical model of three-dimensional shapes using parametric surface representations [8]. Similar to PDMs, shape and gray-level information in the boundary vicinity was incorporated in the model. The method's performance was demonstrated on 3-D segmentation of neuroanatomical structures. A multiscale 3-D shape modeling approach called M-reps was developed by Pizer *et al.* [9]. M-reps support a coarse-to-fine hierarchy and model shape variations via probabilistically described boundary positions with width- and scale-proportional tolerances. Three-dimensional echocardiographic image segmentation using core atoms was reported by Stetten [10]. Davatzikos *et al.* presented a deformable model in which geometric information is embedded via a set of affine-invariant attribute vectors; these vectors characterize the geometric structure around a model point from a local to a global scale, forming an adaptive focus deformable statistical shape model [11]. The methodology was applied to segmentation of neuroanatomical structures.

In all the above-referenced approaches, the models primarily hold information about shape and its allowed variations. The information about image appearance is only considered in close proximity to the object borders. A powerful, model-driven segmentation technique called active appearance model (AAM) was recently introduced by Cootes *et al.* [12]–[14]. An AAM describes the image appearance and the shape of an object in a set of examples as a statistical shape-appearance model. AAMs can be applied to image segmentation by minimizing the difference between the model and an image along statistically plausible shape/intensity variations (analysis by synthesis). AAMs have shown to be highly robust in the segmentation of routinely acquired single-phase single-slice cardiac MR [15] and echo images [16], because they exploit prior knowledge about the cardiac shape, image appearance, and observer preference in a generic way. For a detailed background on AAMs and their application to image segmentation, the reader is referred to [13].

Until now, AAMs have only been applied to 2-D images and to normalized (fixed-phase) 2-D time sequences [15]–[17]. Two-dimensional active appearance motion models [16], [17] have demonstrated the ability of time-continuous segmentation by exploiting temporal coherency in the data. However, these 2-D + time AAMs do not represent a true 3-D approach. Their segmentation ability is limited to cases with fixed numbers of preselected frames; they rely on a priori knowledge of image frame correspondences within each cardiac cycle.

The primary contribution of this paper is the development of a fully three-dimensional active appearance model (3-D AAM) that requires no additional interactively supplied information. A demonstration of its segmentation performance in volumetric or temporal image segmentation of cardiac structures is given below. No 3-D AAM has been reported to date that is capable of successful segmentation of cardiac MR and echocardiographic images. The model's behavior is learned from manually traced segmentation examples during an automated training stage. The shape and image appearance of the cardiac structures are contained in a single model. This ensures a spatially and/or temporally consistent segmentation of three-dimensional cardiac images.

II. METHODS

A. Point Distribution Model Concept

Point distribution models describe populations of shapes using statistics of sets of corresponding landmarks of the shape instances [4], [5], [18]. By aligning N shape samples (consisting of n landmark points) and applying a principal component analysis (PCA) on the sample distribution, any sample $\mathbf{x}$ within the distribution can be expressed as an average shape $\overline{\mathbf{x}}$ with a linear combination of eigenvectors P superimposed

$$\mathbf{x} = \overline{\mathbf{x}} + P\mathbf{b}. \tag{1}$$

In two-dimensional models, $p = \min(2n, N - 1)$ eigenvectors P form the principal basis functions, while in a three-dimensional model; $p = \min(3n, N - 1)$ eigenvectors are formed. (The minimum operator is needed since we frequently have more corresponding shape points than training set samples.) In both cases, the corresponding eigenvalues provide a measure for compactness of the distribution along each axis. By selecting the largest q eigenvalues, the number of eigenvectors can be reduced, where a proportion k of the total variance is described such that

$$\sum_{i=1}^{q} \lambda_i \geq k \cdot \text{Total where Total} = \sum_{i=1}^{p} \lambda_i. \tag{2}$$

B. Representing Shape of 3-D Cardiac Ventricles

Extending the two-dimensional PDM to three dimensions is a nontrivial task. To create a compact and specific model, point correspondences between shapes are required. Even if landmark points are easily identifiable in both models, specifying uniquely corresponding boundary points in between these landmarks is difficult in 3-D. In a 2-D case [15], boundary points may be identified by evenly sampling points on a boundary from one landmark to the next. In a 3-D case, the problem lies in defining a unique sampling of the object surfaces.

For the purpose of ventricular segmentation, a normalized cylindrical coordinate system is defined with its primary axis aligned with the long axis of the heart and the secondary axis aligned with the posterior junction of the right and left ventricles in the basal slice. The cardiac ventricles resemble a cylindrical or paraboloid shape. First, contours are sampled slice-by-slice at even angle increments. To transform the rings in the normalized cylindrical coordinate system, each point on the ring is connected by a straight line to the next adjacent corresponding point on the rings above and below. Starting from the apex slice to the basal slice, a fixed number of slicing planes are placed evenly along the long axis. New points are interpolated where the planes intersect the lines. This yields a set of corresponding boundary points for each sampled left ventricle across the population of ventricles (Fig. 1).

C. Three-Dimensional Point Distribution Models

Aligning shape samples to a common scale, rotation, and translation is important for a compact model to be generated during the PCA stage. Procrustes analysis [19], [20] is used, whereby an arbitrary shape is selected as the initial average shape estimate. All the other shapes are aligned to this average
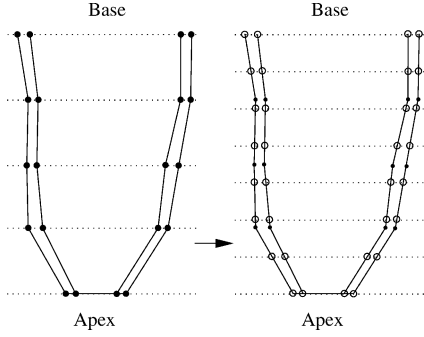
Fig. 1.   A cross-sectional depiction of transforming a cardiac MR stack with manually placed landmarks to a normalized cylindrical coordinate system.

using a least squares minimization. A new average is computed by a simple mean across the corresponding points, and the algorithm repeats until convergence.

For the 2-D case, aligning one shape to another can be solved analytically by minimizing scale, rotation, and translation terms. Extending to 3-D, the minimization of scaling, translations, and rotation differences along the three axes may lead to singularities known as gimbal lock. Assuming that 3-D translation is represented by a separate translation vector $\mathbf{t}$, a quaternion $\mathbf{q}$ representation of scaling and rotation avoids such behavior [21].

A quaternion $\mathbf{q}$ is defined as the linear combination of a scalar term $q_0 \geq 0$ and three right-handed orthonormal vectors ($\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$)

$$\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}. \tag{3}$$

The magnitude of the quaternion is defined as

$$|\mathbf{q}| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} \tag{4}$$

and any unit length quaternion can be written as

$$\mathbf{q} = \cos(\varphi) \cdot \mathbf{u} + \sin(\varphi) \cdot \mathbf{u} \tag{5}$$

where $\mathbf{u}$ is a unit vector and φ represents a rotational twist along the unit vector. Thus, any scaling and rotation in 3-D can be expressed as a quaternion, where scaling is expressed by the magnitude of the quaternion and the 3-D rotation is expressed by the direction of the unit vector $\mathbf{u}$ and rotation φ. The Cartesian rotation matrix is shown in (6) at the bottom of the page.

Together, the position and orientation of a 3-D object can be represented as a seven-element pose vector $(\mathbf{q}|\mathbf{t}) = [q_0, q_1, q_2, q_3, t_i, t_j, t_k]$.

The alignment of two 3-D shape instances is accomplished using a well-known procedure given by Besl *et al.* [22] to optimize for $\mathbf{q}$ and $\mathbf{t}$. Aligning all the shapes is a matter of employing the Procrustes analysis using Besl's procedure to calculate the pose parameters. Once shape alignment is finished, principal component analysis is applied to the 3-D models in a way that is no different from the conventional 2-D application [5].

D. Modeling Volume Appearance

The first part of creating an appearance model of volume is to warp all the sample volumes to the average shape to eliminate shape variation and bring voxel-wise correspondence across all the training samples, such that the voxel intensities can be represented as a shape-free vector of intensity values. Warping an image $\mathbf{I}$ to a new image $\mathbf{I}'$ involves creating a function that maps control points $\mathbf{x}_i$ to $\mathbf{x}'_i$ as well as the intermediate points in between. For the 2-D case, either piecewise affine warping or thin-plate spline warping is adequate. In our models, piecewise warping is preferred because it is significantly faster than thin-plate spline warping.

In 2-D piecewise affine warping, landmark points are used to construct the shape area as a set of triangles. The well-known Delaunay triangulation algorithm is suitable for computing such a triangular mesh and can be found in many computational geometry references. Individual triangular patches are locally warped using barycentric coordinates. Given a triangle with the three corners $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, we can represent any point $\mathbf{x}$ within the triangle as $\mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 + \gamma\mathbf{x}_3$, where $\gamma = 1 - (\alpha + \beta)$ and $\alpha + \beta + \gamma = 0$. In order for a point $\mathbf{x}$ to fall inside a triangle, $0 \leq \alpha, \beta, \gamma \leq 1$ must be true.

Piecewise affine warping is implemented as follows.

For each pixel location $\mathbf{x}'$ in $\mathbf{I}'$:

1) Find the triangle t' that contains $\mathbf{x}'$ by solving α, β, and γ for each triangle and finding the triangle where $0 \leq \alpha$, $\beta, \gamma \leq 1$.
2) Find the equivalent pixel location $\mathbf{x}$ by computing $\mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 + \gamma\mathbf{x}_3$, where $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are the triangle points from the original image.
3) Copy the pixel value in $\mathbf{I}$ located by $\mathbf{x}$ into the warped image $\mathbf{I}'$ located at $\mathbf{x}'$. Some form of pixel interpolation such as bilinear may be used at this stage.

In our 3-D models, piecewise affine warping is extended to tetrahedrons with four corners $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, and $\mathbf{x}_4$. Any point within the tetrahedron is represented as $\mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 + \gamma\mathbf{x}_3 + \delta\mathbf{x}_4$. In a general case, creating a tetrahedral representation of volume is solved using a 3-D Delaunay triangulation algorithm. However, due to the cylindrical nature of the left ventricular (LV) shape, a manually defined volume partitioning in regular tetrahedrons was utilized. Each slice level is constructed of pie-shaped wedges built on four tetrahedrons with exterior profile cubes built with five tetrahedrons. Piecewise affine warping is implemented in a similar fashion as in the 2-D case. Because all volumes are warped to the average volume, barycentric coordinates $\alpha, \beta, \gamma, \delta$ are precomputed for each fixed voxel point, eliminating the time-consuming process of searching for the enclosing tetrahedron for each voxel point during the matching. Due to the regular geometry of the tetrahedrons in our volume partitioning, the barycentric coordinate computation did not become ill-posed.

$$R = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \tag{6}$$

After the warping phase, the shape-free intensity vectors are normalized to an average intensity of zero and an average variance of one to remove the effects of brightness and contrast variations across scans. Next, PCA is applied to the shape-free intensity vectors to create an intensity model. In agreement with the AAM principle, shape information and intensity information are combined into a single active appearance model. Lastly, another PCA is applied to the coefficients of the shape and intensity models to form a combined appearance model [23].

In the equations below, the subscript s corresponds to shape parameters and the subscript g represents intensity (gray-level) parameters. To summarize, the 3-D AAM is created as follows.

1) Let $\mathbf{x}_i$ denote a vector of 3-D landmark points for a given sample i. Compute a 3-D PDM and approximate each shape sample as a linear combination of eigenvectors, where $\mathbf{b}_s = P_s^T(\mathbf{x} - \overline{\mathbf{x}})$ represents the sample shape parameters.
2) Warp each image to the mean shape using a warping such as piecewise affine or thin plate spline warping to create shape-free intensity vectors.
3) Normalize each intensity vector, applying a global intensity transform with parameters $\mathbf{h}_i$, to match the average intensity vector $\overline{\mathbf{g}}$.
4) Perform a PCA on the normalized intensity images.
5) Express each intensity sample as a linear combination of eigenvectors, where $\mathbf{b}_g = P_g^T(\mathbf{g} - \overline{\mathbf{g}})$ represents the sample shape parameters.
6) Concatenate the shape vectors $\mathbf{b}_s$ and gray-level intensity vectors $\mathbf{b}_g$ in the following manner:

$$\mathbf{b} = \begin{pmatrix} W\mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} WP_s^T(\mathbf{x} - \overline{\mathbf{x}}) \\ P_g^T(\mathbf{g} - \overline{\mathbf{g}}) \end{pmatrix} \qquad (7)$$

where the weighting matrix W is a diagonal matrix relating the different units of shape and intensity coefficients.
7) Apply a PCA to the sample set of all $\mathbf{b}$ vectors, yielding the appearance model

$$\mathbf{b} = Q\mathbf{c}. \qquad (8)$$

E. Matching 3-D AAM to Image Data

Matching an appearance model to image data involves minimizing the root mean square (rms) intensity difference between the image data and appearance model instance by modifying the affine transformation, global intensity parameters, and appearance coefficients. A gradient descent method is used that employs the relation between model coefficient changes and changes in the voxel intensity difference between the target image and synthesized model [23]. This relation is derived during a training stage.

Let $\mathbf{t}$ and $\mathbf{q}$ represent the translation and quaternion transformation parameters and $\mathbf{h}$ the intensity transform parameters. As shown above, shape $\mathbf{x}$ is derived in the target image from the appearance coefficient $\mathbf{c}$ and the affine transformation vectors $\mathbf{t}$ and $\mathbf{q}$. Then, shape intensity vector $\mathbf{g}_s$ is sampled from the target volume data after warping the space defined by $\mathbf{x}$ to the mean shape $\overline{\mathbf{x}}$. The model intensity vector $\mathbf{g}_m$ is derived from the appearance coefficients $\mathbf{c}$ with the global intensity corrected via $\mathbf{h}$. The error function E is the rms difference of $\mathbf{g}_s - \mathbf{g}_m$.

Gradient descent optimization requires the partial derivatives of the error function defined by the intensity of the target and synthesized model volume. While it is not possible to create such a function analytically, these derivatives may be approximated using fixed matrices computed by randomly perturbing model coefficients for a set of known training images and observing the resulting difference in error images [23]. Using a set of training images, their corresponding modeling parameters $\mathbf{c}$, $\mathbf{t}$, $\mathbf{q}$, and $\mathbf{h}$ are randomly displaced, thus creating a difference between $\mathbf{g}_s$ and $\mathbf{g}_m$. From the parameter displacements and the resulting difference intensity vectors, gradient approximating matrices A_c, A_t, A_q, and A_h can be determined using reduced-rank multivariate linear regression. Alternatively, the gradient matrices may be built one column at a time by averaging the Gaussian weighted differences between the target and synthesized image of each individual model perturbation. The latter method is preferred for 3-D AAM matching due to lower memory requirements, better representation of high order eigenmodes, and faster computation. This iterative refinement technique of precomputed fixed matrices versus brute-force gradient descent optimization was formulated by Cootes [13] as well as by Baker and Matthews [24]. Formally, the gradient matrices are created as follows.

1) Select an object from the training set with known appearance model parameters $\mathbf{c}_0$, $\mathbf{t}_0$, $\mathbf{q}_0$, and $\mathbf{h}_0$.
2) For each element in the model parameters, $\mathbf{c}$, $\mathbf{t}$, $\mathbf{q}$, or $\mathbf{h}$, perturb a single element by a fixed $\delta\mathbf{p}$ with the rest of $\delta\mathbf{c}$, $\delta\mathbf{t}$, $\delta\mathbf{q}$, and $\delta\mathbf{h}$ assigned to zero. Typically, $\mathbf{c}$ is perturbed within ± 1.5 standard deviation, $\mathbf{t}$ by 3–5 voxels, and $\mathbf{q}$, $\mathbf{h}$ by 10% of their original value.
3) Let $\mathbf{c} = \delta\mathbf{c} + \mathbf{c}_0$. Compute shape $\mathbf{x}$ and texture $\mathbf{g}_m$.
4) Apply an affine transformation to $\mathbf{x}$ by first transforming $\mathbf{x}$ using $\delta\mathbf{t}$ and $\delta\mathbf{q}$, then transforming the result by $\mathbf{t}_0$ and $\mathbf{q}_0$. This cascaded transform is required to maintain linearity.
5) Create the image patch $\mathbf{g}_s$ warped from the target image to the mean shape using shape $\mathbf{x}$.
6) Apply global intensity scaling to $\mathbf{g}_s$ by using $\delta\mathbf{h}$ first and then scaling the result by $\mathbf{h}_0$.
7) Compute $\delta\mathbf{g} = \mathbf{g}_s - \mathbf{g}_m$.
8) Compute the slope, $\delta\mathbf{s} = \delta\mathbf{g}/\delta\mathbf{p}$. Weight the slope by a normalized Gaussian function with the ± 3 standard deviation set to the maximum and minimum model perturbation values.
9) Accumulate the slope with previous slopes for that given element.
10) Go to Step 2) and repeat until all elements and perturbations of each element are sufficiently covered. Place the average slope into the appropriate column in the gradient matrices A_c, A_t, A_q, or A_h.
11) Go to Step 1) and repeat until there is sufficient coverage of displacement vectors.

The corresponding model correction steps are computed as

$$\delta\mathbf{c} = A_c(\mathbf{g}_s - \mathbf{g}_m) \qquad (9)$$
$$\delta\mathbf{t} = A_t(\mathbf{g}_s - \mathbf{g}_m) \qquad (10)$$
$$\delta\mathbf{q} = A_q(\mathbf{g}_s - \mathbf{g}_m) \qquad (11)$$
$$\delta\mathbf{h} = A_h(\mathbf{g}_s - \mathbf{g}_m). \qquad (12)$$

Matching the AAM to the image data is accomplished as follows.

1) Place the mean appearance model ($\mathbf{c}$, $\mathbf{h} = 0$; $\mathbf{t}$, $\mathbf{q}$ defined by the initial model position) roughly on the object of interest and compute the difference image $\mathbf{g}_s - \mathbf{g}_m$.

2) Compute the rms error of the difference image E.

3) Compute the model corrections $\delta\mathbf{c}$, $\delta\mathbf{t}$, $\delta\mathbf{q}$, and $\delta\mathbf{h}$ from the difference image [(9)–(12)].

4) Set $k = 1$.

5) Compute new model parameters as $\mathbf{c} := \mathbf{c} - k\delta\mathbf{c}$, $\mathbf{t} := \mathbf{t} - k\delta\mathbf{t}$, $\mathbf{q} := \mathbf{q} - k\delta\mathbf{q}$, and $\mathbf{h} := \mathbf{h} - k\delta\mathbf{h}$.

6) Based on these new parameters, recompute $\mathbf{g}_s - \mathbf{g}_m$ and find the rms error.

7) If the rms error is less than E, accept these parameters and go to Step 2).

8) Else try setting k to 1.5, 0.5, 0.25, 0.125, etc., and go to Step 5). Repeat steps 5–8 until the error cannot be reduced any further.

III. CASE STUDIES

To investigate the clinical potential of the reported 3-D active appearance model under clinically realistic conditions, AAMs were trained and tested in two substantially different medical imaging modalities: multislice short-axis cardiac magnetic resonance and four-chamber echocardiographic image sequences. While the overall approach is identical in the two reported application areas, some modality- as well as application-specific differences exist and are provided as follows.

A. Cardiac MRI

Data: Cardiac MR image sequences were collected from 38 normal subjects and 18 patients, yielding a total of 56 short-axis 3-D cardiac MR data sets. Patients were selected who were suffering from different common cardiac pathologies (among others, different types of myocardial infarction, hypertrophic cardiomyopathy, arrhythmia). Images were acquired using standard electrocardiography (ECG) gated fast field echo MR pulse sequences on a Philips Gyroscan NT 15 scanner. Slices were acquired in a per-slice manner, under breathhold in end-expiration. End-diastolic images were used in this study. Image resolution was 256×256 pixels, with a field of view of 400–450 mm and slice thickness of 8–11 mm. Between eight and 14 slices were scanned to at least cover the entire left ventricle, depending on LV dimensions and slice spacing.

Independent Standard: The left ventricular endocardium and epicardium were manually traced by an expert observer who was blinded to the results of the computer analysis. Following common practice in clinical quantitative cardiac MR analysis, endocardial (ENDO) contours were drawn behind the papillary muscles and trabeculae, and epicardial (EPI) contours were drawn on the inside of the epicardial fat layer. The apical slice was defined as the last slice with both ENDO and EPI contours visible. Apex slices with merely a small visible muscle cap were excluded. In each slice, a reference point was manually placed at the posterior junction of the left and right ventricular walls to define point correspondence between different samples as required during the AAM training.
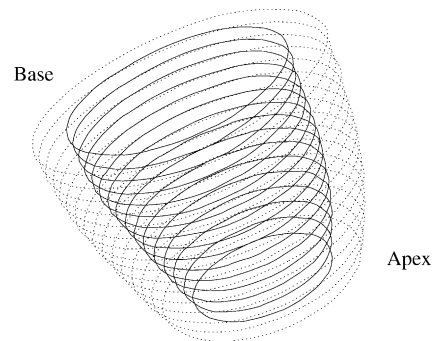

Fig. 2. A wireframe representation of the mean LV shape in the normalized cylindrical coordinate system.

Training Using Leave-One-Out Approach: To maximize the effective size of the training set, validation was performed using a leave-one-out approach [25]. AAM models were trained on image and contour data from 55 subjects, and the model matching performance was subsequently evaluated on the one left-out data set. The training process was repeated 56 times, always leaving out a different data set, which was then used for validation. The model voxel size varied from one data set to another, but in general with $3 \times$ subsampling, the model occupied approximately 6000 voxels and required about 1 h of training and 2–3 min for matching on a 1-GHz Windows PC.

Matching Procedure: In midventricular short-axis MR images, the left ventricle can usually be identified as an approximately circular object [Fig. 3(a)]. This fact is used for automated initialization of the 3-D AAM. A previously validated Hough transform-based method determines a 2-D centroid of the LV long axis for each MR image slice [26]. A 3-D centroid of a line segment fitted through the 2-D centroids of individual MR slices defines the initial position of the 3-D AAM.

To make the 3-D segmentation procedure completely independent from any user interaction regarding the rotation and scale of the heart in the short-axis plane, the matching process was repeatedly performed for a range of five orientations and three scales. This multiple initialization is important because AAM matching may be dependent on initial positioning since gradient descent may contain local minima. The matching result yielding the smallest quadratic intensity error was selected as the final match. The matching procedure resulted in a set of endo- and epicardial contours for each volumetric MR image.

Quantitative Validation: To exclude obvious matching errors from further quantitative analyses, matching results were visually evaluated. A matching result was scored as acceptable when the ENDO and EPI contours in the majority of slices showed good agreement with the image data. The number of unacceptable results yielding matching failures is reported. Studies with a poorly localized apex but with correctly localized midventricular segments were graded as acceptable. In some cases, the 3-D model did not deploy far enough to span over all the slices for which manual contours were available. Then, manual contours were present in extreme apical or basal slices but no automatic contours were identified there. Such segmentations were not excluded, and the numbers of slices missed by the AAM method are reported below.

To quantitatively assess the performance of the 3-D AAM approach, surface positioning errors were determined comparing
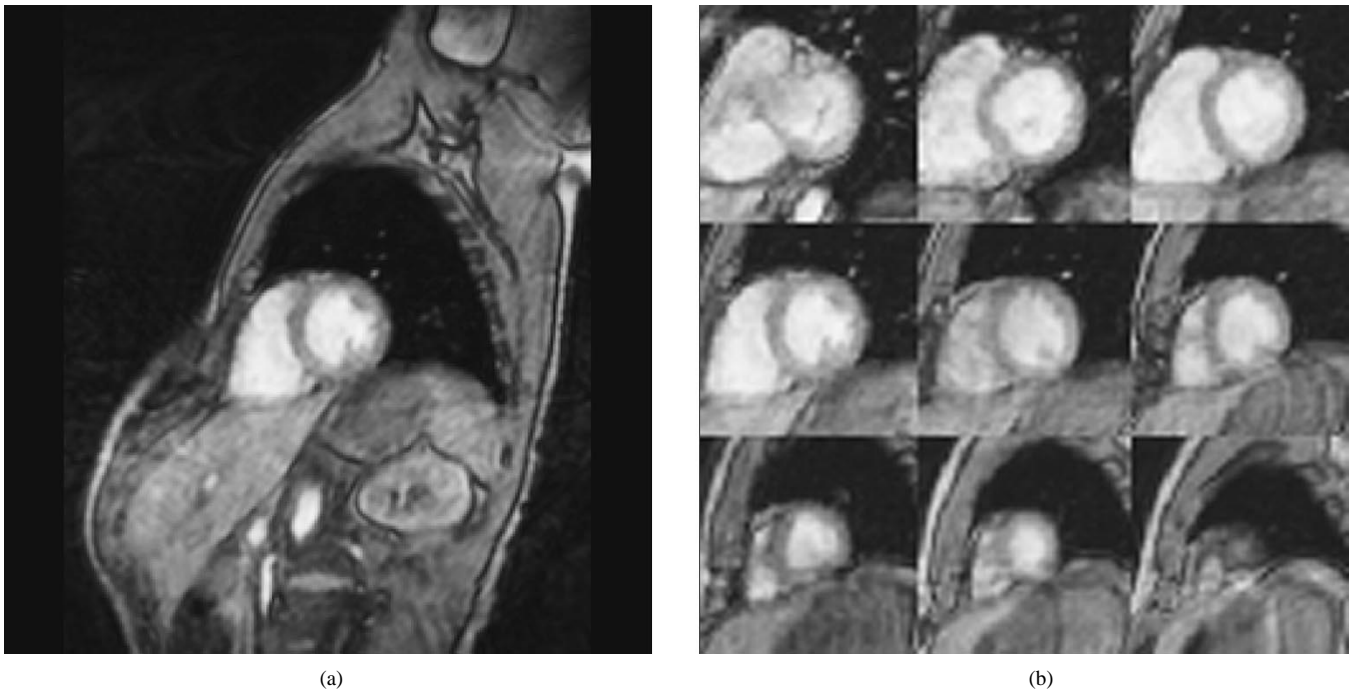
(a)   (b)

Fig. 3. Example cardiac MR images used for validation. (a) LV segmentation was performed in volumetric images consisting of 8–12 full-size MR images like the one shown here. (b) Subimages depicting LV detail in all nine images of this volumetric data set. See Fig. 4 for the segmentation results.
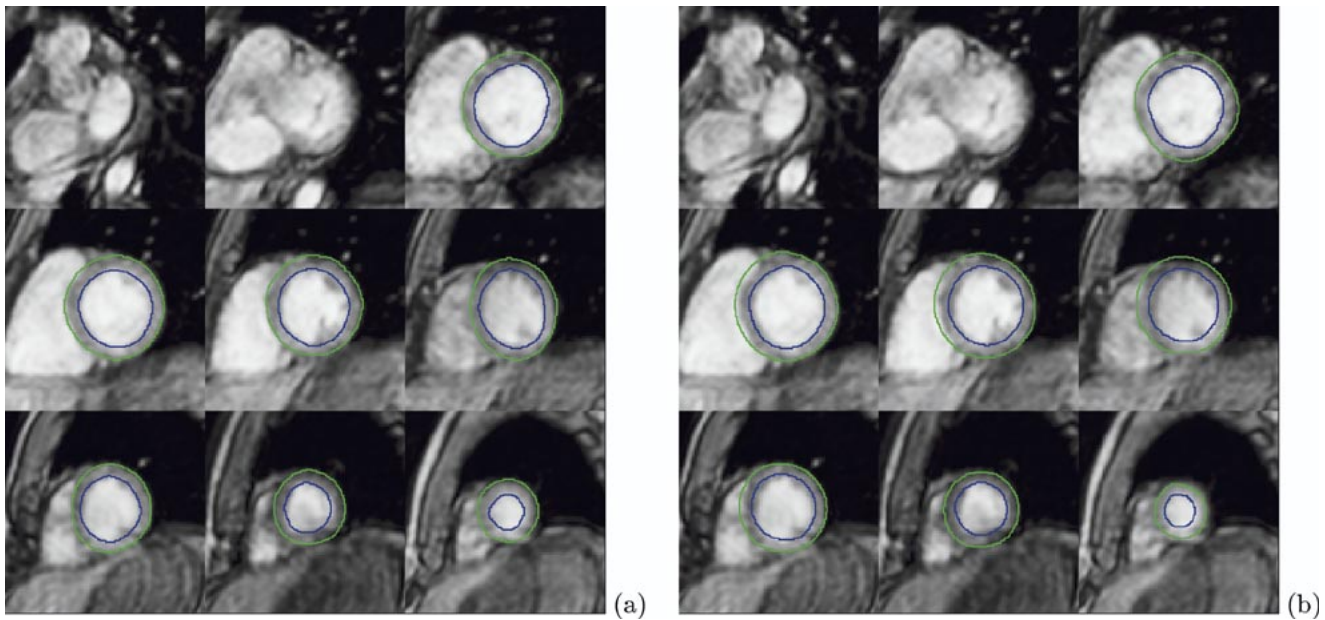

Fig. 4. Segmentation results in testing-set image data. (a) Manually identified contours forming an independent standard. (b) Three-dimensional AAM determined segmentation of the left ventricle. The 3-D AAM segmentation was performed in full-size image volumes; see Fig. 3.

the automatically detected endo- and epicardial surfaces with the independent standard. The average signed and unsigned surface positioning errors were defined by measuring the distances between points along rays perpendicular to the centerline between the respective manual contours and the computer-determined surfaces; 100 rays were used for each contour. Surface positioning errors are expressed in millimeters as mean ± standard deviation. A negative sign of the signed error value means that the automatically determined surface was inside of the observer-defined surface.

Three clinically important measures were calculated and used for performance assessment: LV cavity volume, LV epicardial volume, and LV myocardial mass. The volumetric indexes were determined using all slices for which both manually traced contours and computer-determined surfaces were available and were expressed in cm^3. The LV mass measurements are reported in g. Regression analysis was used to compare the computer measurements with the independent standard.

Results: Fig. 4 shows an example of an automatically analyzed volumetric MR data set. Fig. 5 demonstrates several stages
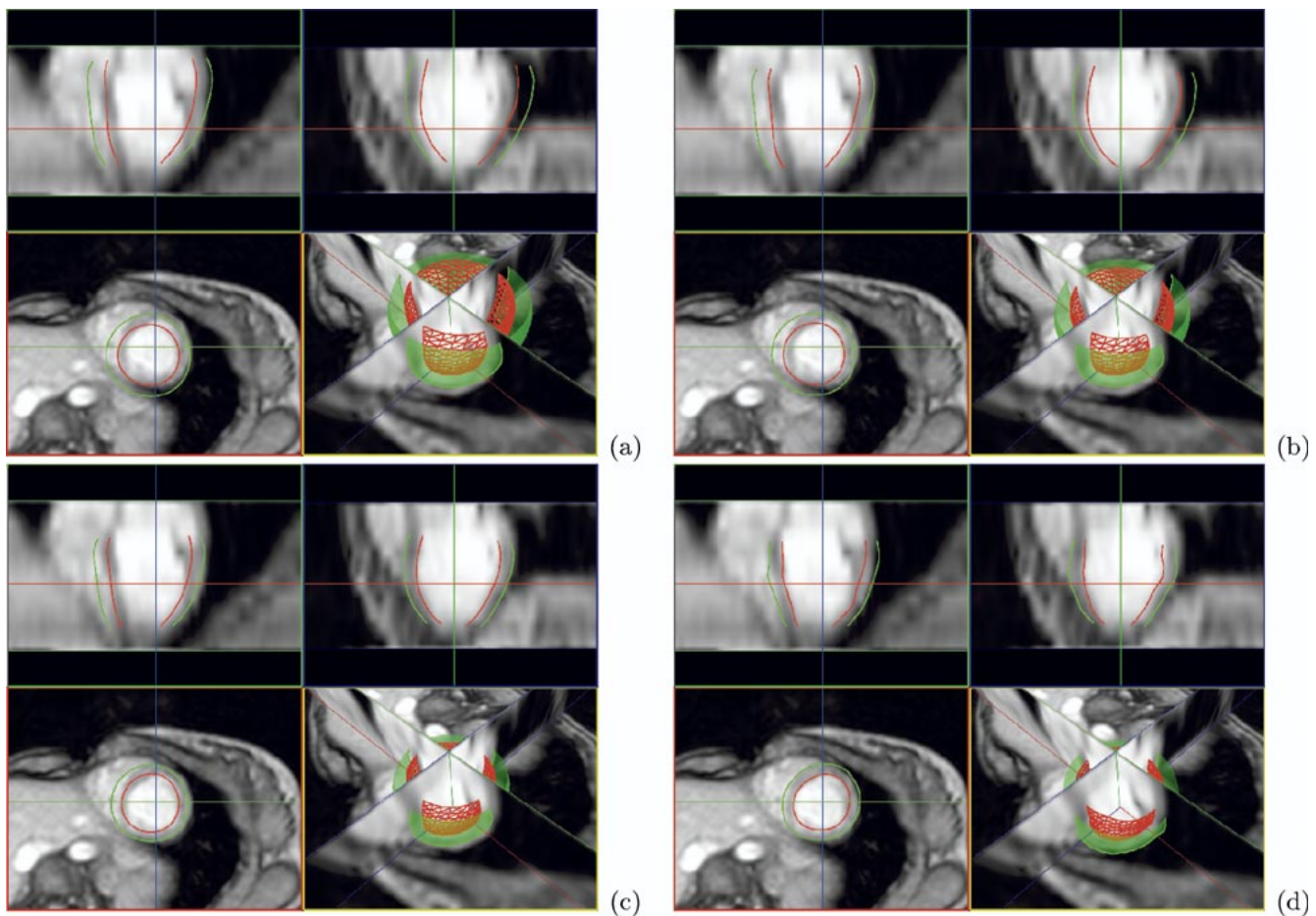
Fig. 5. 3-D AAM matching process. (a) The initial position of the model in the volumetric data set. (b), (c) Stages during the iterative model matching process. (d) The final match. Note the color coding of all frames and the coordinate axes. The color-coded straight lines show position of frames in the other two cutting planes.

of the model matching process, starting with the initial model position and ending with the final fit. The leave-one-out validation yielded 56 fully automatically detected 3-D segmentation results depicting endocardial and epicardial surfaces—with all testing surfaces being fully disjoint from the training sets. In 53 out of 56 cases, the matching procedure resulted in an acceptable match. In three cases (two patients and one normal subject), the matching diverged from a plausible solution because it was attracted by a neighboring structure. These three cases were excluded from further quantitative analyses.

In the 53 volumetric MR images, manually identified contours were available in 391 MR slices. As mentioned above, the 3-D model sometimes does not fully deploy in the longitudinal direction to cover the apical or basal slices. In 28 of 53 data sets, computer-determined contours were present in all slices for which manual contours were available. In 18 data sets, computer-determined contours were missing in one MR slice (apical or basal), and in seven data sets, two MR slices with missing contours were present. Overall, the 3-D AAM identified contours in 359 of 391 MR slices, or in 96.2% of all cases in which manual contours were deemed identifiable. The 359 slices were used for quantitative validation.

Mean signed endo- and epicardial surface positioning errors were -0.46 ± 1.33 mm and -0.29 ± 1.16 mm, respectively,

showing a slightly negative border detection bias. The mean unsigned positioning errors were 2.75 ± 0.86 mm for the endocardial contours and 2.63 ± 0.76 mm for the epicardium, demonstrating small absolute differences from the independent standard (voxel sizes ranged from $1.56 \times 1.56 \times 8$ mm to $1.76 \times 1.76 \times 11$ mm). Fig. 6 shows a very good correlation of the manually identified and 3-D AAM-determined LV endo- and epicardial volumes as well as correlation of computer-determined LV wall mass with the independent standard.

B. Echocardiography

The 3-D AAM segmentation was also applied to endocardial border detection in echocardiographic ultrasound image sequences. In these temporal sequences, the third dimension represents time. Spatial 2-D coordinates were converted from pixels into millimeters by applying the image calibration factor. A consistent 3-D set is formed by converting the time dimension (s) into a spatial dimension (mm) using a fixed ratio ("speed" of 40 mm/s). This ratio was chosen to obtain similar sizes of the 3-D object in all three dimensions. A consistent 3-D voxel space is obtained in which the shape and appearance of the heart over a full cardiac cycle can be modeled as a 3-D object. A 3-D AAM for this object was applied to image
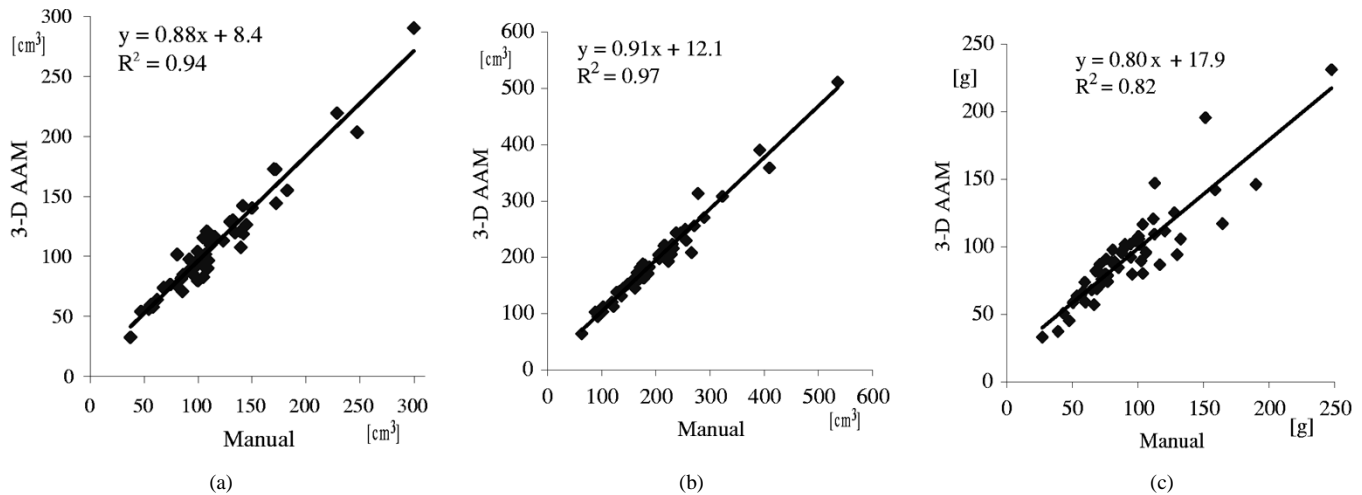
Fig. 6.   Comparison of observer-defined and computer-determined LV measurements in 53 volumetric MR images used for validation. (a) Endocardial volume. (b) Epicardial volume. (c) Myocardial mass.
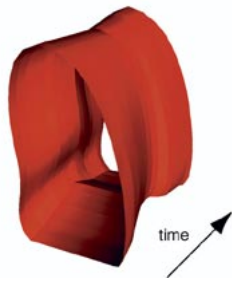

Fig. 7.   Motion of the left ventricular endocardium in ultrasound four-chamber cross-section represented as a 3-D object. Time axis from bottom left to top right.

sequences extending over three cardiac cycles. This approach allows fully automated detection of time-continuous contours in time sequences of arbitrary length without knowledge of end-diastolic (ED) and end-systolic (ES) time points. The model adjusts itself in both time and space and locates a complete cardiac cycle automatically. This is an important difference from the previously reported 2-D + time AAM approach [16].

In a training set, corresponding shape points on the endocardial contour are defined for each frame of one complete cardiac cycle (ED to ED) based on expert-drawn contours. Point correspondence in the third (time) dimension is defined by using the relative cardiac phase of each image. The relative cardiac phase of each image is modeled as a value between zero and two, defined for systolic frames as the frame's relative position between ED and ES (range 0–1) and for diastolic frames as its relative position between ES and following ED incremented by one (range 1–2). For each training case, 17 time slices with equidistant relative phase (starting and ending with ED) are identified by nearest neighbor interpolation. Of these frames, the 2-D shape point coordinates are extended to 3-D spatial coordinates by multiplying the frame's time by the speed ratio. The resulting 3-D shape is represented as a surface of 3-D points (Fig. 7). The image appearance of the heart is modeled as a vector of voxel intensity values in an image volume patch spanned by the manual contour surface, extended equally in all directions. Partitioning of the image volume into regular tetrahedrons was coded

specifically for this shape. The actual image volume was sampled sparsely (factor 5–7) after appropriate Gaussian filtering to limit the dimension of the intensity vector to a practical size of a few thousand voxels. The 3-D AAM provides time-continuous segmentation for one complete cardiac cycle, located automatically in both time and space. An ultrasound-specific initialization procedure was followed, in which the model was allowed to iterate from 25 different initializations (five 3-D size factors ranging between 0.9 and 1.3 compared to the average size and five temporal positions along the time axis). The matching result with the lowest quadratic intensity error was considered the best match.

Data: To allow comparison of 3-D AAM results with the 2-D + time AAM approach, the method was tested on the same data set [16]. Sets of transthoracic echocardiographic four-chamber sequences were acquired at 25 frames/s from 129 unselected infarct patients participating in a clinical trial. These were single-beat (end-diastole to end-diastole) sequences with 15–33 image frames per heartbeat artificially extended to three cardiac cycles as described above. Images were digitized at a resolution of 768×576 pixels with different calibration factors (0.28 to 0.47 mm/pixel). The total data set was split randomly into a training set of 65 patients and a testing set of 64 patients. The model voxel size varied from one data set to another, but in general with $8 \times$ subsampling the model occupied approximately 7000 voxels, required about 1 h of training and 2–3 min for matching on a 1-GHz Windows PC. The larger subsampling of the echocardiographic model reduced the voxel size and complexity to approximately the same as the MR model, resulting in similar training and matching times.

Independent Standard: In all sequences, an expert observer who was blinded to the computer analysis results manually outlined the contours of the endocardium in all frames of the image sequences. To compare the performance of 3-D AAMs with that of human experts, inter- and intraobserver variabilities of manual contour tracing were determined in a subset of 19 randomly chosen image sequences. Manual contour definition was independently repeated by the same expert observer and by another expert observer.
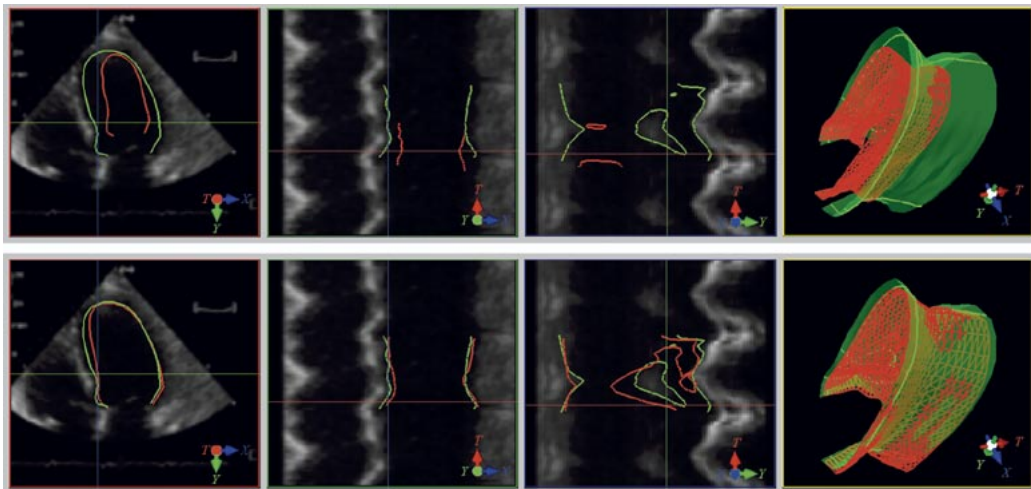
Fig. 8. Example of the 3-D AAM matching process in an ultrasound time sequence. The position of the model within the 3-D image data is shown in red. The independent standard is shown in green. Top row gives the initial positioning of the 3-D AAM; bottom row provides its final match. The left three panels represent image information along three perpendicular planes cutting through the three-dimensional data set. The leftmost column provides an image within the temporal plane; the two middle columns depict image data along the horizontal and vertical cutting planes. Note the color-coding of all frames and the coordinate axes. The color-coded straight lines show position of this frame in the other two cutting planes. In the rightmost panel, a 3-D rendering of model and independent standard is shown.

Quantitative Validation: To compare the automatically detected contours with the observer-identified independent standard, the number of segmentation failures was determined. Failures were identified as segmentations in which the 3-D AAM-defined borders did not agree well with the independent standard (average unsigned spatial distance component >7.5 mm). In the successfully segmented images, unsigned three-dimensional endocardial border positioning errors were defined as unsigned distances between matched model points and image-based 3-D shape points extracted for the testing-set images in a same manner as the training-set shape points. These distances were calculated in 3-D and also split into spatial (xy) and temporal (z) components, where the temporal component can be expressed in mm or milliseconds by inverting the "speed" conversion specified above. Furthermore, endocardial areas were determined for all time slices; regression analysis was used to compare the computer-determined areas with the independent standard.

Results: In 57 of the 64 tested echocardiographic image sequences (success rate 89%), the 3-D AAM-defined borders agreed well with the independent standard. An example of the matching procedure is given in Fig. 8. In the successful 57 temporal sequences, three-dimensional absolute endocardial surface positioning errors were 3.90 ± 1.38 mm; the 2-D spatial component was 3.35 ± 1.05 mm, which compares favorably with two-dimensionally determined (within the same image frame) interobserver variability of 3.82 ± 1.44 mm. The intraobserver variability was 2.32 ± 0.75 mm. The result also compares reasonably well with the previously reported two-dimensional endocardial border positioning errors of 3.35 ± 1.22 mm (success rate 97%) achieved by our less general 2-D + time active appearance motion model implementation [16]. The temporal error of border positioning was 37.0 ± 29.6 ms, less than a single frame duration of 40 ms. Frame-based endocardial area regression over the 57 successful matches is given in Fig. 9.
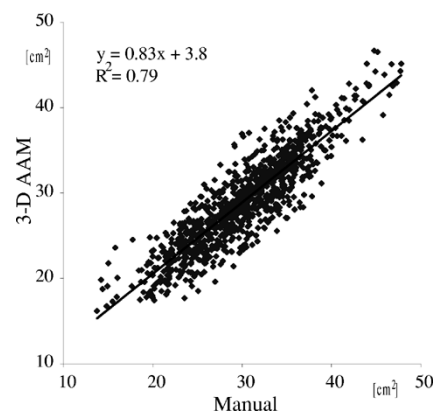

Fig. 9. Comparison of observer-defined and computer-determined LV endocardial areas in all time slices of the 57 echocardiographic four-chamber sequences (969 images) used for validation.

IV. DISCUSSION

A method for three-dimensional segmentation of medical images has been presented and its performance demonstrated in two cardiac image analysis applications. The method allows fully automated segmentation of volumetric or temporal images. Its inherent 3-D character incorporates information context in all three dimensions—a feature frequently not possible with previous approaches. The step leading from existing 2-D active appearance models to a functional 3-D AAM is not an easy one. Mastering the fully 3-D behavior will facilitate the method's extension to analysis of three-dimensional temporal sequences, one of the ultimate goals of medical image analysis. This discussion will focus on three areas: 1) performance of the current method; 2) its limitations with respect to design of 3-D AAMs for clinical applications; and 3) future extensions of the methodology to facilitate four-dimensional image analysis.

A. Performance of the Current Method

The described method is fully automated, with no user interaction. This is a significant strength in comparison to existing

semiautomated approaches to cardiac segmentation that require almost undivided attention of the analysis operator. In addition, the value of the matching function after the 3-D AAM's convergence can identify the situation when the model failed to segment the cardiac structures successfully, thus potentially allowing to incorporate yet another level of automated quality control. Consequently, it should be possible to identify segmentation successes with high confidence and then limit review of the results to a small number of image data sets for which the possibility of imperfect segmentation is indicated.

At present, performance robustness is achieved by multiple initializations; 15 initial positions of the model were used for MR segmentation and 25 different initializations for echocardiographic data. This approach improves the results but also increases computational demands. While the processing times are quite favorable—between 2–3 min for one data set in both applications (this includes the multiple initializations; 1-GHz Windows PC)—the number of model matching processes can be decreased by replacing brute-force multiple initializations with better positioning of the model. For example, information about the ventricular size and orientation is available in MR data from the Hough transform steps. In ultrasound, an estimate of ED and ES frames can be derived from ECG information or image sequence analysis, and there are several methods for estimating the size and position of ventricles automatically. No such information is currently utilized.

The border positioning errors as well as the volume and mass measures revealed a slight but systematic bias toward smaller contours. The presented 3-D AAM assumes the image data to be truly three-dimensional. However, the MR data sets are acquired over several heartbeats as sequences of 2-D slices, not in a volumetric fashion. Individual slices are frequently acquired during separate breath-holds. Differences in inspiration level and thus heart position lead to variable shifts in the LV position between slices. The effect of these shifts is threefold.

1) During the training of the model, unsystematic shifts are modeled in the shape model, decreasing the sensitivity of the model with respect to real shape variation.
2) During the matching, the match for multiple other adjacent slices is greatly disturbed by a single shifted slice.
3) The border positioning and volumetric errors are affected by the shifted slices.

This may well explain a large part of the underestimation of the reported MR volume measures.

For the ultrasound application, results were promising but less convincing than those achieved using our previously reported 2-D + time AAM approach. Area regression (Fig. 9) over the 57 successful matches exhibited a systematic area underestimation of 3.8% ($y = 0.83x + 3.6$, $R^2 = 0.79$). Compared to the 2-D + time AAM ($y = 0.91x + 1.73$, $R^2 = 0.76$; area underestimation 2.9%), a slightly higher systematic area error and a flatter regression line were observed. In part, this can be attributed to the extra degree of freedom with which the 3-D AAM has to cope. The 2-D + time AAM uses a priori knowledge of the phase/time aspect. However, the non-Gaussian distribution of intensity values in ultrasound is likely an even more important problem. Its indication is that model localization in the time dimension is much more accurate than in the spatial

localization. For the 2-D + time AAM approach in ultrasound, intensity distributions were normalized nonlinearly to deal with ultrasound-specific intensity properties [16]. This resulted in a substantially improved accuracy of the border localization. The nonlinear normalization improved both the systematic area underestimation and the slope of the area regression line. The 3-D version of this correction has not yet been developed. After its implementation, further improvements of segmentation accuracy in the echocardiographic images are expected.

Another problem observed in the testing case studies is associated with incomplete deployment of the model in the longitudinal (z) direction in the MR studies. In 32 of 391 MR slices, the contours were not determined by the 3-D AAM due to insufficient pulling force in the longitudinal direction, thus not covering one of the apical or basal slices. Currently, such slices were excluded from quantitative assessments. Missing slices may, however, substantially influence volumetric indexes of complete ventricles. Substantial MR slice thickness and consequently lack of detail in the z direction are the main reason. In the analyzed data sets, there is a lack of a clear "end" of the LV at the basal level. Making the 3-D AAM to cover the entire longitudinal length of the cardiac ventricles requires future attention. In addition, the apical part of the ventricle only asserts a small influence in the volumetric matching error function, due to the relatively small volume of the apex. As a result, the matching procedure tends to "sacrifice" apex accuracy for better mid-ventricular accuracy. Currently, improvements in apex boundary localization performance by nonuniformly weighting the three parts of the ventricle (apical, mid-ventricular, basal) are being investigated.

B. Current Limitations

Point correspondence between shapes is a critical factor for 3-D AAMs to become a general purpose method for segmentation of volumetric data. Due to the limited complexity of the left ventricle shape, simple parameterization of the surface was sufficient in this work. However, developing methodologies for creating point correspondence among a population of more complex 3-D shape instances is an actively researched area.

There are several solutions to developing correspondence in volumetric data. One of the simplest, representing a 3-D shape as a stack of 2-D contours, was used in our implementation. Another popular method is projecting landmark points on a spherical coordinate system, but this method is generally limited to single convex objects [27]. Lorenz *et al.* present a method whereby a 3-D template mesh is created from a shape instance. This template is deformed to other shape instances in the population providing landmark and intermediate points [28]. Duta *et al.* specify an automatic construction of 2-D shape models by approximating shape instances to polygons of a common number of boundary points. These polygons are aligned using Procrustes analysis, and the corresponding points are extracted from the nearest polygon vertices [29]. Although a 2-D method is reported, it may be extended to 3-D models. A framework for automated landmark identification in 2-D was reported by Hill *et al.* [30]. The correspondence algorithm that was developed with PDMs in mind locates a matching pair of sparse polygonal approximations by maximizing a landmark similarity cost function using a greedy algorithm. Extension

to 3-D is under investigation. Guest *et al.* describe point correspondence based on sensitivity of the point being matched to movement [31]. The driving idea is that a perturbation of reliable corresponding points shall not result in large displacements. The method was used for 2-D and 3-D registration of pre- and postoperative facial range scans. Kelemen *et al.* create point correspondence by an area-preserving parameterization followed by object-oriented normalization of its starting point, thus yielding a continuous mapping function between similar objects in 3-D [8]. Lastly, Davis *et al.* create statistical shape representation based on minimizing the total information required to encode and decode the original shape samples using information theory. This is done by creating corresponding points via a parametric representation of each individual shape and optimizing this function with respect to the compactness of the final shape model [32].

As with 2-D AAMs, another limitation of 3-D AAMs is an uncertainty of adequate covering of all pathologies in training sets. Although the AAM technique demonstrated segmentation reliability, there are always concerns using a model-based method built upon a finite set of training data to correctly handle often grossly abnormal patient data. One solution to this problem is to utilize larger training sets to cover all representative pathologies. User interaction is another solution to the limitations of a finite model. Cootes proposes adding a statistical framework to the AAM matching technique, thereby introducing a prior term to the matching phase. Using this model, a user can interactively provide constraints to landmarks improving the accuracy of the matching [33]. Another possible technique is allowing the user to interactively correct segmentation results and adapt the AAM model to incorporate these corrections. Taking advantage of the inherent linearity of AAM models, it is believed that such modifications may be used to improve the model. Lastly, local deformations are often lost due to the overgeneralizing of PCA given a finite number of samples in the training set. We previously reported the hybrid AAM whereby an edge-based technique such as ASM is employed simultaneously with the region-based AAM. After each iteration, the shape and pose results are combined using a weighted average. This method was originally proposed for fully automated segmentation of 2-D cardiac MR and shown to produce better results than AAM alone [15].

To keep data size, memory requirements, and processing time within acceptable limits, subsampling of image data was applied in generating the intensity models (up to a factor of eight for ultrasound). Although appropriate Gaussian filtering was used, this may compromise the overall accuracy. Improvements in speed and accuracy can be expected when a multiscale approach as described by Cootes is employed [12]. However, memory requirements remain a concern due to the 3-D nature of the processed data.

C. Extensions to Allow 4-D Image Analysis

The heart is a dynamic system, making time-independent segmentation inadequate. Applying 3-D AAM segmentation to the full cardiac cycle would require multiple models for different phases because any temporal knowledge of the interrelationship between frames would be lost. Several existing methods have been developed for 3-D + time cardiac segmentation taking into account the temporal relationship between frames. For example, one technique by Jacob *et al.* [34] solves temporal coherency between active shape models in echocardiograms through the use of a Kalman filter, creating a motion model to predict the cardiac cycle in addition to a shape model. An alternative method by Montagnat *et al.* [35] segments cylindrical echocardiographic images using deformable models. Here temporal coherence is accomplished by reinitializing the deformable model using the previous segmentation, while incorporating a 4-D anisotropic diffusion filter that significantly improves the spatial and temporal information between frames.

To extend the 3-D AAM framework to 3-D + time, we propose to incorporate a time element into the model by phase-normalizing objects to a common time correspondence and concatenating shape and texture vectors of individual phases into a single shape and texture vector. Such a technique has been found efficient in 2-D + time AAM and ASM models [16], [17], [36] and is promising as a future extension of 3-D AAMs.

V. CONCLUSION

A three-dimensional active appearance model method for analysis of volumetric cardiac images and temporal image sequences was presented and its performance demonstrated in two substantially different cardiac imaging modality case studies. To our knowledge, this is the first report describing a comprehensive design of a well-validated three-dimensional active appearance model-based segmentation. The model's training from manually traced segmentation examples as well as its segmentation of previously unseen images are fully automated. It carries substantial promise for successful application in a clinical setting.

REFERENCES

[1] A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images: A review," *IEEE Trans. Med. Imag.*, vol. 20, pp. 2–25, Jan. 2001.
[2] B. C. Vemuri and A. Radisavljevic, "Multiresolution stochastic hybrid shape models with fractal priors," *ACM Trans. Graph.*, vol. 13, pp. 177–207, Oct. 1994.
[3] L. H. Staib and J. S. Duncan, "Model-based deformable surface finding for medical images," *IEEE Trans. Med. Imag.*, vol. 15, pp. 720–731, 1996.
[4] T. F. Cootes, D. H. Cooper, C. J. Taylor, and J. Graham, "Trainable method of parametric shape description," *Image Vision Comput.*, vol. 10, no. 5, 1992.
[5] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vision Image Understand.*, vol. 61, pp. 38–59, 1995.
[6] A. Hill, T. F. Cootes, and C. J. Taylor, "Medical image interpretation: A generic approach using deformable templates," *Med. Inform.*, vol. 19, no. 1, pp. 47–59, 1994.
[7] N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images: An improved active shape model," *IEEE Trans. Med. Imag.*, vol. 17, pp. 1049–1062, Dec. 1998.

[8]   A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neurological data sets," *IEEE Trans. Med. Imag.*, vol. 18, pp. 828–839, Oct. 1999.

[9]   S. M. Pizer, S. Joshi, T. Fletcher, M. Styner, G. Tracton, and J. Z. Chen, "Segmentation of single-figure objects by deformable M-reps," in *MICCAI 2001*, W. Niessen and M. Viergever, Eds., Berlin, 2001, pp. 862–871.

[10] G. D. Stetten and S. M. Pizer, "Medial-node models to identify and measure objects in real-time 3-D echocardiography," *IEEE Trans. Med. Imag.*, vol. 18, pp. 1025–1034, Oct. 1999.

[11] D. Shen, E. H. Herskovits, and C. Davatzikos, "An adaptive focus statistical shape model for segmentation and shape modeling of 3D brain structures," *IEEE Trans. Med. Imag.*, vol. 20, pp. 257–270, Apr. 2001.

[12] T. F. Cootes, C. Beeston, G. J. Edwards, and C. J. Taylor, "A unified framework for atlas matching using active appearance models," in *Information Processing in Medical Imaging*, A. Kuba and M. Samal, Eds.   Berlin, Germany: Springer-Verlag, 1999, Lecture Notes in Computer Science, pp. 322–333.

[13] T. F. Cootes. (1999) Statistical models of appearance for computer vision. [Online]. Available: http://www.isbe.man.ac.uk/~bim/Models/app_model.ps.gz

[14] G. Edwards, C. Taylor, and T. Cootes, "Interpreting face images using active appearance models," in *Proc. 3rd Int. Conf. Automatic Face and Gesture Recognition 1998*, Nara, Japan, 1998, pp. 300–305.

[15] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka, "Cardiac segmentation using active appearance models," *IEEE Trans. Med. Imag.*, vol. 20, pp. 415–423, May 2001.

[16] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance models," *IEEE Trans. Med. Imag.*, to be published.

[17] B. Lelieveldt, S. Mitchell, J. Bosch, R. van der Geest, M. Sonka, and J. Reiber, "Time-continuous segmentation of cardiac image sequences using active appearance motion models," in *Information Processing in Medical Imaging*, ser. Lecture Notes in Computer Science.   Berlin, Germany: Springer-Verlag, 2001, vol. 2082, pp. 446–452.

[18] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 2nd ed.   Pacific Grove, CA: PWS, 1998.

[19] C. Goodall, "Procrustes methods in the statistical analysis of shape," *J. Roy. Stat. Soc. B*, vol. 53, no. 2, pp. 285–339, 1991.

[20] F. L. Bookstein, *Morphometric Tools for Landmark Data*.   Cambridge, U.K.: Cambridge Univ. Press, 1991.

[21] S. Altmann, *Rotations, Quaternions and Double Groups*.   Oxford, U.K.: Clarendon, 1986.

[22] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 239–256, Feb. 1992.

[23] T. F. Cootes, G. J. Edwards, and C. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 681–685, June 2001.

[24] S. Baker and I. Matthews, "Equivalence and efficiency of image alignment algorithms," in *Proc. Comput. Vision Pattern Recognition Conf.*, vol. 1, 2001, pp. 1090–1097.

[25] K. W. Bowyer, "Validation of medical image analysis techniques," in *Handbook of Medical Imaging*, M. Sonka and J. M. Fitzpatrick, Eds.   Bellingham, WA: SPIE, 2000, vol. 2, Medical Image Procesing and Analysis, pp. 567–607.

[26] R. J. van der Geest, V. G. M. Buller, E. Jansen, H. J. Lamb, L. H. B. Baur, E. E. van der Wall, A. de Roos, and J. H. C. Reiber, "Comparison between manual and semiautomated analysis of left ventricular volume parameters from short-axis MR images," *J. Comput. Assist. Tomogr.*, vol. 21, no. 5, pp. 756–765, 1997.

[27] C. Brechbuhler, G. Gerig, and O. Kubler, "Parameterization of closed surfaces for 3D shape description," *Comput. Vision Image Understand.*, vol. 62, pp. 154–170, 1995.

[28] C. Lorenz and N. Krahnstover, "3D statistical shape models for medical image segmentation," in *Proc. 2nd Int. Conf. 3-D Digital Imaging and Modeling*, 1999, pp. 414–423.

[29] N. Duta, A. K. Jain, and M.-P. Dubuisson-Jolly, "Automatic construction of 2D shape models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 433–446, May 2001.

[30] C. J. Hill, A. Taylor, and A. Brett, "A framework for automatic landmark identification using a new method of nonrigid correspondence," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 241–251, Mar. 2000.

[31] E. Guest, E. Berry, R. Baldock, M. Fidrich, and M. Smith, "Robust point correspondence applied to two- and three-dimensional image registration," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 165–179, Feb. 2001.

[32] R. Davies, T. F. Cootes, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," in *Proc. 17th Conf. Information Processing in Medical Imaging*, 2001, pp. 50–63.

[33] T. F. Cootes and C. J. Taylor, "Constrained active appearance models," *ICCV*, vol. 1, pp. 748–754, 2001.

[34] G. Jacob, A. Noble, M. Mulet-Parada, and A. Blake, "Evaluating a robust contour tracker on echocardiographic sequences," *Med. Image Anal.*, vol. 3, no. 1, pp. 63–75, 1999.

[35] J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache, "Anisotropic filtering for model based segmentation of 4D cylindrical echocardiographic images," *Pattern Recogn. Lett. (Special Issue on Ultrasonic Image Processing and Analysis)*, vol. 24, no. 4–5, pp. 815–828, Feb. 2003.

[36] G. Hamarneh and T. Gustavsson, "Deformable spatio-temporal shape models: Extending ASM to 2D + Time," in *Proc. 12th Br. Machine Vision Conf.*, 2001, pp. 13–22.