Multi-modal Brain Image Registration Based on Subset Definition and Manifold-to-Manifold Distance

Weiwei Liu, Yuru Pei^(✉), and Hongbin Zha

Peking University, Beijing China
peiyuru@cis.pku.edu.cn

Abstract. Image registration is an important procedure in multi-modal brain image processing. The main challenge is the variations of intensity distributions in different image modalities. The efficient SSD based method cannot handle this kind of variations. And other approaches based on modality independent descriptors and metrics are usually time-consuming. In this article, we propose a novel similarity metric based on manifold-to-manifold distance imposed on the subset of original images. We define a subset for a compact representation of the original image. Manifold learning technique is employed to reveal the intrinsic structure of the sampled data. Instead of comparing the images in the original feature space, we use the manifold-to-manifold distance to measure the difference. By minimizing the distance between the manifolds, we iteratively obtain the optimal registration of the original image pair. Experiment results show that our approach is effective to deal with the multi-modal image registration on the BrainWeb dataset.

Keywords: Multi-modal registration · Subset definition · Manifold-to-manifold distance

1 Introduction

Multi-modal image registration is essential to brain image processing, e.g. the image fusion, which can provide comprehensive information of multi-modal images. Conventional registration methods based on sum of squared difference (SSD) cannot be directly applied since the intensity distributions are obviously different. Some modality-independent metrics such as mutual information (MI) [15] and cross correlation (CC) [1] were proposed for the multi-modal image registration. The registration methods based on MI and CC are generally time consuming. Moreover, the modality-independent feature descriptors can be used for multi-modal image registration via the landmark-based methods. However, the calculation of modality independent descriptors can be complicated. Efforts were also made to figure out a common representation of the reference and moving images and then use L2 norm as the similarity metric to optimize the transformation. One example is the registration method built upon the Laplacian image [22]. However, the nonlinear embedding of the original images often involves a large computation cost considering the image dimensionality.

Fig. 1. The framework of our proposed registration method

We propose our registration framework based on the subset definition and the manifold-to-manifold distance as illustrated in Fig. 1. A subset is defined for the compact representation of the original image. We calculate the low-dimensional manifolds of the extracted subsets and use the manifold-to-manifold distance to estimate the image difference. First, we define a subset via edge detection on the reference image. And the coordinates of the sampled points are transferred to the moving image. Image patches centered on these points from both moving and reference images are used to describe the context information. Second, we use the dimensionality reduction technique to deal with the sampled subset. Considering the complicated structures of medical images, we use Laplacian eigenmaps [2] for the low-dimensional embedding with preservation of locality. Instead of measuring the similarity in the original feature space, we estimate the image difference via the distance between manifolds of the subsets. The registration is optimized iteratively by minimizing the manifold-to-manifold distance.

2 Related Work

Multi-modal image registration plays an important role in the field of brain image processing. It has been studied for years and many approaches have been proposed to solve the problem. One type of solution is to sample a dense point cloud on the surface

of the brain and boundaries between different tissues, and then the alignment between point clouds is estimated with algorithms like ICP [5] and robust point matching (RPM) [16]. A similar but more efficient approach is to focus on some specific positions in the brain, so that the landmarks are sparsely distributed. The landmarks can be defined manually. Han et al. [6] use the regression forest to learn the positions of manually marked points. The extraction of landmarks can also be automatically by analyzing the geometry of the images. Commonly used features in brain images are ridges or crest lines [20]. Some semantic information can effectively improve the performance of the registration method. For instance, the prior of brain segmentation is very helpful information in the task of registration [11, 18]. Features extracted from images under different modalities are often not precisely matched due to the differences in image characteristics. To cope with this limitation, Lee et al. [9] propose a feature alignment method based on Gaussian-weighted distance map by spreading the weights of feature points to their neighborhood. An alternative to feature-based method is regarded as intensity-based method. Instead of extracting feature points, this type of methods uses directly the intensity values of the images. Considering the intensity variations of multi-modal images, the difference of intensity values is not an effective metric to describe the relationship between images to be registered. Therefore, some modality-independent similarity metrics are proposed so that the registration framework can adapt with the demand of multi-modal images. One typical example is the mutual information (MI) [15]. And some efforts have been made to improve its performance and robustness, such as normalized mutual information (NMI) [19] and conditional mutual information (CMI) [14]. Recently, methods that combine geometric and intensity features are proposed to make the intensity based registration more robust. Sulcal information can be used to constraint the registration of brain images [4, 8]. The cortical surface is considered in the processing of brain volume data [13]. The method of MI is also extended by combining the gradient field calculated in the images [12].

Further related work includes some registration approaches based on internal similarity and manifold learning. Instead of exploring the relationship of the original images, some preprocessing steps are involved to make the images comparable. The already mentioned work of Penney et al. [17] considers the internal similarity in local areas to obtain modality-independent descriptors. The similar idea is also used by Heinrich et al. [7] in their work of modality independent neighborhood descriptor (MIND). Vectors that consist of differences between each pixel and its neighbors are organized as the descriptors of local structures. A second group of techniques seek to figure out a common representation for multi-modal images so that mono-modal registration based on L1 and L2 metrics can be applied. Wachinger et al. [21] use entropy image to represent the structures of input images. Locations with obvious anatomic structures have high entropy, while the entropy values of smooth areas are low. Finally, we want to refer to the method of Laplacian image [22]. Manifold learning technique is applied on image patches sampled across the whole image and a new intensity value is assigned to each pixel according to the corresponding coordinate in the low-dimensional space. Given the number of pixels in an image, this processing can be very time consuming.

3 Subset Definition

Due to the high dimensionality of the image data, we define a subset of the original image for a compact representation. In this article, we use the Canny operator [3] to extract points on edges from the brain image. We observe that the extracted points almost cover all the contours of sulcus and gyrus as well as the boundaries between white and grey matters.

The edge points are still densely distributed on the image. We further randomly downsample the detected points so that the final subset contains no more than 1000 samples. Image patches around these points are used as the subset to represent the original image.

We only use the Canny edge detector on the reference image and transfer the coordinates of the sampled points to the moving image. The transferred coordinates on the moving image actually deviate from the real boundaries. We use image patches around the sampled points to depict their context.

4 Metric Based on Manifold-to-Manifold Distance

Our subset definition is still a modality-related representation since the intensity values are directly used. We employ Laplacian eigenmaps [2] to get low-dimensional manifolds of the defined subsets to reveal their intrinsic structure and eliminate the extrinsic structure. The intrinsic structure describes the internal similarity of patches within each image and provides modality-independent information.

Let $\mathcal{M}^r, \mathcal{M}^m$ denote the low-dimensional manifolds of the subsets defined in the reference and moving images respectively. Laplacian eigenmaps starts with the construction of a neighborhood graph that describes the structure of the data. An image patch can be represented with a high-dimensional vector, where each entry corresponds to the intensity value of a pixel. We denote the data point corresponding to the i^{th} image patch as p_i , and its counterpart in the low dimensional space as q_i . The Euclidean norm $\|p_i - p_j\|$ is used to estimate the distance between two patches in the high-dimensional space. This distance is only defined when the patches are close to each other. In this work, we choose the implementation that searches for the k nearest neighbors and connects the corresponding nodes with weights assigned to the edges in the form of heat kernel $W_{ij} = e^{-\|p_i - p_j\|^2 / (2\sigma^2)}$, where σ^2 indicates the variance. Subsequently, the low-dimensional manifold with locality best preserved on average is obtained by minimizing $\sum_{i,j} W_{ij} \|p_i - p_j\|^2$.

In the neighborhood graph, edges only connect nodes that are close to each other, and the number of neighbors should not be too large. But when the value of k is too small, the neighborhood graph may be separated into several isolated subgraph. To avoid the emergence of undefined relationship between any pair of data points, we set $k = 100$ so that there is only one connected subgraph. And this selection does not severely contradict with the assumption of locality.

The dimension of the low-dimensional space is set to 2. The mapping results of Laplacian eigenmaps only guarantee that the local structures are preserved. We need to put the low-dimensional manifolds of the subsets into a mutual space before estimating the distance between the manifolds. Here, we figure out an affine transformation by minimizing the distance

$$\|T \cdot \mathcal{M}^m - \mathcal{M}^r\|^2. \quad (1)$$

where T indicates the transformation imposed on the manifold of the moving image.

Subsequently, the distance between two manifolds is formulated as

$$d(\mathcal{M}^m, \mathcal{M}^r) = \frac{1}{2} \|\hat{T} \cdot \mathcal{M}^m - \mathcal{M}^r\|^2, \quad (2)$$

where $\hat{T}$ is the optimal transformation that minimize the distance in Eq. 1. This manifold-to-manifold distance is used as the image distance measurement in our registration approach.

5 Image Registration

The registration is calculated in an iterative way. In every iteration, the transformation parameters are updated to lower the difference between the reference and moving images measured with the manifold-to-manifold distance. The updated transformation is performed on the moving image. We use the Laplacian eigenmaps to calculate the low-dimensional manifold of the moving image. And then, the distance between the manifolds of the subsets defined on the reference and moving images is calculated. The iteration ends when the manifold-to-manifold distance reaches the convergence. Based on our proposed similarity metric, the problem of image registration can be converted to the minimization of the manifold-to-manifold distance defined in Eq. 2.

$$\hat{T}_I = \operatorname{argmin}_{T_I} d(\mathcal{M}^m(T_I), \mathcal{M}^r). \quad (3)$$

We use T_I to denote the transformation of the original image. $\mathcal{M}^m(T_I)$ indicates the low dimensional manifold of the moving image when transformation T_I is imposed on it.

We consider the rigid transformation on 2D images with 3 registration parameters to be optimized: the rotation r , the translations on x and y directions t_x, t_y , respectively. The registration is optimized with BFGS [10]. The gradient of the objective function can be written as

$$\frac{\partial d}{\partial T_I} = \left(\frac{\partial d}{\partial r}, \frac{\partial d}{\partial t_x}, \frac{\partial d}{\partial t_y} \right)^T, \quad (4)$$

where d is short for $d(\mathcal{M}^m(T_I), \mathcal{M}^r)$. The partial derivatives are estimated as follows,

$$\frac{\partial d}{\partial r} = \left\| \hat{T} \cdot \mathcal{M}^m - \mathcal{M}^r \right\| \frac{\partial \left(\hat{T} \cdot \mathcal{M}^m \right)}{\partial r}. \quad (5)$$

$$\frac{\partial d}{\partial r} = \left\| \hat{T} \cdot \mathcal{M}^m - \mathcal{M}^r \right\| \frac{\partial \left(\hat{T} \cdot \mathcal{M}^m \right)}{\partial t_x} \quad (6)$$

$$\frac{\partial d}{\partial r} = \left\| \hat{T} \cdot \mathcal{M}^m - \mathcal{M}^r \right\| \frac{\partial \left(\hat{T} \cdot \mathcal{M}^m \right)}{\partial t_y} \quad (7)$$

6 Experiments

We apply our method on the brain MRI images provided by BrainWeb database¹, which includes modalities of T1, T2 and PD. The images have been adjusted and they are well aligned. In order to demonstrate the effectiveness of our proposed similarity metric, we show plots of the similarity between slices selected from the T1- and T2-weighted images for rotation and translations in Fig. 2. Three different slices are tested. It can be seen that the image difference measured by the manifold-to-manifold distance approaches the minimum when two images are well aligned.

Fig. 2. Plots of the similarity respect to rotation (a) and translations in x and y directions (b)

¹ <http://www.bic.mni.mcgill.ca/brainweb/>.

We also test our method in registration experiments. The transformation parameters are calculated for all combinations of image modalities. We first make a random rigid transformation on one of the images, with rotations in the range of $\pm 15^\circ$ and translations ± 15 mm. The calculated parameters are compared with them and the errors are shown in Table 1. For each configuration in Table 1, the registration is calculated 100 times with random starting positions. The registration methods proposed in [21] based on the entropy image and Laplacian image as well as the approach based on MI [15] are compared in this table. The experiments prove that the Laplacian eigenmaps for the defined subset can effectively depict the brain images and lead to accurate registration results.

Table 1. Registration errors for all combinations of the three MR modalities. Rotation in degree and translation in mm. The results of the registration based on the Entropy image and Laplacian image are from [21]

Modalities	Metric	Rotation	Translation x	Translation y
T1-T2	MI	3.026	4.785	5.513
	Entropy	2.084	4.539	5.231
	Laplacian	2.584	2.061	2.168
	Ours	1.937	1.596	1.882
T1-PD	MI	3.734	5.921	4.085
	Entropy	2.283	4.782	4.750
	Laplacian	1.750	3.007	1.929
	Ours	1.823	2.014	2.628
T2-PD	MI	2.206	3.042	3.554
	Entropy	1.732	4.296	3.780
	Laplacian	1.171	2.350	1.984
	Ours	1.065	2.136	1.778

In our work, image patches are used to describe the local structures around the detected edge points. The performance of registration is related to the selection of the patch size. If the image patch is too small, it cannot contain enough structural information. The patch size should be large enough so that the subsets defined in the reference and moving images have sufficient overlap, but too large patches also lead to high computational cost. The setting of the patch size is application related. In our experiments, we consider the rotation up to 15° and the translations 15 mm. We set the patch size to 31×31 , and the dimension of the image patch is 961. In Fig. 3, we illustrate the changes of similarity respect to translation in y direction with different patch sizes. When

the patch size is too small, a lot of local minima show up as the translation is obvious. Without a proper initial value, the result is likely to be trapped in the local minima.

Fig. 3. The similarity respect to translation in y direction with patch sizes of 5×5 (left), 15×15 (middle) and 31×31 (right).

7 Conclusion

In this article, we proposed a novel similarity metric for multi-modal brain images based on subset definition and manifold-to-manifold distance, and utilize it to solve the registration of brain images. We use a subset made up of patches around the sampled edge points to depict the structure of the brain images. Then, Laplacian eigenmaps is employed to calculate the low-dimensional manifold of the subset. The image difference is estimated via the manifold-to-manifold distance. Our results demonstrate that this metric provides a good measurement for the image difference and leads to effective image registration.

References

1. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **12**(1), 26–41 (2008)
2. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: *NIPS*, vol. 14, pp. 585–591 (2001)
3. Canny, J.: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 679–698 (1986)
4. Collins, D.L., Le Goualher, G., Venugopal, R., Caramanos, A., Evans, A.C.: Cortical constraints for non-linear cortical registration. In: Höhne, K.H., Kikinis, R. (eds.) *VBC 1996*. LNCS, vol. 1131, pp. 307–316. Springer, Berlin Heidelberg (1996)
5. Feldmar, J., Malandain, G., Declerck, J., Ayache, N.: Extension of the ICP algorithm to non-rigid intensity-based registration of 3D volumes. In: *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, pp. 84–93. IEEE (1996)
6. Han, D., Gao, Y., Wu, G., Yap, P.-T., Shen, D.: Robust anatomical landmark detection for MR brain image registration. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *MICCAI 2014, Part I*. LNCS, vol. 8673, pp. 186–193. Springer, Heidelberg (2014)

7. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, M., Schnabel, J.A.: MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. *Med. Image Anal.* **16**(7), 1423–1435 (2012)
8. Hellier, P., Barillot, C.: Coupling dense and landmark-based approaches for nonrigid registration. *IEEE Trans. Med. Imag.* **22**(2), 217–227 (2003)
9. Lee, H., Lee, J., Kim, N., Kim, S.J., Shin, Y.G.: Robust feature-based registration using a Gaussian-weighted distance map and brain feature points for brain PET/CT images. *Comput. Biol. Med.* **38**(9), 945–961 (2008)
10. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. *Math. Prog.* **45**(1–3), 503–528 (1989)
11. Liu, J.: Segmentation-assisted registration for brain MR images. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, pp. 335–353. Springer, New York (2011)
12. Liu, J., Tian, J.: Registration of brain MRI/PET images based on adaptive combination of intensity and gradient field mutual information Copyright © 2007 J. Liu and J. Tian. This is an open access article distributed under the Creative Commons Attribution License (2007)
13. Liu, T., Shen, D., Davatzikos, C.: Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage* **22**(4), 1790–1801 (2004)
14. Loeckx, D., Slagmolen, P., Maes, F., Vandermeulen, D., Suetens, P.: Nonrigid image registration using conditional mutual information. *IEEE Trans. Med. Imag.* **29**(1), 19–29 (2010)
15. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.* **16**(2), 187–198 (1997)
16. Papademetris, X., Jackowski, A.P., Schultz, R.T., Staib, L.H., Duncan, J.S.: Computing 3D non-rigid brain registration using extended robust point matching for composite multisubject fMRI analysis. In: Ellis, R.E., Peters, T.M. (eds.) *MICCAI 2003*. LNCS, vol. 2879, pp. 788–795. Springer, Heidelberg (2003)
17. Penney, G.P., Griffin, L.D., King, A.P., Hawkes, D.J.: A Novel Framework for Multi-modal Intensity-based Similarity Measures Based on Internal Similarity. *Medical Imaging. International Society for Optics and Photonics*, Bellingham (2008)
18. Pohl, K.M., Fisher, J., Levitt, J.J., Shenton, M.E., Kikinis, R., Grimson, W.L., Wells, W.M.: A unifying approach to registration, segmentation, and intensity correction. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3749, pp. 310–318. Springer, Heidelberg (2005)
19. Pradhan, S., Patra, D.: Nonrigid image registration of brain MR images using normalized mutual information. In: *Proceedings of the Second International Conference on Soft Computing for Problem Solving*, pp. 1069–1077. Springer, India (2014)
20. Subsol, G., Roberts, N., Doran, M., Thirion, J.P., Whitehouse, G.H.: Automatic analysis of cerebral atrophy. *Magn. Reson. Imag.* **15**(8), 917–927 (1997)
21. Wachinger, C., Navab, N.: Entropy and Laplacian images: Structural representations for multi-modal registration. *Med. Image Anal.* **16**(1), 1–17 (2012)
22. Wachinger, C., Navab, N.: Manifold learning for multi-modal image registration. In: *BMVC*, pp. 1–12 (2010)