Multi-Atlas Based Segmentation with Local Label Fusion for Right Ventricle MR Images

Wenjia Bai[1], Wenzhe Shi[1], Haiyan Wang[1], Nicholas S. Peters[2], and Daniel Rueckert[1]

[1] Biomedical Image Analysis Group, Department of Computing, Imperial College London, UK
[2] National Heart and Lung Institute, St Mary's Hospital, Imperial College London, UK

Abstract. Evaluation of right ventricular (RV) function is essential for the diagnosis of cardiovascular diseases. However, to date, it heavily relies on manual segmentation which is time-consuming and dependent on the observer's experience. This paper presents a multi-atlas based segmentation method which labels the RV myocardium and blood pool by ensembling opinions from multiple atlases. It only requires an initial input in form of a few landmarks. Experimental results on a common data set show that the method yields good segmentation results which closely match manual segmentation. The method can provide accurate measurement for clinically important indices used for RV function evaluation.

Keywords: right ventricle segmentation, multi-atlas based segmentation, local label fusion

1 Introduction

Evaluation of right ventricular (RV) function is important for the diagnosis of cardiovascular diseases such as arrhythmogenic right ventricular displasia (ARVD). The evaluation normally involves measurement of the ventricular mass (VM), the end-diastolic blood pool volume (EDV) and end-systolic volume (ESV) using cardiac MR. The right ventricular ejection fraction (EF) can then derived from EDV and ESV. However, RV evaluation is still mainly performed manually in practice, which is time-consuming and dependent on the observer's experience. Unlike left ventricle (LV) segmentation, computer-aided RV segmentation is difficult due to the large variance of the RV shape and the very thin wall thickness. The RV shape is highly variable across subjects, whereas the LV is mostly of a regular circular shape. In addition, the RV wall is only 2-6 mm thick at end-diastole [1], which is much thinner than the LV wall.

Recently, multi-atlas based segmentation has been demonstrated to significantly improve segmentation accuracy compared to single atlas based segmentation [2]. It registers the target image to each of the atlases and combines the propagated labels from multiple atlases to form a consensus segmentation. It accounts for the anatomical shape variability by using a number of atlases which

represent different anatomical shapes and structures. It is robust because segmentation errors associated with single atlas propagation can be averaged out when combining multiple atlases. The consensus segmentation is less likely to be affected, when an individual atlas does not match the target image quite well or when registration errors occur for an individual atlas. Its performance can be further improved by selecting a subset of atlases which look more similar to the target image than the others [3–5], or by using weighted label fusion in which the contribution of each atlas is locally weighted by its similarity to the target image [6–9].

In this paper, we propose a multi-atlas based method for RV segmentation. Atlas selection and local weighted label fusion is used to reduce computational cost and to improve segmentation performance. To account for image-to-atlas registration errors, a number of voxels in a local search region are considered when evaluating the similarity and fusing the atlas labels. The method is evaluated on a common data set provided by the MICCAI 2012 RV Segmentation Challenge.

2 Methods

2.1 Framework

Consider an image $I = \{I(x)|x \in \Omega\}$, where x denotes the voxel and $\Omega \subset \mathbb{R}^3$ denotes the lattice on which the image is defined. The goal of segmentation is to estimate a label map L associated with the image I, in which each voxel is assigned a discrete label representing its tissue class. In a multi-atlas based segmentation method, we have a number of atlases $\{I_n|n = 1, \cdots, N\}$ with corresponding label maps $\{L_n|n = 1, \cdots, N\}$ already known. We register the target image I with each of the atlases and infer a label map from this atlas. Combining the opinions from all the atlases, a fused label map is generated as the segmentation.

2.2 Image Registration

The spatial transformation Φ_n from the target image I to an atlas image I_n is estimated by image registration. Image registration amounts to an optimisation problem, where the similarity metric between the target image I and the transformed atlas image $I_n(\Phi_n)$ is maximised.

$$\Phi_n = \arg\max_{\Phi_n} f(I, I_n(\Phi_n))$$

where f denotes the similarity metric. We use normalised mutual information (NMI) as the similarity metric, which is widely used in the image registration domain because it does not assume any explicit functional relationship between two images and it is efficient to calculate [10]. The transformation Φ_n is initialised by landmark registration. Five landmarks are used, respectively the two RV

insertion points on the basal slice, a point on the anterolateral wall, a point on the inferolateral wall and the RV apex. Then an affine transformation and a non-rigid B-spline transformation are subsequently computed by optimising the image similarity metric. Image registration was implemented in a multi-resolution approach using the IRTK software package[1] [11].

After the non-rigid transformation Φ_n is estimated, the atlas label map L_n is propagated to the target space using shape-based interpolation [12, 13].

$$L'_n(x) = L_n(\Phi_n(x)), x \in \Omega$$

where L'_n denotes the infered label map from the nth atlas. We perform image registration between the target image and each of the atlas images so that finally we have N infered label maps for the target image.

2.3 Local Label Fusion

Each propagated atlas label map $L'_n(x)$ has an opinion for the label estimate at voxel x. The opinions from all the atlases are weighted and combined [6–9]. The weight is given by the similarity between the target voxel intensity $I(x)$ and the atlas voxel intensity $I'_n(x) = I_n(\Phi_n(x))$.

$$P(L(x) = l|L'_n(x), I(x), I'_n(x)) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{[I(x)-I'_n(x)]^2}{2\sigma_i^2}} \cdot \delta_{l,L'_n(x)}$$

where σ_i denotes the parameter for the Gaussian distribution and δ_{l_1,l_2} denotes the Kronecker delta function. The more similar $I(x)$ and $I'_n(x)$ is, the higher weight label $L'_n(x)$ will have.

Since the estimated spatial transformation may not be perfect and there can be slight mis-alignment between the target image and the propagated atlas images after registration, the atlas voxels adjacent to x are also allowed to have an opinion on the label estimate at x. Their opinions are weighted not only by intensity similarity but also by the distance to the target voxel x.

$$\begin{aligned} &P(L(x) = l|L'_n(x + \Delta x), I(x), I'_n(x + \Delta x)) \\ &= \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{[I(x)-I'_n(x+\Delta x)]^2}{2\sigma_i^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_d} e^{-\frac{D(\Delta x)^2}{2\sigma_d^2}} \cdot \delta_{l,L'_n(x+\Delta x)} \end{aligned} \quad (1)$$

where Δx denotes the coordinate offset, $D(\Delta x)$ denotes the physical distance of the offset and σ_d denotes the parameter for distance weighting. The offset Δx is drawn from a local search region S centred at x, for example, 3×3×3 voxels around x.

We use Eq. (1) as the weighting function for label fusion in our method, which considers the intensity similarity between the target image and the atlas

[1] `http://www.doc.ic.ac.uk/~dr/software/`

and also accounts for potential registration errors. Combining the opinions from all the atlases, the label estimate $\tilde{L}$ is given by the label with the highest vote.

$$\tilde{L}(x) = \arg\max_{l} \sum_{n=1}^{N} \sum_{\Delta x \in S} P(L(x) = l | L'_n(x + \Delta x), I(x), I'_n(x + \Delta x)) \quad (2)$$

2.4 Post-Processing

Local label fusion does not guarantee the smoothness of the label map. For example, the blood pool labels may protrude out of the myocardial wall or there may be some myocardial wall labels inside the blood pool (see Figure 1), which is physiologically impossible. To generate a plausible label map, we fit the estimated label map $\tilde{L}$ to a model. In this paper, the atlas that is most similar to the target image in terms of NMI is selected as the model for simplicity[2]. The target label map $\tilde{L}$ is registered to the model label map using label consistency as the similarity metric [12]. The model label map is then warped to the target space as the final segmentation.

3 Experiments and Results

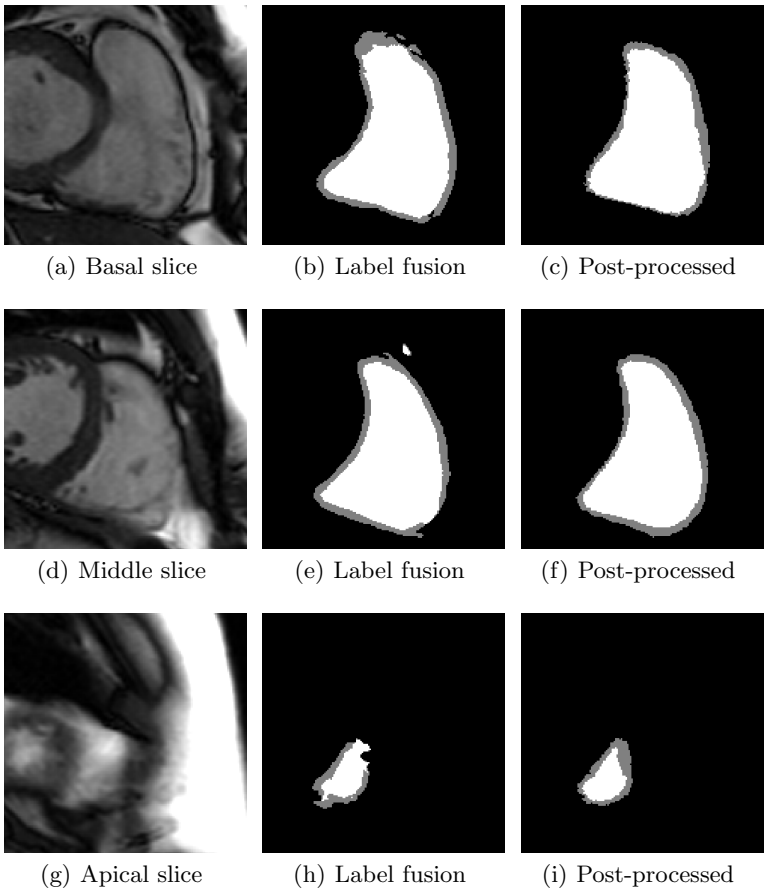
The method was evaluated on the data set provided by the segmentation challenge organisers. Test set 1 consists of MR images of 16 subjects. Each image was segmented using 16 atlases with known label maps from the training set. To reduce computational cost, a subset of 10 atlases that were most similar to the target image after affine registration were selected [3]. Non-rigid registration and local label fusion were performed only on this subset. The parameters σ_i and σ_d were determined by performing segmentation on 6 training images which were randomly picked. The values that achieved the best segmentation accuracy were selected. In our experiments, $\sigma_i = 75$ and $\sigma_d = 0.5$ mm.

The main computational cost is the non-rigid registration, which took about 2 minutes for each pair of images. Considering that 2 frames (ED and ES) needed to be processed and there were a subset of 10 atlases, non-rigid registration would take about 40 minutes for each test subject. Affine registration, atlas selection and local label fusion took just a few minutes. Since we ran the registrations parallel on a workstation server, it took us about 10 minutes to process a test subject for both frames.

3.1 Visual Inspection

Figure 1 illustrates the segmentation on a test image at three different slice locations (basal, middle and apical). Local label fusion results in good segmentation

[2] However, more sophisticated models such as an average atlas or an anatomical surface model can be used as well for the purpose of post-smoothing and may provide a better performance.



(a) Basal slice (b) Label fusion (c) Post-processed

(d) Middle slice (e) Label fusion (f) Post-processed

(g) Apical slice (h) Label fusion (i) Post-processed

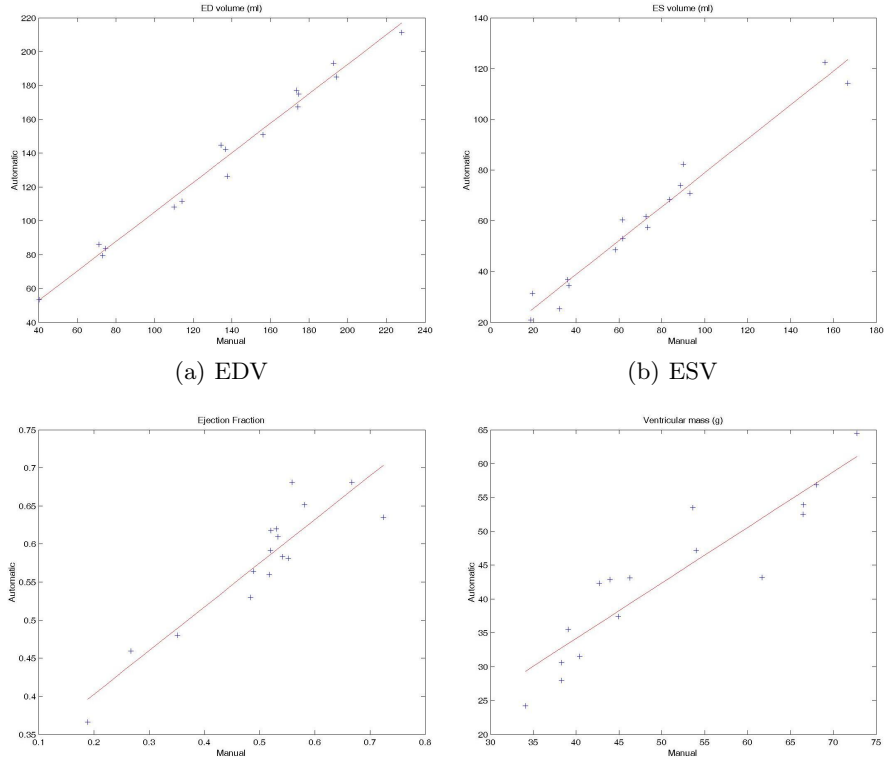
Fig. 1. An example of segmentation at three different slice locations. Top to bottom rows: basal, middle and apical slices. Left to right columns: intensity image, fused label map and post-processed label map. In the label map, the myocardial wall and the blood pool are coloured as grey and white respectively.

Table 1. Dice metric (DM) and Hausdorff distance (HD, unit: mm) between the computer-aided segmentation and manual delineation. The measurements are given for both endocardium and epicardium respectively at end-diastole and end-systole.

		End-diastole	End-systole
Endo	DM	0.86 ± 0.11	0.69 ± 0.25
	HD	7.70 ± 3.74	11.16 ± 5.53
Epi	DM	0.88 ± 0.08	0.77 ± 0.17
	HD	7.93 ± 3.72	11.72 ± 5.44

(a) ED (b) ES

Fig. 2. A subject image at ED and ES respectively. The RV is relatively small at ES. The myocardium becomes very thick and it becomes difficult to discern the blood pool. It can be more difficult to register the image at ES to an atlas image than at ED.



(a) EDV (b) ESV

(c) Ejection fraction (d) Ventricular mass

Fig. 3. Linear regression between measurements given by our segmentation and manual measurements. The X-axis represents the manual measurement, whereas the Y-axis represents the automated measurement.

on the basal and middle slices, as shown in Figure 1(b) and (e), because the blood pool and the myocardium are well defined in the corresponding intensity images. However, the continuity and smoothness of the myocardial wall is not guaranteed. For example, in Figure 1(e), there is no myocardial wall at the bottom right of the label map. After post-processing and fitting to an anatomical model, the anatomical topology is recovered, as shown in Figure 1(f). The segmentation at the apical slice, Figure 1(h), is not satisfying due to the strong bias field in the intensity image and the poor contrast of the myocardium. Again, post-processing improves segmentation.

3.2 Quantitative Evaluation

The Dice overlap metric (DM) between the segmentation and manual labelling was computed. In addition, the Hausdorff distance (HD) between the segmented contours and manual contours was measured. Table 1 reports the measurements for both the endocardium and epicardium respectively at end-diastole (ED) and end-systole (ES) on test set 1 of 16 subjects. The average Dice metric at ED is 0.86 for endocardium and 0.88 for epicardium. It is better than the Dice metric at ES, which is 0.69 for endocardium and 0.77 for epicardium. It may be due to two reasons. First, when we use landmarks to initialise registration, we only pick landmarks on the ED images to minimise human interaction. These landmarks are then used to initialise the affine registrations for both ED and ES. As a result, the ES image registration may be worse than the ED image registration, subsequently resulting in worse label fusion. Second, some subjects have very strong RV contraction. The RV is relatively small at ES, as illustrated in Figure 2. Therefore, image registration between the target image and the atlases can be more difficult at ES than at ED.

The ED and ES volumes, ejection fraction and ED ventricular mass are clinically important indices to assess the RV function. They were measured based on our segmentation and compared to manual measurements. Figure 3 shows the linear regression between the two. The correlation coefficient and coefficients for linear regression ($y = ax + b$) are respectively: EDV ($R = 0.992$, $a = 0.874$, $b = 17.859$), ESV ($R = 0.982$, $a = 0.668$, $b = 12.125$), EF ($R = 0.916$, $a = 0.574$, $b = 0.288$) and VM ($R = 0.914$, $a = 0.821$, $b = 1.352$). The automated measurements are in a good correlation with the manual measurements.

4 Conclusions

In this paper, we proposed a multi-atlas based method for right ventricle segmentation. The atlas label maps are registered to the target image and combined by local label fusion, weighted by intensity similarity and distance. The fused label map is then fit to a model to ensure that the anatomical topology is preserved. The method was tested on a common data set and yielded accurate segmentation. The automated segmentation matches manual segmentation in terms of the Dice metric and Hausdorff distance. Automated measurement of clinical indices

also shows good correlation with manual measurement. The current method requires a few landmarks to initialise image registration. Automatic detection of the right ventricle and initialisation of the transformation would be our future work so that the method can become fully automatic.

References

1. R. Prakash. Determination of right ventricular wall thickness in systole and diastole. echocardiographic and necropsy correlation in 32 patients. *British Heart Journal*, 40(11):1257–1261, 1978.
2. R.A. Heckemann, J.V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115–126, 2006.
3. P. Aljabar, RA Heckemann, A. Hammers, JV Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–738, 2009.
4. J.M.P. Lotjonen, R. Wolz, J.R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, D. Rueckert, et al. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *NeuroImage*, 49(3):2352–2365, 2010.
5. E.M. van Rikxoort, I. Išgum, Y. Arzhaeva, M. Staring, S. Klein, M.A. Viergever, J.P.W. Pluim, and B. van Ginneken. Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus. *Medical Image Analysis*, 14(1):39–49, 2010.
6. X. Artaechevarria, A. Muñoz-Barrutia, and C. Ortiz-de Solorzano. Combination strategies in multi-atlas image segmentation: Application to brain MR data. *IEEE Transactions on Medical Imaging*, 28(8):1266–1277, 2009.
7. I. Išgum, M. Staring, A. Rutten, M. Prokop, M.A. Viergever, and B. van Ginneken. Multi-atlas-based segmentation with local decision fusion - Application to cardiac and aortic segmentation in CT scans. *IEEE Transactions on Medical Imaging*, 28(7):1000–1010, 2009.
8. T.R. Langerak, U.A. van der Heide, A.N.T.J. Kotte, M.A. Viergever, M. van Vulpen, and J.P.W. Pluim. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). *IEEE Transactions on Medical Imaging*, 29(12):2000–2008, 2010.
9. M.R. Sabuncu, B.T.T. Yeo, K. Van Leemput, B. Fischl, and P. Golland. A generative model for image segmentation based on label fusion. *IEEE Transactions on Medical Imaging*, 29(10):1714–1729, 2010.
10. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.
11. D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
12. 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 Transactions on Medical Imaging*, 21(9):1151–1166, 2002.
13. S.P. Raya and J.K. Udupa. Shape-based interpolation of multidimensional objects. *IEEE Transactions on Medical Imaging*, 9(1):32–42, 1990.