A multi-image graph cut approach for cardiac image segmentation and uncertainty estimation

Wenzhe Shi[1], Xiahai Zhuang[2], Robin Wolz[1], Duckett Simon[3], KaiPin Tung[1], Haiyan Wang[1], Sebastien Ourselin[2], Philip Edwards[1], Reza Razavi[3], and Daniel Rueckert[1]

[1] Biomedical Image Analysis Group
Department of Computing
Imperial College London
[2] Centre for Medical Image Computing
Department of Computing
University College London
[3] The Rayne Institute
Kings College London

Abstract. Registration and segmentation uncertainty may be important information to convey to a user when automatic image analysis is performed. Uncertainty information may be used to provide additional diagnostic information to traditional analysis of cardiac function. In this paper, we develop a framework for the automatic segmentation of the cardiac anatomy from multiple MR images. We also define the registration and segmentation uncertainty and explore its use for diagnostic purposes. Our framework uses cardiac MR image sequences that are widely available in clinical practice. We improve the performance of the cardiac segmentation algorithms by combining information from multiple MR images using a graph-cut based segmentation. We evaluate this framework on images from 32 subjects: 13 patients with ischemic cardiomyopathy, 14 patients with dilated cardiomyopathy and 5 normal volunteers. Our results indicate that the proposed method is capable of producing segmentation results with very high robustness and high accuracy with minimal user interaction across all subject groups. We also show that registration and segmentation uncertainties are good indicators for segmentation failures as well as good predictors for the functional abnormality of the subject.

1 Introduction

Magnetic Resonance (MR) imaging can be used to visualize the anatomy and function of the heart in detail. The most commonly available MR images of the heart include multiple stacks of short-axis (SA) and long-axis (LA) MR images. These images are typically acquired as cine sequences showing the heart throughout the entire cardiac cycle. Due to the anisotropic resolution of the images (high in-plane resolution but low out-of-plane resolution) and the fact that different

slices of the stack are acquired during different breath-holds automated segmentation is difficult. On the other hand, 3D volumetric cardiac MR imaging is now becoming feasible [1]. These images have high spatial resolution and are free from inter-slice motion. However, the images have a lower signal-to-noise-ratio and lack of contrast compared to SA and LA MR images. Therefore, combining 3D and cine MR image data, has potential to provide better accuracy and robustness for automated segmentation.

One of the widely recognized technique for cardiac anatomy segmentation is to propagate a pre-constructed atlas to the unseen images using image registration [2, 3]. By using a locally affine registration method (LARM), this technique is able to deal with large shape variations of the heart. Another alternative is voxel based segmentation [4, 5] The method is able to achieve sub-voxel accuracy but requires a good initialization.

Atlas propagation is widely used either for the initialization for cardiac segmentation [2] or as the primary segmentation method [3]. An important but not yet fully explored aspect of such image segmentation is: How can we quantify and visualize the segmentation uncertainty? This question can be further divided into uncertainty arising from the registration [6, 7] and uncertainty about the final segmentation. No matter how robust a segmentation technique is, it is important to have the ability to alert the user if the uncertainty of the segmentation quality is high. High uncertainty can either be a sign of an unreliable segmentation result or of an abnormal cardiac anatomy.

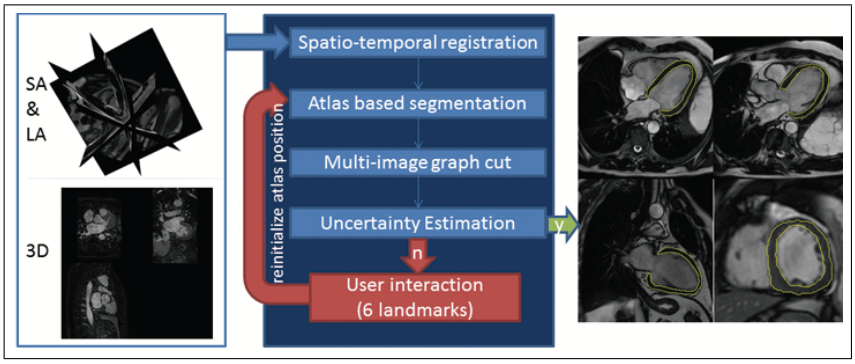

Fig. 1. Work-flow of the automatic segmentation and uncertainty estimation framework.

In this paper we extend an automatic image segmentation technique [8] to a framework that simultaneously uses information from multiple (possibly sparsely sampled) cardiac images. The integration of registration- and intensity-based segmentation has shown the ability to achieve both good robustness and accuracy [8]. In the proposed framework shown in Fig.1 we automatically segment the right ventricle, left ventricle and myocardium simultaneously from high-resolution 3D

MR images (3D) as well as multiple stacks of SA and LA cine MR images. Before image segmentation we transform all images into a common spatial and temporal coordinate system and correct the misalignments between inter- and intra- sequences [9]. We first developed a registration scheme that propagates a probabilistic atlas to the subject's coordinate system and then used a multiple component EM (MCEM) estimation algorithm for an initial segmentation. The segmentation is then refined using a multi-image graph cut algorithm.

We also explore the potential of registration and segmentation uncertainty in improving the robustness: If and only if the uncertainty of the segmentation is high, the system will ask the user to input additional landmarks to help better initialize the atlas-to-subject registration. The landmarks include apex, center of mitral valve, center of left ventricle and two right ventricle insertion points. The next section describes the segmentation framework in detail; Section 3 introduces the idea of using uncertainty in the analysis. Finally, section 4 shows results from 32 patients while section 5 summarizes and concludes the paper.

2 Cardiac segmentation using multiple images

SA and LA cine MR views provide images with high spatial resolution within each slice, but the spatial resolution between slices is poor. Nevertheless, both SA and LA images have high temporal resolution revealing dynamic information about the heart. By contrast, 3D MR images acquired within a single breath-hold provide a static image of the heart with high spatial resolution in all three directions. However, these images are often noisy and provide less good contrast for the myocardium, leading to less accurate delineation. Therefore, we propose to use all three types of MR images within a unified segmentation framework that employs a two-step segmentation technique using registration and intensity-based segmentation.

2.1 Spatio-temporal registration

The images we use for each patient consist of stacks of SA and LA images (acquired as cine images) as well as a 3D anatomical end-systolic volume. The LA image stacks consist of four (4CH), three (3CH) and two chamber (2CH) views. Note that, the SA and LA images are acquired during a separate breath-hold for each slice while the 3D anatomical image is acquired in a single breath-hold. Due to potential differences in the position of the heart (e.g. due to respiration) there is usually some spatial misalignment between the images (inter-sequence misalignment) as well as between individual slices of the SA and LA images (intra-sequence misalignment). In addition there is temporal misalignment between the 3D anatomical image and the SA and LA cine images. In order to use multiple images simultaneously, these misalignments must be corrected. The 3D image provides good spatial resolution to serve as target for accurate slice-to-volume registration [9]. In this framework we first register all available LA and SA image sequences to the 3D image using a 1D temporal registration by

maximizing normalized mutual information as a similarity measure followed by a spatial 3D rigid registration using the same similarity measure. The resulting spatio-temporal (4D) transformation corrects both the inter-sequence misalignment and intra-sequence slice shifts so that all images can be transformed to the same spatio-temporal coordinate system, in this case, the coordinate system of the 3D anatomical MR image.

2.2 Atlas based segmentation

In [8] image registration was used to propagate an atlas constructed from normal population to subjects. A locally affine registration method (LARM) [3] was used to address the large local shape variability of pathological cardiac anatomy, commonly seen across large populations with pathologies. LARM is integrated into the registration process as an intermediate registration step between a global affine registration and a fully non-rigid registration. Compared to traditional registration schemes, LARM is capable of providing a good initial alignment between the images of patients with pathologies and the atlas constructed from normal subjects.The deformation is defined under the following equation 1:

$$T(X) = \begin{cases} G_i(X) & X \in V_i \\ \sum_{i=1}^{i=n} W_i(X) G_i(X) & \text{otherwise} \end{cases} \tag{1}$$

where $G_i(X)$ is region V_i's estimated affine transformation and W_i is the distance between given X and V_i.

After atlas propagation, We [8] used a two-component Gaussian mixture model for the myocardial tissue modelling infarcted and non-infarcted myocardial tissue while being spatially constrained by the probabilistic atlas [2] propagated.

We extend the above method to multi-image atlas propagation using a combined normalized mutual information similarity measure in which the similarity for each image is weighted by the relative number of voxels in the image.

2.3 Multi-image graph cut refinement

The MCEM algorithm [8] segments the 3D, SA and LA images separately although the atlas is propagated to all images simultaneously. We propose to use an energy function based on Markov Random Fields (MRF) in combination with graph-cuts [5] to refine the segmentation across all images at the same time. 4D graph cuts have recently been used to segment image sequences [10, 11]. Here, we have adopted the 4D graph cut approach to utilize information from multiple MR images with different spatial resolution. To differentiate our approach from a 4D graph cut segmentation of image sequences, we refer to it as multi-image graph cut segmentation.

Let I_i be the i-th image of multiple images, segmenting I_i is defined as a process of assigning a label $f_p \in L$ to each voxel $p \in I_i$. An MRF-based energy

function can be formulated as:

$$E(f) = \lambda \sum_{p \in I_i} D_p(f_p) + \sum_{\{p,q\} \in N} V_{p,q}^{intra}(f_p, f_q) + \sum_{\{p,q\} \in M} V_{p,q}^{inter}(f_p, f_q) \quad (2)$$

where N and M are a neighborhood of voxels within an image and across different images respectively and f is the labeling of I_i [5]. The data term $D_p(f_p)$ measures the disagreement between the *a-priori* probabilistic model and the observed data. $V_{p,q}(f_p, f_q)$ is a smoothness term penalizing discontinuities of the segmentation in N or M. The parameter λ governs the influence of the data and smoothness terms. We found heuristically that setting $\lambda = 2$ leads to robust results for myocardium segmentation. Two different smoothness terms are chosen respectively for inter image similarity and intra image similarity since they are intuitively distinguished. For intra image similarity continuity in intensity space is enforced. While for inter image similarity comes from overlap between voxels. And continuity in intensity space is neither granted nor meaningful due to different modalities and strong spatial alignment.

To optimize eq. (2), a graph $G =< V, E >$ with a node $v \in V$ for each voxel p is defined on images. Each edge $e \in E$ consists of connections between node v and two terminal nodes F and B (also called source and sink node) as well as connections between neighboring voxels. The terminal nodes F and B represent the two labels describing foreground and background, respectively. By determining a minimum cut on graph G, the desired segmentation can be obtained[5]. The data term $D_p(f_p)$ is estimated using the MCEM segmentation [8] which generates a probability for each class of each voxel.

The smoothness term between neighboring voxels within an image is defined over a cubic neighborhood N by the following equation:

$$V_{p,q}^{intra} = w_{intra} \frac{1}{\ln(1 + (I_p - I_q)^2) + \epsilon} \quad (3)$$

Here I is intensity and ϵ is a small constant value which compensates for noise when I_p is close to I_q. For neighbouring voxels we define $w_{intra} = 1/d$ where d is the distance between two voxels.

For voxels across different images, a different smoothness term is chosen. We define a smoothness term that depends on the degree of overlap between the voxels instead of the intensity similarity to enforce spatial consistency and address the different modalities between images. We use the Dice metric to compute the amount of overlap between images

$$V_{p,q}^{inter} = w_{inter}(2||S_p \cap S_q||)/(||S_p|| \cup ||S_q||) \quad (4)$$

where S_p and S_q are the voxel volumes of voxel p and q and w_{inter} is a constant weight chosen as 2 from extensive experiment. The result of this equation is real number between 0 and 1 due to different voxel size and position of the images. The smoothness term is defined in a neighborhood M where $V_{p,q}^{inter} > 0$.

By using this multi-image graph cut approach, we connect intra-image voxel neighbors according to their intensity similarity and distance and inter-image

voxels neighbors according to their spatial overlap. This enables us to segment multiple images simultaneously and consistently. Finally the energy function is optimized using graph-cuts and multiple labels are achieved at the same time using the expansion and the swap algorithm[5].

3 Uncertainty definition and evaluation

3.1 Registration Uncertainty

Given two images, I_{atlas} and I_i (the subject's ith image), we can estimate a transformation $\mathbf{T}$ which maps image I_i to I_{atlas} so that a voxel of $I_i(\mathbf{x})$ correspond to $I_{atlas}(\mathbf{T}(\mathbf{x}))$ and their intensity values should be similar. Using a probabilistic formulation for the image registration problem [6], the uncertainty of a transformation $\mathbf{T}$ at point $\mathbf{x}$ can be modeled by the following equation :

$$uc(\mathbf{T}(\mathbf{x})|(I_i(\mathbf{x}), I_{atlas})) = 1 - \frac{p((I_i(\mathbf{x}), I_{atlas})|\mathbf{T}(\mathbf{x}))p(\mathbf{T}(\mathbf{x}))}{p((I_i(\mathbf{x}), I_{atlas}))} \tag{5}$$

We model the likelihood term, $p((I_i(\mathbf{x}), I_{atlas})|\mathbf{T}(\mathbf{x}))$, as a normal distribution of the intensity difference between transformed I_{atlas} and I_i estimated using an EM algorithm after histogram equalization. Similarly, the prior of the transformation, $p(\mathbf{T}(\mathbf{x}))$, is modeled as a Rician distribution of the Jacobian determinant of the transformation [12]. The distribution is estimated based on the inversion technique proposed in [13]. The Rician distribution is a non-negative and asymmetric distribution which approximates the distribution of the Jacobian determinant well for a given transformation. Finally, $p((I_i(\mathbf{x}), I_{atlas}))$ can be modelled as a constant term.

3.2 Segmentation Uncertainty

The uncertainty of a given label from our 4D graph cut segmentation can be modeled by the following equation:

$$uc(L_j|I_i(\mathbf{x})) = 1 - \frac{p(I_i(\mathbf{x})|L_j)p(L_j)}{p(I_i(\mathbf{x}))} \tag{6}$$

where $p(I_i(\mathbf{x})|L_j)$ is the likelihood that intensity of $I_i(\mathbf{x})$ belongs to L_j as estimated by the segmentation method. $p(I_i(\mathbf{x}))$ is modelled as a constant term and

$$p(L_j) = \frac{1}{log(\delta_j + 1 + \epsilon)} \tag{7}$$

Here ϵ is a small constant value and δ is the interquartile range of the multiple component distribution that represents L_j's intensity distribution from our segmentation method. The interquartile range is chosen because it's a robust statistic that conveys the dispersion of a distribution [6] and corresponds well to the intra-region homogeneity of a segmentation. It is robust in the sense that it provides meaningful information even for non-Gaussian distributions like the ones that can be obtained from the MCEM segmentation.

Table 1. Validation results: The Dice overlap measure for the endocardial segmentation (LV) and epicardial segmentation (LV+MYO) results comparing automatic and manual segmentation. $*$ means that the pair-wise t-test is significant at $p < 0.05$ and $\star$ $p < 0.01$

Group	Segmentation	Segmentation using SA only[8]	Proposed segmentation
all	endocardial$*$	0.907 ± 0.032	0.920 ± 0.026
	epicardial$*$	0.908 ± 0.028	0.921 ± 0.024
apex	endocardial$\star$	0.837 ± 0.095	0.900 ± 0.059
	epicardial$\star$	0.838 ± 0.079	0.910 ± 0.052
mid	endocardial	0.918 ± 0.028	0.923 ± 0.024
	epicardial	0.920 ± 0.029	0.925 ± 0.021
basal	endocardial $*$	0.894 ± 0.048	0.916 ± 0.041
	epicardial $*$	0.896 ± 0.036	0.917 ± 0.033

3.3 Uncertainty quantification and user interaction

For each voxel $\mathbf{x}_i$ in all images, its registration and segmentation uncertainty can be evaluated and visualized using eqs. (5) and (6) respectively. We can further define the registration and segmentation uncertainty of a given region L_j by averaging over the region. The quantification of uncertainty can be used to inform the user about how reliable the segmentation results are.

Based on results from the uncertainty analysis, we can design a system that that detects abnormally high uncertainty. In our analysis we have four failure cases in which the global affine registration fails during the atlas registration. The subsequent segmentations also fail. Myocardial registration uncertainty is a good indicator for failed global affine registration (failed cases 0.86 ± 0.14, successful cases 0.39 ± 0.08 $p < 0.0001$). A combination of registration and segmentation uncertainties is better in terms of classification using linear discriminant analysis (LDA) (failed cases 1.73 ± 0.12, successful cases 1.2 ± 0.08 $p \ll 0.0001$). A good threshold for detecting segmentation failures using the combined registration and segmentation uncertainty is 1.59 derived from LDA with accuracy of 100%.

In the cases that we detect a segmentation failure, the user is asked to define 6 landmarks (apex, center of left ventricle, anterior and inferior insertion points of right ventricle, center of right ventricle and center of basal plane). These landmarks are also defined in the atlas. By introducing knowledge about these additional 6 landmarks, the atlas-to-image registration can be initialized more accurately and all segmentations performed correctly.

4 Results

In this paper we used datasets form 32 subjects. Each dataset consists of short axis (SA), long axis (LA) four (4CH), three (3CH) and two chamber (2CH) cine MR image sequence ($2.2 \times 2.2 \times 10, \mathrm{mm}$, 30 phases) and anatomical 3D MR images ($1.1 \times 1.1 \times 1.1, \mathrm{mm}$, one phase).

Manual segmentations were performed by a cardiologist to extract the myocardium and left ventricle in all images after spatio-temporal registration. We

then compared these to the segmentation results obtained via our previous technique [8] which uses SA MR images only and the proposed technique. Both techniques utilize the landmarks from the user for four patients for which the global affine registration fails. For comparisons between the methods we used the Dice metric, $D = (2||S_a \cap S_b||)/(||S_a|| + ||S_b||)$ where S_a and S_b are respectively the manual label segmentation and automatic label segmentation. The results are summarized in tab.1. The results indicate that our proposed segmentation scheme performed better than the original method especially on the basal and the apex segments. The basal and apex are very difficult to segment using SA MR images only due to the large slice thickness and partial volume. In the proposed approach the segmentation in these region is enhanced by adding information from 3D and LA images.

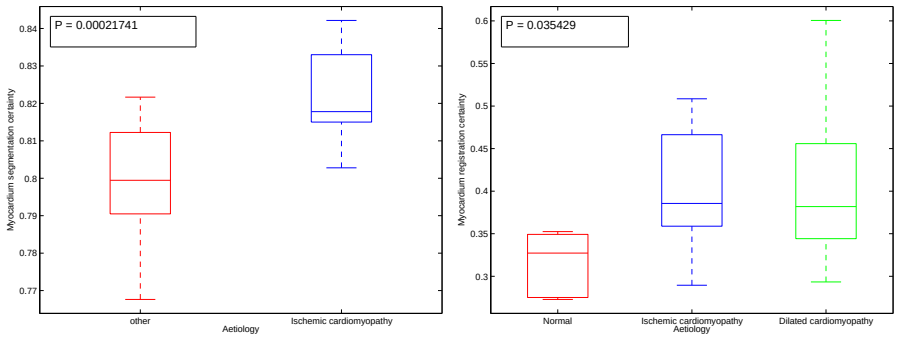

Fig. 2. Left figure shows myocardial segmentation uncertainty and right figure shows myocardial registration uncertainty

Uncertainty is a good indicator for the failure of the global affine registration. However, during our experiment, there is no strong correlation between uncertainty and accuracy of the segmentation if the uncertainty does not rise beyond the threshold used. If the segmentation is considered successful, the uncertainty relates more to abnormality of the patient's cardiac anatomy than the accuracy of the segmentation. This is possibly due to the fact that our segmentation algorithm is designed to segment pathological images well using LARM [3] and MCEM [8]. To examine if the uncertainty correlates to the abnormality of the patients, we assume that segmentation uncertainty which comes from intensity and geometry distribution relates to abnormal intensity like ischemic cardiomyopathy while registration uncertainty which comes from geometry distribution corresponds well to abnormal geometry like dilated cardiomyopathy. Figure.2 shows that myocardial segmentation uncertainty is a very good predictor for separating ischemic cardiomyopathy from the rest of subjects (ischemic subjects 0.82 ± 0.012, other subjects (normal and dealated) 0.80 ± 0.016 $p < 0.001$), meanwhile myocardial registration uncertainty is a good predictor for separating dilated cardiomyopathy from normal subjects (dilated subjects 0.38 ± 0.08,

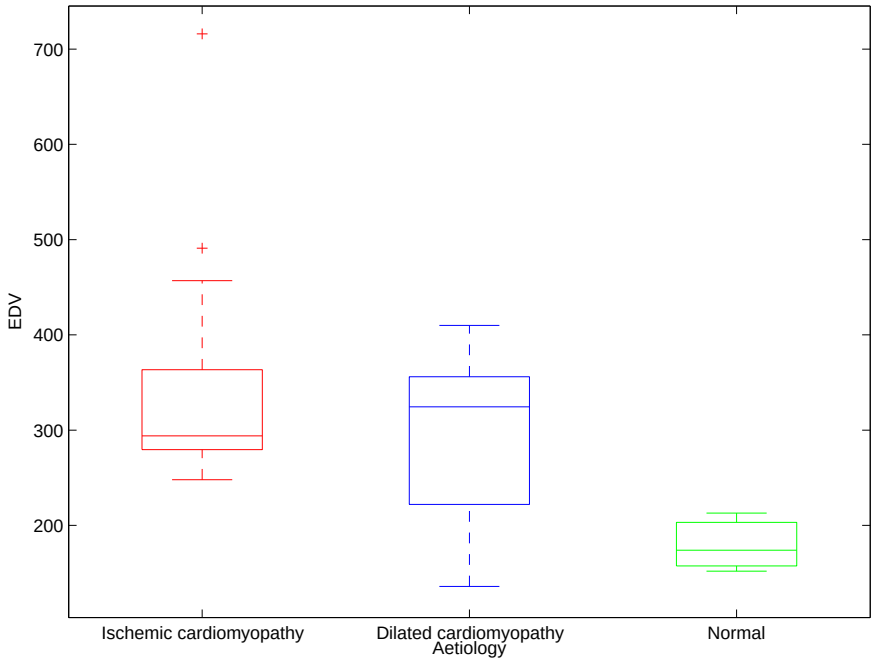

Fig. 3. End Diastolic Volume(EDV) of the subjects from the segmentation, p value between normal and ischemic is 0.025, between normal and dilated is 0.054 and between dilated and ischemic is 0.287.

normal subjects 0.33 ± 0.04 $p < 0.05$) but not from ischemic subjects. Compare to EDV Figure.3, uncertainty outperforms EDV by distinguishing ischemic ($p < 0.001$ against $p < 0.05$) from other and separate dilated ($p < 0.05$ against $p > 0.05$) from normal.

5 Conclusion and Future Work

In this paper we present a novel two-step multiple image segmentation framework using three widely available MR image sequences. Using LARM and MCEM we are able to deal with local shape variations as well as infarcted myocardium. The segmentation is performed simultaneously from all images using a multi-image graph cut approach. The accuracy is significantly improved compared to previous segmentation methods by utilizing both intra- and inter-image information Table.1. We finally define a system that detects segmentation failures using registration and segmentation uncertainties.

Cardiac pathology is not always easily detectable in images, e.g. the transposition of vessels, but likely to be detected by registration and segmentation uncertainty. Since we can define uncertainty for every part of the cardiac anato-

my, it is desirable to investigate if the relationship between uncertainty and abnormality could help to detect these pathologies automatically.

References

1. Uribe, S., Muthurangu, V., et al.: Whole-heart cine MRI using real-time respiratory self-gating. Magnetic Resonance in Medicine **57**(3) (2007) 606–613
2. Lorenzo-Valdés, M., Sanchez-Ortiz, G., Rueckert, D., et al.: Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. Medical Image Analysis **8**(3) (2004) 255–265
3. Zhuang, X., Rhode, K., Ourselin, S., et al.: A Registration-Based Propagation Framework for Automatic Whole Heart Segmentation of Cardiac MRI. IEEE Transactions on Medical Imaging (2010) 1612–1625
4. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging **20**(1) (2001) 45–57
5. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on pattern analysis and machine intelligence (2001) 1222–1239
6. Risholm, P., Pieper, S., Samset, E., Wells, W.: Summarizing and Visualizing Uncertainty in Non-rigid Registration. MICCAI (2010) 554–561
7. Kybic, J.: Bootstrap resampling for image registration uncertainty estimation without ground truth. IEEE Transactions on Image Processing **19**(1) (2009) 64–73
8. Shi, W., Zhuang, X., et al.: Automatic segmentation of different pathologies from cardiac cine mri using registration and multiple component em estimation. In: Functional Imaging and Modeling of the Heart. Volume 6666 of Lecture Notes in Computer Science. (2011) 163–170
9. Chandler, A., Razavi, R., et al.: Correction of misaligned slices in multi-slice MR cardiac examinations by using slice-to-volume registration. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro., IEEE (2006) 474–477
10. Linguraru, M., Pura, J., Chowdhury, A., Summers, R.: Multi-organ Segmentation from Multi-phase Abdominal CT via 4D Graphs Using Enhancement, Shape and Location Optimization. MICCAI (2010) 89–96
11. Wolz, R., Heckemann, R., Aljabar, P., et al.: Measurement of hippocampal atrophy using 4D graph-cut segmentation: Application to ADNI. NeuroImage (2010)
12. Rohlfing, T., Maurer Jr, C., Bluemke, D., Jacobs, M.: Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. IEEE Transactions on Medical Imaging **22**(6) (2003) 730–741
13. Koay, C., Basser, P.: Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. Journal of Magnetic Resonance **179**(2) (2006) 317–322