Multi-atlas Spectral PatchMatch: Application to Cardiac Image Segmentation

Wenzhe Shi[1], Herve Lombaert[3], Wenjia Bai[1], Christian Ledig[1], Xiahai Zhuang[2], Antonio Marvao[1], Timothy Dawes[1], Declan O'Regan[1], and Daniel Rueckert[1]

[1] Biomedical Image Analysis Group, Imperial College London, UK
[2] Shanghai Advanced Research Institute, Chinese Academy of Sciences, China
[3] INRIA, Asclepios Project-Team, Sophia-Antipolis, France

Abstract. The automatic segmentation of cardiac magnetic resonance images poses many challenges arising from the large variation between different anatomies, scanners and acquisition protocols. In this paper, we address these challenges with a global graph search method and a novel spectral embedding of the images. Firstly, we propose the use of an approximate graph search approach to initialize patch correspondences between the image to be segmented and a database of labelled atlases. Then, we propose an innovative spectral embedding using a multi-layered graph of the images in order to capture global shape properties. Finally, we estimate the patch correspondences based on a joint spectral representation of the image and atlases. We evaluated the proposed approach using 155 images from the recent MICCAI SATA segmentation challenge and demonstrated that the proposed algorithm significantly outperforms current state-of-the-art methods on both training and test sets.

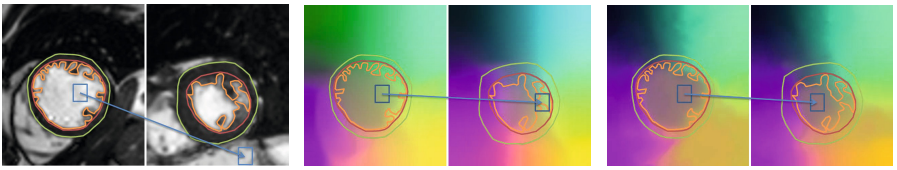
1 Introduction

An important step in the analysis of cardiac magnetic resonance (MR) images is the segmentation of the image into different anatomical structures or regions. One of the most popular segmentation approaches is based on multi-atlas label fusion [1,2,3]. The main components of these methods are atlas selection, atlas propagation and label fusion.

In the atlas propagation step, affine or non-rigid registration methods are commonly used, such as the free-form deformation (FFD) [4] registration or the Demons [5] algorithm. The registration is commonly based on intensity similarities and constrained to ensure one-to-one correspondences between the target image and the atlas. This restriction ensures a realistic deformation that preserves the topology of the atlas structures in the target image. However, it also limits the ability of the registration to capture large or local variations in shape.

To relax the method's dependence on accurate registrations, recent research has focused on patch-based label fusion methods [6,7,8,9,10] including their application to cardiac MR images [7,11]. These approaches compensate registration error by searching for correspondences between the target image and atlas within a limited search window. However, intensity based features, which are

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 348–355, 2014.



(a) (b) (c)

Fig. 1. This figure shows the most similar patches under different representations of the images. Red, green, and orange contours are the manual segmentation of endocardial, epicardial boundaries and trabeculae/papillary muscles respectively. For spectral representations, RGB color represents the first three eigenmodes. (a) shows the intensity images and the correspondence; (b) shows the independent spectral representations and the correspondence; (c) shows the joint spectral representation and the correspondence.

often used as patch selection criteria, can have ambiguous matches especially for larger search windows as shown in Fig. 1a.

PatchMatch [12] is a popular search method to find global patch correspondences between two images based on an approximate graph search. This approach is appealing due to its unique ability to capture large variations while remaining computationally feasible. However, there is no intrinsic regularization and corresponding patches do not necessarily preserve the topology of anatomical structures as shown in Fig. 3d.

In this paper, we aim to search for patch correspondences between the target image and warped atlases without any restriction of the search window size for the purpose of label fusion. There are two challenges, as discussed above: Firstly, cardiac MR images exhibit significant variability in terms of anatomy. This renders the application of conventional registration methods difficult. Secondly, the optimal matching patches, in terms of intensity features, may come from different anatomical regions around the heart. This limits the usefulness of patch selection methods like label fusion or the PatchMatch algorithm [12].

In this paper, we propose a novel method called multi-atlas spectral PatchMatch (MASP) to overcome these limitations. Here we combine spectral matching [13] with multi-atlas PatchMatch (MAPM) [14]. Recent advances in spectral matching [13], based on spectral graph theory [15], allow global correspondences to be established between two graphs by linking them together. By modelling images as graphs, this method can be applied to image registration [16]. However, any alteration in the images will inherently change the graph representation and result in perturbations in shape isometry. These alterations include variation of the objects, changes of the regions of interest, and different views of the scene. This will lead to the change in the spectral representation as shown in Fig.2d and Fig. 1b, thus, affect the performance of matching.

In order to build a consistent spectral embedding of the images, we first represent the images as graphs [16]. However, different from [16], we construct a multi-layered graphical representation of the image and atlases by using the patch correspondence stemming from the PatchMatch (Fig. 2a). We then learn the joint spectral representation of the graph using the principles from [13]

(Fig. 2b). Finally, we estimate the correspondences using MAPM across different images simultaneously based on the spectral representation of the patches (Fig. 2c). Based on the estimated patch correspondences, we then segment the images with the method proposed in [6].

The contribution of this paper is the introduction of a novel joint spectral representation of images and atlases. This representation is intrinsically aware of the image content, in our case, the anatomy of the cardiac region. In contrast to the application of [16], we capture shape properties among all images simultaneously. By using this new representation, we enable a multi-layered graph search strategy and recover unambiguous global patch correspondences between the unseen image and atlases. Our results demonstrate that the proposed algorithm significantly outperforms state-of-art algorithms. We confirm our results through a blinded and external online evaluation.

2 Method

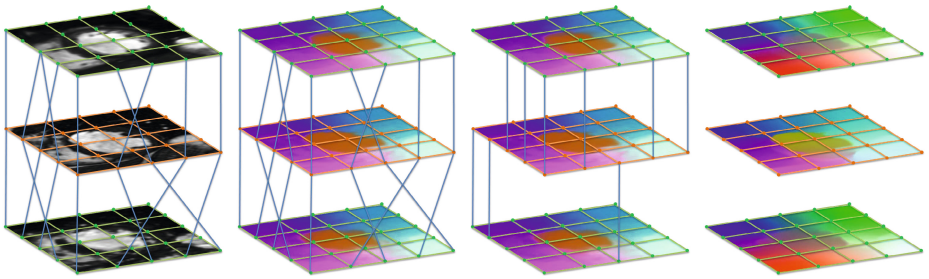


Fig. 2. Algorithm overview – intermediate steps for segmenting an unseen image and two atlases. The number of atlases, vertices and edges has been simplified for better illustration. The unseen image is overlaid with an orange grid and atlases are overlaid with a green grid. (a) shows the image-atlas graph as initialized by PatchMatch; (b) shows the spectral representation of the image-atlas graph; (c) shows the final estimation of the correspondences using the spectral representation and MAPM; (d) shows the spectral representation of individual images.

2.1 Image-Atlas Graph Initialization

The PatchMatch algorithm proposed by Barnes et al. [12] finds corresponding patches between two images or regions without any restriction of the search space. It is based on the assumption that given a good match of one patch, it is very likely that the neighbouring patches share the same or similar correspondences [12]. The first step of our segmentation method is to construct a multi-layered image-atlas graph from the unseen image and atlases using the correspondences estimated with PatchMatch.

Given an unseen image $\mathbf{U}$ and an warped atlas database $\mathbf{A}$ (individual atlases are denoted as $\mathbf{A_k}, K = |\mathbf{A}|$), we find for each point $\mathbf{p} = (x, y, z)$ in image $\mathbf{U}$

a correspondence in each atlas $\mathbf{A_k}$, $\mathbf{C_k}(\mathbf{p}) = (\mathbf{q}, k)$ where $\mathbf{q} = (x', y', z')$ is the closest match in atlas $\mathbf{A_k}$ for a given distance function $D(\mathbf{p}, \mathbf{q})$ between patches $\mathbf{P}$ and $\mathbf{Q}$ centered at $\mathbf{p}$ and $\mathbf{q}$. Here, the distance is the Euclidean distance of the feature vectors. For each patch $\mathbf{P}$, the feature vector consists of the intensity, the intensity gradients in all three directions, and the spatial location x, y, z of each point $\mathbf{x} \in \mathbf{P}$. This is equivalent to embedding each patch of the unseen image and the K atlases in a M dimensional space, where $M = 7|\mathbf{P}|$ is the number of features in the feature vector. The choice of the Euclidean distance ensures the triangle inequality so that the multi-layered image-atlas graph introduced in the next section is embedded in a metric space.

2.2 Image-Atlas Graph Construction

Using the estimated correspondence fields $\mathbf{C_1}, ..., \mathbf{C_K}$, we can construct a single multi-layered undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ from image $\mathbf{U}$ and atlases $\mathbf{A}$ as shown in Fig.2a. The graph is constructed with the vertices $\mathbf{V} = \mathbf{V_U} \cup \mathbf{V_A}$ representing voxels of image $\mathbf{U}$ and atlases $\mathbf{A}$ and the edges $\mathbf{E}$ consisting of the union of $\mathbf{E_U}$, $\mathbf{E_{A_1}}, ..., \mathbf{E_{A_K}}$ and $\mathbf{E_{C_1}}, ..., \mathbf{E_{C_K}}$. Within the unseen image $\mathbf{U}$ and atlases $\mathbf{A}$, we define $\mathbf{E_U}, \mathbf{E_{A_1}}, ..., \mathbf{E_{A_K}}$ such that each voxel is connected to its immediate spatial neighbours within the same image. Between the unseen image and atlases, the edges $\mathbf{E_{C_1}}, ..., \mathbf{E_{C_K}}$ are given by the corresponding fields $\mathbf{C_1}, ..., \mathbf{C_K}$. The distance $D(\mathbf{p}, \mathbf{q})$ between the vertices is defined as in Sec.2.1.

For the multi-layered graph $\mathbf{G}$ the $|\mathbf{V_U} \cup \mathbf{V_A}| \times |\mathbf{V_U} \cup \mathbf{V_A}|$ weighted adjacency matrix has the form:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W_U} & \mathbf{W_{C_1}} & ... & \mathbf{W_{C_K}} \\ \mathbf{W_{C_1}^T} & \mathbf{W_{A_1}} & 0 & 0 \\ ... & 0 & ... & 0 \\ \mathbf{W_{C_K}^T} & 0 & 0 & \mathbf{W_{A_K}} \end{bmatrix}, \tag{1}$$

where $\mathbf{W_U}$ and $\mathbf{W_{A_k}}$ are the adjacency matrices of intra-image edges of the unseen image and atlas $\mathbf{A_k}$. $\mathbf{W_{C_k}}$ are the weighted adjacency matrices of inter image-atlas edges defined by the correspondence field $\mathbf{C_k}$. For each $w_{\mathbf{p},\mathbf{q}}$ where $\mathbf{p}$ and $\mathbf{q}$ are two voxels of the unseen image and/or atlases, we define $w_{\mathbf{p},\mathbf{q}} = \exp(-D(\mathbf{p}, \mathbf{q})^2 / 2\sigma^2)$ if $\exists e_{\mathbf{p},\mathbf{q}} \in \mathbf{E}$ or 0 otherwise. The parameter σ is set to the standard deviation of $D(\mathbf{p}, \mathbf{q})$ $\forall e_{\mathbf{p},\mathbf{q}} \in \mathbf{E}$. The resulting graph is simplified and illustrated in Fig. 2a. Examples are presented in the supplementary material.

2.3 Spectral Feature Extraction

Inspired by [13], we use the image-atlas graph $\mathbf{G}$ to exploit the joint spectral representation of the unseen image and atlases. In this paper, we use the symmetric normalized Laplacian [15] for the spectral embedding. The Laplacian is defined as $\mathbf{L^{norm}} = \mathbf{D^{-1/2}}(\mathbf{D} - \mathbf{W})\mathbf{D^{-1/2}}$ where $\mathbf{D}$ is the diagonal degree matrix [15]. The spectral decomposition of the graph Laplacian $\mathbf{L^{norm}} = \mathbf{S\Lambda S^{-1}}$ provides eigenvalues $\mathbf{\Lambda}$ and associated eigenvectors $\mathbf{S} = (s_{\cdot,0}, s_{\cdot,1}, ..., s_{\cdot,|\Lambda|})$, where $s_{\cdot,i}$ is the i^{th}

column of S. We denote the spectral representation, $\mathbf{S^N} = (s_{\cdot,0}, s_{\cdot,1}, ..., s_{\cdot,N})$, a N-dimensional embedding of the images. Each voxel $\mathbf{x}$ of images in $\mathbf{U}$ and $\mathbf{A}$ has the spectral feature defined as $s^{norm}_{\mathbf{x},(1,...,N)}$, which is a row of $\mathbf{S^N}$ normalized to the intensity range of the unseen image.

2.4 Multi-atlas Spectral PatchMatch

MAPM was proposed to find the corresponding patches between one unseen image and multiple atlases in [14]. Different from PatchMatch as described in Sec.2.1, it finds only one correspondence for each point $\mathbf{p} = (x, y, z)$ in image $\mathbf{U}$ $\mathbf{C}(\mathbf{p}) = (\mathbf{q}, k)$, so that $\mathbf{A_k}(\mathbf{q})$ is the closest match in all atlases $\mathbf{A}$. The MAPM algorithm [14] consists of three different steps which we briefly describe as follows: The correspondence $\mathbf{C}(\mathbf{x}) = (\mathbf{x}, R(K))$ is initialised by a uniform random selection $R(K)$ that selects between $\mathbf{A_1}$ to $\mathbf{A_K}$. After initialization, the correspondence $\mathbf{C}$ is improved by iterating between propagating good matches to its neighbours and searching for better matches across different atlases simultaneously. This process is performed until the sum of all distances between each patch in the unseen image and its correspondence converges.

The advantage of MAPM is that it can find global patch correspondences between an unseen image and multiple atlases simultaneously. The complexity of the algorithm does not increase with the number of atlases [14]. However, using intensity features, the match can be ambiguous and may connect different anatomical structures. In this paper, we propose adding the spectral feature s^{norm} into MAPM. We redefine the distance function as follows,

$$D_s(\mathbf{p}, \mathbf{q}) := D(\mathbf{p}, \mathbf{q}) + \sqrt{\sum_{\mathbf{p}' \in \mathbf{P}} \sum_{i=1}^{N} (s^{norm}_{\mathbf{p}',i} - s^{norm}_{\mathbf{q}',i})^2}, \tag{2}$$

where $\mathbf{q}'$ are corresponding points to $\mathbf{p}'$ in the atlas patch $\mathbf{Q}$. Using this approach, the shape context contained in the spectral features can be used to support MAPM to identify differences between different anatomical structures as shown in Fig. 1. Compared to Spectral Demons [16], this is the first attempt to build a joint spectral representation of multiple images. If we build the image graph independently, large perturbations of the graph will cause fundamental differences in the spectral representations which are difficult to recover [13]. The final label probability $P_L(\mathbf{x})$ at each voxel $\mathbf{x}$ is estimated with the method proposed in [6]. We use a weighting function identical to the one in Sec. 2.2.

3 Application to Cardiac MR Image Segmentation

The proposed framework was applied to cardiac cine MR images. The images used were from healthy and clinical cohorts of the cardiac atlas project [17] and demonstrated a wide range of shape variations. This data set was also used in the MICCAI-SATA 2013 segmentation challenge [18]. A standard acquisition was performed, including an axial stack of cine b-SSFP MR images in the left

ventricular short axis plane. The image dimensions are 192 x 192 x 16 x 30 and the voxel sizes are 1.5 x 1.5 x 6 mm. Ground truth segmentations of the myocardium were provided throughout the cardiac cycle for the training set of 83 subjects. In addition, a test set of 72 subjects were provided without ground truth segmentations.

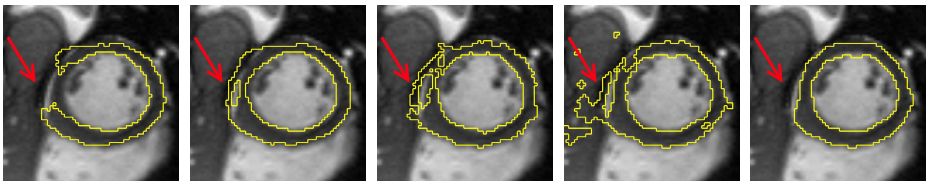
MASP was applied between slices of the unseen image and the atlases to deal with the respiratory motion and large slice thickness. Before label fusion, the atlases were aligned to the unseen image followed by intensity normalization [19]. The initial alignment was obtained using five manual landmarks similar to [11]. In the experiments, the patch size was set to 3 x 3 voxels [11]. To initialize the image-atlas graph, each pixel was connected to its four neighbouring pixels within the same image and one inter image-atlas neighbour from the unseen image to each atlas using PatchMatch [12]. The dimension of spectral features was $N = 3$ [16].

3.1 Results

We performed a leave-one-out cross-validation using the end diastolic (ED) frame of the training set. We compared the Dice metric between the segmentation result and the ground truth using different segmentation techniques including the majority voting, patch-based fusion [7], joint label fusion [9], MAPM [14] and MASP. To ensure fair comparison, all methods were based on the same

Table 1. The mean and standard deviation of Dice metric for different methods. The first, second and third rows show the results using affine registration, the result using FFD registration and the differences between the FFD and affine registration respectively. Paired sample t-test shows a significant difference between the proposed method and the other approaches (p-value < 0.01 indicated by *)

	Majority	Patch based [7]	Joint label [9]	MAPM [12]	MASP
Affine	0.673 (0.098)*	0.736 (0.090)*	0.729 (0.068)*	0.720 (0.050)*	0.800 (0.054)
FFD	0.737 (0.086)*	0.762 (0.083)*	0.761 (0.060)*	0.728 (0.050)*	0.801 (0.048)
Differences	0.065 (0.029)*	0.026 (0.021)*	0.035 (0.022)*	0.008 (0.013)*	0.001 (0.015)



(a) (b) (c) (d) (e)

Fig. 3. This figure illustrates the segmentation results of different methods using FFD. (a) shows majority voting; (b) shows patch-based segmentation [7]; (c) shows joint label fusion [9]; (d) shows the MAPM [14]; (e) shows the proposed MASP.

affine registration and refined by FFD registration. To estimate the sensitivity of the methods against different registrations approaches, we presented results based on both affine and FFD registration. For each unseen image, 20 atlases were selected from a pool of 82 images after affine registration as in [3].

The results are summarized in Tab. 1. MASP yields a significant improvement in the segmentation accuracy. MASP is also significantly less sensitive to the initial registration as shown in the third row. In addition, the difference when using MASP with different initialisation methods (affine, FFD) was not significant. The intensity based methods have difficulties in distinguishing the difference between liver and myocardium (Fig.3b,Fig.3c and Fig.3d). In particular, MAPM did not preserve the topology of the myocardium. In contrast, MASP can recover large variation in shape while preserving the topology (Fig.3e). This is due to the fact that MASP uses a global search strategy which is, due to spectral features, aware of the anatomical context.

We also submitted our result based on affine registration on the test set to the SATA segmentation challenge. To label the myocardium consistently across temporal frames, we used graph cuts [21] with the label probability $P_L(\mathbf{x})$ as the data term and linked between neighboring frames as the smoothness term. The results are summarized in the leader board[1] under the entry name MASP_AREG_GC with a mean Dice metric of 0.807 and a mean Hausdorff distance of 1.287mm.

4 Conclusion

In this paper we developed MASP: a segmentation method for, but not limited to, cardiac MR images. By creating a multi-layered image-atlas graph, we used spectral embedding to estimate spectral features of the unseen image and multiple atlases simultaneously. Compared to the intensity features, the spectral features contain shape context and facilitate the estimation of an unambiguous correspondence field. We have shown that our method is less sensitive to the initial registration and robust to the presence of large shape variations.

Acknowledgement. This research was supported by British Heart Foundation grant PG/12/27/29489.

References

1. Rohlfing, T., Brandt, R., Menzel, R., Maurer, C.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage 21(4), 1428–1442 (2004)
2. Heckemann, R., Hajnal, J., Aljabar, P., Rueckert, D., Hammers, A.: Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 33(1), 115–126 (2006)
3. Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., Rueckert, D.: Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage 46(3), 726–738 (2009)

[1] `http://masi.vuse.vanderbilt.edu/submission/leaderboard.html`

4. Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D.: Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18(8), 712–721 (1999)
5. Thirion, J.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis 2(3), 243–260 (1998)
6. Rousseau, F., Habas, P.A., Studholme, C.: A supervised patch-based approach for human brain labeling. IEEE Transactions on Medical Imaging 30(10), 1852–1862 (2011)
7. Coupé, P., Manjķn, J., Fonov, V., Pruessner, J., Robles, M., Collins, D.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. Neuroimage 54(2), 940–954 (2011)
8. Asman, A.J., Landman, B.A.: Non-local statistical label fusion for multi-atlas segmentation. Medical Image Analysis 17(2), 194–208 (2013)
9. Wang, H., Suh, J., Das, S., Pluta, J., Craige, C., Yushkevich, P.: Multi-atlas segmentation with joint label fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(3), 611–623 (2013)
10. Liao, S., Gao, Y., Lian, J., Shen, D.: Sparse patch-based label propagation for accurate prostate localization in CT images. IEEE Transactions on Medical Imaging 32(2), 419–434 (2013)
11. Bai, W., Shi, W., O'Regan, D.P., Tong, T., Wang, H., Jamil-Copley, S., Peters, N.S., Rueckert, D.: A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images. IEEE Transactions on Medical Imaging 32(7), 1302–1315 (2013)
12. Barnes, C., Shechtman, E., Goldman, D.B., Finkelstein, A.: The generalized patch-Match correspondence algorithm. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part III. LNCS, vol. 6313, pp. 29–43. Springer, Heidelberg (2010)
13. Lombaert, H., Sporring, J., Siddiqi, K.: Diffeomorphic spectral matching of cortical surfaces. In: Gee, J.C., Joshi, S., Pohl, K.M., Wells, W.M., Zöllei, L. (eds.) IPMI 2013. LNCS, vol. 7917, pp. 376–389. Springer, Heidelberg (2013)
14. Shi, W., et al.: Cardiac image super-resolution with global correspondence using multi-atlas patchMatch. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part III. LNCS, vol. 8151, pp. 9–16. Springer, Heidelberg (2013)
15. Chung, F.R.: Spectral graph theory, vol. 92. AMS Bookstore (1997)
16. Lombaert, H., Grady, L., Pennec, X., Ayache, N., Cheriet, F.: Spectral log-demons: Diffeomorphic image registration with very large deformations. International Journal of Computer Vision (2013)
17. Fonseca, C., Backhaus, M., Bluemke, D., Britten, R., Do Chung, J., Cowan, B., Dinov, I., Finn, J., Hunter, P., Kadish, A., et al.: The cardiac atlas project: an imaging database for computational modeling and statistical atlases of the heart. Bioinformatics 27(16), 2288–2295 (2011)
18. Andrew, A., Alireza, A.A., Hongzhi, W., Brian, A., Simon, K.W., Bennett, L.: MICCAI 2013 segmentation algorithms, theory and applications (SATA) challenge results summary (2013)
19. Nyúl, L., Udupa, J., et al.: On standardizing the MR image intensity scale. Magnetic Resonance in Medicine 42(6), 1072 (1999)
20. Lin, X., Cowan, B.R., Young, A.A.: Automated detection of left ventricle in 4D MR images: Experience from a large study. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 728–735. Springer, Heidelberg (2006)
21. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(11), 1222–1239 (2001)