A General Learning Framework for Non-rigid Image Registration

Guorong Wu[1], Feihu Qi[1], and Dinggang Shen[2]

[1] Depart. of Computer Science and Engineering, Shanghai Jiao Tong University, China
{grwu, fhqi}@cs.sjtu.edu.cn
[2] Section of Biomedical Image Analysis, Depart. of Radiology, University of Pennsylvania
Dinggang.Shen@uphs.upenn.edu.cn

Abstract. This paper presents a general learning framework for non-rigid registration of MR brain images. Given a set of training MR brain images, three major types of information are particularly learned, and further incorporated into a HAMMER registration algorithm for improving the performance of registration. First, the best features are learned from different types of local image descriptors for each part of brain, thereby the learned best features are consistent on the correspondence points across individual brains, but different on non-correspondence points. Moreover, the statistics of selected best features is learned from the training samples, and used to guide the feature matching during the image registration. Second, in order to avoid the local minima in the registration, the points hierarchically selected to drive image registration are determined by the learned consistency and distinctiveness of their respective best features. Third, deformation fields are adaptively represented by B-splines, with more control points placed on the regions with large shape variations across individual brains or on the regions with consistent and distinctive best features. Also, the statistics of B-splines based deformations is captured and used to regularize the brain registration. Finally, by incorporating all learned information into HAMMER registration framework, promising results are obtained on both real and simulated data.

1 Introduction

Deformable registration has been extensively investigated for decades, and various algorithms, based on either intensity or features, have been proposed [1-4]. However, almost all registration algorithms employ the same set of local image features, expecting to obtain good registration for all brain regions simultaneously. It is worth noting that some features might be adequate for registration of a certain part of brain, but not necessarily good for other parts of brain. We previously proposed to learn the best-scale geometric moment invariants (GMIs) for each brain location [5], thereby each point in the brain has its own best-scale GMIs to distinguish itself from others. After integrating these learned best-scale GMIs into the HAMMER registration algorithm [4], we obtained more accurate and consistent warping results. *However*, it requires the tissue segmentation of brain images, before using it for brain registration. Also, it only uses the same types of features, i.e., GMIs, for guiding the registration.

G.-Z. Yang et al. (Eds.): MIAR 2006, LNCS 4091, pp. 219 – 227, 2006.

In recent years, many local image descriptors have been proposed in the computer vision area, for object detection and recognition. The performances of various local image descriptors, including SIFT, steerable filters, and moment invariants, have been extensively evaluated, and it is concluded that SIFT has the best performance [6]. It is worth noting that, once different types of features, i.e., local features, global features, and spatial features, are combined and appropriately selected by Adaboosting, a better classification rate can be achieved [7]. Similarly, since the brain registration can be modeled as a correspondence detection problem and the complexity of brain anatomy varies highly across different parts of brain, different types of features should be adaptively selected, i.e., by Adaboosting, for different parts of brain, in order to significantly improve the performance of registration in all parts of the brain.

Traditionally, deformations between two brains are represented by the displacement fields on brain points, and regularized by a smoothness constraint, such as a Laplacian term. Once a number of brain images have been registered, the deformations between these brain images can be statistically captured, i.e., by PCA on each wavelet band of deformations [8], and further used as a statistical constraint to regularize the deformations estimated by a registration algorithm [8].

In this paper, we present a general learning framework for non-rigid registration of MR brain images. Three major types of information used in the HAMMER registration algorithm are right now all learned from a given set of training brain samples. First, the best features used to consistently distinguish the corresponding points across individual brains are learned and selected from different types of local image descriptors, separately for each part of brain. Also, the statistics of selected best features is further learned, and used to guide the feature matching during the image registration. Second, the hierarchical selection of brain points for image registration is based on the learned consistency and distinctiveness of the best features on each point. Third, deformation fields are efficiently represented by B-splines, with adaptive control points particularly placed more on the regions with large anatomical variations or on the regions with consistent and distinctive best features. Furthermore, the statistics of B-splines based deformations is also captured and used to regularize the brain registration. Once all necessary information is learned and appropriately incorporated into the HAMMER registration framework, we obtain the promising results on both real and simulated data.

2 Method

2.1 Learning Best Features

This subsection presents a systematic learning-based method for boosting the best features on each brain point. First, three types of local image features are briefly introduced, followed by an example that demonstrates the importance of using the best features in brain registration. Then, two criteria to boost the best features on each point are created, based on the distinctiveness and consistency of each feature. Finally, our method for selecting best features based on Adaboosting is summarized.

Best Features: For each point v in the image, different kinds of local image descriptors [6] can be applied, to extract local features from a sphere of certain radius

s. Without loss of generality, three complementary local descriptors are used in this paper, i.e., RIFT [9], SPIN [9], and local histogram [10], respectively. RIFT, standing for *Rotation Invariant Feature Transform*, is the generalization of SIFT descriptor, and can be computed efficiently without determination of the dominant orientation. SPIN is another rotation-invariant descriptor, which captures the distribution of image intensity in the neighborhood of a particular point. Local histogram is a simple and effective feature, which has been investigated for registration in [10].

All of these local descriptors capture the high-dimensional histogram-like features. In order to obtain a compact representation of these features for image registration, regional features, i.e., mean and variance, are calculated from each local descriptor. In particular, seven regional features, i.e., three for RIFT, two for both SPIN and local histogram, are used in our method. On the other hand, as we found in [5], scale is also an important factor for selecting suitable features. Thus, four scales, i.e., 4, 8, 12, and 16mm, are used; note that, for each scale, seven regional features are calculated. Therefore, there are totally 28 (4x7) features for each point, which can be represented as $G=\{g(i)|i=1...N\}$, where N=28 and three different types of features are included.

No single type of features in G can be used as a universal signature to accurately establish correspondences for each part of brain, as demonstrated in Fig. 1. Note that one type of features might be effective for some parts of brain, but not necessarily effective for others. Therefore, it is important to capture the best features for each part of brain separately. For a particular scale (i.e., 8mm), the similarity between each reference point (indicated by red crosses) in Fig. 1(a) and all points in Fig. 1(b), measured by RIFT, SPIN, and local histogram, are shown in Fig. 1(c)-(e), respectively. Dark red denotes high similarity, and deep blue denotes low similarity. It can be observed that no single type of features can reliably establish the correspondences for these two selected reference points. On the other hand, by using the best features that will be selected by the methods described next, we can successfully distinguish each reference point, as shown in Fig. 1(f). This example shows the importance of boosting the best features separately for each part of brain, in order to distinguish all parts of brain simultaneously.

Criteria: Two criteria are proposed to select the best features from G. First, for each brain point v, its best features should make itself distinctive from non-correspondence points, thus reducing the ambiguity in correspondence detection during the image registration. For example, for a sulcal root point in Fig. 2 (denoted as red dot in (a) and red crosses of its corresponding points in (b) and (d)), its best features, along with the best features on all points of the red sphere N_1, should be maximally different from those of points in the ring N_2. For convenience, we call all points in N_1 as positive samples, and all points in N_2 as negative samples. The histogram of each feature $g(i)$ in the positive sample set N_1, $h(g(i),N_1)$, and in the negative sample set N_2, $h(g(i),N_2)$, can be estimated, respectively. Then, the Jenson-Shannon (JS) divergence of these two histograms can be used to measure the distinctiveness of the feature $g(i)$, denoted as $JS(g(i))$.

Second, the best features should be consistent on the corresponding points in the different individual brains, thus facilitating the correspondence detection across individuals by just looking for the points with similar features. In other words, the best features of a reference point (i.e., a red dot in Fig. 2(a)) should be similar to those of the corresponding points in other individuals (i.e., red crosses in Fig. 2(b)-(d)).

That means, we require the histogram of the feature $g(i)$ in the positive sample set N_1, $h(g(i),N_1)$, being compact, i.e., entropy of $g(i)$, denoted as $E(g(i))$, being minimal.

Selection: By combining these two criteria, we can rank each feature $g(i)$ based on an integrated measurement, i.e., $T(g(i))=JS(g(i))-E(g(i))$, and select the one with the largest measurement as the best feature. After selecting the first best feature, we can use this feature to classify each positive or negative sample, and the misclassified samples will be increased with weights, *according to Adaboosting algorithm*, while the correctly classified samples will be decreased with weights. By updating the weight for each sample, the histograms of each remaining feature $g(i)$ in the positive and the negative sample sets, i.e., $h(g(i),N_1)$ and $h(g(i),N_2)$, should be updated accordingly. Thus, the integrated measurement, $T(g(i))$, of each remaining feature $g(i)$ need to be recalculated, and a remaining feature with the largest integrated measurement should be selected as the second best feature. By similarly repeating the above step, we can obtain an expected number of best features, i.e., 6 best features in this study.

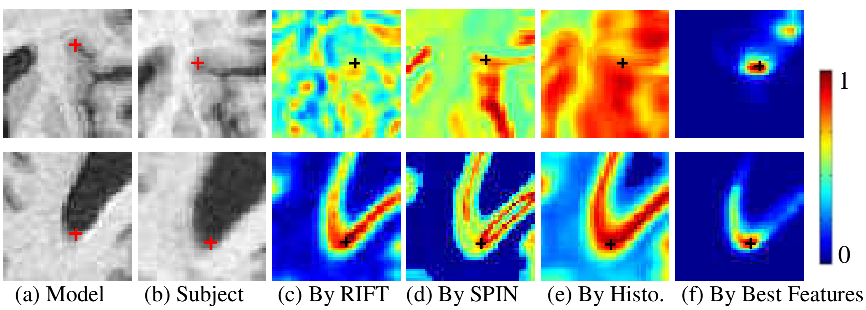


Fig. 1. The similarities between each reference point (red crosses in (a)) and all points in the subjects (b), measured by RIFT, SPIN, local histograms, and our best features, are shown in (c)-(f), respectively. It confirms the important of using the best features to distinguish each brain point. *This figure is best viewed with color.*

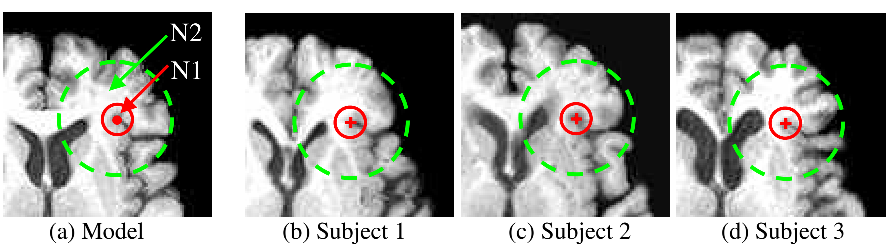


Fig. 2. A reference point (red dot) and its corresponding subject points (red crosses) are shown in (a) and (b)-(d), respectively. Our best feature selection criteria require, (1) features in N_1 and N_2 should be different, (2) features on reference point and its corresponding points should be similar.

Fig. 3 displays the best features, learned by our proposed method. In this experiment, 6 best features are selected for each brain point, from 28 features. The average scale used by 6 best features on each brain point is shown in Fig. 3(b), which indicates generally the large scales used for uniform brain regions and small scales for complex brain regions, similar to the best scales selected for GMIs [5]. In order to visually appreciate the actual features finally selected for a certain part of brain, i.e., within the red block, the selection results on all 28 features, obtained from 4 scales and 3 types, are displayed in Fig. 3(c). Red indicates the locations where the particular features are selected. It can be observed that RIFT features, which are based on edge orientation, are selected for the points around the boundary. On the other hand, local histogram based features are picked for the points in uniform regions, generally using large scales as expected. Overall, these results indicate that our learning-based method well utilizes the characteristics of each type of features.

It is worth noting that, the statistics of the selected best features for each brain point can be also estimated from all training brain images. Therefore, this information can be used to statistically measure the similarity between the best features of two points under comparison during the registration procedure, thus helping look for correct correspondences.

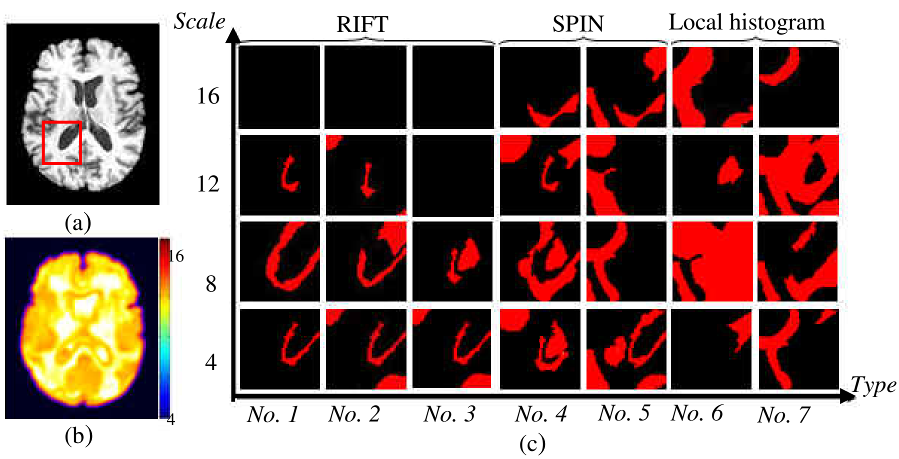


Fig. 3. The selected best features for the template brain in (a). The average scale used to calculate the best features is displayed in (b). To visually appreciate the actual features finally selected for each point in the red block of (a), the selection results on all 28 features, from 4 scales and 3 types, are displayed in (c). *This figure is best viewed with color.*

2.2 Learning to Hierarchically Select Points for Driving Image Registration

After learning the best features, each brain point can be best distinguished by its best features. However, different brain points have different distinctiveness even using their best features. For example, the root of sulcus and the crown of gyrus can be relatively reliably distinguished, and thus can be more comfortably used to drive the brain registration.

In this paper, we also design a learning strategy to determine the distinctiveness of each point v in image matching, from a set of training brain samples. Two criteria are used: (1) consistency of best features of each brain point across the correspondences of individual brains, i.e., $C(v)$; (2) distinctiveness of best features of this point from those of non-correspondences, i.e., $M(v)$. $C(v)$ and $M(v)$ can be similarly defined by the entropies as in Section 2.1. By simply combining these two measures, i.e., obtaining an overall measure $C(v)*M(v)$, we can rank points according to this overall measure, and hierarchically select brain points, called active points, to drive the brain registration. In this way, the efficiency of the registration algorithm and the robustness to local minima can be both improved.

2.3 Learning Statistics of Brain Deformations

There are two main goals in this subsection. First, B-splines [11] are adopted to efficiently represent the brain deformations. Second, the adaptive placement of control points in different parts of brain and the statistics of parameters on control points are both learned from a set of brain deformation fields, captured between a selected template and a number of training brain images. By incorporating this learned statistical representation of brain deformations into the HAMMER registration algorithm, the estimated deformation fields for new subjects can be statistically regularized.

Adaptive placement of control points: Control points in the B-splines are adaptively placed in the brain, according to the variation of brain deformations D(v) and the distinctiveness of best features $C(v)*M(v)$ on each brain point v, which can be integrated into a single measurement, i.e., $a(v)=C(v)*M(v)*D(v)$. This can be implemented in the following recursive way. First, divide the whole brain into a number of non-overlapping cubes, each with the size of 32x32x32 mm^3, and control points are placed in the corners of each cube. Second, check each cube to see whether the total measurement, i.e., the summation of $a(v)$ within this cube, is over a certain threshold; if yes, then this cube will be equally divided into 8 smaller cubes, and new control points are placed on the corners of new smaller cubes. Third, repeat the second step to create smaller and smaller cubes, until no one can be further divided or the size of cube is reaching the pre-selected minimal size.

PCA on parameters of control points: Since we only have a limited number of deformation fields, we perform PCA independently on each cube, i.e., all initial cubes with the size of 32x32x32 mm^3. This learned statistics on parameters of control points is then used to constrain and regularize the brain deformation fields during the registration procedure.

2.4 Summary of Learning-Based Registration Method

All learned information is incorporated into the framework of HAMMER registration algorithm, to improve the performance of brain registration. The learned best features are used for each template brain point to look for its correspondence in the subject brain, while the points with distinctive and consistent best features are hierarchically selected to drive the brain registration mainly. Moreover, each iteratively estimated deformation field will be statistically constrained by the learned statistics of parameters of control points in the B-splines.

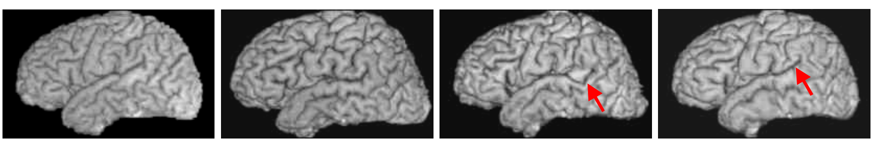
Note that all training brain images used in this paper were carefully skull-stripped and tissue-segmented, and their deformations *with respect to* the template were estimated by the HAMMER registration algorithm. More accurate registration could be achieved by first extensively labeling and landmarking a number of images, and then applying a high-dimensional warping algorithm adequately constrained by these manual labels and landmarks. In addition, the brain deformation estimated by different registration algorithms can be jointly used, thus allowing our learning-based method to integrate the merits of different registration algorithms.

3 Results

Our learning-based registration method has been evaluated on both real and simulated data, and its performance is compared with that of HAMMER algorithm. All of the following experiments are performed on 3D images, with a same set of parameters. It takes about 3 hours for registering a pair of brain images, with size of 256x256x124.

3.1 Experiments on Real Data

For real data, our learning-based registration method achieves the visual improvement in some parts of brain regions, i.e., a region indicated by the red arrows in Fig. 4 (c)



(a) Model (b) Subject (c) By HAMMER (d) By our method

Fig. 4. Visual improvement achieved by our method, particularly in the area indicated by the red arrows

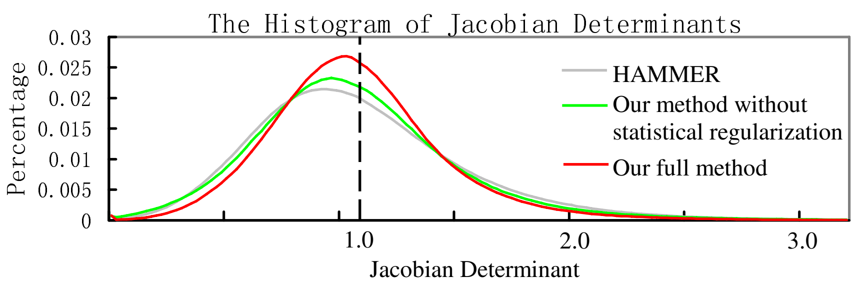


Fig. 5. Comparison of the histograms of Jacobian determinants of deformation fields, estimated by three different methods. *This figure is best viewed with color.*

and (d). In order to further show the performance of our learning-based method in statistically regularizing the B-splines based brain deformations, we compute and compare the histograms of Jacobian determinants of deformation fields, respectively estimated by HAMMER, our registration method without statistical regularization, and our full method. As shown in Fig. 5, the Jacobian determinants are more tightly distributed around the degree of 1.0 by our full method, indicating the advantage of regularizing deformation fields with prior statistical knowledge.

3.2 Experiments on Simulated Data

Simulated data [13] is used to quantitatively evaluate the performance of our registration method. For example, the average deformation estimation error can be calculated between the true deformation and the one estimated by a registration method. The average deformation error is 0.98mm by HAMMER, and 0.86mm by our method, indicating about 12% reduction in registration error achieved by our method. Note that, in our method, the testing samples are different from the training samples.

Notice that, for same simulated data, the average deformation estimation error can be decreased to 0.66mm by our previous method using the learned best-scale GMIs on the tissue-segmented images [12]. This indicates (1) the possibility of using the tissue segmentation results to improve the brain registration; (2) for direct registration of original MR brain images, a larger number of best features should be selected from more types of various features. Both of these will be our future research directions.

4 Conclusion

A general learning framework has been presented for non-rigid registration of MR brain images. Given a set of training brain images, best features for distinguishing each brain point can be learned. Also, the points used to hierarchically drive the image registration can be also learned, based on their consistency and distinctiveness. Moreover, the statistics of parameters of control points in B-splines can be further learned, and used to statistically regularize each iteratively estimated brain deformations. By incorporating all learned information into a HAMMER registration algorithm, we obtained the promising registration results on both real and simulated data. To further improve the performance of this learning-based registration method, we will learn the best features from more types of local descriptors in the future.

References

1. Y. Wang and L. H. Staib, “Elastic model-based non-rigid registration incorporating statistical shape information,” MICCAI’98, 1999. 1496: p. 1162-1173.
2. J. C. Gee, M. Reivich, and R. Bajcsy, “Elastically deforming 3D atlas to match anatomical brain images,” Journal of Computer Assisted Tomography, 1993. 17: p. 225-236.
3. G. E. Christensen and H. J. Johnson, “Consistent Image Registration”, IEEE Trans. on Med. Imaging, 2001. 20(7): p. 568-582.
4. D. Shen and C. Davatzikos, “HAMMER: Hierarchical attribute matching mechanism for elastic registration,” IEEE Trans. on Med. Imaging, 2002. 21(11): p. 1421-1439.

5. G. Wu, F. Qi, and D. Shen, "Learning Best Features for Deformable Registration of MR Brains," MICCAI'05, October, 2005. p. 179-187.
6. K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *CVPR'03*, 2003. p. 257-263.
7. W. Zhang *et al*, "Object Class Recognition Using Multiple Layer Boosting with Heterogeneous Features," CVPR'05, 2005. p. 323-330.
8. Z. Xue, D. Shen, and C. Davatzikos, "Statistical-Constrained High-Dimensional Warping Using Wavelet-Based Priors," MMBIA' 06, NY, USA, June, 2006.
9. S. Lazebnik, C. Schmid, and J. Ponce, "A Sparse Texture Representation Using Local Affine Regions," IEEE Trans. on PAMI, 2005. 27(8): p. 1265-1278.
10. D. Shen, "Image Registration by Hierarchical of Local Spatial Intensity Histograms," *MICCAI'04*, France, September, 2004.
11. D. Rueckert et al, "Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images," IEEE Trans. on Med. Imaging, 1999. 18(8): p. 712-721.
12. G. Wu, F. Qi, and D. Shen, "Improve Brain Registration Using Machine Learning Methods," ISBI'06, Arlington, USA, April 2006.
13. Z. Xue, D. Shen, B. Karacali, and C. Davatzikos, "Statistical Representation and Simulation of High-Dimensional Deformations: Application to Synthesizing Brain Deformations," MICCAI'05, Palm Springs, USA, October, 2005.