Automatic Segmentation of the Myocardium in Cine MR Images Using Deformable Registration

Marie-Pierre Jolly[1], Christoph Guetter[1], Xiaoguang Lu[1], Hui Xue[1], and Jens Guehring[2]

[1] Siemens Corporate Research, Image Analytics and Informatics, Princeton, NJ, USA
[2] Siemens AG, Healthcare Sector, Erlangen, Germany
marie-pierre.jolly@siemens.com

Abstract. This paper proposes a system to automatically segment the left ventricle in cardiac MR cine images. Individual frames are segmented using a shortest path algorithm and temporal consistency is enforced through the backward and forward deformation fields of an inverse consistent deformable registration. In addition, a segmentation of the mitral valve plane is obtained from long axis images. This algorithm was applied to 95 datasets as part of the STACOM'11 4D LV Segmentation Challenge. We analyze the results and evaluate the strengths and weaknesses of our system.

1 Introduction

Cardiovascular disease is the leading cause of death in the western world and there is a need to efficiently diagnose the health of the heart and the myocardium. Magnetic resonance imaging (MRI) is a possible means to observe the behavior of the heart. Physicians are most interested in the left ventricle (LV) because it pumps oxygenated blood to the rest of the body. MRI cine data consists of slices of the heart over time. In order to quantify measures such as ejection fraction, LV volume over time, myocardial mass, and myocardial thickening, they need a precise outline of the myocardium in all slices and all temporal phases. This can be extremely time consuming and physicians would like to rely on automatic software for this task. This paper proposes a system to automatically segment the LV endocardium and epicardium in all images of a cardiac MRI cine study.

The main difficulties in segmenting the myocardium are: a) the presence of papillary muscles and trabeculation in the blood pool (both in the LV and in the right ventricle) that contribute to partial voluming effect between the blood and muscle; b) there are often no clear edges between the myocardium and the liver; c) if there is fat around the heart, the fat/lung edges are stronger than the myocardium/fat edges; d) the myocardium becomes blurry in the apex slices; e) the cut between the LV and the left atrium in the base slice is very subtle (the muscle becomes thinner in the left atrium). Given all these difficulties, there has been a large number of publications on LV segmentation in cine MRI images [11],[12].

O. Camara et al. (Eds.): STACOM 2011, LNCS 7085, pp. 98–108, 2012.

Having studied all this work and having implemented some of the techniques, we believe that a pure 4D segmentation approach is currently not ideal since it is very difficult to build a model that is general enough to cover all possible shapes and dynamics of the LV. The segmentation usually results in surfaces that are too smooth and do not follow the true contours accurately enough. In addition, MR slices are very far apart (8-10mm) compared to the in-slice resolution (1-2 mm) such that the 3D segmentation becomes very anisotropic. Finally the slices might be mis-registered due to different breath-hold positions, which creates additional problems in fitting the model. Of course, the opposite approach of segmenting each image individually results in little cohesion between images, and contours that are not smooth over time.

We have chosen to segment all frames in one slice using deformable registration, taking advantage of the strong temporal correlation between frames. The main idea in our algorithm is to use an inverse consistent deformable registration to register all frames to the first frame in one slice. Then, the segmentation can be applied to any frame and propagated to any other frame in the sequence through the forward and backward deformation fields. We model the gray levels of the different regions in the heart area, including the partial voluming region, to help with difficulty a) and c). In addition, to try to overcome difficulty e), we use the long axis slices to recover a mitral valve base plane which is used to cut the short axis slices and compute a more accurate volume of the LV. The following sections describe individual steps of the algorithm in more detail.

2 Left Ventricle Segmentation

The segmentation of the left ventricle is divided into 4 steps: 1) detection of the left ventricle in the images; 2) mitral valve base plane segmentation; 3) slice segmentation through inverse consistent deformable registration and shortest path recovery; and 4) propagation between slices to segment the whole LV. Each of these steps will be described in the following sections.

2.1 LV Blood Pool Detection

The method for automatic detection of the LV blood pool is described in [6]. We threshold the first harmonic of the Fourier transform in each slice to detect the beating heart. Bright connected components are then extracted in each slice and characterized by their shape, motion, connectivity over time, etc. A graph is built by creating a node for each connected component. Pairs of nodes corresponding to connected components on neighboring slices are linked based on their similarity in shape, temporal behavior, distance, etc. The graph is then partitioned using isoperimetric clustering [3] and the corresponding connected components form 3D objects. The cluster that is roundest and "shrinks nicely" over time corresponds to the connected components inside the blood pool. Even though this process does not generate a blood pool region on all slices, it is a good starting point for the rest of the algorithm.

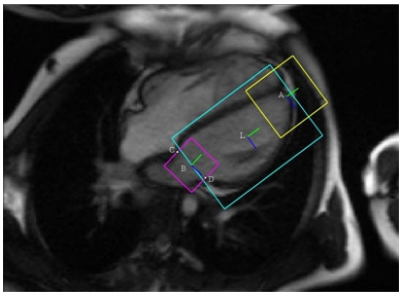

Fig. 1. Boxes to define the mitral valve area, the apex area, and the combined area

2.2 Mitral Valve Base Plane Segmentation

If long axis slices were acquired, they can be used to generate a plane approximation for the mitral valve in the following manner. We use the algorithm proposed in [9] to detect the mitral valve leaflet anchor points and the apex point in the end-diastolic (ED) and end-systolic (ES) frames (roughly estimated from the approximate blood pool). The mitral valve area, the apex area, and the combined area are each represented by 2D bounding boxes (as illustrated in Fig. 1 with 5 degrees of freedom (2 translations, 2 scale factors, and 1 rotation angle) to be estimated in the current image. We use a marginal space search strategy [14] where each position, orientation, and scale detector is a probabilistic boosting tree based on Haar wavelet-like and steerable features. The combined area is used to provide additional constraints to the search. Mitral valve points and apex points are then generated on frames other than ED and ES through linear interpolation and the mitral valve plane is fitted using a least squares approach by combining all mitral valve points on multiple long axis slices.

2.3 Slice Segmentation

The segmentation of the left ventricle is described in more details in [7]. The main idea behind the algorithm is the use of an inverse consistent deformable registration [4]. The registration computes a dense deformation field between any two frames in a slice without having to explicitly register every possible pair of frames. This is achieved by making the registration inverse consistent so that forward and backward deformation fields are recovered during the registration process, by alternately updating each deformation field at each step of the gradient descent minimization. So all frames are registered to an arbitrary keyframe (say the 1^{st} frame) using the consecutive strategy illustrated in Fig. 2. Then, the deformation field between frames i and j is obtained by compounding the deformation field between frames 1 and j and the inverse deformation field between frames 1 and i.

The core of the algorithm is illustrated in Fig. 3 for a given slice. Each frame is examined one at a time (this corresponds to the different rows in Fig. 3).

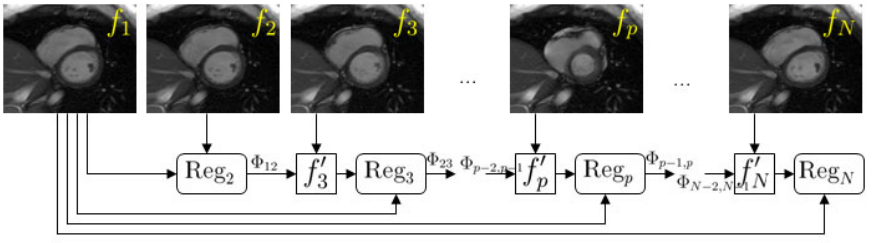

Fig. 2. Consecutive registration process

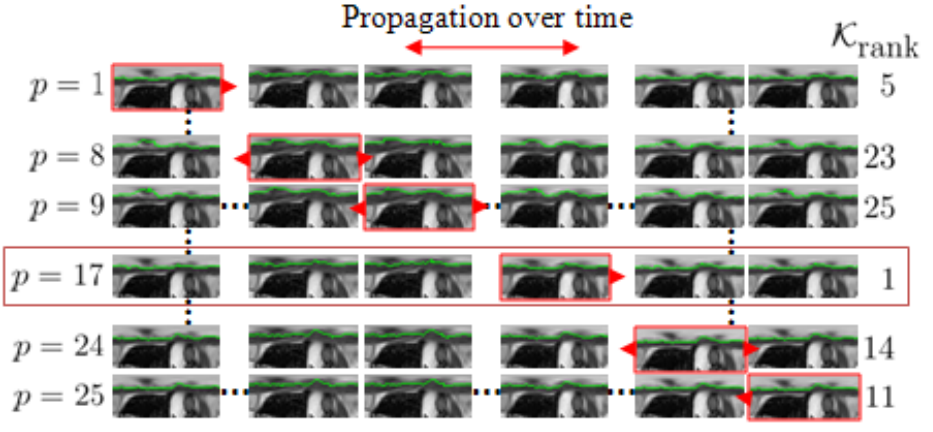

Fig. 3. Slice segmentation algorithm: for each frame $p = 1, ..., P$, recover a contour using Dijkstra's algorithm in polar space, propagate the contour to all other frames, and repeat. Choose the combination of contours (in this case recovered from frame 17) with the lowest cost.

For a frame p, the contour C_p is recovered in polar space using a minimum path algorithm as described later. The contours C_q in the other frames $q = 1, ..., P, q \neq p$ are generated using the deformation fields converted to polar space by $C_q(C_p) = \Phi_{1q}(\Phi_{1p}^{-1}(C_p))$ where Φ_{ij} is the deformation field between frames i and j (this corresponds to the different columns in Fig. 3). Then, the energy of this series of contours is given by $\mathcal{K}(p) = \sum_q E(C_q(C_p))$ where $E(C)$ is the edge cost associated with contour C. This same process is applied to all phases $p = 1, ..., P$ and the final segmentation is the one whose energy is lowest: $\mathcal{K} = \min_p \mathcal{K}(p)$. Once the best sequence has been recovered, the best polar contour in the best frame is converted to Cartesian space and propagated to the other frames using the forward and backward deformation fields in Cartesian space. In addition, the convex hull of the endocardium contour is generated to further enforce that it goes behind the papillary muscles.

The recovery of a contour in one frame is based on gray level properties of the image as described in [7]. The histogram of the LV region is analyzed to

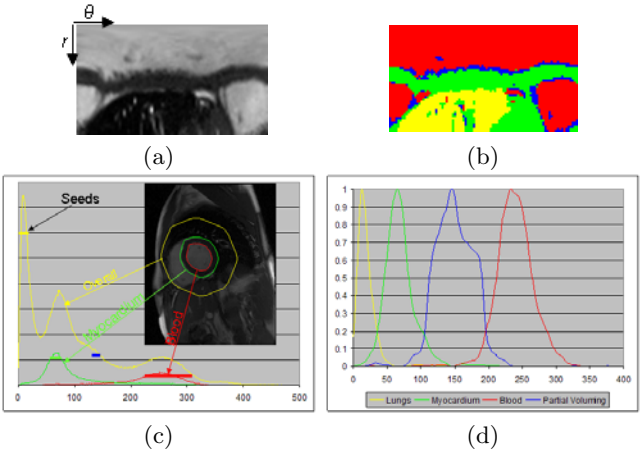

Fig. 4. Gray level analysis: (a) Polar image; (b) Multiseeded fuzzy connectedness region labeling; (c) Original histograms; (d) Final histograms (view in color)

find seeds for a multi-seeded fuzzy connectedness algorithm [5] and generate gray level properties of the different regions (blood, muscle, air, and partial voluming) as illustrated in Fig. 4. Then, the edge costs for the endocardium and the epicardium are computed using the Deriche filter edge detector on the original images and on the region probability images. The contour with smallest edge cost is recovered in polar space using Dijkstra's algorithm.

2.4 Segmentation of the Entire LV

The first slice to be segmented is the slice where the detected blood pool is the roundest. These contours are then propagated to the previous slice by applying the deformable registration and the deformed contours are used as priors to segment the previous slice. This process is repeated all the way to the base slice, where the base slice is identified as the slice closest to the mitral valve anchor points detected in the long axis slices. In the base slice, there is an additional process to identify when the myocardium cuts through the aortic valve and replace those segments of the endocardium and epicardium contours by straight lines (see Fig. 5(e)).

The initial slice contours are also propagated to the next slice toward the apex in the same way. An apex point is extracted in the long axis slices using the algorithm in [9] in conjunction with the extraction of the mitral valve leaflet anchor points. When reaching slices close to this apex point, tests are performed to determine if the apex has been reached and the downward propagation should stop. These tests include large shrinking of the contours compared to the previous slice, inconsistency in the modeling of the regions gray level properties (in which case the last segmentation is removed before the propagation is stopped), etc.

The last step of the workflow consists of intersecting the mitral valve base plane with the short axis slices and removing the parts of the contours that are above the plane (see Fig. 5(e)).

3 Experiments

We have tested our algorithm on the 100 validation datasets from the STACOM'11 4D LV Segmentation Challenge run by the Cardiac Atlas Project (CAP) [2]. All the data came from the DETERMINE [8] cohort which consists of patients with coronary artery disease and prior myocardium infarction. The data were acquired using steady-state free precession (SSFP) MR imaging protocols with thickness ≤ 10 mm, gap ≤ 2 mm, TR 30-50 ms, TE 1.6 ms, flip angle 60°, FOV 360 mm, and 256×256 image matrix. The data were acquired at multiple sites using scanners from different vendors.

The ground truth images provided in this challenge were defined by an expert using an interactive guide point modeling segmentation algorithm [13]. A finite element model of a heart was fitted with this algorithm to the cardiac MR images and the expert refined the fitted model following the wall motion throughout the cardiac cycle by using guide points, interactively. To get binary images of the myocardium, software from the CAP group was applied to calculate the intersection between the fitted heart model and the original MRI slices.

We were able to process 95 of the 100 datasets. Five of the datasets (DET0006901, DET0044401, DET0008401, DET0012301, and DET0014301) violated our assumptions, namely that all short axis slices should have the same number of phases, and all images within one slice and all images in the short axis matrix should have the same geometry (number of rows, number of columns, and pixel size). We generated contours for all 95 datasets and obtained the binary images of the myocardium by automatically painting all pixels inside the epicardium and outside the endocardium. Since we did not segment the long axis images (we only used them to compute the mitral valve plane), we reported the binary images for the short axis images only. In 4 of the 95 cases, the automatic detection of the LV did not work (DET0007901, DET0011501, DET0012001, and DET0013801). For these cases, the system allows the user to move the center of the segmentation to the blood pool and rerun the segmentation.

The validation performed by the CAP team produced the following measures from the binary images of the ground truth and segmented myocardium: sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), dice, and jaccard (see `http://en.wikipedia.org/wiki/Sensitivity_and_specificity` for a very good description of these statistical measures). In addition, we calculated ED volume, ES volume, and mass, and the CAP team provided us with the difference between our values and the ground truth. Fig. 5 shows some segmentation examples. The speed of the system was measured on a dual core laptop (2.93GHz with 4GB of RAM) for an average size dataset with 18 slices (14 short axis and 4 long axis) and 20 phases. The segmentation takes less than 3 minutes: 2 minutes to compute the deformation fields for the entire dataset, 5 seconds

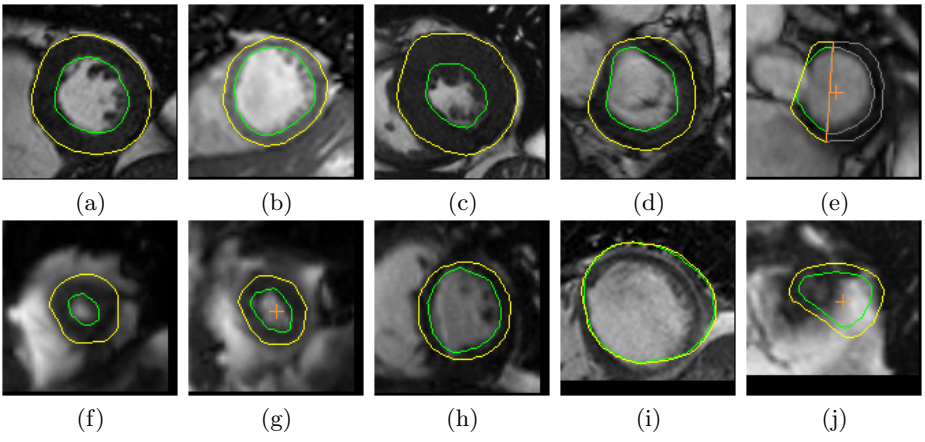

Fig. 5. Examples of segmentation results: (a) mid ventricular slice; (b) mid ventricular slice at ED; (c) mid ventricular slice at ES; (d) slice close to the valve; (e) base slice with mitral valve base plane cut and aortic valve straight line; (f) apex slice; (g) apex slice; (h) inaccurate segmentation; (i) poor segmentation; (j) poor apex segmentation

to detect the heart and the blood pool, 3 seconds to segment the base plane, 1 second for the segmentation of each slice, and 1 second to register neighboring slices and define the contour priors for the next slice. It can be seen that the registration is the bottleneck but since our implementation is parallelized, the timings on a system with more than 2 processors are actually faster.

We first present the overall results, focusing on the 95 datasets that were processed. There were a total of 28213 short axis images, 25481 of which contained ground truth. For 934 images with ground truth, our algorithm did not produce any segmentation result. This happens in the base and apex slices when the algorithm did not propagate up or down enough. Reciprocally, our algorithm segmented 861 images which should not have been segmented, because the algorithm propagated too far up or down. Considering that wrong images being segmented would require the user to delete contours or manually add contours, our system achieves very good results, with a sensitivity of 0.96, specificity of 0.68, accuracy of 0.94, PPV of 0.97, and NPV of 0.67 for predicting a segmentation on a particular image.

We now focus on the results for the images where both ground truth and segmented myocardium were available. The statistics over all segmentations are given in Table 1. Given that the images are so much larger than the myocardium, the size of the true negative (TN) set is so large, and the specificity, accuracy, and NPV are all very close to 1 and not very interesting to study. We will therefore concentrate on the other measures which are much more meaningful. The median values for the sensitivity (0.71), PPV (0.83), dice (0.77), and jaccard (0.62) indicate that the algorithm performs well. It is important to note that these numbers are sensitive to the pixels on the boundary of the region. We are not comparing the endocardium and epicardium regions. Instead, we are comparing

Table 1. Overall results for all 95 datasets

measure	avg	std	min	1st	med	3rd	max
sensitivity	0.621	0.272	0.000	0.443	0.714	0.837	1.000
specificity	0.998	0.002	0.980	0.997	0.998	0.999	1.000
accuracy	0.993	0.005	0.965	0.991	0.994	0.996	1.000
PPV	0.748	0.226	0.000	0.696	0.826	0.895	1.000
NPV	0.995	0.004	0.976	0.994	0.996	0.998	1.000
dice	0.656	0.253	0.000	0.545	0.766	0.840	0.953
jaccard	0.533	0.242	0.000	0.375	0.620	0.724	0.911

the myocardium region which is relatively thin. Therefore, a few pixels included or excluded at the boundary of the segmentation on the endocardium side or on the epicardium side do not change the overall appearance of the segmentation much, but change these statistical measures.

Fig. 6 shows box plots of the dice coefficient distribution for each dataset. The box plots for sensitivity, PPV, and jaccard look very similar to this one. It can be seen that, aside from 5 datasets where the algorithm performed very poorly (DET0007901, DET0023801, DET0011501, DET0011601, and DET0013801), and 2 datasets where the results were not very good (DET0030301 and DET0014901), in general, the segmentation is in good agreement with the ground truth. In all cases of bad segmentation, the algorithm had trouble with the modeling of the gray levels for the different regions. In some cases, it resulted in a very thin myocardium as in Fig. 5(i), and in other cases, parts of the myocardium were missed as in Fig. 5(h). Dataset DET0007901 is particularly difficult because alternating slices have very different brightness levels. Since our algorithm uses the gray level distributions of the previous slice to initialize the modeling in the current slice, the algorithm breaks.

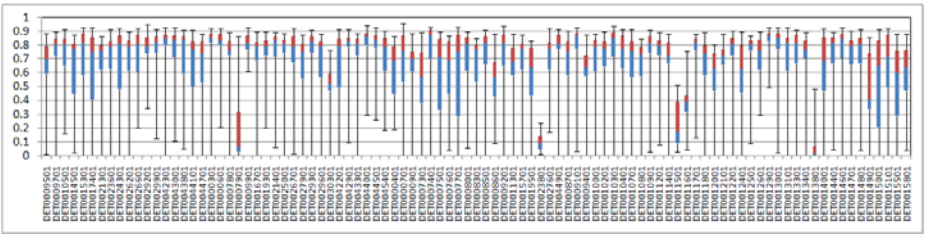

Fig. 6. Box plots of the dice coefficient distribution for each dataset

Fig. 7 shows box plots of the dice coefficient distribution for different slice positions in the dataset. Again, box plots for sensitivity, PPV, and jaccard look very similar. For each image, the relative slice position (between 0 and 1) was calculated with respect to the slices containing ground truth. It can clearly be seen that the segmentation results get worse toward the base and toward the

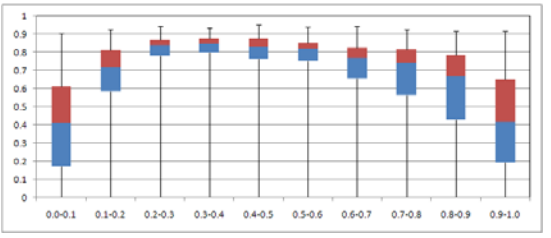

Fig. 7. Box plots of the dice coefficient distribution over various relative slice positions

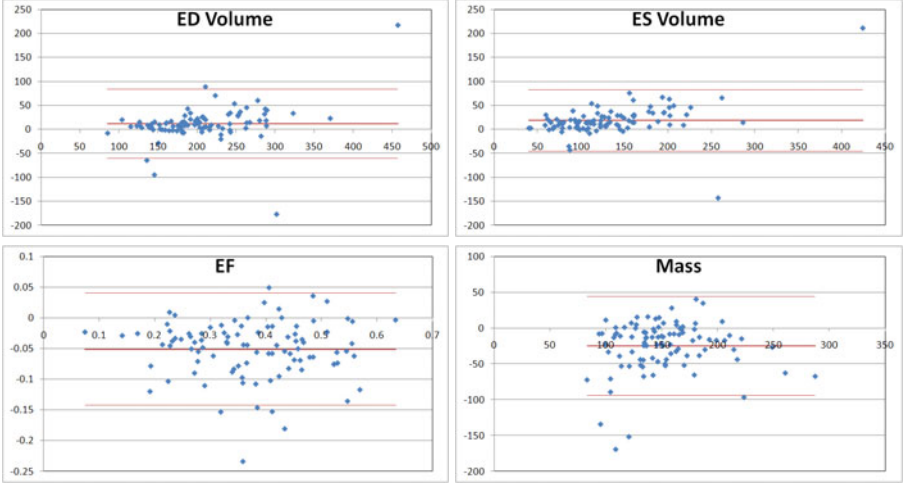

Fig. 8. Bland-Altman plots for the ED volume, ES volume, ejection fraction, and mass

apex. Going toward the apex, the algorithm sometimes has difficulties shrinking the shape of the left ventricle enough as seen in Fig. 5(j). In the base, the discrepancies are due to two factors, namely the thickness of the myocardium, and the exact positions of the mitral valve and aortic valve. Indeed, since the contours are cut by the mitral valve plane and parts are erased (see Fig. 5(e)), if the plane is estimated wrong, the cuts might be completely off and the overlap between the ground truth and segmented myocardium might be very small.

Finally, we compared the clinical measurements. Fig. 8 shows Bland-Altman plots for the ED volume, ES volume, ejection fraction, and mass computed from the ground truth and the segmentation results. Our volumes are calculated by multiplying the area of the contours by half the distance to the previous slice plus half the distance to the next slice. When the contour intersects with the base plane, each individual pixel area multiplied by its distance to the base plane is added to the entire volume. It can be seen that the volumes are slightly over-estimated, by 11.67ml for ED and 18.48ml for ES. The two outliers correspond to datasets DET0011601 and DET0023801 which we have already mentioned

for having poor segmentation results. Even though the volumes are over estimated, the ejection fraction turns out to be very accurate, underestimated by only 0.05%. The mass is underestimated by 25g. We believe that these numbers are clinically acceptable. The bias in volume and mass are likely due to differences in methods used to calculate the measurements from the segmented slices.

4 Conclusions

We have presented a system to automatically segment the left ventricle in cardiac cine MR images. The system segments short axis images, and uses long axis images, when available, to extract the mitral valve leaflet anchor points and generate a mitral valve base plane. The segmentation of the short axis images is performed one slice at a time using an inverse consistent deformable registration algorithm to enforce an implicit smoothness constraint over time. The contours themselves are recovered using a minimum cost path algorithm.

We have tested our system on 95 datasets from the STACOM'11 4D LV Segmentation Challenge validation set and have analyzed the results. This exercise was very useful to understand the strengths and weaknesses of our system. We were aware that the base and apex areas are more challenging and have some ideas on how to improve this by combining the slice based segmentation with a model-based approach [10]. Also, we know that the weakest link in the slice segmentation is the gray level modeling. We might be able to model the different regions distributions better using linear combinations of discrete Gaussians [1].

References

1. El-Baz, A., Gimel'farb, G.: EM based approximation of empirical distributions with linear combinations of discrete Gaussians. In: ICIP, pp. 373–376 (2007)
2. Fonseca, C.G., Backhaus, M., Bluemke, D.A., Britten, R.D., Chung, J.D., Cowan, B.R., Dinov, I.D., Finn, J.P., Hunter, P.J., Kadish, A.H., Lee, D.C., Lima, J.A.C., Medrano-Gracia, P., Shivkumar, K., Suinesiaputra, A., Tao, W., Young, A.A.: The Cardiac Atlas Project – an imaging database for computational modeling and statistical atlases of the heart. Bioinformatics (2011)
3. Grady, L., Schwartz, E.L.: Isoperimetric partitioning: A new algorithm for graph partitioning. SIAM Journal on Scientific Computing 27(6), 1844–1866 (2006)
4. Guetter, C., Xue, H., Chefd'hotel, C., Guehring, J.: Efficient symmetric and inverse-consistent deformable registration through interleaved optimization. In: ISBI, pp. 590–593 (2011)
5. Herman, G.T., Carvalho, B.M.: Multiseeded segmentation using fuzzy connectedness. IEEE Trans. Pattern Analysis and Machine Intelligence 23(5), 460–474 (2001)
6. Jolly, M.-P.: Automatic Recovery of the Left Ventricular Blood Pool in Cardiac Cine MR Images. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 110–118. Springer, Heidelberg (2008)
7. Jolly, M.P., Guetter, C., Guehring, J.: Cardiac segmentation in MR cine data using inverse consistent deformable registration. In: ISBI (2010)

8. Kadish, A.H., Bello, D., Finn, J.P., Bonow, R.O., Schaechter, A., Subacius, H., Albert, C., Daubert, J.P., Fonseca, C.G., Goldberger, J.J.: Rationale and design for the Defribrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation (DETERMINE) trial. J. Cardiovascular Electrophysiology 20(9), 982–987 (2009)
9. Lu, X., Georgescu, B., Jolly, M.-P., Guehring, J., Young, A., Cowan, B., Littmann, A., Comaniciu, D.: Cardiac Anchoring in MRI Through Context Modeling. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 383–390. Springer, Heidelberg (2010)
10. Lu, X., Wang, Y., Georgescu, B., Littman, A., Comaniciu, D.: Automatic Delineation of Left and Right Ventricles in Cardiac MRI Sequences using a Joint Ventricular Model. In: Metaxas, D.N., Axel, L. (eds.) FIMH 2011. LNCS, vol. 6666, pp. 250–258. Springer, Heidelberg (2011)
11. Petitjean, C., Dacher, J.N.: A review of segmentation methods in short axis cardiac MR images. Medical Image Analysis 15, 169–184 (2011)
12. Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A.J., Wright, G.A.: Evaluation framework for algorithms segmenting short axis cardiac MRI. MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge (2009)
13. Young, A.A., Cowan, B.R., Thrupp, S.F., Hedley, W.J., Dell'Italia, L.J.: Left ventricular mass and volume: Fast calculation with guide-point modeling on MR images. Radiology 202(2), 597–602 (2000)
14. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. In: ICCV (2007)