Interactive Registration and Segmentation for Multi-Atlas-Based Labeling of Brain MR Image

Qian Wang[1], Guorong Wu[2], Min-Jeong Kim[2], Lichi Zhang[1], and Dinggang Shen[2](✉)

[1] Med-X Research Institute, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
{wang.qian,lichizhang}@sjtu.edu.cn

[2] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
{grwu,minjeong_kim,dinggang_shen}@med.unc.edu

Abstract. In the conventional multi-atlas-based labeling methods, atlases are registered with each unlabeled image, which is then segmented by fusing the labels of all registered atlases. The registration is typically ignorant about the segmentation while the segmentation of each individual unlabeled image is independently considered, both of which potentially undermine the accuracy in labeling. In this work, we propose the interactive registration-segmentation scheme for multi-atlas-based labeling of brain MR images. First, we learn the distribution of all images (including atlases and unlabeled images) and register them to their common space in the groupwise manner. Then, we segment all unlabeled images simultaneously, by fusing the labels of the registered atlases in the common space as well as the tentative segmentation of the unlabeled images. Next, the (tentative) labeling feeds back to refine the registration, thus all images are more accurately aligned within the common space. The improved registration further boosts the accuracy to determine the segmentation of the unlabeled images. According to our experimental results, the iterative optimization to the interactive registration-segmentation scheme can improve the performances of the multi-atlas-based labeling significantly.

1 Introduction

It is needed by many studies that certain medical images should be labeled into different anatomical regions-of-interest (ROIs), in order to facilitate the following region-based analysis. Manual labeling, though probably accurate with well-trained experts, costs high especially for the large-scale population of images. On the contrary, automatic labeling method shows the advantage in reducing the needs of human interactions. Therefore, the technique is highly desirable in medical image analysis and has been intensively investigated recently.

To *label* (also known as to *segment* or to *parcellate*) brain MR images can be attained in many different ways. Among them, atlas-based segmentation provides

H. Zha et al. (Eds.): CCCV 2015, Part I, CCIS 546, pp. 240–248, 2015.
DOI: 10.1007/978-3-662-48558-3_24

an efficient solution that yields comparable accuracy with respect to manual labeling. Specifically, after *registering* the atlas to an unlabeled image, the labeling information associated with the atlas then propagates to the unlabeled image and further *segments* it. If image registration is accurate, the well-established correspondences between the two images can guarantee that the propagated segmentation is highly reliable.

Instead of using a single atlas, the multi-atlas-based labeling strategy is more preferred in recent studies to alleviate the potential errors in registering a single atlas. To this end, the unlabeled image can be parcellated by fusing contributions from several registered atlases. For example, after all atlases are spatially normalized with the unlabeled image via registration, the majority voting scheme determines the to-be-estimated label at a specific voxel of the unlabeled image as the most frequent label from the same voxel locations of all registered atlases. In this way, the labeling accuracy is usually higher since the potential errors in registering certain atlases with the unlabeled image become less influential.

A typical scenario for multi-atlas-based labeling is to segment a population of images, in which only a few atlases are pre-labeled. Both *registration* and *segmentation*, the two major components in multi-atlas-based labeling, can thus be specially adapted for better performances. Though the labeling can propagate by registering atlases with each unlabeled image directly, more advanced registration schemes are beneficial for the sake of accurate labeling. In [12] , for instance, the population of images is embedded into a high-dimensional manifold where similar images are closely distributed. Higher labeling accuracy can be achieved, since the labeling is always propagated between similar images and the related registration is more reliable.

After (roughly) registering atlases with a certain unlabeled image, the segmentation of the unlabeled image can be determined in various ways [1,5,8,11]. Most methods in the literature apply the mono-directional label fusion by propagating the labeling from the atlases to the unlabeled image only. However, recent studies [8,4] show that the segmentation of an unlabeled image can also benefit from other unlabeled images. Specifically, a certain unlabeled image can finally be segmented from fusing both the labels of atlases and the tentative segmentation of other unlabeled images. In this way, not only the consistency across the segmentation of each unlabeled image and the atlases, but also the intrinsic consistency among all unlabeled images, are well preserved.

Though the multi-atlas-based labeling can split into *registration* followed by *segmentation*, the two components are usually independently considered. The interaction between them, which could improve the labeling accuracy, is mostly ignored. To this end, we propose a novel multi-atlas-based labeling method, which applies the *interactive registration-segmentation scheme* and thus significantly differs from other methods in the literature. The proposed method iteratively refines the registration of all images and then the segmentation of the unlabeled images. In particular, the groupwise registration [9,10,13] learns the distribution of the entire image population that consists of all atlases as well as unlabeled images, and deforms all images to a common space. We then segment the unlabeled images simultaneously in the common space, by fusing both the labeling of the registered atlases and the tentative segmentation of the unlabeled images. The (tentative) labeling of all images feeds back to further improve the registration accuracy in two folds: (1) the labeling helps better understand the distribution of the entire image population after all images are roughly aligned in the common space; (2) the labeling is regarded as additional image descriptors and can be used to guide the registration of all images directly. In this way, the segmentation contributes to register all images more accurately in the common space, which in turn leads to better estimation of the segmentation.

2 Method

We propose to solve the multi-atlas-based labeling problem via interactive registration and segmentation. Specifically, we register the image population, including atlases and unlabeled images, to a common space in the groupwise manner (Section 2.1). Then, the segmentation of each unlabeled image is derived from fusing not only the labeling of the registered atlases, but also the tentative segmentation of other unlabeled images (Section 2.2). Further, the tentative segmentation guides to register all images more accurately in the common space (Section 2.3). The proposed methodology will be summarized in Section 2.4.

2.1 Image Registration via Minimum Spanning Tree

We register all atlases and unlabeled images to their common space in the groupwise manner, by taking advantage of the distribution of the entire image population. In particular, all images are first embedded into a fully connected graph, where the nodes indicate individual images and the edge linking each pair of images records their in-between distance, i.e., the sum of square differences (SSD) of intensities. A minimum spanning tree (MST) is then extracted from the graph. The root of the tree is determined to represent the geometric median image in the population, from which the sum of distances to other images is the minimal. All images are connected with the root of the tree either directly or via other images/nodes. More detailed explanations on the construction of the MST can be found in [2,3,4]. Note that the node at the root of the tree, or the median image of the population, will act as the common space to which all images in the population are registered.

The learned MST helps register all images in the population to the root in a recursive manner. In particular, given each non-root image, its path that traverses along edges to the root of the tree can be easily identified. If the parent node of the image under consideration is the root, the direct registration (i.e., via diffeomorphic Demons [6]) will be computed immediately. Otherwise, the non-root image will utilize the deformation belonging to its parent node as an initialization and further refine to generate its own deformation towards the root. The recursive callbacks can eventually deform all images to the common space. Compared with the direct registration of two images that might be very different in anatomies, the MST provides robust initialization in estimating the deformation field.

Due to the essentially high-dimensional image data and the limited size of the image population, the estimation of the MST might be inaccurate. Here, we build the tree from an augmented population that consists of more simulated images. The simulated images are derived by perturbing the pre-determined median image in five steps [4]: (1) A set of images is directly registered with the median image; (2) All deformations are then inverted; (3) Principal component analysis (PCA) is applied to capture the variation within all inverted deformations; (4) By perturbing coefficients in the learned PCA model, a set of deformations can be simulated; (5) All simulated deformations are applied to warp the median image and generate a set of simulated images in the final. We follow the same setting in [4] to specify the number of the simulated images to be twice the size of the original image population. The augmented population, including atlases, unlabeled images, and simulated images, leads to the MST that better captures the distribution of the image population.

2.2 Segmentation via Label Fusion

After all images are registered to the common space, we are able to segment the unlabeled images given the atlases. We apply the local voting strategy for stochastic label fusion. Denoting the m-th $(m = 1, \cdots, M)$ registered atlas as I_m and L_m as its label, the label for the n-th $(n = 1, \cdots, N)$ unlabeled image $\mathcal{I}_n$ at x, or $\mathcal{L}_n(x)$, can be assigned with the label l at the following likelihood

$$p(\mathcal{L}_n(x) = l) = \sum_{m=1}^{M} w(I_m, \mathcal{I}_n, x)\delta(L_m(x), l). \tag{1}$$

In the equation above, $\delta(L_m(x), l)$ returns 1 if and only if $L_m(x) = l$; otherwise 0. The weight $w(I_m, \mathcal{I}_n, x)$ indicates the contribution of I_m to label $\mathcal{I}_n$ by L_m, and obviously relates to the similarity between I_m and $\mathcal{I}_n$ at x. By using $d(I_m, \mathcal{I}_n, x)$ to denote the distance of the two respective intensity patches centered at x of both I_m and $\mathcal{I}_n$ (with the size $3 \times 3 \times 3$ in voxel), we define $w(I_m, \mathcal{I}_n, x) = \exp(-d^2(I_m, \mathcal{I}_n, x)/2\sigma^2)$ as σ relates to the standard deviation of all patch-to-patch distances. The exact label of $\mathcal{L}_n(x)$ is determined as the value l of the maximal likelihood in the final.

To segment a certain unlabeled image consistently with the entire population, the tentative labeling of other unlabeled images should also participate into the local voting. Therefore, the likelihood in labeling $\mathcal{I}_n(x)$ can be calculated by

$$p(\mathcal{L}_n(x) = l) := \sum_{m=1}^{M} w(I_m, \mathcal{I}_n, x)\delta(L_m(x), l) + \sum_{k=1}^{N} w(\mathcal{I}_k, \mathcal{I}_n, x)\delta(\mathcal{L}_k(x), l). \tag{2}$$

Stable solution to the above can be iteratively attained [4]. Eq.2 implies that the label $\mathcal{L}_n(x)$ complies with both the atlases and other unlabeled images. Also note that the simulated images, though participating into registration (Section 2.1), are not included in label fusion. All simulated images are instantiated by perturbing the median image, which would arbitrarily dominate the segmentation result with the simulated images included in label fusion.

2.3 Interactive Registration and Segmentation

Though the unlabeled images can be segmented in Section 2.2, the tentative segmentation is derived from the yet imperfect registration as in Section 2.1. On the contrary, the tentative segmentation is capable of feeding back for more accurate registration of all images in the common space, which can further improve the performance in segmentation. In our method, the registration benefits from the tentative segmentation in two folds:

1. The initial MST is estimated prior to the non-rigid registration. The high variation among all images, as well as the simple image distance measure (i.e., SSD of intensities), may lead to improperly estimated MST and thus limit the registration accuracy. On the other hand, after all images are roughly registered to the common space, the distribution of the entire image population is relatively compact and can be better learned by considering the consistency of the segmentation of all images. That is, the MST can be updated by considering the tentative segmentation.

2. The registration should also favor the consistency within the segmentation of individual images, which is only pursuit in the segmentation part of conventional methods though. In particular, we regard the tentative segmentation of images as additional image descriptors other than intensities. Besides to minimize the intensity inhomogeneity, the registration aims to directly eliminate the labeling inconsistency as well. In particular, we require the registration to align the boundaries of corresponding labels of individual images. The estimated deformation fields are then applied to register all images more accurately in the common space.

Update MST. To learn the MST for representing the image distribution, we propose to measure the image-to-image distance by the inconsistency between their segmentation, after all images are (roughly) registered to the common space. For any two (atlas or unlabeled) images I_m and I_n, we calculate their overall Dice overlap ratio upon all labels

$$r(I_m, I_n) = (2 \sum_x \delta(L_m(x), L_n(x)))/(||I_m|| + ||I_n||), \tag{3}$$

where $||\cdot||$ computes the size of the labeled volume. The distance of the two images is then derived by $\exp(-r^2(I_m, I_n)/(2\beta^2))$, as β is manually specified.

Given pairwise distances of all images, we are then able to build a new MST. Note that the root of the updated MST is still kept as the median image that is previously selected in the initial MST (Section 2.1). In this way, the common space in registration is fixed, though each non-root image will further refine its own deformation field. Moreover, the updated MST consists of only atlases and unlabeled images, while the simulated images are not incorporated. We argue that, after the initial registration in Section 2.1, the atlases and the unlabeled images distribute tightly in the nearby of the median image. Thus we do not need the simulated images to help update the tree.

Update Registration. We directly apply the (tentative) segmentation of all images to refine the deformation fields, in order to compensate for the inconsistency within the labeling. After deforming all images towards the median image, we first extract the boundaries of all labels for every warped image (i.e., by applying the Canny edge detector on the labeling map). The boundary voxels for a certain image then form a discrete pointset, which should be aligned with the label boundaries of other images. Next, we apply a Gaussian kernel to smooth the detected boundaries and convert the discrete boundary pointset into a continuous volume of Gaussian mixture [7] in the image space. Finally, the volumes of Gaussian mixtures for the labeling of a pair of images can be easily registered, i.e., via the diffeomorphic Demons [6]. The newly updated MST is also applied in the above, as the registration upon the label boundaries is performed in the recursive manner (c.f. in Section 2.1).

Note that all images further refine their registration to the common space after being warped following their previously estimated deformation fields. Therefore, we concatenate the previous deformation field of each image and its new deformation for refinement into a single field, which warps the image from its original space to the common space directly. To compensate for potential errors in the above, the concatenated deformation functions as an initialization, and is further adjusted by minimizing the intensity inhomogeneity between the specific image and the median image designating the common space. In particular, the registration adjustment is also achieved through the diffeomorphic Demons [6], yet with the high-resolution optimization only and very limited number of iterations. In this way, both the image intensity and the tentative segmentation contribute to update the registration.

2.4 Summary: The Interactive Registration-Segmentation Pipeline

We summarize the proposed method as follows:

1. Estimate the MST to organize all images in the population;
2. Register all images to the root of the tree and deform them to the common space;
3. Segment all unlabeled images in the common space via label fusion;
4. Go to Step 1 if not converged, continue otherwise;
5. Pull the segmentation result back to the space of each unlabeled image.

The solution above is in the iterative fashion. We impose a fixed number (i.e., 4) of iterations to be the convergence criterion, as the observed improvement upon the labeling accuracy becomes tiny after 4 iterations in our experiment. Then, after inverting the estimated deformations, the segmentation of each unlabeled image can be warped to the original image space in Step 5. The simple pullback of the segmentation might be contaminated by the errors in deforming the labels. Therefore, for each unlabeled image in Step 5, we register all atlases and other unlabeled images back to its original image space, and apply the label fusion (Section 2.2) for the final determination of the segmentation. The registration above can be efficiently solved, since the deformations between all images and the common space are known already. Moreover, note that our method is reduced to be MABMIS [4], if only a single iteration is allowed.

3 Experimental Results

We apply the proposed method to the NIREP dataset and compare with MABMIS [4] in order to demonstrate the importance of the interaction between registration and segmentation. The NIREP dataset consists of 16 images, each of which comes with 32 labeled ROIs. All images are resampled to the isotropic size of 256 × 256 × 256 and properly pre-processed (including bias correction, skull-stripping, etc.).

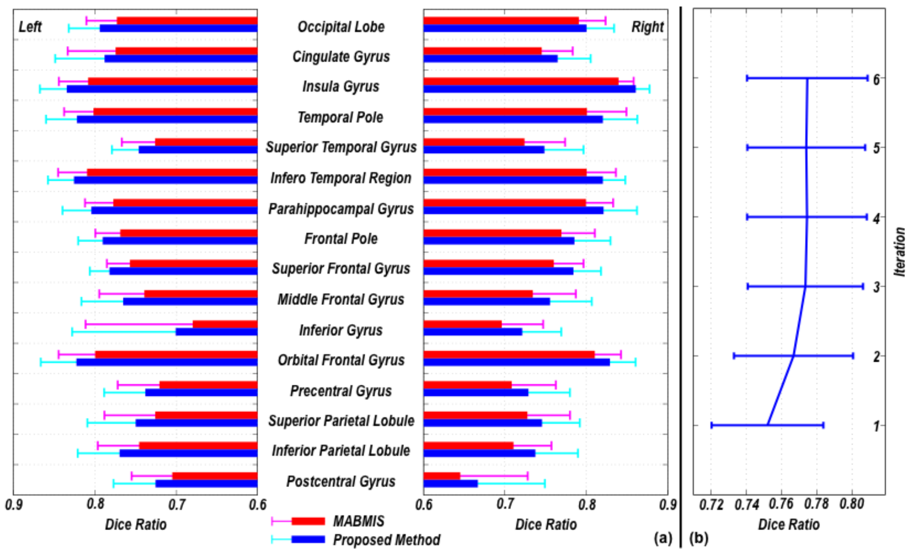

Fig. 1. (a) The average Dice ratios, as well as standard deviations, of 32 ROIs in the NIREP dataset yielded by MABMIS and the proposed method; (b) The iteration changes of the overall Dice ratio produced by our method.

We randomly partition all images into two equally sized subsets. By taking a certain subset of images as atlases in turn, we are able to label the other subset. The accuracy of the estimated segmentation is then evaluated against the ground truth (i.e., the manual segmentation). The average Dice ratios on all 32 labels, as well as the standard deviations, are plotted in Fig. 1(a). In particular, our method scores the overall Dice ratio at $77.45 \pm 3.39\%$ upon all labels, higher (+2.14%) than $75.31 \pm 3.17\%$ of MABMIS. The results imply that the interactive registration-segmentation scheme lead to improved accuracy in multi-atlas-based labeling.

We also plot the iterative changes of the overall Dice ratio, as well as the standard deviations, in Fig. 1(b). Clearly the interaction between registration and segmentation boosts the labeling accuracy within limited number of iterations. By allowing 4 iterations in our experiment, it typically costs around 4 hours to label 8 images given another 8 atlases (single thread, Intel Core i5 CPU, 3.1GHz, 8G memory).

4 Discussion

In this work, we propose a novel multi-atlas-based labeling method for brain MR images, by utilizing the interactive registration-segmentation scheme. Different from most conventional methods, we allow the (tentative) segmentation of previously registered images to feed back, which results in better registration of all images and thus more accurate labeling as confirmed by the experimental result. Compared to [4], our method costs higher computation time in iterative optimization. However, significant improvement upon the speed performance is expected if introducing parallelization into our implementation. Moreover, large-scale study will also be conducted in the future to evaluate the performance of our method more comprehensively.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (NSFC) Grants (61473190, 61401271, 81471733).

References

1. 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)
2. Hamm, J., Ye, D.H., Verma, R., Davatzikos, C.: Gram: A framework for geodesic registration on anatomical manifolds. Medical Image Analysis **14**(5), 633–642 (2010)
3. Jia, H., Wu, G., Wang, Q., Kim, M., Shen, D.: iTree: Fast and accurate image registration based on the combinative and incremental tree. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1243–1246. IEEE (2011)
4. Jia, H., Yap, P.T., Shen, D.: Iterative multi-atlas-based multi-image segmentation with tree-based registration. Neuroimage **59**(1), 422–430 (2012)
5. 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)
6. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage **45**(1), S61–S72 (2009)

7. Wang, F., Vemuri, B.C., Rangarajan, A., Eisenschenk, S.J.: Simultaneous nonrigid registration of multiple point sets and atlas construction. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(11), 2011–2022 (2008)
8. Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(3), 611–623 (2013)
9. Wang, Q., Chen, L., Yap, P.T., Wu, G., Shen, D.: Groupwise registration based on hierarchical image clustering and atlas synthesis. Human Brain Mapping **31**(8), 1128–1140 (2010)
10. Wang, Q., Wu, G., Yap, P.T., Shen, D.: Attribute vector guided groupwise registration. NeuroImage **50**(4), 1485–1496 (2010)
11. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging **23**(7), 903–921 (2004)
12. Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative: Leap: learning embeddings for atlas propagation. NeuroImage **49**(2), 1316–1325 (2010)
13. Ying, S., Wu, G., Wang, Q., Shen, D.: Hierarchical unbiased graph shrinkage (hugs): A novel groupwise registration for large data set. NeuroImage **84**, 626–638 (2014)