

Label Inference with Registration and Patch Priors

Siqi Bao and 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, we present a novel label inference method that integrates registration and patch priors, and serves as a remedy for labelling errors around structural boundaries. With the initial label map provided by nonrigid registration methods, its corresponding signed distance function can be estimated and used to evaluate the segmentation confidence. The pixels with less confident labels are selected as candidate nodes to be refined and those with relatively confident results are settled as seeds. The affinity between seeds and candidate nodes, which consists of regular image lattice connections, registration prior based on signed distance and patch prior from the warped atlas, is encoded to guide the label inference procedure. For method evaluation, experiments have been carried out on two publicly available data sets and it only takes several seconds for our method to improve the segmentation quality significantly.

1 Introduction

Due to poor contrast condition and intensity inhomogeneity in brain magnetic resonance (MR) images, it is challenging to provide a reliable segmentation result. Manual labelling is tedious and time-consuming, which also suffers from inter- and intra-labeler variability [13]. Various automatic labelling methods have been proposed and atlas-based segmentation approaches become widely used owing to the relatively high accuracy. With manually labelled atlas, the label map for the target image can be propagated from the atlas based on nonrigid registration [9]. To obtain a reasonable deformation field, smoothness or regularization term is conventionally enforced during registration [6,14]. However, because of the anatomical variability among subjects, the enforcement of regularization can lead to labelling errors near the object surfaces.

Patch-based method is first introduced for image denoising [4] and later employed in medical image segmentation. In [11], with the similarity values calculated from a kernel function as weights, the small patches inside a region weighted vote for the labels of the target image. In this process, no nonrigid registration is required and all information provided by the small patches will be utilized. To reduce the adverse impact from dissimilar patches, an extension is proposed in [3], by first ranking these small patches based on structure similarity and then combining the selected ones together for the final labelling.

As pointed out in [15], patch-based methods can fail to provide accurate labels close to boundaries and one label inference approach based on Gaussian processes has been proposed by estimating contour-driven distribution over label maps. However, as a result of poor contrast and weak boundaries, the contour is difficult to extract in brain MR images. In this paper, we present a novel label inference method integrated with registration and patch priors, to help correct the label errors around structural boundaries. Experiments on two public data sets indicate that the proposed label refinement method (label inference) can improve the segmentation quality significantly and efficiently.

2 Methodology

In this paper, label inference is formulated on an undirected weighted graph, with nodes selected automatically from the target image based on confidence evaluation. Besides the typical image lattice connections, registration prior based on signed distance and patch prior from the warped atlas image are encoded to assist the refinement procedure. Under the framework of Random Walker, label inference can be viewed and solved as the discrete Dirichlet problem.

2.1 Confidence Evaluation

In Fig. 1, the segmentation result for the left hippocampus provided by one benchmark registration method ANTs is shown in blue. As compared with the ground truth displayed in red, it can be observed that label errors mainly lie around the boundary of the subcortical structure. In other words, the segmentation results for pixels or voxels close to the structural border have low confidence level as compared with those far away from the perimeter. In this paper, a new label inference method is proposed to improve the atlas-based extraction of one subcortical structure in brain MR images. Signed Distance Function (SDF) [12] is employed as shape representation and utilized to help evaluate the segmentation confidence of the warped label map. The SDF for one binary image can be constructed by calculating the Euclidean distance between one pixel and its nearest pixel on the object boundary. Each pixel in the image has its corresponding signed distance and negative or positive value indicates inside (foreground) or outside (background) the object respectively.

For one pixel located around the boundary, the absolute value of its signed distance approaches 0 and its confidence of the current label is relatively low. In our approach, we assume that the segmentation confidence of one pixel is proportional to the absolute value of its signed distance. Given one positive value ρ , the pixels with signed distance $-\rho < d < \rho$ are regarded as candidate nodes, whose labels have low confidence level and need to be refined. The labels for the pixels with signed distance $d \geq \rho$ or $d \leq -\rho$ are regarded as confident results and these pixels can be viewed as foreground or background seeds.

We use x_i to represent the probability that one node belongs to the foreground ($x_F = 1$ for the foreground seed and $x_B = 0$ for the background seed). By

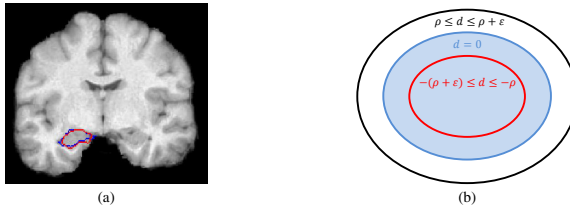


Fig. 1. (a) Comparison of segmentation results from ANTs (Blue) and manually labelled ground truth (Red). (b) Blue curve: boundary of the object with signed distance $d = 0$; Red curve: the layer inside the object with signed distance $-(\rho + \epsilon) \leq d \leq -\rho$; Black curve: the layer outside the object with signed distance $\rho \leq d \leq \rho + \epsilon$.

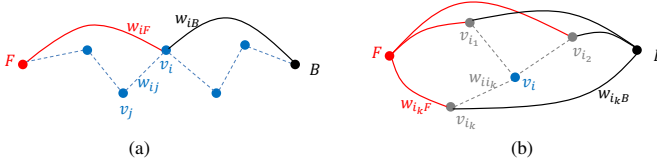


Fig. 2. (Colour Image) (a) Illustration of encoding lattice connections (blue dashed lines) and registration prior (red and black lines). (b) Illustration of encoding patch prior from the atlas (gray dashed lines) and rater reliability (red and black lines). Virtual nodes v_{ik} ($k = 1, \dots, K$) are associated with candidate node v_i .

encoding the relationship between candidate nodes and seeds into one graph $G = (V, E)$, we try to seek the optimal solution for x_i and improve the current segmentation quality. V refers to the set of nodes, which consists of nodes selected from the target image (foreground seeds V_F , background seeds V_B and candidate nodes V_C) and virtual nodes V_V generated from the warped atlas. The foreground seeds V_F are obtained by selecting pixels with signed distance $-(\rho + \epsilon) \leq d \leq -\rho$, and the background seeds V_B are chosen from the pixels with signed distance $\rho \leq d \leq \rho + \epsilon$, as shown in Fig. 1(b). To balance the influence from foreground and background seeds on candidate nodes, we randomly select background seeds to make the numbers of two kinds of seeds equal. As for the virtual nodes V_V , they act as mediators between candidate nodes and seeds, which will be illustrated in detail later. $E \subseteq V \times V$ refers to the set of edges connecting two nodes i and j , with w_{ij} as edge weight.

2.2 Registration Prior

With the foreground and background seeds settled, we can propagate the influence of seeds to candidate nodes through image lattice. For the candidate node $v_i \in V_C$, we consider the affinity between v_i and its neighbors, with the edge weight w_{ij} assigned with the common Gaussian function,

$$\forall v_j \in \mathcal{N}(v_i), \quad w_{ij} = \exp(-\beta_1(I_T(v_i) - I_T(v_j))^2), \tag{1}$$

where $\mathcal{N}(v_i)$ refers to the 6-nearest neighbors of v_i in 3-D images, $I_T(\cdot)$ is the intensity value for the pixel in the target image and β_1 is one tuning parameter. In Fig. 2(a), the blue dashed lines refer to the edges connecting candidate nodes (blue dots) and some of their 6-connected lattice neighbors, with edge weights which are defined in Equation (1).

Besides the regular lattice connections, registration prior is also exploited to guide the label inference procedure according to the signed distances of candidate nodes. For one node with a smaller signed distance, more impacts are supposed to be received from the foreground seed and vice versa. To involve the registration prior in the graph, direct edges between the candidate node and seeds are appended, as the red and black lines shown in Fig. 2(a). The red dot is one foreground seed ($F \in V_F$) and the black dot is one background seed ($B \in V_B$). We define w_{iF} and w_{iB} using the sigmoid function,

$$w_{iF} = \frac{1}{1 + e^{d_i}}, \quad w_{iB} = \frac{1}{1 + e^{-d_i}}, \quad (2)$$

in which d_i is the signed distance for candidate node v_i . With the definitions above, $0 < w_{iF} < 1$, $0 < w_{iB} < 1$, $w_{iF} + w_{iB} = 1$ can be inferred. To some extent, w_{iF} can represent the probability implied by registration prior that v_i belongs to the foreground, similarly for w_{iB} . For the candidate nodes located at the boundary ($d_i = 0$), the impacts of registration prior from the foreground and background seeds are equal and then other prior knowledge can dominate the label inference of these nodes, in such way the boundary errors can be suppressed. In our approach, the SDF can help choose candidate nodes to be refined and assist the utilization of the registration prior accordingly.

2.3 Patch Prior

Based on the deformation field determined in the registration procedure, the label map of the atlas can be deformed to provide the segmentation result for the target image. At the same time, the intensity profile of the atlas can also be warped based on the deformation field. Due to regularization or smoothness term utilized during registration, the warped intensity profile of the atlas differs with that of the target image. As such, it is reasonable to see that the warped label map is inaccurate for the target image at some pixels, especially around the structural boundaries. In this paper, we propose to use patch prior from the warped atlas to assist the label inference procedure of the target image and to relax the constraint brought by regularization during registration.

For one candidate node v_i in the target image, the intensity information around it, in terms of patch, $\mathcal{P}_1(v_i)$ is taken into consideration. Similar patches for $\mathcal{P}_1(v_i)$ are searched inside a larger patch $\mathcal{P}_2(v_i)$ in the warped atlas image. In 3-D MR image segmentation, the sizes of \mathcal{P}_1 and \mathcal{P}_2 are $(2r_1 + 1)^3$ and $(2r_2 + 1)^3$ respectively, where $r_2 > r_1 > 0$. The similarity between two patches is defined as follows,

$$\forall \mathcal{P}_1(v_{i_k}) \subset \mathcal{P}_2(v_i), \quad S(v_i, v_{i_k}) = \sum_{x \in \mathcal{P}_1(v_i), y \in \mathcal{P}_1(v_{i_k})} (I_T(x) - I_A(y))^2, \quad (3)$$

where I_T and I_A are the intensity profiles of the target image and warped atlas respectively, v_{i_k} is one pixel in the warped atlas such that patch $\mathcal{P}_1(v_{i_k})$ is inside the larger patch $\mathcal{P}_2(v_i)$. After performing sorting on the similarity values, we can pick up K patches $\mathcal{P}_1(v_{i_k})$ ($k = 1, \dots, K, K \geq 2$) which are similar with $\mathcal{P}_1(v_i)$, to decrease the adverse impact from dissimilar patches. The central pixels v_{i_k} of these similar patches are added as virtual nodes to the graph, connected with corresponding candidate node v_i , as shown in Fig. 2(b). The edge weight w_{ii_k} between candidate node v_i and one of its virtual nodes v_{i_k} is defined based on the patch similarity,

$$\forall v_{i_k} \in V_V(v_i), \quad w_{ii_k} = \frac{1}{Z(v_i)} S(v_i, v_{i_k})^{-\beta_2}, \tag{4}$$

where $V_V(v_i)$ is the set of virtual nodes associated with v_i . As experimentally demonstrated in [1], the above weight gives better performance as compared with other related similarity metrics. β_2 is one tuning parameter and $Z(v_i)$ is a normalization constant to guarantee $\sum_{k=1}^K w_{ii_k} = 1$.

To deal with the manual labelling inconsistency [13] for the prior provided by the warped atlas, we introduce one term α to encode this kind of uncertainty. The edge weights between the virtual point and the seeds are given as follows,

$$w_{i_k F} = \alpha L(v_{i_k}), \quad w_{i_k B} = 1 - \alpha L(v_{i_k}), \tag{5}$$

where α is one tuning parameter related to the rater reliability, $L(\cdot)$ is the label of a node, with 1 or 0 standing for foreground or background respectively.

2.4 Label Inference

The basic energy function for target image segmentation [10] is given as follows,

$$\begin{aligned} E(x) &= E_{unary}(x) + E_{binary}(x), \\ &= \sum_{v_i} (w_{iF}^q |x_i - 1|^p + w_{iB}^q |x_i - 0|^p) + \sum_{e_{ij}} w_{ij}^q |x_i - x_j|^p, \end{aligned} \tag{6}$$

where the first unary term measures the data cost of each node independently and the second term is the pair-wise potential through edges. By assigning various values to p and q , the general energy function can be deformed to different models, like Graph Cut, Random Walker, Power Watershed, etc. Considering shrunk surface emerged in Graph Cut and rough boundary acquired by Power Watershed, we herein choose to model label inference under the framework of Random Walker, with p and q set to 2. By combining the two graphs shown in Fig. 2, the objective function of the label inference problem can be written as,

$$\begin{aligned} \min_x \quad & \sum_{v_i} [w_{iF}^2 (x_i - 1)^2 + w_{iB}^2 x_i^2] + \sum_{e_{ij}} w_{ij}^2 (x_i - x_j)^2 + \\ & \sum_{v_{i_k}} [w_{i_k F}^2 (x_{i_k} - 1)^2 + w_{i_k B}^2 x_{i_k}^2] + \sum_{e_{ii_k}} w_{ii_k}^2 (x_i - x_{i_k})^2, \\ \text{s.t.} \quad & x_F = 1, \quad x_B = 0. \end{aligned} \tag{7}$$

The first term involves the registration prior based on signed distance and the second term is the nodal affinity through 6-connected image lattice. As for the third and fourth terms, the patch prior from the warped atlas is introduced as virtual points, setting up intermediate connections between candidate nodes and seeds. According to Equation (7), label inference is a discrete Dirichlet problem and can be solved as the Laplace equation with Dirichlet conditions using Graph Analysis Toolbox [7]. With the unique solution obtained for x_i , the labels of candidate nodes can be updated, $L(v_i) = 1$ if $x_i \geq \frac{1}{2}$ and $L(v_i) = 0$ otherwise.

It is worth noting that the proposed label refinement method is not limited to single-atlas based segmentation and it can be extended to multi-atlas based segmentation with minor changes in the following two steps. Since multiple atlases can produce different label maps, majority voting is first carried out to obtain the initial segmentation result for the target image. As for the selection of virtual nodes, we will take the patch similarity values from all atlases into consideration and select K similar patches after ranking.

3 Experiments

The performance of our label inference (LI) method has been evaluated on two publicly available data sets – IBSR (www.cma.mgh.harvard.edu/ibsr) and LPBA40 (www.loni.ucla.edu/Atlases/LPBA40). IBSR consists of 18 T1-weighted MR brain images with 84 labeled structures and LPBA40 includes 40 subjects with 56 structures delineated. Given various modes existed in a large population [2], we first divided each data set into subgroups based on Affinity Propagation (AP) clustering [5] using Mutual Information and selected the central image of each cluster as the atlas. For the rest of the images in each subgroup, histogram matching with Insight Toolkit (www.itk.org) and affine transformation using FLIRT [8] were carried out as pre-processing.

In the experiments, our LI method was applied to the initial label maps provided by two benchmark methods – ANTs and IRTK, which rank high in the evaluation of 14 nonrigid registration methods [9]. During implementation, the intensity distance $(I_T(v_i) - I_T(v_j))^2$ in w_{ij} has been normalized to $[0, 1]$ and the parameter settings used in the evaluation are listed as follows, $\beta_1 = 5$, $\beta_2 = 0.5$, $\rho = 2$, $\varepsilon = 1$, $\alpha = 0.9$ and $K = 2$.

Dice Coefficient (DC) was employed to assess the labelling accuracy on subcortical structures and each structure was divided into left/right, with DC values calculated respectively. Quantitative segmentation results measured with DC on two data sets are listed in Table 1 and Table 2, with the highest value written in bold. It can be observed that with the assistance of LI, the segmentation accuracy of ANTs and IRTK can be improved considerably (1.8% and 2.9% on IBSR, 3.0% and 2.9% on LPBA40).

To provide a comprehensive evaluation, we tested the performance of our extended label inference (ELI) method in multi-atlas based segmentation. The parameter settings were remained unchanged except for K , which was set to the number of atlases. To fuse the label maps generated by multiple atlases,

Table 1. Label Inference on IBSR data set

| Methods | Thalamus | Caudate | Putamen | Pallidum | Hippocampus | Amygdala | Average |
|---------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| ANTs | 0.870-0.868 | 0.822-0.794 | 0.851-0.856 | 0.780-0.790 | 0.741-0.756 | 0.677-0.660 | 0.789±0.028 |
| ANTs+LI | 0.884-0.883 | 0.851-0.824 | 0.864-0.878 | 0.789-0.805 | 0.769-0.779 | 0.684-0.672 | 0.807±0.026 |
| IRTK | 0.855-0.840 | 0.792-0.759 | 0.786-0.763 | 0.684-0.666 | 0.687-0.678 | 0.609-0.573 | 0.724±0.106 |
| IRTK+LI | 0.873-0.863 | 0.825-0.780 | 0.811-0.793 | 0.717-0.696 | 0.722-0.715 | 0.638-0.599 | 0.753±0.097 |

Table 2. Label Inference on LPBA40 data set

| Methods | Putamen | Caudate | Hippocampus | Average |
|---------|--------------------|--------------------|--------------------|--------------------|
| ANTs | 0.794-0.803 | 0.756-0.751 | 0.779-0.782 | 0.777±0.040 |
| ANTs+LI | 0.824-0.832 | 0.790-0.785 | 0.805-0.807 | 0.807±0.040 |
| IRTK | 0.777-0.778 | 0.766-0.763 | 0.773-0.761 | 0.770±0.042 |
| IRTK+LI | 0.811-0.810 | 0.794-0.790 | 0.799-0.788 | 0.799±0.044 |

Table 3. Extended Label Inference on IBSR data set

| Methods | Thalamus | Caudate | Putamen | Pallidum | Hippocampus | Amygdala | Average |
|----------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| ANTs+MV | 0.894-0.896 | 0.839-0.826 | 0.884-0.888 | 0.825-0.826 | 0.794-0.807 | 0.742-0.719 | 0.828±0.023 |
| ANTs+WV | 0.894-0.899 | 0.850-0.837 | 0.887-0.889 | 0.819-0.821 | 0.793-0.805 | 0.737-0.709 | 0.828±0.016 |
| ANTs+ELI | 0.906-0.910 | 0.870-0.859 | 0.893-0.901 | 0.831-0.839 | 0.814-0.825 | 0.757-0.738 | 0.845±0.017 |
| IRTK+MV | 0.873-0.868 | 0.820-0.802 | 0.847-0.831 | 0.767-0.752 | 0.753-0.750 | 0.702-0.662 | 0.786±0.053 |
| IRTK+WV | 0.874-0.870 | 0.820-0.805 | 0.854-0.839 | 0.768-0.756 | 0.756-0.758 | 0.699-0.663 | 0.788±0.044 |
| IRTK+ELI | 0.892-0.887 | 0.857-0.840 | 0.873-0.863 | 0.794-0.781 | 0.789-0.787 | 0.730-0.694 | 0.816±0.039 |

Table 4. Extended Label Inference on LPBA40 data set

| Methods | Putamen | Caudate | Hippocampus | Average |
|----------|--------------------|--------------------|--------------------|--------------------|
| ANTs+MV | 0.862-0.858 | 0.828-0.821 | 0.841-0.835 | 0.841±0.028 |
| ANTs+WV | 0.861-0.857 | 0.834-0.828 | 0.840-0.836 | 0.843±0.029 |
| ANTs+ELI | 0.867-0.863 | 0.845-0.841 | 0.847-0.842 | 0.851±0.025 |
| IRTK+MV | 0.845-0.838 | 0.832-0.825 | 0.827-0.814 | 0.830±0.035 |
| IRTK+WV | 0.844-0.837 | 0.833-0.828 | 0.830-0.817 | 0.832±0.035 |
| IRTK+ELI | 0.859-0.852 | 0.845-0.841 | 0.840-0.828 | 0.844±0.030 |

we utilized classic majority voting (MV), with the results of intensity weighted voting (WV) provided for reference. As shown in Table 3 and Table 4, our ELI method can obtain better DC values consistently. All experiments were run on a 3.30 GHz, Dual-Core CPU with 20 GB RAM. Given the average computation time consumed by pair-wise deformation, 32 minutes for ANTs and 22 minutes for IRTK, it only takes around 9 seconds to finish the label inference procedure, which demonstrates that the proposed method can improve the segmentation quality efficiently.

4 Conclusion

In atlas-based image segmentation, the quality of labels around structural boundaries is usually poor. To deal with this problem, in this paper, we propose a novel method to refine these labels. By employing signed distance function to evaluate the initial label map, we can pick up the nodes whose labels need to be refined and select nodes with confident results as seeds. Registration prior based on signed distance and patch prior generated from the warped atlas are encoded in the label inference procedure, together with the nodal affinity through lattice in the target image. Experimental results on two public data sets indicate that the proposed method can improve the labelling quality effectively and efficiently.

References

1. Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de Solorzano, C.: Combination strategies in multi-atlas image segmentation: Application to brain mr data. *IEEE TMI* 28(8), 1266–1277 (2009)
2. Blezek, D.J., Miller, J.V.: Atlas stratification. *MedIA* 11(5), 443–457 (2007)
3. Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. *NeuroImage* 54(2), 940–954 (2011)
4. Coupé, P., Yger, P., Prima, S., et al.: An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. *IEEE TMI* 27(4), 425–441 (2008)
5. Frey, B.J., Dueck, D.: Clustering by passing messages between data points. *Science* 315(5814), 972–976 (2007)
6. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., et al.: Dense image registration through mrfs and efficient linear programming. *MedIA* 12(6), 731–741 (2008)
7. Grady, L., Schwartz, E.: The graph analysis toolbox: Image processing on arbitrary graphs. *CAS/CNS Technical Report Series* (021) (2003)
8. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17(2), 825–841 (2002)
9. Klein, A., Andersson, J., et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *NeuroImage* 46(3), 786–802 (2009)
10. Lézoray, O., Grady, L.: *Image Processing and Analysis with Graphs: Theory and Practice*. CRC Press (2012)
11. Rousseau, F., Habas, P.A., Studholme, C.: A supervised patch-based approach for human brain labeling. *IEEE TMI* 30(10), 1852–1862 (2011)
12. Sethian, J.A.: *Level set methods and fast marching methods*. CUP (1999)
13. Shattuck, D.W., Mirza, M., Adisetiyo, V., et al.: Construction of a 3d probabilistic atlas of human cortical structures. *NeuroImage* 39(3), 1064–1080 (2008)
14. Sotiras, A., Komodakis, N., Glocker, B., Deux, J.-F., Paragios, N.: Graphical models and deformable diffeomorphic population registration using global and local metrics. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I*. LNCS, vol. 5761, pp. 672–679. Springer, Heidelberg (2009)
15. Wachinger, C., Sharp, G.C., Golland, P.: Contour-driven regression for label inference in atlas-based segmentation. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013, Part III*. LNCS, vol. 8151, pp. 211–218. Springer, Heidelberg (2013)