Automatic Extraction of 3D Dynamic Left Ventricle Model from 2D Rotational Angiocardiogram

Mingqing Chen[1], Yefeng Zheng[1], Kerstin Mueller[2,3], Christopher Rohkohl[2], Guenter Lauritsch[2], Jan Boese[2], Gareth Funka-Lea[1], Joachim Hornegger[3], and Dorin Comaniciu[1]

[1] Image Analytics and Informatics, Siemens Corporate Research, Princeton, NJ, USA
[2] Healthcare Sector, Siemens AG, Forchheim, Germany
[3] Pattern Recognition Lab, University Erlangen-Nuremberg, Germany
yefeng.zheng@siemens.com

Abstract. In this paper, we propose an automatic method to directly extract 3D dynamic left ventricle (LV) model from sparse 2D rotational angiocardiogram (each cardiac phase contains only five projections). The extracted dynamic model provides quantitative cardiac function for analysis. The overlay of the model onto 2D real-time fluoroscopic images provides valuable visual guidance during cardiac intervention. Though containing severe cardiac motion artifacts, an ungated CT reconstruction is used in our approach to extract a rough static LV model. The initialized LV model is projected onto each 2D projection image. The silhouette of the projected mesh is deformed to match the boundary of LV blood pool. The deformation vectors of the silhouette are back-projected to 3D space and used as anchor points for thin plate spline (TPS) interpolation of other mesh points. The proposed method is validated on 12 synthesized datasets. The extracted 3D LV meshes match the ground truth quite well with a mean point-to-mesh error of $0.51 \pm 0.11mm$. The preliminary experiments on two real datasets (included a patient and a pig) show promising results too.

1 Introduction

The real time angiocardiogram on a C-arm system is the workhorse imaging modality for many cardiac interventions. Nowadays it is becoming desirable to create a dynamic 3D left ventricle (LV) model automatically, which could be served to 1) quantitatively evaluate cardiac function, such as LV volume, LV ejection fraction (EF) and regional wall motion during the intervention, 2) provide visual guidance during intervention by overlaying this model onto 2D fluoroscopy, and 3) reduce motion artifacts for 3D image reconstruction [1]. A direct approach is to segment the LV from cine CT volumes, which are derived by electrocardiogram (ECG) gated reconstruction [2]. However, the X-ray source/detector needs to sweep 5-6 times to capture enough 2D projection data for each cardiac phase. The patient is required to hold breath during the whole

G. Fichtinger, A. Martel, and T. Peters (Eds.): MICCAI 2011, Part III, LNCS 6893, pp. 471–478, 2011.

procedure of 30 s (which is difficult to a very sick patient) to remove the respiratory motion. Furthermore, longer acquisition time incurs more radiation dose, which is a big concern nowadays.

Alternatively, one can generate a dynamic mesh directly from 2D projections. There are some previous studies using biplane and multi-view systems to recover thorax and lung [3,4] and the LV surface [5,6,7]. In Lotjonen et al.'s work, a prior model, which is a representative of average anatomy, is projected onto each 2D angle. The virtual projection is matched with input image using free form deformation [3,4]. Veistera et al. used triangulated mesh of the LV segmented from an MR image as prior model and projected the mesh onto each 2D projection. Manually annotated contour is used to deform the mesh. Moriyama et al. [6] proposed an iterative framework to recover LV mesh from multi views by fitting a 4D surface model based on B-splines. However, the geometry is limited to a star shape due to the spherical coordinates used.

In this paper, we propose a fully automatic method to fit a dynamic 3D LV model using projection images generated by a single rotation of C-arm. The proposed method has several advantages: 1) Since the prior LV model is segmented from an ungated volume reconstructed from the 2D projections, the 3D to 2D mapping is straightforward. 2) The detection of LV blood pool boundary is efficient. The searching is confined to the profiles of silhouette points of projected mesh, without filtering the entire image. 3) Thin plate spline (TPS) interpolation is used to compute the deformation of the whole mesh. The inclusion of smoothing factor in TPS allows correction of outliers in estimated deformation vectors.

2 Method

The main steps of our approach are illustrated in Fig. 1 with the intermediate results shown in Fig. 2. Firstly a 3D CT volume is reconstructed using all 2D projections. Though with significant cardiac motion artifacts, the image quality is good enough for automatic extraction of a rough 3D LV endocardium mesh model using the approach proposed in [8] (Fig. 2a). We then project this static mesh onto each 2D image (Fig. 2b). The silhouette is extracted from the projected mesh and adjusted along the normal direction to the optimal position

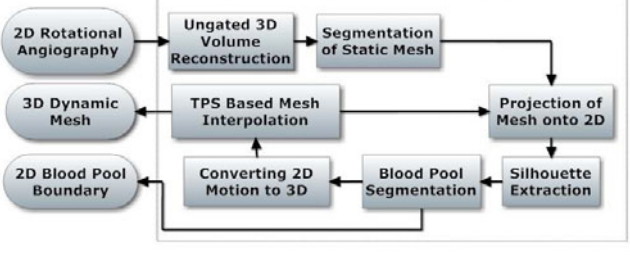

Fig. 1. Flowchart of the extraction of 3D dynamic left ventricle meshes

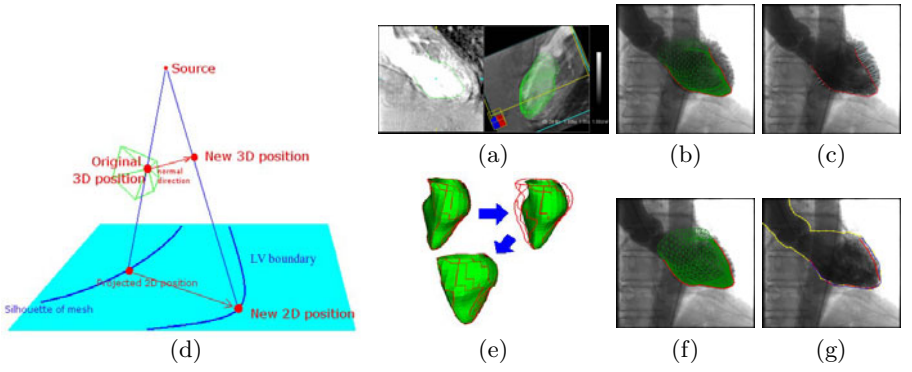

Fig. 2. Intermediate results of the proposed method. (a) Ungated 3D reconstruction volume overlaid with a static left ventricle (LV) mesh. (b) Projected static mesh (green) and its silhouette (red). (c) Deforming the silhouette to fit the LV blood pool boundary (blue). (d) Converting 2D deformation to 3D. (e) Warping the whole mesh using the silhouettes as anchor points. (f) Re-projection of deformed mesh. (g) Comparison of the mesh silhouette (red), segmented (blue) and annotated (yellow) LV blood pool.

using a learning based boundary detector (Fig. 2c). The 2D deformation vector is then back-projected to 3D (Fig. 2d). The patient image we are working on has 133 projections, which is acquired in 5 s. For a typical heart rate of 70 bpm, each cardiac phase has only five projections. Thus the mesh deformation is determined by points from five silhouettes, which is used as anchor points for TPS interpolation (Fig. 2e). In this way a sequence of 3D meshes are generated. The estimated dynamic mesh can further be re-projected to corresponding 2D angiographic images for further refinement again. The iterative procedure converges in a few iterations.

2.1 Mesh Silhouette Extraction

On a 2D projection image, the LV blood pool filled with contrast agent is the most salient image boundary. The LV mesh is deformed to make its silhouette match the blood pool boundary. An efficient approach is proposed to extract the silhouette of a projected mesh based on the the connectivity of mesh edges. It should be noted that the silhouette is not only composed with full length of original mesh edges. Two edges not coplanar in 3D may intersect in a 2D projection and only part of each is on the silhouette (Fig. 3). The general procedure is as follows. 1) Select candidate mesh edges for the silhouette; 2) Split the candidate edges at the intersection points; and 3) Edge following to extract silhouette. The efficiency of our approach comes from two facts: 1) The first step is accomplished by topological analysis of the projected mesh. For a triangulated mesh (e.g., our heart model), an edge cannot be on the silhouette if it is shared by two triangles on the opposite sides of this edge. Thus most edges are excluded for further investigation in this step. 2) The second step determines whether an intersection point exists or not for every remaining edge pairs. The computation of

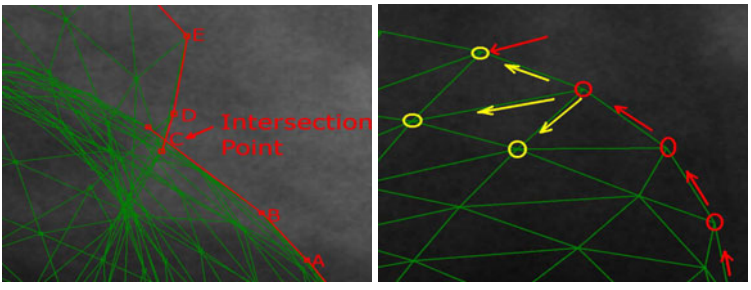

Fig. 3. Mesh silhouette extraction. **Left**: Splitting candidate silhouette edges (red) at the intersection (point C). **Right**: Edge following, which consistently follows the right-most edge.

intersection is time consuming. It is sped up by performing a quick condition check based on bounding box, which uses light operations to reject majority of the non-intersecting edge pairs. Thus the time consuming intersection computation is applied to only a few edge pairs.

Finally, an edge following procedure is taken to extract the silhouette from the split edges (Fig. 3). It is initialized with the right-most point of the projected mesh, which is guaranteed to be on the silhouette. Initialized with the upward search direction, it iteratively follows the right-most edge until coming back to the starting point. The whole path composes the extracted mesh silhouette.

2.2 Blood Pool Segmentation

During the training stage, the manually annotated LV blood pool boundary is given as input. For each boundary point, the steerable features [8] of pixels along the normal direction of this point are taken to extracted for a probabilistic boosting tree (PBT) classifier, which is a combination of the decision tree and AdaBoost classifier.

Let's assume the silhouette is composed of N points. During the segmentation stage, for each silhouette point n, a profile with predefined length K along the normal direction is extracted. A probability value is computed for each pixel on the profile based on PBT. Thus an $N \times K$ 2D array $A(n, k_n)$ can be computed, where n denotes the silhouette point index and k_n represents the pixel index along profile n. It is desirable to deform the silhouette contour to pixels with maximized summation of probability value, under some smoothness constraint that the difference of displacements between neighboring points is restricted to be smaller than λ. Mathematically the optimization problem has the form as:

$$max \sum_{n=1}^{N} A(n, k_n), \ under\ the\ constraint \begin{cases} |k_1 - k_N| < \lambda \\ |k_{n+1} - k_n| < \lambda, \end{cases} \tag{1}$$

It can be solved with an optimal graph search algorithm [9], where a 1D loop surface can be detected in a 2D graph.

2.3 Back-Projection of Deformation Vectors

To estimate the 3D motion, we assume that each mesh point is moving along its surface normal. It is a reasonable assumption since the LV motion is dominated by the contraction and dilation of the chamber. For each silhouette point $\mathbf{M}$, we project its 3D normal vector $\mathbf{n}$ onto 2D to compute the 2D normal vector $\mathbf{n_P}$

$$\mathbf{n_P} = \frac{(\mathbf{P}(\mathbf{M}^+ + \mathbf{n}^+) - \mathbf{P}\mathbf{M}^+)^-}{\|(\mathbf{P}(\mathbf{M}^+ + \mathbf{n}^+) - \mathbf{P}\mathbf{M}^+)^-\|}, \tag{2}$$

where $\mathbf{P}$ is a 4×3 projection matrix for 3D to 2D mapping. Superscript $^+$ denotes representation of original vectors using the homogeneous coordinate by adding "1" as another dimension of perspective plane, while superscript $^-$ represents using the original vector (For 2D, (x, y, s) is converted as $(x/s, y/s)$). The distance of 3D displacement (t) can be calculated by solving the following equation,

$$\mathbf{P}(\mathbf{M}^+ + (\mathbf{nt})^+) = \mathbf{M}'^+_{\mathbf{P}}, \tag{3}$$

where $\mathbf{M}'^+_{\mathbf{P}}$ is the 2D homogeneous coordinate of blood pool boundary, which is also the projection of the mesh point's new location $\mathbf{M}'$. By expanding Eqn. (3), we get

$$\frac{P_{11}(M_x + n_x t) + P_{12}(M_y + n_y t) + P_{13}(M_z + n_z t) + P_{14}M_w}{(P_{31}(M_x + n_x t) + P_{32}(M_y + n_y t) + P_{33}(M_z + n_z t) + P_{34}M_w} = M'_{Px}, \tag{4}$$

where P_{ij} the element at the i^{th} row and j^{th} column of projection matrix $\mathbf{P}$. (M_x, M_y, M_z, M_w) is the coordinate of $\mathbf{M}^+$, and M'_{Px} is the x coordinate of the 2D blood pool boundary point.

2.4 TPS Interpolation

The 3D positions of non-silhouette mesh points need to be interpolated. The thin plate spline (TPS) [10] is a popular coordinate interpolation approach by minimizing a physical energy function,

$$E = \sum_{i=1}^{k} \|f(x_i) - y_i\|^2 + \lambda \iint_{R^2} \left[(\frac{\partial^2 z}{\partial x^2})^2 + 2(\frac{\partial^2 z}{\partial x \partial y})^2 + (\frac{\partial^2 z}{\partial y^2})^2 \right] dxdy, \tag{5}$$

where the first term is the interpolation error of the anchor points and the second term is the bending energy of the transformation of a "thin plate." The parameter λ is tuned appropriately to control the balance between exact mapping of anchor points and rigidness of deformation.

3 Experiments

We first validate the proposed method on synthesized datasets using the 3D LV meshes extracted from 12 4D CT scans, on which we know both the 2D and 3D ground truth. The whole system is also tested on two real datasets, including a patient and a pig.

3.1 Validation on Synthesized Data

In this experiment, we validate the basic assumptions (e.g., the 3D motion direction along the surface normal and TPS interpolation) used in the proposed method without considering the blood pool segmentation error. The dynamic LV mesh sequence extracted from 12 4D CT scans are used as ground truth. The imaging geometry of one real patient data is used, which has 133 projections. Originally the 4D CT has 10 phases. The dynamic meshes are resampled to 26 phases to mimic the five cardiac cycles during image acquisition. The averaged mesh of the dynamic sequence is computed as the prior model. The proposed method is applied to estimate the dynamic sequence, which is compared to the real sequence to measure the accuracy.

LV volumes and point-to-mesh error are used to measure the difference between computed meshes and ground truth. Fig. 4 shows the convergence of LV volume error on 12 sequences through five iterations of the algorithm, where the error is reduced quickly in the first two iterations. The initial volume error is about 30 ml. We achieve a mean volume error of 3.27 ± 0.71 ml and relative volume error of $4.0 \pm 4.4\%$. As shown by the volume curve in Fig. 4, the estimated LV volume matches the ground truth quite well. The estimated ejection fraction (EF) of 43.7% is close to the true value of 45.8%. For the point-to-mesh error, the initial mean error is 2.19 ± 0.56 mm, and the final error is significantly reduced to 0.51 ± 0.11 mm, which is smaller than the mesh resolution (the mean mesh edge length is 4.29 mm).

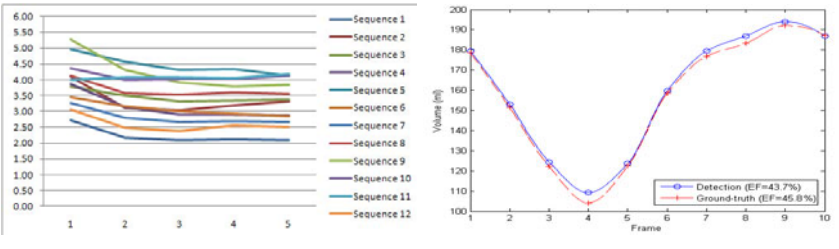

Fig. 4. Validation on synthesized data. **Left**: LV volume error (ml) of 12 mesh sequences plotted against the number of iterations; **Right**: The estimated LV volume curve and ejection fraction (EF) against the ground truth for one sequence.

3.2 Validation on Real Data

The proposed method is also validated on two real datasets, including a patient and a pig. For the patient dataset, about five cardiac cycles were captured in the image sequence with 133 projection images. The frame rate of the pig dataset was set much higher, resulting in 395 2D projection images, spanning 13 cardiac cycles. The ECG signal is recorded to assign a cardiac phase to each 2D projection. Since we do not have ground truth of the 3D mesh, we measure the distance of the mesh silhouette to the annotated blood pool boundary. The error of the automatically segmented blood pool (using the approach presented in Section 2.2) is also validated, with the mean error of 2.35 ± 1.32 mm on the patient

data. The error of the re-projected mesh silhouette is slightly higher, 2.84 ± 1.44 mm. The 2D angiographic images are very noisy and the LV could be occluded by other highly contrasted structures, e.g., the aorta and the diaphragm. To study the inter-observer variability, two clinicians were asked to annotate the blood pool independently. The mean error between the two manual annotations is $1.60 \pm 0.87mm$. Fig. 5 shows the extracted LV mesh and the projected silhouette on two frames. Our approach is also tested on a pig dataset, where the image contrast is much lower and there are strong confounding boundaries nearby such as the LV epicardium. The mesh silhouette error after four iterations is $5.60 \pm 2.05mm$.

Using the estimated 3D cardiac motion, we perform motion-compensated reconstruction [11]. As shown in Fig. 6 on the patient dataset, the trabeculations of the LV is clearly visible after motion compensated.

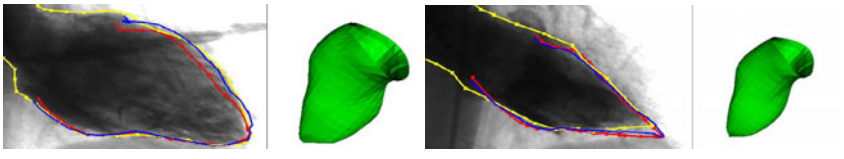

Fig. 5. Estimated 3D mesh (green mesh) and the re-projected mesh silhouette (red contour) on two angiographic images. For comparison, the automatically segmented blood pool boundary (blue) and manual annotation (yellow) are also shown.

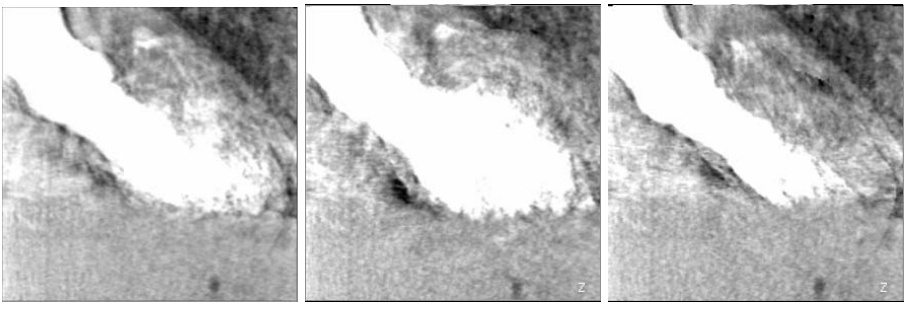

Fig. 6. Reconstructed CT volumes using the estimated 3D motion for the patient dataset. **Left**: Ungated reconstruction. **Middle** and **Right**: Motion-compensated reconstruction for the end-diastolic and end-systolic phases, respectively.

4 Conclusion

We proposed a novel method to automatically extract the 3D dynamic LV model from 2D rotational angiographic images captured in a single rotation of a C-arm system. The estimated dynamic model can be used to extracted several important measurements (e.g., volume and EF) of the LV function and it can also

be overlaid onto 2D fluoroscopic images to provide visual guidance during cardiac interventions. The experiment on synthesized data validated the appropriateness of the proposed method. The feasibility study on two real datasets (a patient and a pig) showed promising results and we plan to test on more real datasets in the future. Our approach can be extended to extract other organs with cyclic motion (e.g., other heart chambers and the aorta).

Acknowledgments. We would like to thank to Drs. Patrick W. Serruys, Carl Schultz, Peter de Jaegere, and Robert van Geuns from Thorax Center, Erasmus MC, Rotterdam, The Netherlands for acquiring data.

References

1. Prummer, M., Hornegger, J., Lauritsch, G., Wigstrom, L., Girard-Hughes, E., Fahrig, R.: Cardiac C-arm CT: A unified framework for motion estimation and dynamic CT. IEEE Trans. Medical Imaging 28(11), 1836–1849 (2009)
2. Lauritsch, G., Boese, J., Wigstrom, L., Kemeth, H., Fahrig, R.: Towards cardiac C-arm computed tomorgraphy. IEEE Trans. Medical Imaging 25, 922–934 (2006)
3. Lotjonen, J., Magnin, I., Reinhardt, L., Nenonen, J., Katila, T.: Automatic reconstruction of 3D geometry using projections and a geometric prior model. In: Taylor, C., Colchester, A. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 192–201. Springer, Heidelberg (1999)
4. Lotjonen, J., Magnin, I., Nenonen, J., Katila, T.: Reconstruction of 3-D geometry using 2-D profiles and a geometric prior model. IEEE Trans. Medical Imaging 18(10), 992–1002 (1999)
5. Veistera, H., Lotjonen, J.: Reconstructing 3D boundary element heart models from 2D biplane fluoroscopy. Proc. Functional Imaging and Modeling of the Heart, 17–23 (2001)
6. Moriyama, M., Sato, Y., Naito, H., Hanayama, M., Ueguchi, T., Harada, T., Yoshimoto, F., Tamura, S.: Reconstruction of time-varying 3D left ventricular shape from multiview x-ray cineangiocardiograms. IEEE Trans. Medical Imaging, 773–785 (2002)
7. Chhatriwalla, A., Kramer, C., Peixoto, A., Samady, H.: Measurement of left ventricular mass by contrast ventriculography. Texas Heart Inst. J. 31(7), 323–327 (2008)
8. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. IEEE Trans. Medical Imaging 27(11), 1668–1681 (2008)
9. Li, K., Wu, X., Chen, D.Z., Sonka, M.: Optimal surface segmentation in volumetric images—A graph-theoretic approach. IEEE Trans. Pattern Anal. Machine Intell. 28, 119–134 (2006)
10. Bookstein, F.: Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Machine Intell. 11(6), 567–585 (1989)
11. Schäfer, D., Borgert, J., Rasche, V., Grass, M.: Motion-compensated and gated cone beam filtered back-projection for 3-D rotational X-ray angiography. IEEE Trans. Medical Imaging 25(7), 898–906 (2006)