

# Learning Best Features and Deformation Statistics for Hierarchical Registration of MR Brain Images

Guorong Wu <sup>1,2</sup>, Feihu Qi <sup>1</sup>, and Dinggang Shen <sup>2\*</sup>

<sup>1</sup> Department of Computer Science and Engineering  
Shanghai Jiao Tong University, Shanghai, China 200240  
{grwu, fhqi}@sjtu.edu.cn

<sup>2</sup> Section of Biomedical Image Analysis, Department of Radiology  
University of Pennsylvania, Philadelphia, PA 19104  
Dinggang.shen@uphs.upenn.edu

**Abstract.** A fully learning-based framework has been presented for deformable registration of MR brain images. In this framework, the entire brain is first adaptively partitioned into a number of brain regions, and then the best features are learned for each of these brain regions. In order to obtain overall better performance for both of these two steps, they are integrated into a single framework and solved together by iteratively performing region partition and learning the best features for each partitioned region. In particular, the learned best features for each brain region are required to be identical, and maximally salient as well as consistent over all individual brains, thus facilitating the correspondence detection between individual brains during the registration procedure. Moreover, the importance of each brain point in registration is evaluated according to the distinctiveness and consistency of its respective best features, therefore the salient points with distinctive and consistent features can be hierarchically selected to steer the registration process and reduce the risk of being trapped in local minima. Finally, the statistics of inter-brain deformations, represented by multi-level B-Splines, is also hierarchically captured for effectively constraining the brain deformations estimated during the registration procedure. By using this proposed learning-based registration framework, more accurate and robust registration results can be achieved according to experiments on both real and simulated data.

## 1 Introduction

Deformable registration for medical images is very important for many clinical applications. So far, many registration algorithms [1-4], based on image intensities or features, have been developed. Generally, the same type of features is used for all points in the entire images, and each point is equally treated during the whole registration procedure. For regularizing the deformations estimated between two images under registration, simple smoothness constraints such as Laplacian

---

\* Corresponding author. Email: dinggang.shen@uphs.upenn.edu (Dinggang Shen).

smoothness term are typically employed. As indicated next, many previous methods are limited at several aspects.

First, the same type of image features is not always effective in distinguishing different parts of images. For example, in registering MR brain images of different subjects, the cortical regions might need very different types of features to distinguish themselves, compared to the uniform white matter regions. Also, the distinctiveness of features is highly related to the size of neighborhood used for the calculation of features. As clearly demonstrated in [5, 6], the best-scale Geometric Moment Invariants (GMIs) show better performances in brain image registration, compared to GMIs of the identical scale used in the HAMMER algorithm [4]. Therefore, it is important to investigate the best scales as well as the best types of features for different brain regions, in order to best differentiate each brain region during the registration procedure.

Second, equally treating each point during the registration procedure might eventually undermine the registration performance. Actually, some brain regions, such as roots of sulci and crowns of gyri, are more reliable to be distinguished, compared to other regions. Therefore, it is important to develop methods for differentiating those reliable regions and hierarchically using them to steer the brain image registration.

Third, simple constraints on the smoothness of voxel-wise deformation fields might be not effective for registration. Actually, the statistics of deformations among different brain images can be captured from a set of training samples, by performing PCA on each band of wavelet coefficients (wavelet-PCA) of deformations [7] or performing PCA on the parameters of control points of B-Splines [8]. It has been shown that, by using this statistical constraint on deformation fields, the performance of registration algorithm can be greatly improved [7, 9]. It is worth noting that B-splines can efficiently represent the deformation fields, while wavelet-PCA can effectively capture both global and local statistical information of deformations. Therefore, it is important to integrate the advantages of both methods [7, 8], for efficiently and effectively capturing the deformation statistics.

Accordingly, a learning-based registration framework for MR brain images is presented in this paper. To overcome the first limitation, a systematic approach is proposed to adaptively learn the best scale and the best types of image features for each brain region. This is achieved by requiring the learned best features consistent across the correspondences of individual brains, while distinctive from those of neighboring points. The former actually requires the consistency of learned best features, while the latter requires the saliency of learned best features. Both consistency and saliency measures can be integrated and used as an important guide for selecting the reliable points to primarily drive the image registration especially during the initial registration stages, thus overcoming the second limitation of treating different points equally in the registration. Moreover, to integrate the advantages of both methods [7, 8] mentioned above, multi-level B-splines [10] with adaptively placed control points are employed to efficiently represent the deformation fields in multiple levels, and thus different levels of deformation statistics (i.e., both global and local) can be captured by performing separate PCAs on each level of control points.

Compared to the method presented in [11], the steps of learning best features and deformation statistics have been significantly improved. In particular, the partition of

brain region and the selection of best features are integrated into a single framework in this paper, since these two steps intimately depend on each other. This formulation potentially overcomes the limitation by *ad hoc* brain region partition method used in [11]. Moreover, multi-level B-Splines, instead of single-level B-Splines, are used to represent the deformation fields, thereby facilitating the adaptive placement of control points and capturing the deformation statistics from coarse to fine levels.

Promising results have been obtained on both real and simulated data. For real data, good registration results have been achieved in warping subject brains to the template space. On the other side, our method out-performs HAMMER in the aspect of registration consistency when independently warping longitudinal data. For simulated data, the average deformation error is decreased from 0.92 mm by HAMMER to 0.72 mm by the proposed method.

## 2 Method

### 2.1 Learning the Best Features

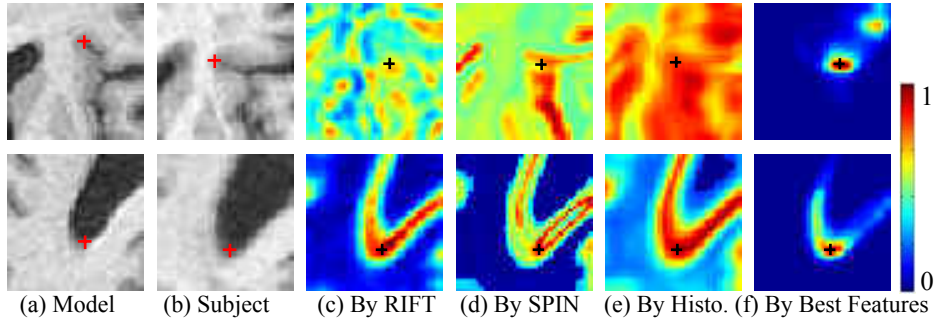
In this subsection, we will first demonstrate the importance of using best feature in point matching. Then, we present a systematical approach of adaptively learning the best scales and best types of features for different brain regions, by iteratively repeating the steps of brain region partition and best features determination, according to some heuristic criteria.

**Importance of Learning Best Features:** For each point  $v$  in an image  $I$ , different local image descriptors can be used to calculate local features from its local spherical neighborhoods with different scales. Without loss of generality, three typical image descriptors, i.e., RIFT [12], SPIN [12], and local spatial histograms [13], are employed. RIFT, standing for *Rotation Invariant Feature Transform*, is a generalization of famous SIFT descriptor [14], and it can be efficiently calculated without the determination of dominant orientation of the local patch. SPIN is another rotation invariant descriptor, which can capture the distribution of image intensities along the radial direction of the point  $v$ . Local spatial histograms capture the overall intensity distributions within neighborhoods of different scales, thus preserving spatial information simultaneously [13].

In order to obtain the compact feature representation on each point, the statistical features, i.e., mean and variance, are collected for each of these three local image descriptors. Therefore, totally six features, i.e., two for each local image descriptor, are obtained. Since scale is highly related to the distinctiveness of features, each local image descriptor is calculated from spherical neighborhoods of four scales, i.e., 4, 8, 12, and 16mm, around the point  $v$ . Thus, totally  $M=6 \times 4=24$  features, i.e.,  $G(v)=\{g_k(v)|k=1 \dots M\}$ , are obtained for each point  $v$ , and can be used as attributes for this point. Notice that  $G(v)$  includes six types of features at four different scales.

It is worth noting that no single type of features can be used as a universal signature to reliably distinguish all brain regions, as demonstrated by Fig. 1. For example, by using RIFT, SPIN, and local-histogram features calculated in a scale of

8mm, the similarities between the red-crossed points in Fig. 1(a) and all points in Fig. 1(b) can be calculated, as color-coded and shown in Figs. 1(c)-(e), respectively. Dark red denotes high similarity, while deep blue denotes low similarity. It can be observed that the corresponding points can not be distinguished. On the contrary, by using the best features that will be selected by the method described next, we can successfully distinguish correspondences, as shown in Fig. 1(f). This example demonstrates the importance of learning the best features for each brain region, in order to maximally distinguish all brain regions simultaneously.



**Fig. 1.** The similarities between each red-crossed point in (a) and all points in (b), measured by RIFT, SPIN, local histograms, and our best features, respectively. The color-coded similarity maps in (c)-(f) indicate the importance of using best features for image registration.

**Method:** To learn the best features for different brain regions, the step of adaptive brain region partition and the step of best features determination should be completed simultaneously. This is because the partition of brain regions depends on the best features selected for each brain location, while the determination of best features also depends on the partition of brain regions since each brain region is reasonable to have the same set of best features. In particular, we iteratively perform (1) the partition of brain regions based on the tentatively estimated best features, and (2) the selection of best features for each tentatively partitioned brain region, until the algorithm converges.

It is worth noting that, the statistics of the selected best features for each brain point can also be estimated from all training samples. 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.

**Partition of Brain Regions:** Brain partition is achieved by adaptively grouping the neighboring points with the similar image features into various regions. Notice that the similarity between image features of neighboring points highly depends on the particular features used, which will be determined by a learning step detailed in the next. For convenience, a weighting vector  $\omega(v) = (\omega_1(v), \dots, \omega_k(v), \dots, \omega_M(v))$  is designed for each point  $v$ , to represent whether a particular feature  $g_k(v)$  is selected (if  $\omega_k(v)=1$ ) or not (if  $\omega_k(v)=0$ ). After brain partition, all points in the same brain region should have the same weighting vector, since they are required to have the same set of best features.

To achieve brain partition, we can first calculate for each point  $v$  the difference of tentatively selected best features between this point and the points in its neighborhood  $N(v)$ . The difference degree  $s(v)$  can be defined as follows:

$$s(v) = \sum_{\text{all samples}} \left( \sum_{u \in N(v)} \sum_{k=1}^M \omega_k(v) (g_k(v) - g_k(u))^2 \right) \quad (1)$$

Notice that  $s(v)$  will be small if features are similar in the neighborhood, while large if features are very different in the neighborhood. By calculating the feature difference measure  $s(v)$  for all points in the brain image, we can obtain a map of feature difference measures,  $\{s(v) | v \in I\}$ . Then, the partition of brain regions can be straightforwardly completed by a region competition algorithm [15], or a graph cutting algorithm [16], or watershed segmentation algorithm [17]. In this paper, we use watershed segmentation algorithm to partition the entire brain into various brain regions, i.e.,  $\{R_i\}$ .

**Selection of Best Features:** As mentioned, it is required that all points in the same brain region,  $R_i$ , have the same set of best features. But the selected best features for each brain region  $R_i$  should make all points in this region as distinctive as possible, for facilitating the feature-based image registration. The distinctiveness can be measured by both saliency and consistency of features on each point  $v$  in  $R_i$ . The saliency evaluates whether the selected best features on this point  $v$  are different from those on nearby points. The consistency evaluates whether the selected best features on this point are statistically similar across its corresponding points of individual subjects.

For convenience, the saliency and the consistency can be first measured for each feature  $g_k(v)$  on each point  $v$ . Jensen-Shannon (*JS*) divergence and entropy (*E*) can be used to formulate the saliency  $Sal(g_k(v))$  and the consistency  $Con(g_k(v))$  of each feature  $g_k(v)$ , respectively:

$$Sal(g_k(v)) = JS(H(g_k; N_1(v)), H(g_k; N_2(v))), \quad Con(g_k(v)) = -E(H(g_k; N_1(v))) \quad (2)$$

where  $N_1$  denotes a small neighborhood around the point  $v$  and  $N_2$  denotes the ring neighborhood outside of  $N_1$ .  $H(g_k; N(v))$  denotes for the histogram of feature  $g_k$  in the neighborhood  $N(v)$ , calculated from all samples with different weights that will be iteratively updated in each round of *adaboosting* as described next.

Then, the saliency and the consistency of all features on each point  $v$  can be obtained as follows:

$$Sal(G(v)) = \sum_{k=1}^M \omega_k(v) \cdot Sal(g_k(v)), \quad Con(G(v)) = \sum_{k=1}^M \omega_k(v) \cdot Con(g_k(v)) \quad (3)$$

where  $\omega_k$  represents whether feature  $g_k$  is currently selected.

Thus, the overall saliency and consistency measurement of a brain region,  $R_i$ , can be obtained by integrating all saliency and consistency measures over all points in the region, as given next:

$$E(R_i) = \sum_{v \in R_i} \left( \sum_{k=1}^M \omega_k(v) \cdot (Sal(g_k(v)) + \lambda \cdot Con(g_k(v))) \right) \quad (4)$$

where  $\lambda$  is a parameter used to balance between saliency and consistency measurements, and it is set to 1.0 in our study.  $E(R_i)$  is a function of weights  $\{\omega_k\}$ . By finding a suitable set of weights  $\{\omega_k\}$  for all points in the region, we can minimize  $E(R_i)$  with the result of obtaining the best features for this region.

To more effectively optimizing  $E(R_i)$ , we might also need to treat different point samples within neighborhoods  $N_1(v)$  and  $N_2(v)$  adaptively. *First*, given the current weight for each point sample, we can conveniently calculate the integrated saliency and consistency measurement for each feature according to the definition of  $E(R_i)$ . This measurement can be used to rank all features, thus allowing us to select the top ranked features as best features. *Second*, by considering the points in  $N_1(v)$  as positive samples and points in  $N_2(v)$  as negative samples for the point  $v$ , we classify both positive and negative samples by using each of the best features currently selected. The *adaboosting* strategy can be used to update the weights for these samples, i.e., weights on correctly classified samples will be decreased while weights on wrongly classified samples will be increased. By repeating the above two steps until convergence, we can finally select a set of best features, i.e., selecting the top ranked features in each round of adaboosting. In our study, four best features are selected for each brain region.

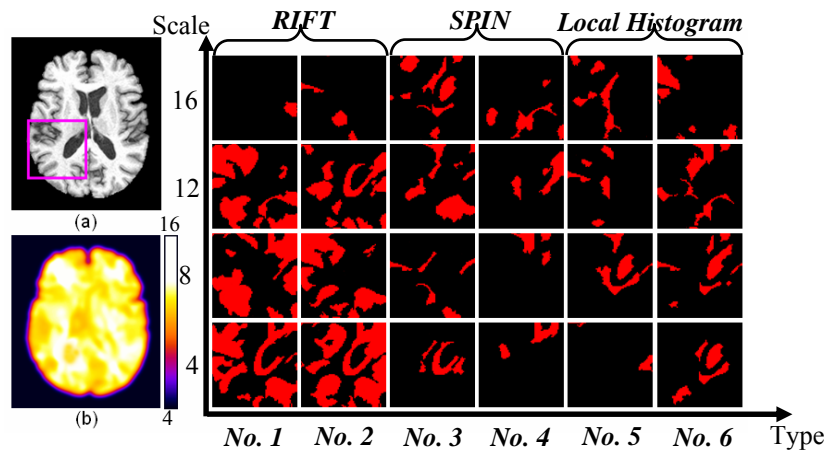
**Results:** In order to evaluate the selected best features by our method, Fig. 2 shows the detailed distributions of various features selected in a cross-sectional brain image. Fig. 2(b) provides the average scales used by the selected best features for the template in Fig. 2(a). As expected, small scales are used for boundary areas, while large scales are used for uniform regions. The right panel of Fig. 3 shows the selection result for a rectangular region in Fig. 2(a), with red color designating the selection of particular feature at particular position. It can be observed that RIFT features, which are based on edge orientation, are often selected for the points around the boundaries. Local histogram based features, often calculated from large scales, are selected for the points in the uniform regions. Overall, these results indicate that our learning-based method can well utilize the characteristics of each type of features.

## 2.2 Learning Active Points

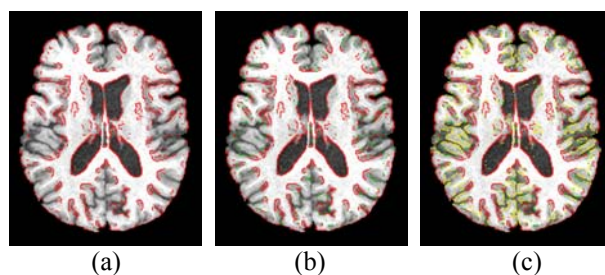
Two main measurements are considered, when selecting the reliable and salient points, called as active points, to hierarchically steer the image registration. The first measurement is the saliency of a point  $v$ , i.e.,  $Sal(G(v))$ , which is defined in Eq. (3). The large saliency value means that the point  $v$  can be easily distinguished from its nearby points, thus reducing the ambiguity in image matching. The second measurement is the consistency of the best features of the point  $v$  across different individuals, i.e.,  $Con(G(v))$ , which is also defined in Eq. (3). The large consistency measurement means that it is relatively easy to find the correspondence of point  $v$  in the other brain images. By combining these two measurements into a single measurement, i.e.,  $A(v)=Con(G(v))+\lambda \cdot Sal(G(v))$ , we can rank all brain points according to their overall measurements,  $\{A(v)\}$ .

Fig. 3 demonstrates the hierarchical selection of active points during image registration. In the initial stage of registration, only the most reliable and salient points are selected, as shown in Fig. 3(a). It can be observed that most active points locate at roots of sulci, crowns of gyri, and boundaries of ventricles, all of which can be distinguished reliably from others. With the progress of registration, more and more points are selected as active points and gradually added into the registration process, as shown by green and yellow in Figs. 3(b) and 3(c). By hierarchically focusing these

active points during the registration procedure, the registration performance can be greatly improved, as demonstrated latter by our experimental results.



**Fig. 2.** 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 pink block of (a), the selection results on all 24 features, from 4 scales and 6 features of 3 types, are displayed in the right panel.



**Fig. 3.** Hierarchical selection of active points for a template brain. Initial active points are shown in red (a) and gradually added ones are shown in green (b) and yellow (c), respectively.

### 2.3 Learning Deformation Statistics

We propose to build up the statistical model on deformation parameters to regularize the deformation fields estimated during the registration procedure. There are three major properties with this model. *First*, multi-level B-Splines are used to represent the deformation fields from coarse to fine levels, thus potentially overcoming the limitation of using single-level B-Splines, i.e., high dimensionality. *Second*, at each level, the control points are adaptively placed according to both deformation approximation degree and overall saliency & consistency measurement around the control points. *Third*, the statistical model is built up at each level, for capturing the statistics of deformations at each particular level, thus facilitating both efficient and

effective representation of deformation fields. All of these three properties are explained in detail below.

**Multi-Level Representation:** Let  $\Omega = \{(x,y,z) | 0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$  be the domain of deformation fields in the  $xyz$  coordinate system, and  $\Phi(\delta) = \{\varphi_{i,j,k} | 0 \leq i < N_x, 0 \leq j < N_y, 0 \leq k < N_z\}$  be the mesh of control points with uniform spacing  $\delta$  overlaid on  $\Omega$ . The deformation  $d(x,y,z)$  of each point  $(x,y,z)$  in  $\Omega$  can be represented as  $d_\delta(x,y,z)$  by a set of cubic B-spline functions  $B$ :

$$d_\delta(x,y,z) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u)B_m(v)B_n(w)\varphi_{i+l,j+m,k+n}, \quad \varphi_{i+l,j+m,k+n} \in \Phi(\delta) \quad (5)$$

where  $i = [x/\delta] - 1$ ,  $j = [y/\delta] - 1$ ,  $k = [z/\delta] - 1$ ,  $u = x/\delta - [x/\delta]$ ,  $v = y/\delta - [y/\delta]$ , and  $w = z/\delta - [z/\delta]$ .  $[x/\delta]$  returns the maximum integer less than  $x/\delta$ , and  $B_r, B_s, B_t$  are uniform cubic B-spline basis functions as defined next:

$$B_0(u) = (1-u)^3/6, \quad B_1(u) = (3u^3 - 6u^2 + 4)/6, \quad B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6, \quad B_3(u) = u^3/6 \quad (6)$$

The parameters on  $\varphi_{i,j,k}$  can be estimated from all known deformations  $\{d(x,y,z)\}$  of the points  $\{(x,y,z)\}$  within the neighborhood  $P_{i,j,k}(\delta)$ :

$$P_{i,j,k}(\delta) = \{(x,y,z) | (i-2) \cdot \delta \leq x < (i+2) \cdot \delta, (j-2) \cdot \delta \leq y < (j+2) \cdot \delta, (k-2) \cdot \delta \leq z < (k+2) \cdot \delta\}$$

In order to represent deformation field  $d(x,y,z)$  from coarse to fine levels, multi-level B-Splines are used. Using a mesh of control points with large uniform spacing, i.e.,  $\delta_0$ , the deformation field  $d(x,y,z)$  can be represented as  $d_{\delta_0}(x,y,z)$ , with the residual deformation field as  $e(x,y,z) = d(x,y,z) - d_{\delta_0}(x,y,z)$ . The residual deformation field  $e(x,y,z)$  can be represented by a mesh of control points with small uniform spacing, i.e.,  $\delta_0/2$ , resulting in a representation of  $e(x,y,z)$  as  $f_{\delta_0/2}(x,y,z)$  and a new residual as  $e(x,y,z) = e(x,y,z) - f_{\delta_0/2}(x,y,z)$ . Notice that  $f_{\delta_0/2}(x,y,z)$  can be estimated from  $e(x,y,z)$  similarly according Eq. (5). By repeatedly halving the spacing, i.e.,  $\delta_0/h$  (with  $h=4,8,\dots$ ), we can obtain various new representation  $f_{\delta_0/h}(x,y,z)$  for new residuals. Accordingly, the deformation field  $d(x,y,z)$  can be represented as:

$$d(x,y,z) \approx d_{\delta_0}(x,y,z) + \sum_{h=2,4,\dots} f_{\delta_0/h}(x,y,z) \quad (7)$$

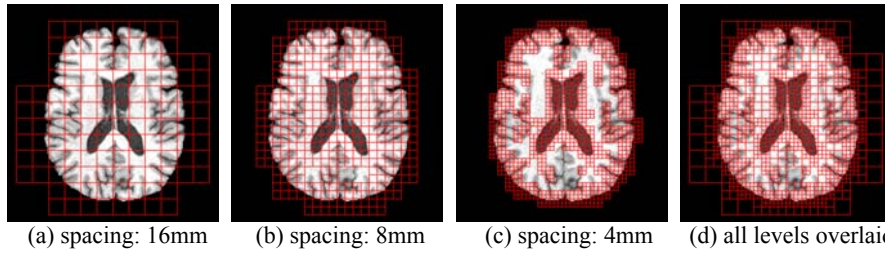
In our case, total four levels of B-Splines with spacing  $\{\delta = \delta_0/h \mid h=1,2,4,8\}$  are used. Notice that, in the application stage, the parameters on  $\varphi_{i,j,k}$  of each level will be estimated from the given deformations on the active points, which actually provides a new way of statistically interpolating the entire deformation fields by multi-level B-Splines, using the deformations available only on a small number of active points.

**Adaptive Placement of Control Points:** The control points of B-Splines can be adaptively placed at each level according to two criteria described next, by taking the advantage of multi-level representations. *First*, if the residual error around a particular control point is below a certain threshold, it is not necessary to split the current cube into smaller cubes in the next levels. *Second*, if the overall measurement of image saliency and consistency around a particular control point is higher than a certain value, it is worth splitting the current cube into smaller cubes, considering the saliency and importance of images in the current cube. By applying these two criteria, about 50% of control points can be discarded in the coarsest level, with most of them as background points. Also, nearly 85% of the control points are not necessarily kept in the finest level, since deformations on those locations have been well represented

by the previous levels. Accordingly, the dimensionality of our statistical model can be significantly reduced. Fig. 4 shows the adaptive lattices at three levels (a)–(c), using the spacing of control points 16mm, 8mm, and 4mm, respectively. Fig. 4(d) shows the overlay of all control points from three different levels, indicating dense control points in the complex areas and sparse control points in the uniform regions.

**Statistical Regularization:** Due to significant dimensionality reduction at each level, it is feasible to use a PCA to learn the statistics of  $\Phi(\delta)$  at each level, based on a relatively small number of training samples. The learned statistics can be used to reasonably constrain the deformation fields estimated during the image registration procedure. In particular, the regularization of tentatively estimated deformation fields can be achieved as follows:

- The parameter of each  $\varphi_{i,j,k}$  on  $\Phi(\delta)$  at current level is computed from the residual error within neighborhood  $P_{i,j,k}$ .
- The statistical model of this level is applied to statistically constraining the parameters on  $\Phi(\delta)$ .
- Update residual error, and repeat the above two steps until the total number of levels is reached.



**Fig. 4.** Adaptively placed control points at three different levels, with the spacing of 16mm, 8mm, and 4mm, respectively. The overlay of all control points from all levels, as shown in (d), indicates the control points are dense in the complex areas and sparse in the uniform regions.

## 2.4 Summary of our Learning-based Registration Framework

Our registration framework includes two parts, i.e., training stage and application stage. In the training stage, the best features are learned from a pool of features by *adaboosting*, and active points are hierarchically selected according to the overall saliency and consistency measures. The multi-level statistical models are learned from the deformation fields represented by multi-level B-Splines.

In the application stage, the most salient points are initially selected as active points to drive the registration. The deformations on other brain points are statistically interpolated by the multi-level B-Splines with learned parameters. With the progress of registration, other less salient points are gradually added and used as active points to start driving the registration. In the procedure of point matching, the learned best features are used to achieve better correspondence detection. In the end of each iterative registration, the tentatively estimated deformation fields are represented by multi-level B-splines, and the parameters of control points at each level are statistically constrained by their respective statistical models learned at that level.

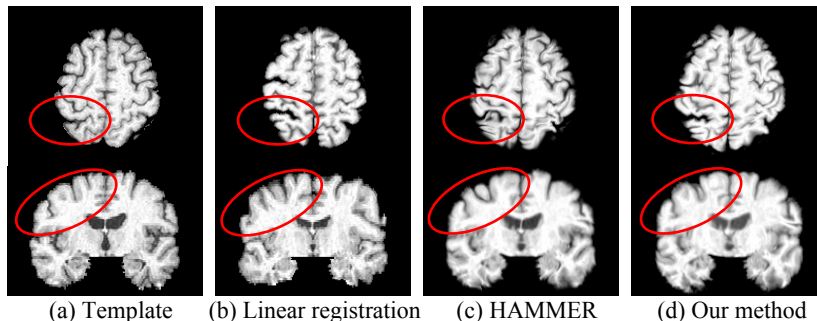
### 3. Experiments

Our learning-based registration method has been extensively evaluated on both real and simulated data, and its performances are also compared with those of the HAMMER algorithm. It's worth noting that our learning-based registration method is developed for achieving similar performance as HAMMER's, but avoiding tissue segmentation as required before the HAMMER registration algorithm. For real data, our learning-based registration method obtained the similar accuracy of registration as HAMMER's, but it is more robust in registering the longitudinal data as demonstrated next. For simulated data, our learning-based registration method produced 0.72 mm registration error, which is less than 0.92 mm produced by HAMMER.

For all experiments described next, total 18 MR brain images are used to learn both the best features and the active points, while 100 deformation fields from [18] are used to learn the deformation statistics. It's worth noting that all results reported in the next are obtained from the testing samples, which are not included in the training set.

#### 3.1 Experiments on Real Data

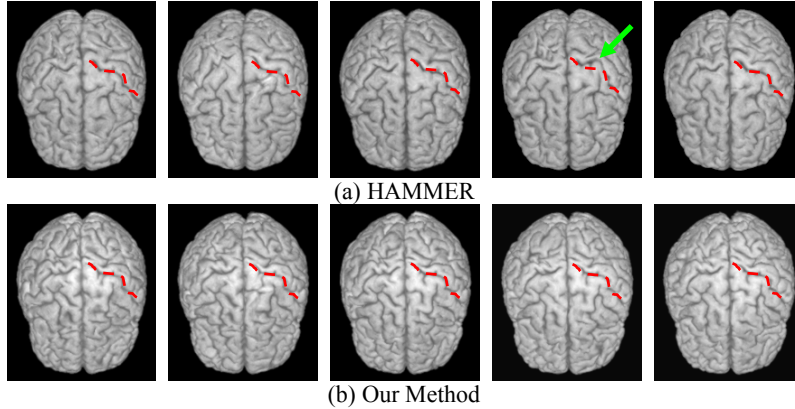
In the first experiment, both our learning-based registration method and HAMMER are tested on a number of real brain images obtained from our datasets. According to visual inspection, most registration results by these two methods are very similar. However, in some special cases, as shown in Fig. 5, our learning-based registration method produces better results than HAMMER, i.e., circled regions in Fig. 5.



**Fig. 5.** Results produced by linear method, HAMMER, and our method, respectively.

In the second experiment, the consistency of both methods in registering different time-point images of the same subject is evaluated. Since the warping consistency is very important to measure longitudinal brain changes, it is expected that the corresponding points in different time-point images of the same subject are consistently warped to the same location in the template space. Fig. 6 shows the warping results of five different year images of the same subject, produced by HAMMER 6(a) and our method 6(b), respectively. For better comparison of warping results on the cortex, 3D renderings of all warped different-year images are provided, with the dotted curves in the same spatial locations placed as landmarks to facilitate the visual comparison. For the region indicated by a green arrow in the year 4 image

of Fig 6(a), its warping result is inconsistent with those in other years. On the contrary, the warping results on this region are relatively consistent by our method.



**Fig. 6.** The warped results by HAMMER and our method are shown in (a) and (b) respectively, where the warping in the year 4 seems inconsistent by HAMMER in the region indicated by a green arrow.

### 3.2 Experiments on Simulated Data

The accuracy of our learning-based registration method is quantitatively evaluated by simulated data. For our method, the average registration error is 0.72mm, while, for HAMMER, its average registration error is 0.92mm. This indicates nearly 25% of error reduction by our method.

## 4. Conclusion

A fully learning-based framework has been developed for deformable registration of MR brain images. Learning methods have been designed to learn (1) the best features to enhance the accuracy of correspondence detection, (2) the active points to hierarchically steer the image registration, and (3) the multi-level deformation statistics to constrain the deformations. By using this learning-based registration framework, both registration accuracy and robustness have been achieved by our method. Our future work will include the investigation of more image features into our framework and the extension of our framework to other registration problems.

## 5. Reference

- [1] Y. Wang and L. H. Staib, "Elastic model-based non-rigid registration incorporating statistical shape information," *Lecture Notes in Computer Science: MICCAI'98*, vol. 1496, pp. 1162-1173, 1999.

- [2] J. C. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming 3D atlas to match anatomical brain images," *Journal of Computer Assisted Tomography*, vol. 17, pp. 225-236, 1993.
- [3] G. E. Christensen and H. J. Johnson, "Consistent Image Registration," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 568-582, 2001.
- [4] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 1421-1439, 2002.
- [5] G. Wu, F. Qi, and D. Shen, "Learning Best Features for Deformable Registration of MR Brains," *Medical Image Computing and Computer-Assisted Intervention ?MICCAI 2005*, pp. 179-187, 2005.
- [6] G. Wu, F. Qi, and D. Shen, "Learning-Based Deformable Registration of MR Brain Images," *IEEE Trans. on Medical Imaging*, vol. 25, pp. 1145-1157, 2006.
- [7] Z. Xue, D. Shen, and C. Davatzikos, "Statistical Representation of High-Dimensional Deformation Fields with Application to Statistically-Constrained 3D Warping Medical Image Analysis," *Medical Image Analysis*, vol. 10, pp. 740-751, 2006.
- [8] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3D statistical deformation models of the brain using non-rigid registration," *IEEE Trans. on Medical Imaging*, vol. 22, pp. 1014-1024, 2003.
- [9] Z. Xue and D. Shen, "Statistically-Constrained Deformable Registration of MR Brain Images," *Fourth IEEE International Symposium on Biomedical Imaging (ISBI 2007), Metro Washington, D.C., USA, 2007*.
- [10] S. Lee, G. Wolberg, and S. Y. Shin, "Scattered Data Interpolation with Multilevel B-Splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, pp. 228-224, 1997.
- [11] G. Wu, F. Qi, and D. Shen, "A General Learning Framework for Non-rigid Image Registration," *International Workshop on Medical Imaging and Augmented Reality (MIAR'06), Shanghai, China*, pp. 219-227, 2006.
- [12] S. Lazebnik, C. Schmid, and J. Ponce, "A Sparse Texture Representation Using Local Affine Regions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1265-1278, 2005.
- [13] D. Shen, "Image Registration by Local Histogram Matching," *Pattern Recognition*, vol. 40, pp. 1161-1172, 2007.
- [14] D. Lowe, "distinctive Image Features form Scale-Invariant Key-points," *International Journal of Computer Vision*, vol. 2, pp. 91-110, 2004.
- [15] S. Zhu and A. Yuille, "Region Competition: Unifying Snake/balloon, Region Growing and Bayes/MDL/Energy for multi-band Image Segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 884-900, 1996.
- [16] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 888-905, 2000.
- [17] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Transactions in Pattern Analysis and Machine Intelligence*, vol. 13, pp. 583-598, 1991.
- [18] Z. Xue, D. Shen, B. Karacali, J. Stern, D. Rottenberg, and C. Davatzikos, "Simulating Deformations of MR Brain Images for Evaluation of Registration Algorithms," *Neuroimage*, vol. 33, pp. 855-866, 2006.