

Learning Best Features for Deformable Registration of MR Brains*

Guorong Wu¹, Feihu Qi¹, and Dinggang Shen²

¹ Department of Computer Science and Engineering,
Shanghai Jiao Tong University, Shanghai 200030, China
{grwu, fhqi}@cs.sjtu.edu.cn

² Section of Biomedical Image Analysis, Department of Radiology,
University of Pennsylvania, Philadelphia, PA 19104
Dinggang.Shen@uphs.upenn.edu

Abstract. This paper presents a learning method to select best geometric features for deformable brain registration. Best geometric features are selected for each brain location, and used to reduce the ambiguity in image matching during the deformable registration. Best geometric features are obtained by solving an energy minimization problem that requires the features of corresponding points in the training samples to be similar, and the features of a point to be different from those of nearby points. By incorporating those learned best features into the framework of HAMMER registration algorithm, we achieved about 10% improvement of accuracy in estimating the simulated deformation fields, compared to that obtained by HAMMER. Also, on real MR brain images, we found visible improvement of registration in cortical regions.

1 Introduction

Deformable registration is very important for medical image analysis. So far, various methods have been proposed [1-7], either based on feature matching or intensity similarity. HAMMER registration algorithm [8] uses an attribute vector, instead of only intensity, as a signature of each point, for reducing the ambiguity in correspondence matching during the image registration procedure. Each attribute vector includes image intensity, edge type and a number of geometric moment invariants (GMIs) calculated in certain neighborhoods for reflecting the anatomy around that point. However, GMIs are calculated from the fixed sizes of neighborhood around each point at each resolution, regardless of whether this point is localizing in the complicated cortical regions or in the simple uniform regions. Thereby, it might be difficult to obtain the distinctive GMIs for every image point, by using *identical* neighborhood size for the whole image.

Recently, in computer vision area, Kadir and Brady [9] studied the implicit relationship between scale and saliency, and found that scale is intimately related to the problem of determining saliency and extracting relevant descriptions. They also proposed an effective method to detect the most salient regions in image, by considering the entropy of local image regions over a range of scales, in order to select regions with highest local saliency in both spatial and scale spaces. Based on [9], Huang *et al*

* This work is supported by NSFC (National Science Foundation of China) 60271033.

[10] proposed to align images under arbitrary poses, by finding the correspondences between salient region features. Although best features have been studied for active shape models [11, 12], however, to our knowledge, it seems that no previous non-rigid registration method considered the relationship between scale and saliency and used this relationship to guide image matching and correspondence detection during the deformable registration procedure.

This paper presents a learning-based method to compute GMIs from the best scales, for significantly reducing the ambiguity in image registration. Each image location will have its own best scale to calculate its GMIs, and simultaneously best scales are made smooth spatially. It is required that, for each point, its GMIs computed from its best-scale neighbor be *similar* across the corresponding points in the training samples, and also be *different* from GMIs of nearby points in all training samples. Entropy used in [9] is adopted here to measure this requirement, and the best scales are obtained by solving an energy minimization problem. Finally, by incorporating those best-scale GMIs into the original HAMMER algorithm, we achieved about 10% improvement in estimating the simulated deformation fields, compared to that obtained by HAMMER.

2 Method

2.1 Attribute Vector with Best Scales

Attribute vector is defined for each point \mathbf{x} in the image I , and it is designed as distinctive as possible, in order to distinguish this point from others in its neighborhood, $N_{\mathbf{x}}$. In HAMMER registration algorithm [8], each attribute vector includes edge type, image intensity, and GMIs. GMIs are computed from a spherical neighborhood around point \mathbf{x} , with radius of $S_{\mathbf{x}}$ that is identical at all image locations. Images, i.e., brain MR images, are usually spatially complicated, thus different regions usually need features computed from its best scale $S_{\mathbf{x}}$ [9], to distinguish itself from others. For example, for brain MR images, the point \mathbf{x} in cortical regions requires a different best scale $S_{\mathbf{x}}$ to compute the distinctive GMIs $\vec{G}(S_{\mathbf{x}})$, compared to the points in uniform regions. Therefore, it is significant to obtain a best scale $S_{\mathbf{x}}$ for each image point \mathbf{x} , based on particular image content around that point.

A machine learning based method is proposed to select best scales in the template space, in order to capture the distinctive attributes for robustly establishing correspondences. Three criteria are used to select best scales. *First*, the GMIs of a point \mathbf{x} , computed from the best-scale neighborhood, should be different from the GMIs of the nearby points in its neighborhood $R_{\mathbf{x}}$, thereby this point \mathbf{x} can be easily recognized. *Second*, the resulted GMIs of a point \mathbf{x} should be statistically similar to the GMIs of its corresponding points in training samples, if a set of training samples is available. *Third*, the selected best scales should be spatially smooth.

Entropy of GMIs is used to measure the above requirements, by following an idea of using entropy for salient region detection [9]. Thus, the *first* criterion requires that the entropy of GMIs in the neighborhood $R_{\mathbf{x}}$, $E_1(\mathbf{x}, S_{\mathbf{x}})$, be maximized, and the *second* criterion requires that the entropy of GMIs over the corresponding points in training samples, $E_2(\mathbf{x}, S_{\mathbf{x}})$, be minimized. Entropy can be computed from the histograms of GMIs [9]. The *third* criterion requires that the differences between $S_{\mathbf{x}}$ and scales S_y of

its small neighborhood r_x , $E_3(\mathbf{x}, S_x) = \sum_{y \in r_x} (S_x - S_y)^2$, be minimized. Therefore, we can obtain best scales for all image points, by using a gradient-based algorithm to minimize the following cost function:

$$E = \sum_{\mathbf{x}} (-E_1(\mathbf{x}, S_x) + \alpha E_2(\mathbf{x}, S_x) + \beta E_3(\mathbf{x}, S_x)) \quad (1)$$

where α and β are two weights. Notably, if there is no training samples available, then we just use the best scale selection method [9], with spatial smoothness constraint, to compute the best scales based on the template image itself.

A learning-based method for selecting the best scale S_x can be summarized next:

- Select a set of brain samples, such as 18 brains we used.
- Use a linear registration algorithm [13] to linearly align those samples to a selected template, thereby obtaining the linearly aligned brain samples.
- Use HAMMER algorithm [8] to register template with each linearly aligned brain sample, thereby obtaining the correspondences of each template point in any brain samples.
- For each template point \mathbf{x} and their corresponding points in training samples, compute their GMIs of different scales S_x , from the linearly aligned brains.
- Determine best scales for all template points jointly, by minimizing the cost function in equation (1).

For increasing the robustness of registration, the registration algorithms are usually implemented in a multi-resolution fashion [8]. Thus, we need to select best scales separately for each resolution, by performing the same best-scale selection method at each resolution.

Smooth maps of best scales are obtained, from fine to low resolutions, as shown in Fig 1 with the smallest scale (radius) 4 and the largest scale 24. The resulted best scales are actually adaptive to the brain anatomy, such as small scales selected in rich edge regions like cortex, and scales increased gradually from exterior to interior brain regions with the largest best scales selected for the uniform regions like white matter (WM) region. Notably, in low resolution, even a small best scale on cortex will capture a large region in the fine resolution (Fig 2), thereby providing the possibility of distinguishing between precentral and postcentral gyri. Also, since the registration algorithm is implemented in a multi-resolution fashion, the registration results from low and middle resolutions will make the two images approximately aligned, thereby local features, based on small best scales selected for cortex, can be used to refine the registration in cortex during the high-resolution registration stage. Fig 2 shows the best scales selected for seven points on ventricular corners, sulcal roots, gyral crowns, and putamen boundary, in three different resolutions, respectively. For convenience, both low and middle resolution images have been upsampled to the same size of high resolution image. The size of circle denotes the value of the best scale. Also, best scales ranged from 4 to 8 are displayed by solid circles, best scales ranged from 8 to 15 displayed by densely-dashed circles, and best scales over 15 displayed by sparsely-dashed circles.

Advantages of using best-scale GMIs. By employing a learning-based best scale selection method described above, we can use adaptive scale to compute GMIs for

each point, thus making it distinctive among its neighboring points as well as similar to its correspondences in other brains. For example, for a template point on sulcal root in Fig 3(a), as indicated by the cross, it is similar to its true correspondence indicated by the asterisk and the false correspondence indicated by the dot in subject (Fig 3(b)), if local image content is compared. Therefore, by measuring the similarities of this template point with all points in the subject image (Fig 3(b)) by the attribute vectors computed from neighborhoods of fixed scale (or size) such as $S_x=3,4,7$ used for low, middle, and high resolution images in HAMMER algorithm, it is not easy to establish correct correspondences, since multiple peaks are existing in the similarity map, as color-coded and shown in Fig 3(c). Red represents the most similar points, which include the false correspondence indicated by the dot in Fig 3(b). *Importantly*, by using our learning-based best scale selection method, we can determine the best scales for this template point at different image resolutions (i.e., $S_x=7,14,8$, respectively for low, middle and high resolutions), and further obtain for this template point all GMIs computed from different resolution images using the selected best scales. The best scales selected in low and middle resolutions ($7 \times 4=28$, $14 \times 2=28$) actually correspond

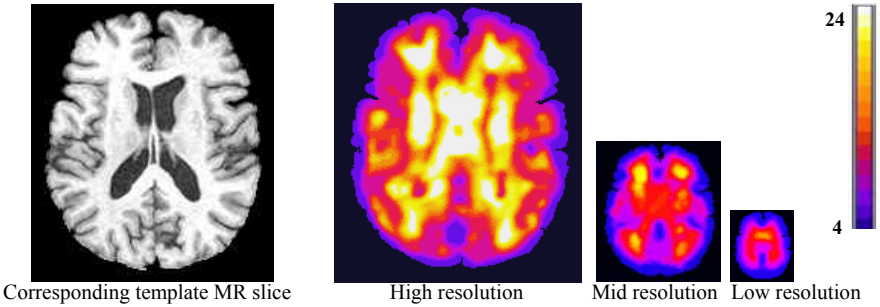


Fig. 1. Best scales selected for the image at three different resolutions, and further color-coded according to the color bar on the right. *This figure is best viewed with color.*

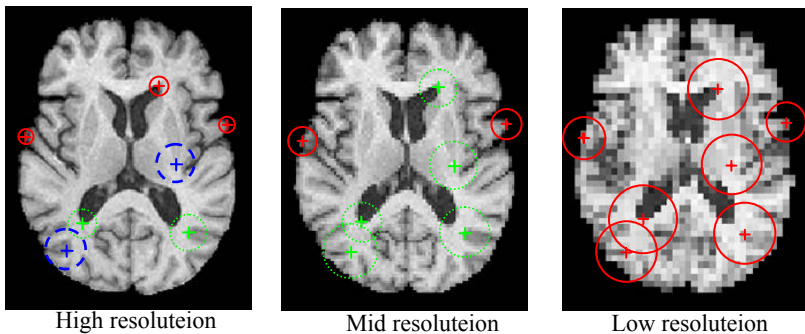


Fig. 2. Best scales of seven selected points in three different resolutions. For convenience, the low and middle resolution images (b,c) were zoomed to the same size of the original image. Here, best scales ranged from 4 to 8 are displayed by solid circles, best scales ranged from 8 to 15 displayed by densely-dashed circles, and best scales over 15 displayed by sparsely-dashed circles. *This figure is best viewed with color.*

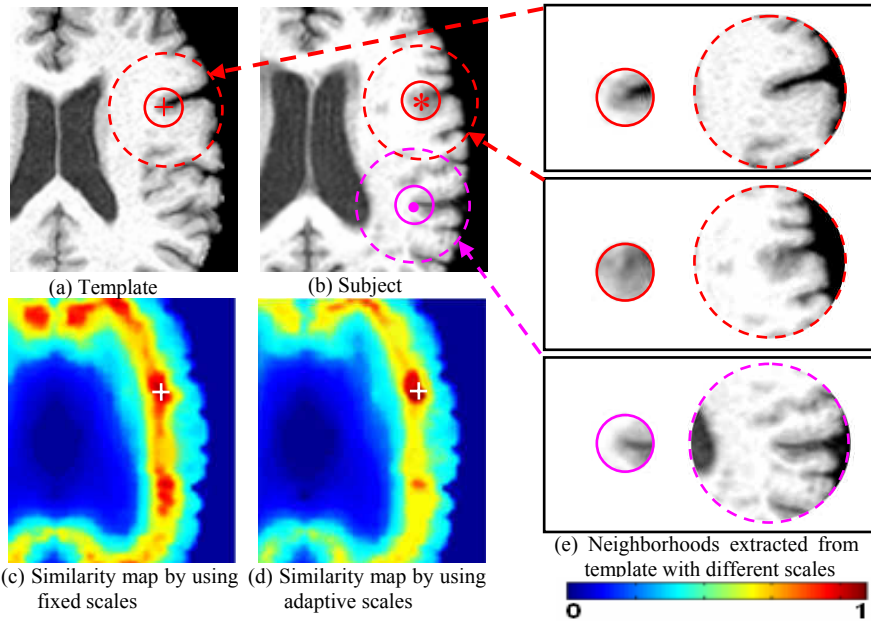


Fig. 3. Advantages of using adaptive scales to compute GMIs for correspondence detection. The similarity of a template point indicated by the cross in (a), is compared to any point in the subject (b), by respectively using GMIs with fixed scales (c) and with learned adaptive scales (d). The color-coded similarity map in (d) indicates distinctive correspondences, compared to the similarity map in (c) which has multiple peaks, with one peak corresponds to the false correspondence which is indicated by the dot in (b). *This figure is best viewed with color.*

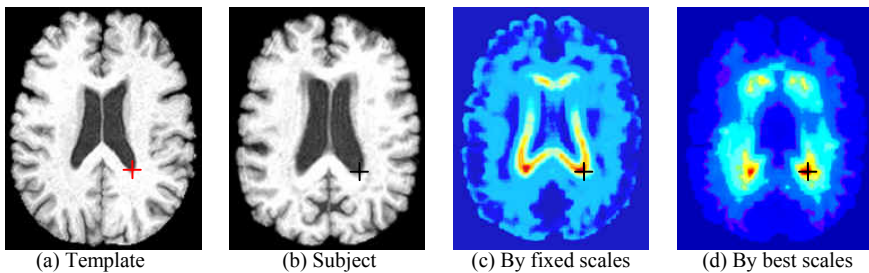


Fig. 4. Similar performances of using fixed scales and learned best scales for distinguishing some brain points, such as ventricular corners. The point in (b), as indicated by black cross, is a detected correspondence of the point in (a), by comparing the GMIs of either fixed scales or learned best scales. The red denotes similar, and blue denotes different. *This figure is best viewed with color.*

to big regions around this template point in the fine resolution, such as big circled images in the right panel of Fig 3. Thus, by using new attribute vectors, we can easily distinguish this template point from two candidate points in Fig 3(b), which has been clearly demonstrated by a color-coded similarity map in Fig 3(d).

Although it is possible to distinguish correspondences for many brain points even using the GMIs with a fixed scale, it may be less distinctive, compared to the method of using learned best scales. Fig 4 shows an example of detecting correspondences in subject image (Fig 4(b)), for a template point in Fig 4(a). According to a color-coded similarity map in Fig 4(c), the method of using fixed scales to compute GMIs can distinguish corresponding points, while it is less distinctive, compared to our method of using learned best scales, as indicated by a color-coded similarity map in Fig 4(d).

2.2 Image Registration by Matching Best-Scale Attributes

All image registration strategies developed in HAMMER algorithm [8], such as definitions of attribute vector similarity and energy function, and hierarchical driving point selection, are adopted by our registration algorithm, except using best-scale GMIs to replace the fixed-scale GMIs for image matching. Notably, the best scale for each location is determined in the template space, thus it is easy to compute GMIs for each *template* point by using the pre-calculated best scale. However, for *subject* image, it is not direct to use appropriate best scales to compute GMIs, since subject is in its own space. To overcome this problem, we will first align the subject to the template space by a linear registration algorithm [13], and then compute in advance the GMIs of all scales (used in the template space) for each subject point. When matching two template and subject points during the deformable registration procedure, we use the particular GMIs included in the attribute vector of the template point as standard and take the corresponding GMIs from the subject point, to measure their similarity. For saving time to compute GMIs of all possible best scales for each subject point, we limit the number of scales used as best scales, such as selecting best scales from a small set of scales $\{i | i = 4 * j, j = 1, 2, 3, \dots, 6\}$.

HAMMER algorithm selects the initial driving points at sulcal roots, crown gyri, and certain areas of strong boundary, and then gradually adds more driving points according to a simple criterion. Here, we adopt the saliency definition in [9], and similarly define the salient measure for each brain point as follows. Given the best scale S_x for a point \mathbf{x} , its salient measure can be defined by the entropy of GMI vectors in its neighborhood N_x , i.e. $E_1(\mathbf{x}, S_x)$, multiplied by a weight that penalizes self-similarity of $E_1(\mathbf{x}, S_x)$ around the best scale S_x [9]. Notably, the definition of $E_1(\mathbf{x}, S_x)$ is the same as that in equation (1).

3 Results

The proposed method has been evaluated by both real and simulated MR brain images with comparison of HAMMER algorithm [8]. All experiments are performed in PC (Pentium 4, 3.0GHz), by using the same set of parameters.

3.1 Experiment on Real MR Brain Images

The proposed registration algorithm has been used to register 18 brain images, and the results by our method are further compared with those by HAMMER algorithm. The average brain produced from 18 normalized brains by our method is visually very similar to that obtained by HAMMER algorithm. However, when we further check

individual registration results, we find the visual improvement by our method in the areas such as cortical regions, although two methods perform equally well on most parts of brain regions. Fig 5 shows two examples, which compare the template with the results obtained by both methods, indicating that our method can align cortical regions more accurately.

3.2 Experiment on Simulated Brain Images

Simulated data is used to quantitatively evaluate the performance of our method. Our simulated data is created by using an elastic warping method [14] to warp a selected brain to be similar to five real brains, respectively, thereby obtaining five simulated brains that are actually five deformed versions of the selected brain. Besides, the regions of precentral gyrus and superior temporal gyrus have been manually labeled in this selected brain, thereby the labels of these two regions can be warped together by the same deformation fields during the simulation procedure. Thus, by using our proposed registration method, we can estimate deformations between the selected brain and each of its deformed brains, and further bring two labeled regions in the simulated brains to the original space of the selected brain. Then, we can measure the overlay degree of the labeled regions. Our method achieves average overlay percentage 88.29%, which is very close to that by HAMMER algorithm on the same dataset (88.48%). The average volume error by our method is 5.18%, while it is 6.67% by HAMMER; this indicates 28.8% of volume error reduction by our method. Moreover, we compare the deformations estimated by our method with those simulated, thus obtaining a histogram of estimation errors as shown by red bars in Fig 6. This result is compared with that obtained by HAMMER algorithm, whose histogram of estimation errors is shown as blue bars in Fig 6. Obviously, the result obtained by our method is

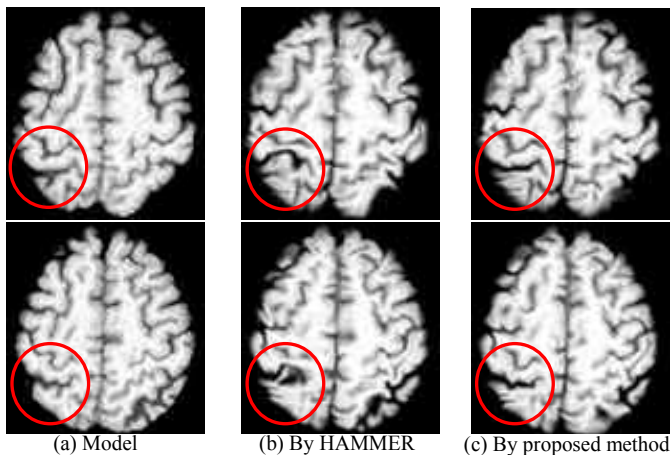


Fig. 5. Visual improvement in registering some brain images by the proposed method, particularly in the cortical regions circled.

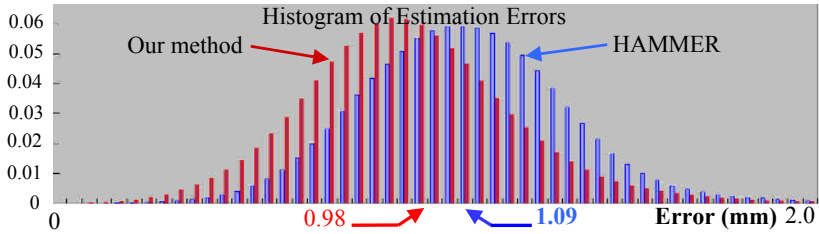


Fig. 6. Performances of estimating simulated deformations by our method (red) and by HAMMER algorithm (blue). The average error is 0.98 mm by our method, and 1.09 mm by HAMMER algorithm, which indicates 10.1% of improvement by our method.

much better, since its histogram is shifted to left, i.e., small errors. The average deformation error is 0.98mm by our method, and 1.09mm by HAMMER, indicating 10.1% of improvement by our method.

4 Conclusion

We have presented a learning based method to adaptively select best-scale GMIs for different image locations, thereby achieving higher registration accuracy by incorporating the selected best-scale GMIs into the HAMMER registration framework. Our learning method requires simultaneously the similarity of corresponding points in the training samples and the difference of a point to its nearby points, in terms of GMIs. It further requires the spatial smoothness of best-scale map. All of these requirements are formulated by a single entropy-based energy function, thereby solved by an energy optimization method. Importantly, our learning method can also be used to learn best features from others, i.e., wavelet-based features [15].

References

1. Y. Wang and L. H. Staib, "Elastic model-based non-rigid registration incorporating statistical shape information," *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 Trans on Med. Imaging*, vol. 20, pp. 568-582, 2001.
4. B. M. Dawant, S. L. Hartmann, and S. Gadamsetty, "Brain Atlas Deformation in the Presence of Large Space-occupying Tumours," *MICCAI'99*, vol. 1679, pp. 589-596, 1999.
5. P. Thompson and A. W. Toga, "A surface-based technique for warping three-dimensional images of the brain," *IEEE Trans on Med. Imaging*, vol. 15, pp. 402-417, 1996.
6. I. Wells, William M., P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, pp. 35-51, 1996.
7. Studholme C, Hill DLG, and H. DJ, "Multiresolution voxel similarity measures for MR-PET registration," *Proc. IPMI, Ile de Berder, France.*, pp. 287-298, 1995.

8. D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans on Med. Imaging*, vol. 21, pp. 1421-1439, 2002.
9. T. Kadir and M. Brady, "Saliency, scale and image description," *International Journal of Computer Vision*, vol. 45(2), pp. 83-105, 2001.
10. X. Huang, Y. Sun, D. Metaxas, F. Sauer, and C. Xu, "Hybrid image registration based on configural matching of scale-invariant salient region features," *IEEE Workshop on Image and Video Registration, Washington D.C.*, July, 2004.
11. B. v. Ginneken, A. F. Frangi, J. J. Staal, B. M. t. H. Romeny, and M. A. Viergever, "Active Shape Model Segmentation With Optimal Features," *IEEE Trans. on Medical Imaging*, vol. 21, pp. 924-933, 2002.
12. S. Li, L. Zhu, and T. Jiang, "Active Shape Model Segmentation Using Local Edge Structures and AdaBoost," *MIAR*, pp. 121-128, 2004.
13. M. Jenkinson, P.R. Bannister, J.M. Brady, and S. M. Smith, "Improved optimisation for the robust and accurate linear registration and motion correction of brain images," *Neuro-Image*, vol. 17, pp. 825-841, 2002.
14. C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Comp. Vision and Image Understanding*, vol. 66, pp. 207-222, 1997.
15. S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. on PAMI*, vol. 11, pp. 674--693, July 1989.