Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · April 2008

DOI: 10.1117/12.770710

5 authors, including:

Yefeng Zheng
Tencent
145 PUBLICATIONS **4,614** CITATIONS
SEE PROFILE

Bogdan Georgescu
Siemens
192 PUBLICATIONS **6,383** CITATIONS
SEE PROFILE

Adrian Barbu
Florida State University
109 PUBLICATIONS **2,220** CITATIONS
SEE PROFILE

Michael Scheuering
Siemens
45 PUBLICATIONS **1,297** CITATIONS
SEE PROFILE

Four-Chamber Heart Modeling and Automatic Segmentation for 3D Cardiac CT Volumes

Yefeng Zheng[a], Bogdan Georgescu[a], Adrian Barbu[b*],
Michael Scheuering[c], and Dorin Comaniciu[a]

[a]Integrated Data Systems Department, Siemens Corporate Research, Princeton, NJ, USA
[b]School of Computational Science, Florida State University, USA
[c]Siemens Medical Solutions CTE-PA, D-91301 Forchheim, Germany

ABSTRACT

Multi-chamber heart segmentation is a prerequisite for quantification of the cardiac function. In this paper, we propose an automatic heart chamber segmentation system. There are two closely related tasks to develop such a system: heart modeling and automatic model fitting to an unseen volume. The heart is a complicated non-rigid organ with four chambers and several major vessel trunks attached. A flexible and accurate model is necessary to capture the heart chamber shape at an appropriate level of details. In our four-chamber surface mesh model, the following two factors are considered and traded-off: 1) accuracy in anatomy and 2) easiness for both annotation and automatic detection. Important landmarks such as valves and cusp points on the interventricular septum are explicitly represented in our model. These landmarks can be detected reliably to guide the automatic model fitting process. We also propose two mechanisms, the rotation-axis based and parallel-slice based resampling methods, to establish mesh point correspondence, which is necessary to build a statistical shape model to enforce priori shape constraints in the model fitting procedure.

Using this model, we develop an efficient and robust approach for automatic heart chamber segmentation in 3D computed tomography (CT) volumes. Our approach is based on recent advances in learning discriminative object models and we exploit a large database of annotated CT volumes. We formulate the segmentation as a two-step learning problem: anatomical structure localization and boundary delineation. A novel algorithm, Marginal Space Learning (MSL), is introduced to solve the 9-dimensional similarity transformation search problem for localizing the heart chambers. After determining the pose of the heart chambers, we estimate the 3D shape through learning-based boundary delineation. Extensive experiments demonstrate the efficiency and robustness of the proposed approach, comparing favorably to the state-of-the-art. This is the first study reporting stable results on a large cardiac CT dataset with 323 volumes. In addition, we achieve a speed of less than eight seconds for automatic segmentation of all four chambers.

Keywords: Heart modeling, heart segmentation, 3D object detection

1. INTRODUCTION

Cardiac computed tomography (CT) is an important imaging modality for diagnosing cardiovascular disease and it can provide detailed anatomic information about the cardiac chambers, large vessels or coronary arteries. Segmentation of cardiac chambers is a prerequisite for quantitative functional analysis and various approaches have been proposed in the literature.[1,2] Except for a few instances,[3,4] most of the previous research focuses on the left ventricle (LV) segmentation. However, complete segmentation of all four heart chambers, as shown in Fig. 1, can help to diagnose diseases in other chambers, *e.g.*, left atrium (LA) fibrillation, right ventricle (RV) overload or to perform dyssynchrony analysis.

There are two major tasks to develop an automatic segmentation system: heart modeling (shape representation) and automatic model fitting (detection or segmentation). These two tasks are closely related. Different

*Adrian Barbu contributed to this work when he was with the Integrated Data Systems Department of Siemens Corporate Research.

Further author information: Send correspondence to Yefeng Zheng, yefeng.zheng@siemens.com.

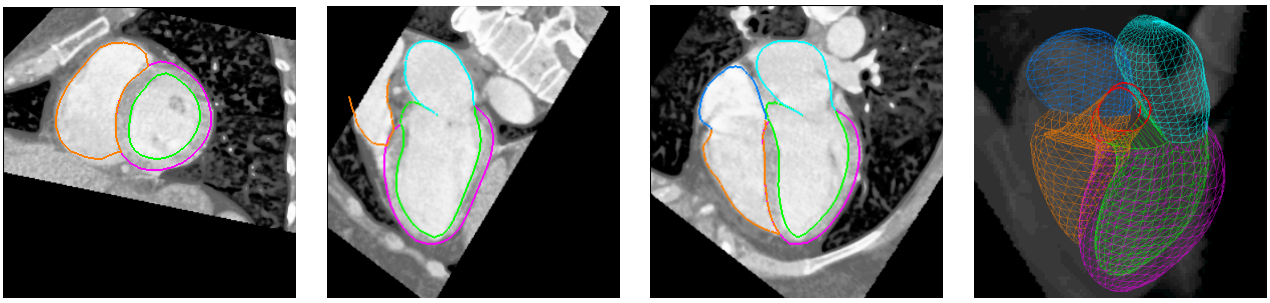


Figure 1. Complete segmentation of all four chambers in a CT volume with green for the left ventricle (LV) endocardial surface, magenta for LV epicardial surface, cyan for the left atrium (LA), brown for the right ventricle (RV), and blue for the right atrium (RA). The first three columns show three orthogonal cuts from the 3D volume data and the last column shows the triangulated surface mesh embedded in the volume.

detection methods need different kinds of priori information about the object, therefore, different shape representations. In the literature, 3D active shape models (ASM)[5] and active appearance models (AAM)[6,7] are very popular. ASM only needs the boundary information, therefore, uses the surface-based representations. However, AAM needs the information inside the object, therefore, volume-based representations, such as tetrahedrons, are exploited.

On the other hand, for heart modeling, we also need to consider the easiness of post-editing. No automatic segmentation is perfect, post-editing by a doctor is necessary in a real clinical application. Due to the complexity of cardiac anatomy, it is not trivial to represent the anatomy accurately while keep the model simple enough for easy post-editing. A flexible model is necessary to capture the heart chamber shape at an appropriate level of details.

1.1 Related Work on Heart Modeling

Though, both volumetric meshes and surface meshes are proposed for heart modeling, the latter is more popular and well suited for automatic heart chamber detection and segmentation. The image intensity inside a heart chamber is uniform. A large intensity jump is often observed on the boundary between different parts. Therefore, object boundary contains most information for heart chamber segmentation. Furthermore, much fewer vertices are required to represent a surface mesh comparing to a volumetric representation, so it is faster to deform a surface mesh during shape detection. Another advantage for a surface representation is that visualization of a surface mesh is much easier.

In 3D heart segmentation, surface representations are widely used in the snake-based deformable models.[3,8–12] Deformable models use little shape prior information, therefore they are likely to be stuck in local optima. Statistical shape models are often used to enforce shape constraints to make the system more robust. However, to build a statistical shape model, it is necessary to establish the point correspondence among a group of shapes. A few automatic approaches were proposed to build a statistical 3D shape model. Pair-wise based approaches[13,14] are easy and fast. By randomly selecting a reference shape and its mesh representation, shape registration is performed between reference and each other shape in the group. Then, the reference mesh is warped toward each individual shape to achieve a consistent sampling for all training shapes. Davies *et al.*[15] proposed a group-wise based method. They formulated the selection of pseudo-landmarks as an optimization problem to minimize the minimum description length (MDL). The problem of this approach is that the number of variables (the positions of pseudo-landmarks on each training shape) for optimization is large if we want to get a dense representation. The optimization process is slow and likely to converge to a local optimum.

Another approach is to establish correspondence among shapes during the manual labeling process. It is relatively easy for 2D curves. Generally, we only need to label a few landmarks, such as points with a high curvature. Uniform sampling between two neighboring landmarks on the curve will result in a dense point correspondence. However, it is difficult to manually label the correspondence in 3D since much more points are involved and there is no natural ordering of mesh points. Fortunately, for a few simple shapes, such as a tube or a parabola, we can consistently resample the surface to establish correspondence.[16]

1.2 Related Work on Heart Segmentation

Given a heart chamber model, we need to develop a scheme to fit the model onto an input volume automatically. Since the heart is a non-rigid shape, the model fitting (or heart chamber segmentation) procedure can be divided into two steps: object localization and boundary delineation. Most of the previous approaches focus on boundary delineation based on active shape models (ASM),[5] active appearance models (AAM),[6,7] and deformable models.[3,8–12] There are a few limitations inherent in these techniques: 1) Most of them are semi-automatic and manual labeling of a rough position and pose of the heart chambers is needed. 2) They are likely to get stuck in local strong image evidence. Other techniques are straightforward extensions of 2D image segmentation to 3D.[17–19] The segmentation is performed on each 2D slice and the results are combined to get the final 3D segmentation. However, such techniques cannot fully exploit the benefit of 3D imaging in a natural way. Lorenzo-Valdés *et al.*[20] proposed a registration based approach, but its performance is not clear for large datasets.

Object localization is required for an automatic segmentation system and discriminative learning approaches have proved to be efficient and robust for solving 2D problems. In these methods, shape detection or localization is formulated as a classification problem: whether an image block contains the target shape or not.[21] To build a robust system, a classifier only has to tolerate limited variation in object pose. The object is found by scanning the classifier exhaustively over an range of possible locations, orientations, scales or other parameters in an image. This searching strategy is different from other parameter estimation approaches, such as deformable models, where an initial estimate is adjusted (*e.g.*, using the gradient descent technique) to optimize a predefined objective function.

Exhaustive searching makes the system robust under local minima. However, there are two challenges to extend the learning based approaches to 3D. First, the number of hypotheses increases exponentially with respect to the dimensionality of the parameter space. For example, there are nine degrees of freedom for the anisotropic similarity transformation[†], namely three translation parameters, three rotation angles, and three scales. Suppose we search n discrete values for each dimension, the number of tested hypotheses is n^9 (for a very coarse estimation with a small n=5, n^9=1,953,125). The computational demands are beyond the capabilities of current desktop computers. Due to this limitation, previous approaches often constrain the search to a lower dimensional space. For example, only the position and isotropic scaling (4D) is searched in the generalized Hough transformation based approach.[22] Hong *et al.*[16] extended the learning based approach to a 5D parameter space for semi-automatic segmentation. The second challenge is that we need efficient features to search the orientation and scale spaces. Haar wavelet features can be efficiently computed under translation and scaling.[21] However, to search for rotation parameters, one either has to rotate the feature templates or rotate the volume (the latter is time consuming). The efficiency of image feature computation becomes more important when combined with a huge number of test hypotheses.

1.3 Overview of Our Approach

In this paper, we propose a new surface-based four-chamber heart model, which has the following advantages:

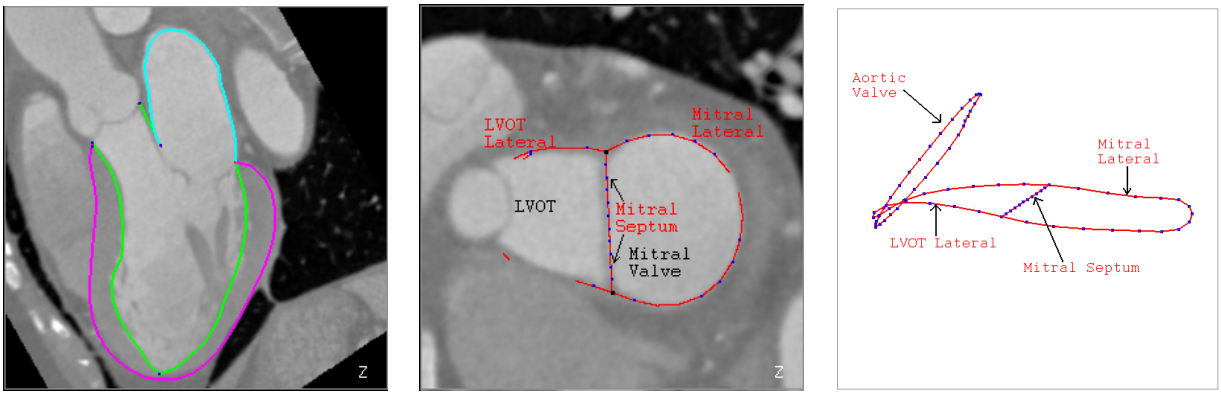
1. It is more accurate in anatomy comparing to the previously used closed surface mesh model.[16] Each chamber is represented as a mesh with one or few holes for the inflow/outflow tracts.

2. We propose two approaches, the rotation-axis based and parallel-slice based methods, to enforce the mesh point correspondence during mesh editing. No time-consuming and error-prone 3D shape registration is necessary.

3. Important landmarks (*e.g.*, valves, cusp points on the interventricular septum) are explicitly represented in our model as control points. These landmarks can be detected reliably to guide the automatic model fitting process.

4. Our model is flexible. Chambers are coupled at atrioventricular valves and it is easy to extract each chamber from the whole heart model.

[†]The ordinary similarity transformation allows only isotropic scaling. In this paper, we search for anisotropic scales to cope better with the non-rigid deformation of the shape.



(a) (b) (c)

Figure 2. Control points for delineating the left ventricle (LV) and left atrium (LA) meshes. (a) LV/LA mesh embedded in a CT volume. (b) Control points around the mitral valve. Since the curves are 3D, only parts of mitral lateral and LVOT lateral are visible on a specific plane. (c) Control points of mitral valve and aortic valve in a 3D view.

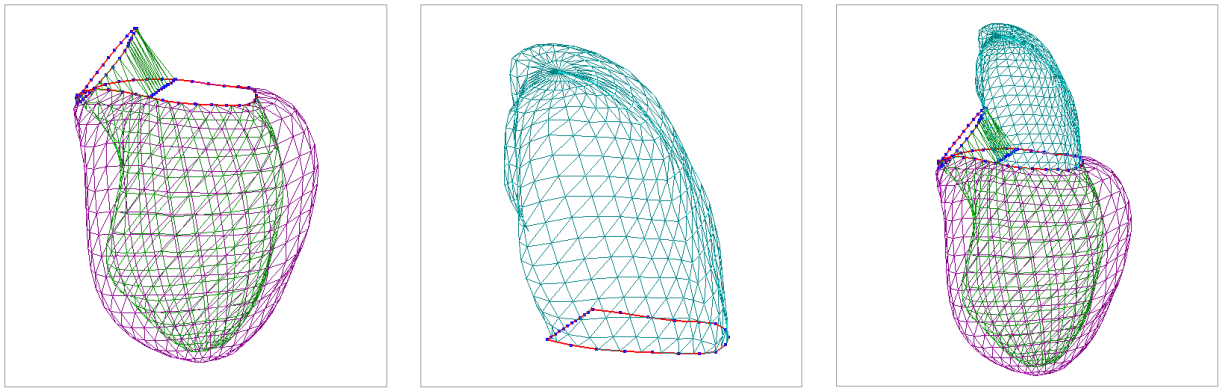


Figure 3. LV/LA meshes with green for LV endocardial surface, magenta for LV epicardial surface, and cyan for LA. LVOT is visualized with the same color as LV endocardial surface. From left to right are LV, LA, and the combined meshes, respectively.

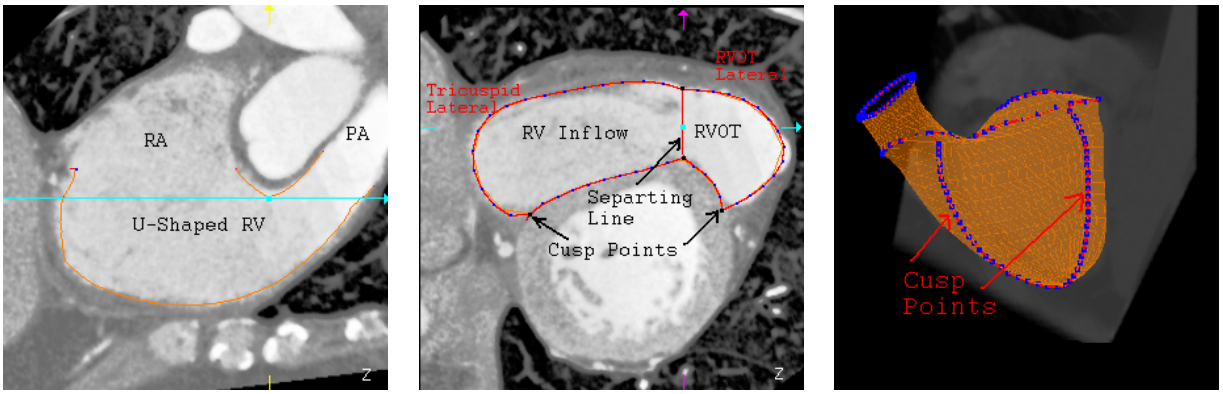
Using the proposed model, we present an efficient, robust, and fully automatic segmentation method for 3D cardiac CT volumes. Our approach is based on recent advances in learning discriminative object models and we exploit a large database of annotated CT volumes. We formulate the segmentation as a two-step learning problem: anatomical structure localization and boundary delineation. A novel algorithm, Marginal Space Learning (MSL), is introduced to solve the 9-dimensional similarity transformation search problem for localizing the heart chambers. After chamber pose estimation using MSL, we detect the control points and use them to warp the initial estimate of the shape. At last, we deform the mesh to fit the image boundary using a learning-based boundary delineation scheme. Extensive experiments on multi-chamber heart segmentation demonstrate the efficiency and robustness of the proposed approach.

2. FOUR-CHAMBER HEART MODELING

In this section, we first describe our four-chamber heart model, and then present our consistent resampling techniques to establish point correspondence, which is necessary to build a statistical shape model.

2.1 LV and LA Models

The major part of LV has a simple shape, which is roughly rotation symmetric. Previously, a closed mesh is often used to represent LV by excluding left ventricular outflow tract (LVOT).[16] There are two limitations with such simplification: 1) It is hard to delineate the interface between LVOT and the major body of LV since there



(a) (b) (c)

Figure 4. Control points for delineating right ventricle (RV) and right atrium (RA). (a) U-shaped silhouette of RV, RA, and pulmonary artery (PA). The divergence plane is indicated by the line in the middle, which passes the divergence of RV inflow and outflow tracts. (b) Control points (tricuspid lateral, RVOT lateral, and ventricular septum cups) labeled on the divergence plane. (c) RV mesh embedded in the volume with interventricular septum cusps indicated.

is no distinctive image features on the interface. 2) Similarly, it is hard to label the basal area around the mitral valve in a consistent way.

In this paper, we explicitly model the mitral valve annulus by opening a hole at that part of the mesh. The new model can be annotated more consistently and it will benefit the automatic detection process. We want to model both the endo- and epi-cardial borders of LV. The commissure contour of both borders corresponds to the mitral valve annulus on one side and the aortic valve level (lying at the bottom edge of the Valsalva sinuses) on the other side. Three curves are explicitly annotated around this area, namely, mitral lateral, mitral septum, and LVOT lateral, as shown in Fig. 2. These three curves define two closed regions, one for the interface between LV and LA, and the other for LV main body and LVOT, as shown in Fig. 2c. The aortic valve is approximated as a plane, which cuts the valve at the bottom of the Valsalva sinuses. A degenerated tube is used to represent the LVOT with one end defined by the aortic valve and the other end defined by the region enclosed by mitral septum and LVOT lateral control points. LA is represented as an open mesh with the hole enclosed by the mitral septum and mitral lateral control points.

Fig. 3 shows the LV/LA mesh with the control points. In these figures, the LV endocardial, epicardial, and LA meshes are represented with 545 points and 1056 triangles each. The LVOT mesh is represented with 64 points and 64 triangles.

2.2 RV and RA Models

RV has a more complex shape comparing to the other chambers. It is a truncated ventricle with separate inflow and outflow tracts. The shape of the short axis intersection of the RV main body is a crescent. The chamber poorly approximates to any convenient geometric shape. Since we know how to consistently sample a mesh for only a few limited shapes, we split the RV shape into three pieces, each has a simple shape and can be consistently sampled using the techniques discussed in Section 2.3. As shown in Fig. 4a, the silhouette of RV, RA, and the pulmonary artery is U-shaped with the inflow and outflow tracts of RV as two arms of the character 'U'. Using a plane (indicated by a cyan line in Fig. 4a) passing the divergence point of the inflow and outflow tracts, we can split RV into three parts with RV main body lying below the cutting plane. This plane is named as RV divergence plane in the following. Similar to the LV model, control points (tricuspid lateral and RVOT lateral) are annotated on the RV divergence plane to define the inflow and outflow connection, as shown in Fig. 4b.

The shape of the short axis view of RV main body is a crescent, as shown in Fig. 4b. Two cusp points on the intersection are important landmarks, and an automatic detection algorithm should be able to detect them reliably. If we do not treat the cusp points specially, the mesh will round-off at these locations after imposing the smoothness constraint. Therefore, we represent these landmark points explicitly in our model.

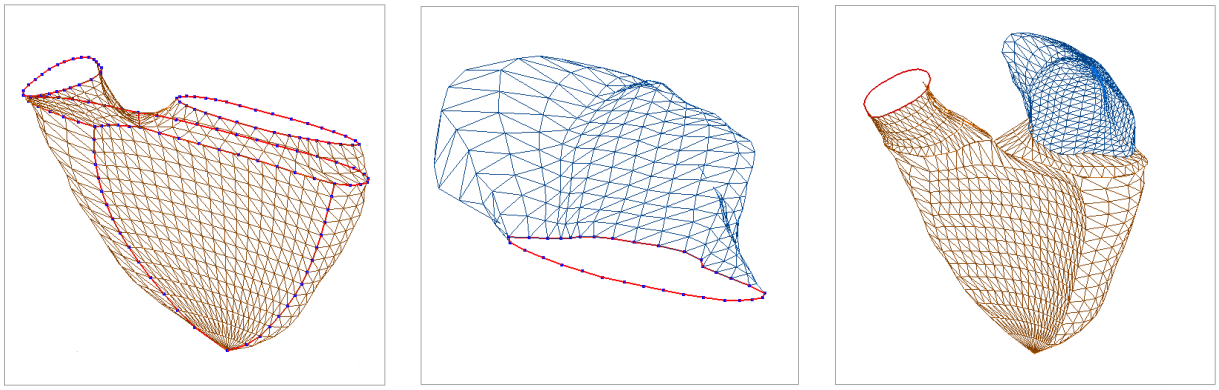
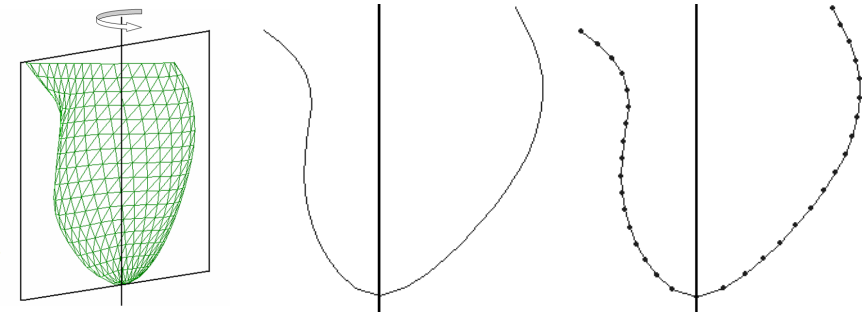


Figure 5. RV/RA meshes with brown for RV and blue for RA. From left to right are RV, RA, and the combined meshes, respectively.



(a) (b) (c)

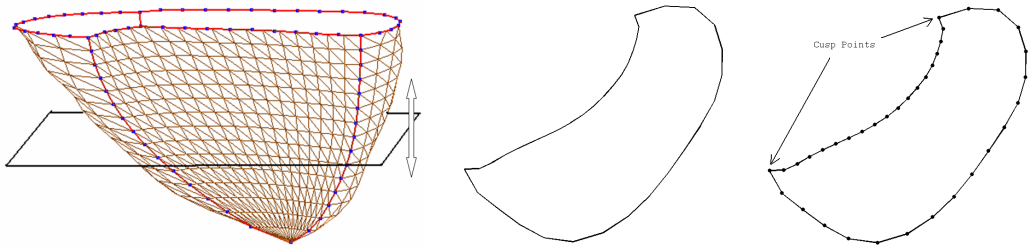
Figure 6. Rotation-axis based resampling method demonstrated for LV endocardial mesh resampling. (a) LV endocardial mesh with its long axis. A cutting plane passing the long axis is also illustrated. (b) The intersection of the mesh with the cutting plane. (c) After uniform resampling (each dot represents a sampling point). The axis separates the intersection into two parts and each is resampled independently.

Similarly, the tricuspid valve and the pulmonary valve are also approximated as a plane with control points explicitly defined to delineate the shape. Similar to LA, RA is represented as an open mesh with the hole defined by the tricuspid valve. Fig. 5 shows the mesh with the control points. In our model, RV is represented with 813 points and 1568 triangles; RA is represented with 545 points and 1056 triangles.

2.3 Establishing Point Correspondence

During manual labeling of ground-truth, we edit the mesh to fit the image data. In this process, there is no point correspondence between two independently labeled meshes. Since our automatic segmentation algorithm exploits a statistical shape model, we need to establish point correspondence. For a few simple shapes, such as a tube or a parabola, we can consistently resample the surface to establish the correspondence. We propose two resampling techniques to establish the point correspondence: the rotation-axis based method for simple parabola-like shapes such as LV, LA, and RA, and the parallel-slice based method for RV, which has a more complicated shape. Comparing to other approaches,[13–15] our solution is easy and each shape is processed independently. In our application, we found the statistical shape model built by the proposed approach is good enough to enforce priori shape constraints, as shown in Fig. 10. If more accurate point correspondence is demanded, other optimization based approaches can be exploited.[13–15] Since optimization based approaches are likely to converge to a local optima, the quality of the final solution heavily depends on the initialization. Our method can provide a quite good initialization for those optimization based approach.

Since it is easy to consistently resample a 2D curve (open or closed), we use a few planes to cut the 3D mesh to get a set of 2D intersection curves. The resulting 2D curves are uniformly sampled to get a point set with build-in correspondence. Using different methods to select cut planes, we develop two resampling schemes, the rotation-axis based method and parallel-slice based method, for a few simple geometric shapes.



(a) (b) (c)

Figure 7. Parallel-slice based resampling method for RV main body. (a) RV mesh. A cutting plane perpendicular to the RV long axis is also illustrated. (b) The crescent-shaped intersection of the mesh with the cutting plane. (c) After uniform resampling (each dot represents a sampling point). Two cusp points separate the intersection into two parts and each is resampled independently.

The rotation-axis based method is appropriate for a roughly rotation symmetric shape, which contains a rotation axis. Cutting the mesh with a plane passing the axis, we get a 2D intersection. As shown in Fig. 6b, the rotation axis separates the intersection into two parts, which is uniformly resampled independently. We then rotate the plane around the axis to get another intersection and resample the intersection in the same way. Repeat the above process, we achieve a set of pseudo-landmarks. This approach is applicable as long as the intersection contains one and only one contour. Since the correspondence achieved by this approach is invariant under the similarity transformation (*e.g.*, translation, rotation, and scaling), it is an reasonable approximation for the contraction and expansion motion, which is the major motion component of a beating heart. We use this approach to resample LV, LA, and RA. This method is also used to resample LVOT, and RV inflow and outflow tracts, which can be approximated as tubes.

The rotation axis based method is not applicable to RV. First, RV is not a rotation symmetric shape. Second, the intersection of the mesh and a cutting plane passing through RV long axis has a complicate shape and may contain multiple contours since RV shape is not convex. Instead, a parallel-slice based method is developed to resample RV, where we use a plane perpendicular to the RV long axis to cut the 3D mesh, as shown in Fig. 7a. As long as the intersection contains one and only one contour, we can consistently resample the 2D contours. The shape of RV short-axis intersection is a crescent containing two cusp points, which split the contour into two parts. We uniformly resample two parts independently, as shown in Fig. 7c. Similarly, we use a set of parallel planes to cut the mesh. To simplify the operation, the distances between neighboring planes are uniform. By triangulating the resampled points, a new mesh can be generated.

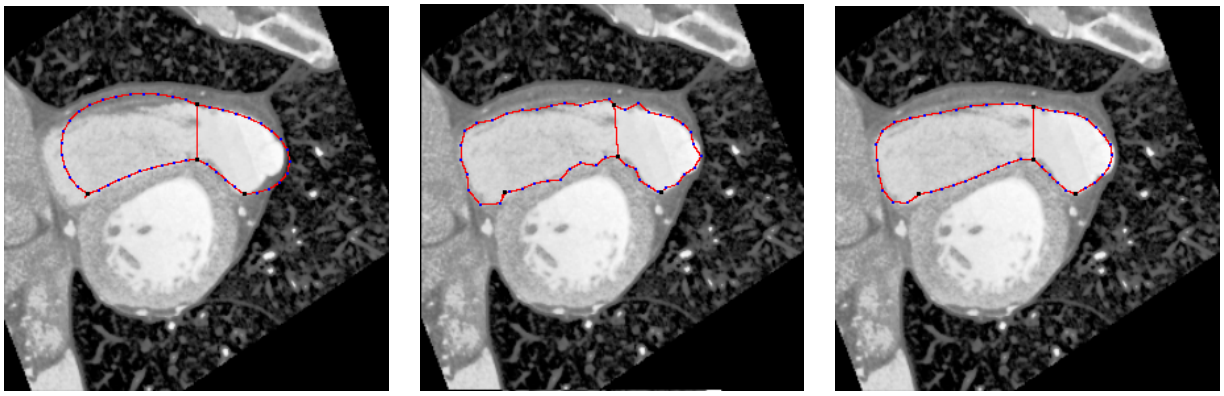
3. AUTOMATIC HEART CHAMBER DETECTION AND SEGMENTATION

Our automatic model fitting approach is based on recent advances in learning discriminative object models and we exploit a large database of annotated CT volumes. We formulate the segmentation as a two-step learning problem: anatomical structure localization and boundary delineation. In this section, we give an overview of our approach. More details are available in our previous publication.[23]

3.1 Chamber Pose Estimation

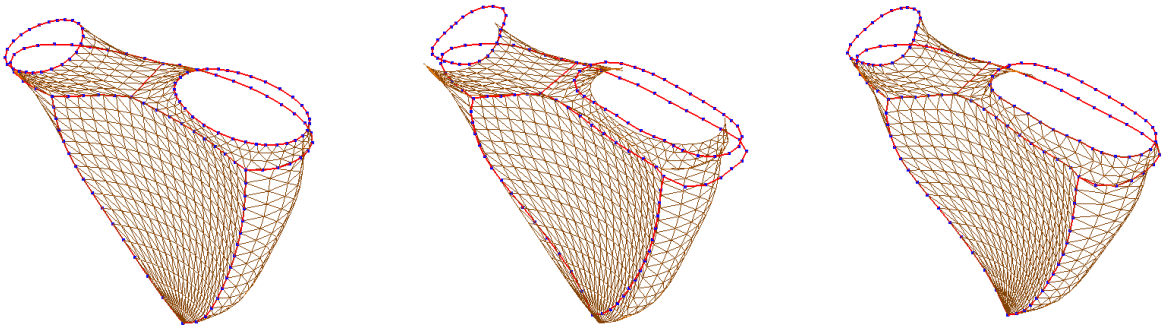
Object localization is required for an automatic segmentation system and discriminative learning approaches have proved to be efficient and robust for solving 2D problems. To address the challenges for extending learning based approach to 3D object detection (as discussed in Section 1.2), we exploit two simple but elegant techniques, Marginal Space Learning (MSL) and steerable features.[23] The idea for MSL is not to learn a classifier directly in the full similarity transformation parameter space but to incrementally learn classifiers on projected sample distributions. As the dimensionality increases, the valid (positive) space region becomes more restricted by previous marginal space classifiers. In our case, we split the estimation into three problems: translation estimation, translation-orientation estimation, and full similarity transformation estimation.

Besides reducing the searching space significantly, there is another advantage using MSL: we can use different features or learning methods in each step. For example, in the translation estimation step, since we treat



(a) (b) (c)

Figure 8. Example of non-rigid deformation estimation for control points, tricuspid lateral and RVOT lateral, on RV divergence plane. (a) Detected mean shape. (b) After boundary adjustment. (c) Final result by projecting the adjusted shape onto a shape subspace (25 dimensions).



(a) (b) (c)

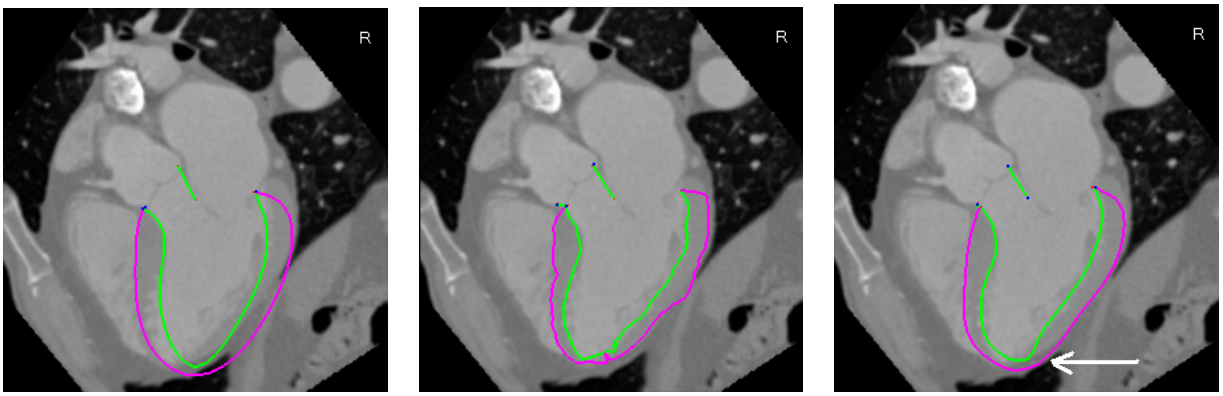
Figure 9. Using detected control points to warp the RV mesh. Red indicates control points and brown shows the RV mesh. (a) Aligned mean shape using the estimated RV pose. (b) After control point refinement. (c) Warped mesh.

rotation as an intra-class variation, we can use the efficient 3D Haar features.[24] In the translation-orientation and similarity transformation estimation steps, we introduce the steerable features. Steerable features constitute a very flexible framework where the idea is to sample a few points from the volume under a special pattern. We extract a few local features, such as voxel intensity and gradient, for each sampling point. To evaluate the steerable features under a specified orientation, we only need to steer the sampling pattern and no volume rotation is involved.

3.2 Non-Rigid Deformation Estimation

After the first stage, we get the position, orientation, and scale of the object. We align the mean shape with the estimated transformation to get a rough estimate of the object shape. Fig. 10a shows the aligned mean shape for LV. We then deform the shape to fit the object boundary. Active shape models (ASM) are widely used to deform an initial estimate of a non-rigid shape under the guidance of the image evidence and the shape priori. Non-learning based generic boundary detector in the original ASM[25] does not work in our application due to the complex background and weak edges. Learning based methods can exploit more image evidences to achieve a robust boundary detection.[26] Specifically, we use steerable features to train a boundary detector.[23]

Control points in our mesh representation have different image characteristics and should be treated specially. In the following, we use RV as an example to illustrate the detection of control points. MSL is used to detect the pose for each valve. After that, we get an aligned mean shape for the control points on a valve. The boundary detectors are then used to move each landmark point along the normal direction to the optimal position, where



(a) (b) (c)

Figure 10. Example of non-rigid deformation estimation for LV with green for the endocardial surface and magenta for the epicardial surface. (a) Detected mean shape. (b) After boundary adjustment. (c) Final delineation by projecting the adjusted shape onto a shape subspace (50 dimensions).

the estimated boundary probability is maximized. After deformation, we project the deformed shape into a shape subspace. Fig. 8 shows the detection of control points on the RV divergence plane. The aligned mean shape under the estimated pose is shown in Fig. 8a. After shape deformation, the control points fit the boundary well, but the contour is not smooth (Fig. 8b). As shown in Fig. 8c, the statistical shape model is very effective to enforce the priori shape constraints by projecting the deformed shape onto a shape sub-space.

The refined control point detection results can be used to warp a mesh to make it fitting the image better. Fig. 9a shows the mean shape aligned with the detected RV pose. Fig. 9b shows the refinement of the control points, which fit the data more accurately, but inconsistent with the mesh. Using the original control points and the refined control points, we can estimate the non-rigid deformation and use it to warp the mesh points. Here, 3D thin-plate-spline (TPS) model[27] is used to do the warping. As shown in Fig. 9c, the mesh points and the control points are consistent again after warping.

After TPS warping, the mesh points are closer to the chamber boundary, but still do not fit the image quite well in general. Again, we train a boundary detector for each mesh surface. The boundary detectors are then used to adjust each mesh point. Fig. 10a shows the aligned left ventricle (LV) for heart chamber segmentation in a cardiac CT volume. Fig. 10b shows the adjusted shape of LV. Shape constraint is enforced by projecting the adjusted shape onto a shape subspace to get the final result,[25] as shown in Fig. 10c. The arrow in the figure indicates the region with better boundary delineation.

4. EXPERIMENTS

In this section, we demonstrate the performance of the proposed method for multi-chamber localization and delineation in cardiac CT volumes. We collected and annotated 323 cardiac CT volumes from 137 patients with various cardiovascular disease. The number of patients in our experiment is significantly larger than those reported in the literature, for example, 10 by Schramm *et al.*,[22] 13 by Ecabert *et al.*,[3] and 18 by Jolly *et al.*[17] The imaging protocols are heterogeneous with different capture ranges and resolutions. A volume contains 80 to 350 slices and the size of each slice is 512×512 pixels. The resolution inside a slice is isotropic and varies from 0.28 mm to 0.74 mm, while slice thickness varies from 0.4 mm to 2.0 mm for different volumes. In the following experiments, each chamber is processed independently. The detected meshes from different chambers may cross each other, natural constraints are imposed as post-processing to solve the conflict. Four-fold cross validation is performed to evaluate our algorithm. Special care is taken to prevent volumes from the same patient appear in both the training and test sets.

The accuracy of boundary delineation is measured with the point-to-mesh distance, E_{p2m}. For each point on a mesh, we search for the closest point on the other mesh to calculate the minimum distance. We calculate the point-to-mesh distance from the detected mesh to the ground-truth and vice verse to make the measurement

Table 1. Mean and variance (in parentheses) of the point-to-mesh error (in millimeters) for the segmentation of heart chambers on 323 volumes based on cross validation.

	Initialization	Baseline ASM[25]	Proposed Approach
LV Endocardium	3.23 (1.17)	2.37 (1.03)	**1.29 (0.53)**
LV Epicardium	3.05 (1.04)	2.78 (0.98)	**1.33 (0.42)**
LA	2.78 (0.98)	1.89 (1.43)	**1.32 (0.42)**
RV	2.93 (0.75)	2.69 (1.10)	**1.55 (0.38)**
RA	3.09 (0.86)	2.81 (1.15)	**1.57 (0.48)**

symmetric. After we get the position, orientation, and scale of the object, we align the mean shape with the estimated transformation to initialize the non-rigid boundary delineation. The mean E_{p2m} error of the initialization ranges from 2.78 mm to 3.23 mm, as shown in Table 1. By deforming the mean shape to fit the boundary, we can reduce the error by a half. The mean E_{p2m} error ranges from 1.29 mm to 1.57 mm for different chambers. LV and LA have smaller errors than RV and RA since the contrast of the blood pool in the left side of a heart is consistently higher than the right side due to the using of contrast agents (as shown in Fig. 11).

We also compare our approach to the baseline ASM using non-learning based boundary detection scheme.[25] The same detected mean shape is used to initialize the deformation. As shown in Table 1, the baseline ASM only slightly reduces the error for weak boundaries (such as LV epicardial, RV, and RA surfaces). It performs much better for strong boundaries, such as LV endocardial and LA surfaces, but it is still significantly worse than the proposed method.

Fig. 11 shows several examples for heart chamber segmentation using the proposed approach. The second row shows a volume with low contrast, our segmentation result is quite good. Our approach is robust even under severe streak artifacts as shown in the third example.

Our approach is fast with an average speed of 7.9 seconds for automatic segmentation of all four chambers on a normal desktop computer (3.2 GHz dual-core CPU and 3 GB memory). It is faster than other reported results, *e.g.*, 3 seconds for LV only using a semi-automatic approach by Hong *et al.*,[16] 15 seconds for non-rigid deformation by Berg *et al.*,[4] 50 seconds for heart localization by Schramm *et al.*,[22] and 2-3 minutes for a 3D AAM based approach by Mitchell *et al.*[7]

5. CONCLUSION

In this paper, we proposed a novel four-chamber surface mesh model for a heart. In heart modeling, the following two factors are considered and traded-off: 1) accuracy in anatomy and 2) easiness for both annotation and automatic detection. To more accurately represent the anatomy, those important landmarks such as valves and cusp points on interventricular septum are explicitly represented in our model. These landmarks can be detected reliably to guide the automatic model fitting process. We enforce the mesh point correspondence during mesh editing. Therefore, it is straightforward to build a statistical shape model, which is used to guide the automatic segmentation process.

Using this model, we develop an efficient and robust approach for automatic heart chamber segmentation in 3D CT volumes. The efficiency of our approach comes from the two new techniques named marginal space learning and steerable features. Robustness is achieved by using recent advances in learning discriminative object models and exploiting large volumetric images databases. All major steps in our approach are learning-based, therefore minimizing the number of underlying model assumptions. According to our knowledge, this is the first study reporting stable results on a large cardiac CT dataset.

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. Medical Imaging* **20**(1), pp. 2–25, 2001.
2. A. F. Frangi, D. Rueckert, and J. S. Duncan, "Three-dimensional cardiovascular image analysis," *IEEE Trans. Medical Imaging* **21**(9), pp. 1005–1010, 2002.

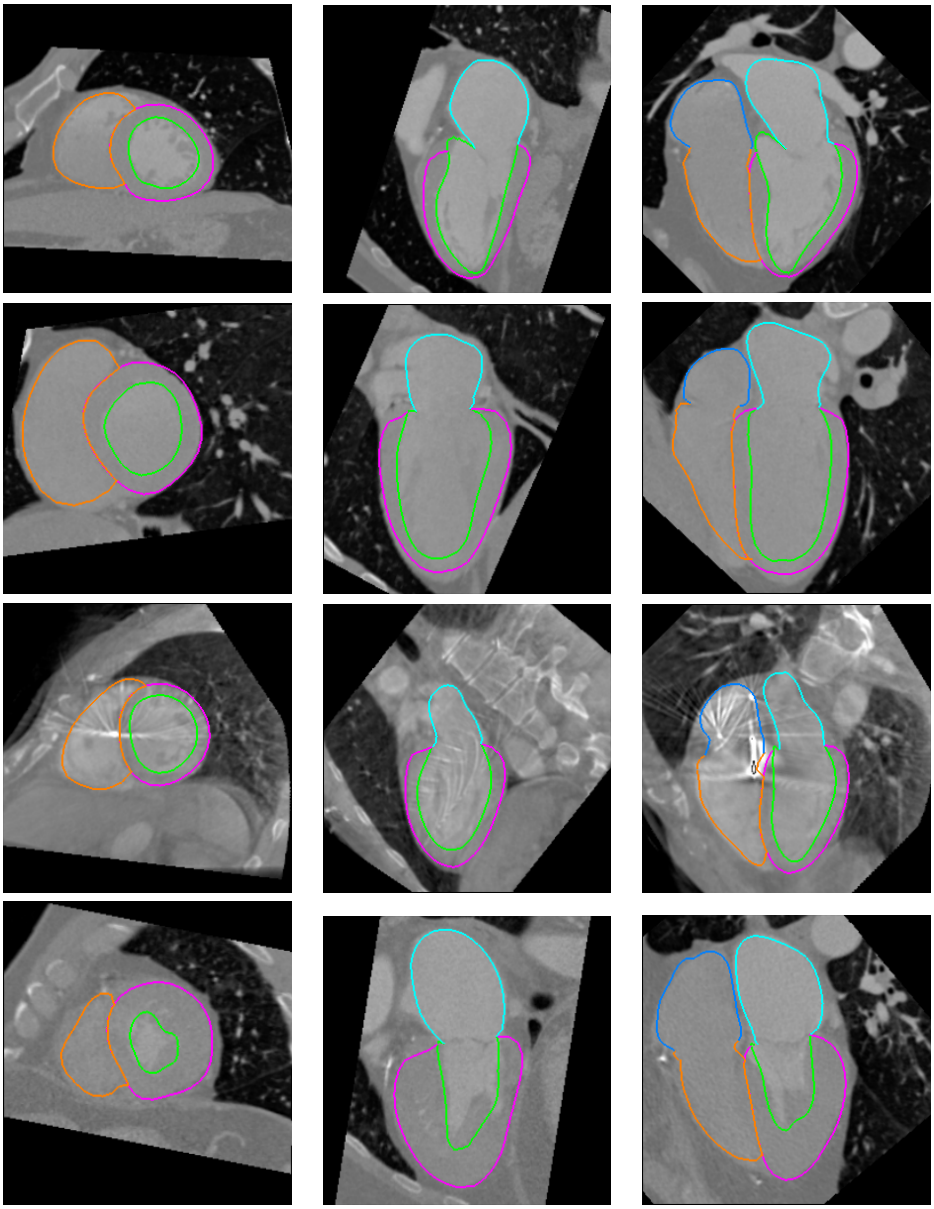


Figure 11. Examples of heart chamber segmentation in 3D CT volumes with green for LV endocardial surface, magenta for LV epicardial surface, cyan for LA, brown for RV, and blue for RA. Each row represents three orthogonal views of a volume.

3. O. Ecabert, J. Peters, and J. Weese, "Modeling shape variability for full heart segmentation in cardiac computed-tomography images," in *Proc. of SPIE Medical Imaging*, pp. 1199–1210, 2006.
4. J. von Berg and C. Lorenz, "Multi-surface cardiac modelling, segmentation, and tracking," in *Proc. of Functional Imaging and Modeling of the Heart*, pp. 1–11, 2005.
5. H. C. van Assen, M. G. Danilouchkine, A. F. Frangi, S. Ordas, J. J. M. Westernberg, J. H. C. Reiber, and B. P. F. Lelieveldt, "SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data," *Medical Image Analysis* **10**(2), pp. 286–303, 2006.
6. A. Andreopoulos and J. K. Tsotsos, "A novel algorithm for fitting 3-D active appearance models: Application to cardiac MRI segmentation," in *Proc. Scandinavian Conf. Image Analysis*, pp. 729–739, 2005.
7. S. C. Mitchell, J. G. Bosch, B. P. F. Lelieveldt, R. J. van Geest, J. H. C. Reiber, and M. Sonka, "3-D

active appearance models: Segmentation of cardiac MR and ultrasound images," *IEEE Trans. Medical Imaging* **21**(9), pp. 1167–1178, 2002.
8. Z. Bao, L. Zhukov, I. Guskov, J. Wood, and D. Breen, "Dynamic deformable models for 3D MRI heart segmentation," in *Proc. of SPIE Medical Imaging*, pp. 398–405, 2002.
9. C. Corsi, G. Saracino, A. Sarti, and C. Lamberti, "Left ventricular volume estimation for real-time three-dimensional echocardiography," *IEEE Trans. Medical Imaging* **21**(9), pp. 1202–1208, 2002.
10. O. Gerard, A. C. Billon, J.-M. Rouet, M. Jacob, M. Fradkin, and C. Allouche, "Efficient model-based quantification of left ventricular function in 3-D echocardiography," *IEEE Trans. Medical Imaging* **21**(9), pp. 1059–1068, 2002.
11. T. McInerney and D. Terzopoulos, "A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis," *Computerized Medical Imaging and Graphics* **19**(1), pp. 69–83, 1995.
12. K. Park, A. Montillo, D. Metaxas, and L. Axel, "Volumetric heart modeling and analysis," *Communications of the ACM* **48**(2), pp. 43–48, 2005.
13. A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans. Medical Imaging* **21**(9), pp. 1151–1166, 2002.
14. C. Lorenz and N. Krahnstover, "Generation of point based 3D statistical shape models for anatomical objects," *Computer Vision and Image Understanding* **77**(2), pp. 175–191, 2000.
15. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Trans. Medical Imaging* **21**(5), pp. 525–537, 2002.
16. W. Hong, B. Georgescu, X. S. Zhou, S. Krishnan, Y. Ma, and D. Comaniciu, "Database-guided simultaneous multi-slice 3D segmentation for volumetric data," in *Proc. European Conf. Computer Vision*, pp. 397–409, 2006.
17. M.-P. Jolly, "Automatic segmentation of the left ventricle in cardiac MR and CT images," *Int. J. Computer Vision* **70**(2), pp. 151–163, 2006.
18. G. I. Sanchez-Ortiz, G. J. T. Wright, N. Clarke, J. Declerck, A. P. Banning, and J. A. Noble, "Automated 3-D echocardiography analysis compared with manual delineations and SPECT MUGA," *IEEE Trans. Medical Imaging* **21**(9), pp. 1069–1076, 2002.
19. I. Wolf, M. Hastenteufel, R. D. Simone, M. Vetter, G. Glombitza, S. Mottl-Link, C. F. Vahl, and H.-P. Meinzer, "ROPES: A semiautomated segmentation method for accelerated analysis of three-dimensional echocardiographic data," *IEEE Trans. Medical Imaging* **21**(9), pp. 1091–1104, 2002.
20. M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert, "Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration," in *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention*, pp. 642–650, 2002.
21. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 511–518, 2001.
22. H. Schramm, O. Ecabert, J. Peters, V. Philomin, and J. Weese, "Towards fully automatic object detection and segmentation," in *Proc. of SPIE Medical Imaging*, pp. 11–20, 2006.
23. Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features," in *Proc. Int'l Conf. Computer Vision*, 2007.
24. Z. Tu, X. S. Zhou, A. Barbu, L. Bogoni, and D. Comaniciu, "Probabilistic 3D polyp detection in CT images: The role of sample alignment," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1544–1551, 2006.
25. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—their training and application," *Computer Vision and Image Understanding* **61**(1), pp. 38–59, 1995.
26. B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans. Medical Imaging* **21**(8), pp. 924–933, 2002.
27. F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformation," *IEEE Trans. Pattern Anal. Machine Intell.* **11**(6), pp. 567–585, 1989.