

A UNIFIED FRAMEWORK FOR ATLAS-BASED SEGMENTATION WITH FORWARD DEFORMATION AND LABEL REFINEMENT

Siqi Bao, Albert C. S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Hong Kong.

ABSTRACT

In this paper, a novel unified framework for atlas-based segmentation is proposed, consisting of two main components: forward deformation and label refinement. A newly designed distance constraint on mesh edges is enforced with contrast sensitivity in forward deformation based on Markov random field. With the edge distance constraint, the object shapes in the atlas and the target images can remain similar during deformation. Considering the shape variations caused by individual difference, we then develop a label refinement process embracing patch registration and label fusion to compensate the small variations around the structural surfaces. As the anatomical correspondences determined in forward deformation can differ from that in label refinement, the conventional one-to-one correspondence constraint can be relaxed in our framework. Experiments on two publicly available databases IBSR and LPBA40 demonstrate that our method can obtain better performance as compared with other state-of-the-art methods.

Index Terms— Segmentation, Deformation, Patch, MRF

1. INTRODUCTION

Segmentation of subcortical structures in brain magnetic resonance (MR) images plays a significant role in clinical diagnosis, surgical planning and therapeutic assessment. While manual labeling is time consuming and sensitive to inter- and intra-rater inconsistencies. A variety of atlas-based segmentation methods have been proposed to first perform nonrigid registration between the target image and an atlas template, and then propagate the labels from the atlas to the target image [1]. However, the automated segmentation of MR images remains a challenging task due to intensity inhomogeneity, bias field and similar intensity profiles among different tissues.

In the conventional methods of pair-wise registration [2, 3], only uni-directional deformation between the atlas and the target images is performed. To improve the quality of anatomical and matching correspondences between two images, inverse consistency [4] and symmetric diffeomorphism [5] methods are proposed to enforce a special constraint that

the forward deformation from the atlas to the target should be consistent with the inverse deformation. In both cases, it is assumed that there exists a one-to-one correspondence implicitly or explicitly. Given the variations among different subjects, using strict one-to-one correspondence can lead to a less satisfactory segmentation quality.

Distinct with the traditional uni- or bi-directional nonrigid registration, a novel integrated scheme composing of forward deformation and label refinement is proposed in this paper. The main contributions of our method are listed as follows. First, besides conventional intensity information, in the forward deformation, the shape constraint in terms of edge distance are encoded to keep the shapes of the atlas and the target image similar. Second, to compensate small shape variations, patch-based label fusion is conducted around the surface of the segmented target structure in the label refinement process. Third, under this new scheme, as the registration directions of forward deformation and label refinement are opposite, the anatomical correspondences established in each step can be different and the restrictive one-to-one correspondence constraint can be relaxed.

2. FORWARD DEFORMATION

In this section, under the framework of Markov random field, we present the modeling of the forward deformation process, which incorporates contrast sensitivity and shape constraint in the form of mesh edge distance.

2.1. Background

Markov random field (MRF) is an undirected graph in which the attribute of one node is influenced by its connected nodes. Its general form is given as follows [6],

$$E = \sum_{p \in P} D_p(l_p) + \sum_{p \in P} \sum_{q \in \mathcal{N}(p)} V_{p,q}(l_p, l_q), \quad (1)$$

where P is defined as a set of nodes, l_p is a discrete label for node p , and $\mathcal{N}(p)$ is its neighborhood system. The first unary term D_p is the sum of data cost and the second term $V_{p,q}$

represents the whole pair-wise potential in a neighborhood system $\mathcal{N}(p)$.

In this paper, meshes are employed to reconstruct the surface of the object structure in the atlas and the vertices of these meshes are used as control points P for deformation. Then $\mathcal{N}(p)$ refers to the set of adjacent vertices which are in the same triangle plane as p . We define the deformation region as $[0, \pm 1, \pm 2, \dots, \pm r]^d$, where r represents the maximum scope that one node can deform and d is the dimension of this region. Each discrete displacement vector $\vec{d}(l_q)$ in the deformation region can be represented by a unique label l_q in the discrete label space. Under the MRF framework, as such, the registration procedure of finding the optimal displacement vector for each node can turn into seeking an optimal label for each node.

Inspired by the good performance of normalized cross correlation (NCC) in nonrigid image registration [7], we utilize it to estimate the data cost term D_p in Equation (1). For one control point $p \in P$, the similarity between two patches (the patch surrounding p in the atlas and the patch centered at $p + \vec{d}(l_q)$ in the target image) is assessed with NCC and encoded to D_p after normalized to $[0, 1]$. The lower the value of D_p , the more similar between two patches.

2.2. Regularization and Edge Distance Constraint

As for the second term in Equation (1), the difference between two adjacent displacement vectors is conventionally used as a regularization term to enforce smoothness [3, 8]. The pair-wise energy for vector difference is defined as follows,

$$V_{p,q}^{\mathbf{R}}(l_p, l_q) = \frac{|\vec{d}(l_p) - \vec{d}(l_q)|}{\max_{l_p, l_q} |\vec{d}(l_p) - \vec{d}(l_q)|}, \quad (2)$$

where $\vec{d}(l_p)$ and $\vec{d}(l_q)$ are the corresponding displacement vectors of l_p and l_q respectively, as shown in Fig. 1. The numerator measures the difference between these two vectors and the denominator represents the maximum difference among all possible combinations of $\vec{d}(l_p)$ and $\vec{d}(l_q)$ inside the deformation region.

For segmenting deep structures in magnetic resonance brain images, due to the poor contrast in the structure boundaries, the intensity information alone is inadequate to provide a reliable segmentation. In this paper, we propose to incorporate the local shape information in the form of Euclidean distance e and the corresponding pair-wise energy is defined as follows,

$$V_{p,q}^{\mathbf{E}}(l_p, l_q) = \frac{|e(l_p, l_q) - e_a|}{\max_{l_p, l_q} |e(l_p, l_q) - e_a|}, \quad (3)$$

where p and q are two neighboring vertices on meshes, e_a is the edge distance between them in the atlas, $e(l_p, l_q)$ is the Euclidean distance in the target image with deformation labels l_p and l_q respectively, as illustrated in Fig. 1. Besides

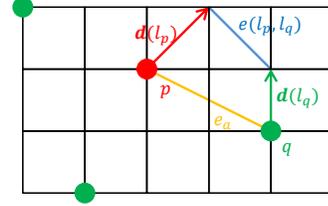


Fig. 1. One node p (red point) and its neighbors q (green points) shown on image lattice; Displacement vector $\vec{d}(l_p)$ for node p and $\vec{d}(l_q)$ for node q respectively; Euclidean edge distance between p and q at the initial stage e_a (origin line) and $e(l_p, l_q)$ (blue line) after deformation.

the constraints on the number of vertices and the connectivity of edges in topology preservation, $V_{p,q}^{\mathbf{E}}$ also enforces the distance conservation of edges during deformation. With this term, the shape information on mesh edge distance is maintained before and after deformations.

Although the regularization term $V_{p,q}^{\mathbf{R}}$ can enforce the similarity between two adjacent displacement vectors, after several iterations, the distance between two nodes can still be different from that at the initial stage. By contrast, the edge distance constraint term $V_{p,q}^{\mathbf{E}}$ helps maintain the distance between two nodes and is not significantly affected by the number of iterations. This new term is essential to our iterative MRF strategy which will be described later.

2.3. Contrast Sensitivity

As shown in Fig. 2, the image contrast condition varies in brain MR images. For the high contrast region, an accurate segmentation can be obtained with the intensity prior. However, for the low contrast region, we need to rely more on the shape prior to find a reliable segmentation result. To effectively enforce the edge distance constraint with different regional contrast conditions, we use the Root Mean Square (RMS) to assess the regional contrast,

$$\alpha_p = \sqrt{\frac{1}{xyz} \sum_{i=1}^x \sum_{j=1}^y \sum_{k=1}^z (I_{ijk} - \bar{I})^2}, \quad (4)$$

where I_{ijk} is the intensity value of a pixel located at (i, j, k) within a $x \times y \times z$ region centered at node p and \bar{I} is the mean intensity value of this region. The RMS-based contrast is first normalized to $[0, 1]$ and then combined with the data cost term. Fusing contrast sensitive data cost term, the edge distance constraint and the regularization term, Equation (1) can be rewritten as,

$$E = \sum_{p \in P} \alpha_p * D_p(l_p) + \frac{1}{|\mathcal{N}|} \sum_{p \in P} \sum_{q \in \mathcal{N}(p)} (V_{p,q}^{\mathbf{E}}(l_p, l_q) + V_{p,q}^{\mathbf{R}}(l_p, l_q)). \quad (5)$$

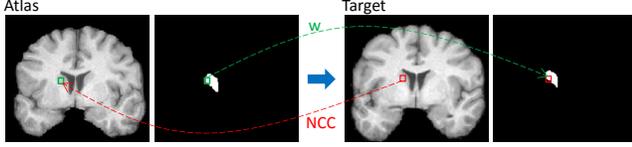


Fig. 2. Flowchart of forward deformation (blue arrow) and label refinement. Red arrow and green arrow: patch registration and label assignment in the label refinement process respectively; Red square and green square: a surrounding patch of a control point in the target image and the most similar patch found in the atlas.

As $\alpha_p * D_p$, $V_{p,q}^E$ and $V_{p,q}^R$ have been normalized between 0 and 1, $|\mathcal{N}|$, which represents the average number of p 's neighbors, is introduced as a normalization constant.

To minimize the energy function in Equation (5), TRW-S [9] is chosen as the optimizer, since it does not have any constraint on the MRF formulation and can guarantee low energy bound during optimization. However, TRW-S will become slow when the size of deformation region (the number of displacement labels) is large. To speed up the optimization procedure, an iterative strategy is exploited in our approach, by taking a small deformation region into consideration each iteration. By combining TRW-S and the iterative strategy, we can find one optimal displacement label for each node in a large deformation field efficiently.

3. LABEL REFINEMENT

According to the clustering results in shape analysis [10], it can be observed that although the shapes of the same structure among different subjects are relatively similar, there still exists small differences. With the forward deformation as described in Section 2, a similar shape as in the atlas can be obtained for the target image. Considering the small shape variations among different subjects, the propagated label maps based on forward deformation are still not accurate enough as a quality segmentation. In this paper, an additional step is introduced to refine the estimated labels based on the concept of patches and patch registration.

Under the assumption that similar appearance is induced from similar anatomical structure, we perform label refinement based on the matching of patch appearances. In Fig. 2, an intensity image and its corresponding label map of one subcortical structure are shown for atlas and target respectively. With the forward deformation (blue arrow), we can get a similar shape as in the atlas for the target image. In the label refinement process, for each control point $p \in P$ in the target image, a patch covering p (red square) is taken into consideration and the most similar patch centered at point $p + \vec{d}(l_p^*)$ (green square) can be obtained by minimizing $D_p(l_p)$.

Based on the correspondence determined in patch registration, the corresponding labels are assigned from the la-

bel map of the atlas to the target image (green arrow) with weights. The definition of weight $w(p, l_p^*)$ is given as follows,

$$w(p, l_p^*) = 1 - D_p(l_p^*). \quad (6)$$

It can be inferred that the voting weight in Equation (6) is proportional to the similarity between two patches. During label assignment, instead of propagating the label of one point from the atlas to the target image, for each pair of corresponding patches, we assign a set of labels for one center point and its nearest neighbors from the atlas to the target image,

$$\forall p \in P, \vec{L}(p) = \text{sgn}(w_i(p, l_p^*) \vec{L}(p + \vec{d}(l_p^*))), \quad (7)$$

where $\vec{L}(p)$ refers to the labels for point p and its 6 nearest neighbors in 3D image, with the values of +1 or -1 standing for that pixel is inside or outside structure respectively.

It is worth noting that this proposed label refinement method differentiates from the label consistency or diffeomorphism methods [4, 5]. Instead of enforcing the constraint that the deformation from the atlas to the target should be consistent with the inverse deformation, our method allows the forward deformation and patch registration in label refinement to be different. Moreover, this kind of label inconsistency can benefit the segmentation result, by relaxing the restrictive constraint of one-to-one correspondence. Based on the similar shape acquired in the forward deformation, the label refinement can compensate the slight shape variations by refining the estimated labels around the surface.

In medical image analysis, it is common that quality manually labeled atlases may be limited while unlabeled target images can be plenty. As such, our method is further extended to collect useful information from unlabeled target images based on group-wise registration [11]. With one single atlas provided and multiple target images to be segmented, other than treating each target image independently, the single atlas and other segmented target images are regarded as new atlases to help with the label refinement process. For each control point $p \in P$ in the target image, the new atlas i ($i = 1, \dots, n$) can provide a set of refined labels $\vec{L}(p + \vec{d}(l_p^{i*}))$ and Equation (7) needs to be rewritten as,

$$\forall p \in P, L(p) = \text{sgn}\left(\sum_{i=1}^n w_i(p, l_p^{i*}) \vec{L}(p + \vec{d}(l_p^{i*}))\right). \quad (8)$$

In [12], a semi-supervised segmentation method is proposed, in which the segmented target images are also used as atlases to assist further segmentation. Distinct with this semi-supervised method, we perform label refinement around the structural surface rather than conducting another round of nonrigid registrations between the new atlases and the target image.

4. EXPERIMENTS

In the experiments, the proposed unified framework has been evaluated on two publicly available MR brain image data sets

– IBSR¹ and LPBA40², which has 18 and 40 subjects respectively. As implied in [13], several modes can exist in a large population and one atlas should be selected for each mode. In the experiments, we first divided each data set into subgroups with the Affinity Propagation (AP) clustering [14] based on mutual information and selected the center image of each cluster as the atlas. Considering the intensity inconsistency among input images, we then conducted histogram matching in each subgroup with the Insight Toolkit³ and all images were registered to the center image based on the affine transformation provided in FSL toolbox⁴.

In the evaluations, our method has been compared with three state-of-the-art methods: SyN [5], FNIRT and Dense MRF [8]. According to the published evaluation of 14 non-rigid registration methods [1], SyN based on symmetric diffeomorphism is selected as one of the best methods. FNIRT based on a linear combination of cubic B-splines, is a convenient and widely used nonrigid registration method provided in FSL toolbox. As our method is formulated under the MRF framework, we also compared with the Dense MRF, which applies dense image registration with FastFD as the optimizer.

The settings of parameters for the compared methods followed that in [1, 8]. As for the parameter settings of our method, the size of the *NCC* and *RMS* region was $7 \times 7 \times 7$. Since iterative strategy was exploited in the optimization process, the deformation region for each iteration was small and set to $[0, \pm 1]^3$ and the number of iterations was 5. Due to the fact that we used triangular meshes to reconstruct the surface of each subcortical structure, the average number of neighbors for each node $|\mathcal{N}|$ was set to 6.

Dice Coefficient (DC) was used to assess the accuracy of segmentation, defined as $DC = \frac{2|A \cap B|}{|A| + |B|}$, where *A* and *B* are two regions of a specific structure in two images and $0 \leq DC \leq 1$. As we focus on segmenting deep structures in brain MR images, quantitative segmentation results of subcortical structures measured with DC on two data sets are listed in Table 1 and Table 2 (highest value written in red). Each subcortical structure was divided into left/right and the Dice Coefficient was calculated respectively. As shown in Table 1, the results on IBSR indicate that for most of the subcortical structures, our method can obtain the highest segmentation accuracy. The DC values listed in Table 2 reveal that the proposed method performs consistently better than the other three methods on LPBA40.

In the method evaluations, all methods were run on a 3.30 GHz, Dual-Core CPU with a 20 GB RAM. Different from the compared methods which perform image registration and label propagation for the whole image volume, our method conducts segmentation for one particular structure each time. The average running time for our method to segment one sub-

Table 1. Segmentation results on IBSR data set.

Method Structure	FNIRT	SyN	Dense MRF	Our Method
Thalamus	0.86-0.85	0.85-0.85	0.86-0.86	0.87-0.87
Caudate	0.77-0.76	0.78-0.76	0.81-0.79	0.81-0.80
Putamen	0.80-0.80	0.83-0.83	0.85-0.85	0.86-0.87
Pallidum	0.73-0.73	0.75-0.76	0.76-0.77	0.78-0.79
Hippocampus	0.69-0.70	0.71- 0.74	0.72-0.74	0.72-0.74
Amygdala	0.65-0.65	0.63- 0.66	0.65-0.63	0.65-0.63
Average	0.75±0.03	0.76±0.04	0.77±0.02	0.78±0.02

Table 2. Segmentation results on LPBA40 data set.

Method Structure	FNIRT	SyN	Dense MRF	Our Method
Putamen	0.77-0.79	0.76-0.76	0.78-0.79	0.81-0.82
Caudate	0.69-0.70	0.71-0.72	0.74-0.75	0.77-0.76
Hippocampus	0.75-0.75	0.74-0.72	0.76-0.77	0.77-0.78
Average	0.74±0.03	0.73±0.03	0.77±0.04	0.79±0.03

cortical structure is around 3 minutes. As for the compared methods, the average running time to label one target image is around 2 minutes for Dense MRF, 6 minutes for FNIRT and 32 minutes for SyN. The experimental results reveal that our proposed framework can segment the particular subcortical structure efficiently.

5. CONCLUSION

In this paper, a novel unified atlas-based segmentation framework is presented, including forward deformation and label refinement. Under the framework of Markov random field, the shape information captured in the form of edge Euclidean distance and contrast sensitivity term are encoded in the formulation of forward deformation. We further perform the label refinement process to update the labels around the surface of the segmented target structure and to compensate the small shape variations. Experiments on IBSR and LPBA40 have been carried out and results demonstrate that our method can outperform other state-of-the-art methods in terms of accuracy and efficiency.

6. ACKNOWLEDGMENT

This work was supported in part by the K. S. Lo Foundation and the Research Grants Council of Hong Kong General Research Fund under Grant 16203115.

¹<https://www.nitrc.org/projects/ibsr>

²<http://www.loni.ucla.edu/Atlases/LPBA40>

³<http://www.itk.org/>

⁴<http://fsl.fmrib.ox.ac.uk/fsl>

7. REFERENCES

- [1] Arno Klein, Jesper Andersson, et al., “Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration,” *NeuroImage*, vol. 46, no. 3, pp. 786–802, 2009.
- [2] Dinggang Shen and Christos Davatzikos, “Hammer: hierarchical attribute matching mechanism for elastic registration,” *IEEE TMI*, vol. 21, no. 11, pp. 1421–1439, 2002.
- [3] Aristeidis Sotiras, Nikos Komodakis, et al., “Graphical models and deformable diffeomorphic population registration using global and local metrics,” in *MICCAI*, vol. 5761, pp. 672–679. Springer, 2009.
- [4] Gary E Christensen and Hans J Johnson, “Consistent image registration,” *IEEE TMI*, vol. 20, no. 7, pp. 568–582, 2001.
- [5] Brian B Avants, Charles L Epstein, et al., “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain,” *MedIA*, vol. 12, no. 1, pp. 26–41, 2008.
- [6] Yuri Boykov, Olga Veksler, and Ramin Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE PAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [7] Brian B Avants, Nicholas J Tustison, et al., “A reproducible evaluation of ants similarity metric performance in brain image registration,” *NeuroImage*, vol. 54, no. 3, pp. 2033–44, 2011.
- [8] Ben Glocker, Nikos Komodakis, et al., “Dense image registration through mrfs and efficient linear programming,” *MedIA*, vol. 12, no. 6, pp. 731–741, 2008.
- [9] Vladimir Kolmogorov, “Convergent tree-reweighted message passing for energy minimization,” *IEEE PAMI*, vol. 28, no. 10, pp. 1568–1583, 2006.
- [10] E. Konukoglu, B. Glocker, et al., “Wesd-weighted spectral distance for measuring shape dissimilarity,” *IEEE PAMI*, vol. 35, no. 9, pp. 2284–2297, 2012.
- [11] Serdar K Balci, Polina Golland, et al., “Free-form b-spline deformation model for groupwise registration,” in *MICCAI Workshop on SR*. 2007, pp. 23–30, Springer.
- [12] Tobias Gass, Gábor Székely, and Orcun Goksel, “Semi-supervised segmentation using multiple segmentation hypotheses from a single atlas,” in *MICCAI Workshop on MCV*, vol. 7766, pp. 29–37. Springer, 2013.
- [13] Mert R Sabuncu, Serdar K Balci, et al., “Discovering modes of an image population through mixture modeling,” in *MICCAI*, vol. 5242, pp. 381–389. Springer, 2008.
- [14] Brendan J Frey and Delbert Dueck, “Clustering by passing messages between data points,” *Science*, vol. 315, no. 5814, pp. 972–976, 2007.