Image Registration by Hierarchical Matching of Local Spatial Intensity Histograms

Dinggang Shen

Section of Biomedical Image Analysis, Department of Radiology,
University of Pennsylvania, Philadelphia, PA 19104
dgshen@rad.upenn.edu

Abstract. We previously presented a HAMMER image registration algorithm that demonstrated high accuracy in superposition of images from different individual brains. However, the HAMMER registration algorithm requires pre-segmentation of brain tissues, since the attribute vectors used to hierarchically match the corresponding pairs of points are defined from the segmented images. In many applications, the segmentation of tissues might be difficult, unreliable or even impossible to complete, which potentially limits the use of the HAMMER algorithm in more generalized applications. To overcome this limitation, we use local spatial intensity histograms to design a new type of attribute vector for each point in an intensity image. The histogram-based attribute vector is rotationally invariant, and more importantly it captures spatial information by integrating a number of local histograms that are calculated from multi-resolution images. The new attribute vectors are able to determine corresponding points across individual images. Therefore, by hierarchically matching new attribute vectors, the proposed registration method performs as successfully as the previous HAMMER algorithm did in registering MR brain images, while providing more general applications in registering images of other organs. Experimental results show good performance of the proposed method in registering MR brain images and CT pelvis images.

1 Introduction

Deformable registration of images has been an active topic of research for over a decade [1,10-15]. We previously presented a HAMMER image registration method that demonstrated high accuracy in superposition of images from different individual brains [2,3]. However, the HAMMER algorithm requires that the images be segmented, before image registration can be performed, since the attribute vectors that are used to hierarchically match the corresponding pairs of points are defined from the segmented image. For images in certain modalities and from certain organs, the segmentation of tissues may be difficult, unreliable or even impossible to complete, which unavoidably limits the applications of the HAMMER algorithm.

It would be very attractive to use the deformation techniques developed in the HAMMER algorithm to directly register two intensity images, without pre-segmentation of images. One immediately available method uses wavelet-based feature extraction to characterize the geometric features around each point [4]. However, wavelet features are computationally expensive, and actually not invariant to rotations of the image. Importantly, in image matching and classification, more features will not always produce better results. It is particularly true in brain

matching, where the relatively high variability of brain structures makes some features, such as detailed features, vary dramatically across individual brains, thus confusing the image matching. Accordingly, features used for image matching are not necessarily very detailed, but they must be robust to structural variations across individuals.

In this paper, we design a new type of attribute vector for each point in an intensity image, based on local spatial intensity histograms. The histogram-based attribute vectors are very fast to compute, and also invariant to image rotation. Importantly, since local histograms are calculated from the intensity image at multiple resolutions, the new attribute vector captures sufficient spatial image information [5], thereby enabling the discrimination of the corresponding points across the individual images. By hierarchically matching new attribute vectors, the proposed registration method performs as successfully as the previous HAMMER algorithm did in registering MR brain images, while providing more generalized applications in the images of other organs or other modalities. Experimental results show good performance of the proposed method in registering MR brain images and CT pelvis images.

2   Method

2.1   Histogram-Based Attribute Vector

Definition: An attribute vector is defined for each point in the image, and used to best characterize the geometric features around that point, in order to reduce the ambiguities in determining the matching pairs of points during the image registration procedure. In this study, local intensity histograms of multi-resolution images around each point are computed and further used as attributes of that point. Moreover, boundary information, i.e., edgeness, is also extracted from each resolution image and used as additional spatial attribute, for discriminating boundary points from others. Therefore, the attribute vector $\mathbf{a}(v)$ of a point v in an image $f(v)$ includes both histogram-based attributes and boundary information, all of which are calculated from multi-resolution images respectively, as detailed next.

Histogram-based attributes are established by the following three steps. *Firstly*, the original image $f(v)$ is down-sampled by a factor of s, resulting in several down-sampled images, $f_s(v_s)$, where $v_s = [v/s]$ and thus $f_s(v_s) = f(v)$ when $s = 1$. For a point v in the original image $f(v)$, its correspondence in the down-sampled image $f_s(v_s)$ is v_s. Gaussian filter is used here to down-sample an image, and a total of three resolution levels, i.e., $s = 1,2,4$, are used. *Secondly*, for each resolution image $f_s(v_s)$, a local histogram $\mathbf{h}_s(v_s)$ of intensities in a spherical region of point v_s is computed. The radius of the spherical region is set to be identical across different resolutions. Therefore, for each point v in the original image $f(v)$, we can obtain several local histograms, i.e., $\{ \mathbf{h}_s(v_s) | s = 1,2,4 \}$, which capture different levels of spatial image information. *Thirdly*, the shape features, such as regular geometric moments [6], are respectively extracted from each histogram $\mathbf{h}_s(v_s)$, and then used as geometric attributes for the point v. Importantly, by extracting shape features from

histograms, we can obtain a relatively shorter vector of attributes (including both mean and variance of intensities in the spherical region) for each point in the image, thereby facilitating the fast and efficient matching of the corresponding points during the image registration procedure. For convenience, let $\mathbf{a}_s^{\text{Hist}}(v)$ represent a vector of low-frequency moments, obtained from a histogram $\mathbf{h}_s(v_s)$. Notably, vector $\mathbf{a}_1^{\text{Hist}}(v)$ captures relatively local features, while vector $\mathbf{a}_4^{\text{Hist}}(v)$ captures relatively global features.

Boundary attributes $b_s(v_s)$, used to measure the edge strongness, are computed from each resolution image $f_s(v_s)$ by the Canny edge detector [7], and encoded as boundary attributes in the attribute vector. For each point v in the original image $f(v)$, its corresponding boundary features in the three resolution images are $b_s(v_s)$, where $s = 1,2,4$, since $v_s = [v/s]$ is a corresponding point of v in the down-sampled image $f_s(v_s)$. For consistent representation of attributes, we use $b_s^{\text{Bound}}(v) = b_s(v_s)$ to represent the boundary attribute obtained at resolution s.

Therefore, the attribute vector of a point v can be finally represented as

$$\mathbf{a}(v) = \left[[\mathbf{a}_1^{\text{Hist}}(v) \quad b_1^{\text{Bound}}(v)], \quad [\mathbf{a}_2^{\text{Hist}}(v) \quad b_2^{\text{Bound}}(v)], \quad [\mathbf{a}_4^{\text{Hist}}(v) \quad b_4^{\text{Bound}}(v)] \right],$$

which includes three different levels of geometric features, with $[\mathbf{a}_1^{\text{Hist}}(v) \quad b_1^{\text{Bound}}(v)]$ as local features, $[\mathbf{a}_2^{\text{Hist}}(v) \quad b_2^{\text{Bound}}(v)]$ as middle-level features, and $[\mathbf{a}_4^{\text{Hist}}(v) \quad b_4^{\text{Bound}}(v)]$ as global features. Each attribute has been normalized between 0 and 1. By comparing the similarity of attribute vectors, we can determine the correspondences for points in the images. The similarity of two attribute vectors, $\mathbf{a}(u)$ and $\mathbf{a}(v)$, of two points, u and v, are defined as follows:

$$m(\mathbf{a}(u), \mathbf{a}(v)) = \prod_s \left(\left(1 - \left| b_s^{\text{Bound}}(u) - b_s^{\text{Bound}}(v) \right| \right) \cdot \prod_i \left(1 - \left| \mathbf{a}_{s,i}^{\text{Hist}}(u) - \mathbf{a}_{s,i}^{\text{Hist}}(v) \right| \right) \right),$$

where $\mathbf{a}_{s,i}^{\text{Hist}}$ is the i-th element of $\mathbf{a}_s^{\text{Hist}}$.

It is worth noting that the histogram-based attributes are invariant to rotational transformations of the image. Furthermore, by normalizing histograms both globally and locally, we can make the histogram-based attributes also robust to intensity inhomogeneities in the image. For example, by normalizing the global histogram, we can make any individual image have intensity distribution similar to that of a model image. Currently, histogram normalization is completed by first linearly transforming an individual's histogram to best match the model's histogram, and then using the optimally-estimated linear transformation parameters to map the intensities of the individual image. In the future, we plan to use a nonlinear histogram normalization method, by firstly using an elastic method to non-linearly establish the correspondences between the histograms of the model and the individual, and then using the established correspondences of intensities to map the intensities in the individual image.

Discrimination ability: In the problem of image matching and registration, it is important to make sure that the corresponding points in the different individuals have the similar attribute vectors. To demonstrate this, in Fig 1, the attribute vector of a

circled point in the left image is compared with the attribute vectors of all points in the middle image. According to the similarity map in the right, where white represents the high similarity, the circled point in the left image is similar only to its corresponding point, as indicated by a circle, in the middle image. Therefore, this example visually demonstrates that our histogram-based attribute vectors are able to determine the correspondences across different individuals.

It is worth noting the importance of including different levels of spatial geometric features into a single attribute vector for image matching and registration. Otherwise, it is impossible to discriminate different points by using only local/global attributes.
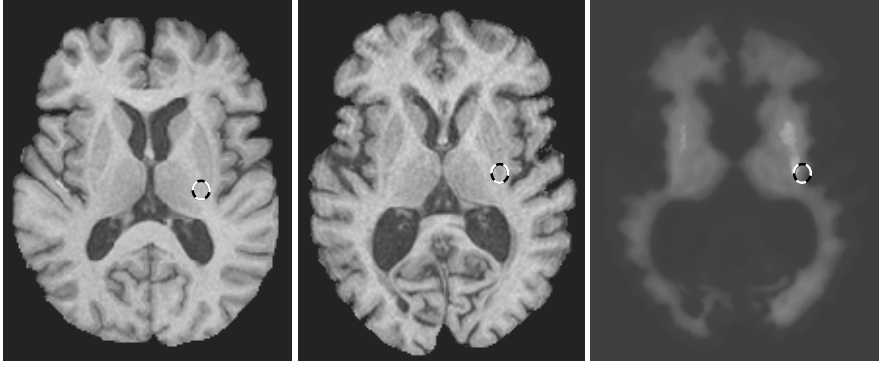

Fig. 1. Similarity of the corresponding points in the two individuals. The attribute vector of a circled point in the left image is compared with the attribute vectors of all points in the middle image. The resulting similarity map is shown in the right. The white represents high similarity, while the black denotes low similarity. The circled point in the middle image is a detected point with the highest degree of similarity.

2.2 Energy Function

Our energy function is formulated very similar to that of the HAMMER algorithm, i.e., it is designed to match the corresponding attribute vectors in the two images. Therefore, the optimization of our energy function can be completed by employing the exactly same deformation strategies developed in the HAMMER algorithm. Let $\mathbf{a}_T(u)$ denote the attribute vector of a point u in the template image $T(u)$, and let $\mathbf{a}_S(v)$ represent the attribute vector of a point v in the subject image $S(v)$. The template image $T(u)$ is deformed to match with the subject $S(v)$ by a displacement field $d(u)$, or equally a forward transformation $h(u) = u + d(u)$. The backward transformation from the subject to the model is $h^{-1}(u)$, which is the inverse of the forward transformation $h(u)$. The following is the energy function that our image registration algorithm will minimize:

$$E = \sum_u \omega_T(u) \left(\frac{\sum\limits_{z \in n(u)} \varepsilon(z)\left(1 - m\left(\mathbf{a}_T(z), \mathbf{a}_S(h(z))\right)\right)}{\sum\limits_{z \in n(u)} \varepsilon(z)} \right)$$

$$+ \sum_v \omega_S(v) \left(\frac{\sum_{z \in n(v)} \varepsilon(z) \big(1 - m\big(\mathbf{a}_T(h^{-1}(z)), \mathbf{a}_S(z)\big)\big)}{\sum_{z \in n(v)} \varepsilon(z)} \right) + \beta \sum_u \left\| \nabla^2 d(u) \right\|.$$

There are three energy terms in this energy function. The first energy term evaluates the match of template with subject, by using forward transformation $h(\cdot)$; while the second energy term evaluates the match of subject with template, by using backward transformation $h^{-1}(\cdot)$. Apparently, our energy function requires the consistent transformations that give identical mapping between two images, regardless of which of the two images is treated as the template [8,2,3].

The first energy term is defined as the weighted summation of neighborhood matching degrees of all points u in the template image. $\omega_T(u)$ is used as a weight for the point u, which can be adaptively adjusted by boundary attributes during the image registration procedure. Notably, this design allows the hierarchical selection of active points to focus on, thus enabling the approximation of a very-high-dimensional (equal to the number of points in the two images) cost function by a significantly lower-dimensional function of only the active points. This latter function has fewer local minima, because it is a function of the coordinates of active points, for which relatively unambiguous matches can be found. Accordingly, we can speed up the performance of image registration and also reduce the chances of local minima. For a point u, the degree of its neighborhood match is defined as the similarity of all attribute vectors in the neighborhood, $n(u)$. This design thereby allows the neighborhood matching during the image registration, which effectively increases the robustness to potentially false matches of active points. Here, z is a neighboring point of u, and its attribute vector $\mathbf{a}_T(z)$ is compared with the attribute vector $\mathbf{a}_S(h(z))$ of its corresponding point $h(z)$ in the subject; the similarity is defined as $m(\cdot,\cdot)$, thereby the difference is $1 - m(\cdot,\cdot)$. The term $\sum_{z \in n(u)} \varepsilon(z)$ is used for normalization. Notably, the design of the second energy term is the same as the first.

The third energy term is used to make sure that the resulting displacement fields $d(\cdot)$ be smooth, by requiring the total Laplacian value of displacement fields to be as small as possible. The parameter β controls the smoothness of the deformation fields.

3   Experimental Results

The performance of the proposed registration method is evaluated by using MR brain images and CT pelvis images. All experiments are performed on the volumetric images.

3.1   MR Brain Images

The brains images used in this study are obtained from our project, the Baltimore Longitudinal Study of Aging (BLSA) [9]. These images of elderly subjects pose

several difficulties in image matching, including reduced tissue contrast, significant atrophy, and motion artifacts.

Averaging 18 individual brains: The sharpness of the average image of the normalized individuals is often used as a visual display of the accuracy of the normalization algorithm. We selected the 18 individual brains used in our previous HAMMER paper [2]. Notably, the ventricles and also other structures in these 18 brains are of various shapes and sizes [2]. By normalizing these 18 brains to the space of a selected model, we can obtain an average image of these normalized 18 brains, as shown in Fig 2. By comparing this average image with the model as in Fig 2, we can observe the close similarity of these two images. Moreover, we can see that the average image is very clear, for example, in the regions of ventricles, caudate nucleus and lenticular nucleus. Therefore, the accuracy of our image registration method in registering MR brain images can be confirmed in part through this simple visual verification.

We also compare the average image of the proposed registration method, with that obtained by our previous HAMMER warping algorithm. As shown in Fig 2, the two average images possess almost the same level of sharpness, indicating the comparable accuracy of the proposed method to that of HAMMER method. Importantly, the registration method proposed in this paper does not require tissue segmentation, thereby making our registration method independent of the tissue-segmentation methods that may produce segmentation errors. Notably, for certain images in certain modalities, the segmentation of tissues may be difficult, unreliable or even impossible
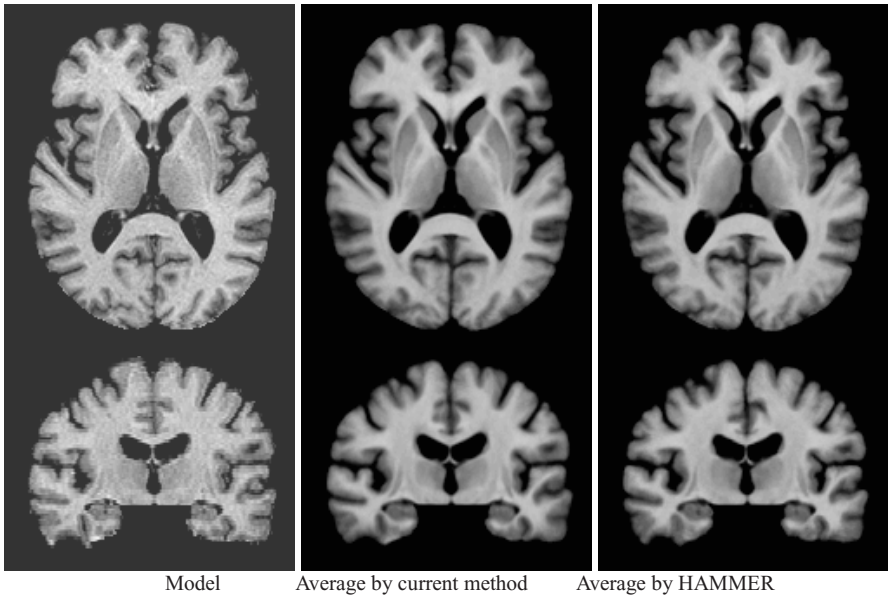

Model   Average by current method   Average by HAMMER

Fig. 2. Demonstration of the accuracy of the proposed method in averaging 18 individual brains. The average brain obtained by the proposed method is compared with that obtained by the HAMMER algorithm, indicating almost the same level of registration accuracy for the two methods.

to complete. Therefore, methods that directly register the intensity images, such as our proposed method, have the potential for success in more generalized applications, as is a major goal of this study.

3.2 CT Pelvis Images

Our registration method is also applicable in registering CT images of human pelvises, as shown in Fig 3. Before registration, the shapes of two individual pelvises and as well as their internal structures are very different, according to the cross-sectional images and 3D renderings in Figs 3a and 3c. After image registration, the two individual images become very similar, not only in their global shapes but in their representation of internal tissues, as shown in Fig 3b.
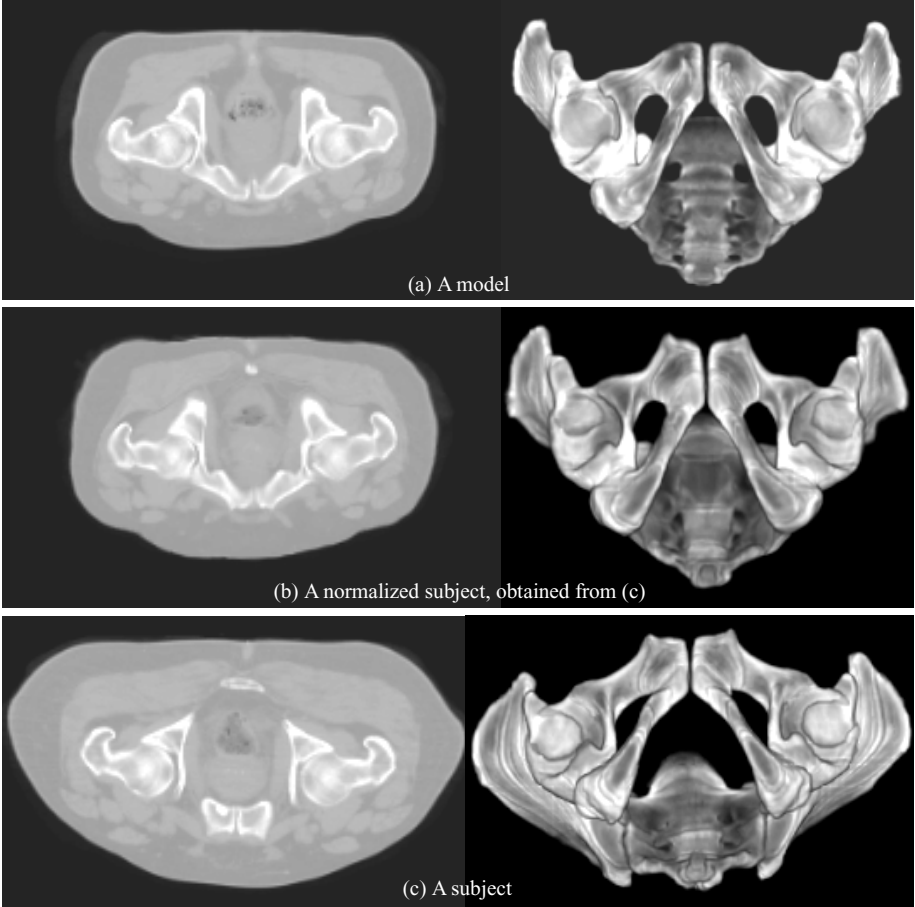

(a) A model

(b) A normalized subject, obtained from (c)

(c) A subject

Fig. 3. Demonstration of the proposed method in registering CT images of pelvises. A subject in (c) is normalized to the space of a model in (a), resulting in a spatially normalized image in (b), which is very similar to the model. Both cross-sectional images and 3D renderings are provided.

4   Conclusion

We have presented a method for the direct registration of intensity images that generalizes our previous HAMMER algorithm and eliminates the need to segment images before registration, thereby making the algorithm applicable to a variety of image sources and image modalities. Attribute vectors are used to characterize the local anatomy of each image point in a hierarchical fashion, in order to robustly match the corresponding points during the image registration procedure. The attribute vector in the HAMMER algorithm was defined from tissue-segmented images. In this paper, it is directly computed from the intensity images, by using the local spatial intensity histograms of multi-resolution images as geometric attributes. The histogram-based attribute vector is fast to compute and invariant to rotations of the image. Most importantly, our new type of attribute vector can distinguish between different points in the image, as long as a sufficient number of spatial features are captured from the local spatial histograms of a sufficient number of multi-resolution images.

References

1. T. McInerney, D. Terzopoulos. "Deformable models in medical image analysis: a survey", *Medical Image Analysis*, 1(2):91-108, 1996.
2. D. Shen, C. Davatzikos. "HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration". *IEEE Trans. on Medical Imaging*, Nov 2002.
3. D. Shen, C. Davatzikos. "Very High Resolution Morphometry Using Mass-Preserving Deformations and HAMMER Elastic Registration". *NeuroImage*, 18(1): 28-41, Jan 2003.
4. Z. Xue, D. Shen, C. Davatzikos. "Correspondence Detection Using Wavelet-Based Attribute Vectors", *MICCAI'03,* Canada, Nov 2003.
5. E. Hadjidemetriou, M. Grossberg, and S.K. Nayar. "Spatial information in multiresolution histograms", *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
6. D. Shen and H.H.S. Ip. "Generalized affine invariant image normalization". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(5):431-440, May 1997.
7. J. Canny. "A computational approach to edge detection", *IEEE Trans. on PAMI*, Vol 8, No. 6, Nov 1986.
8. G.E. Christensen and H.J. Johnson. "Consistent Image Registration", *IEEE Trans. on TMI*, 20(7), 2001, pp. 568-582.
9. S.M. Resnick, A.F. Goldszal, C. Davatzikos, S. Golski, M.A. Kraut, E.J. Metter, R.N. Bryan, and A.B. Zonderman. "One-year age changes in MRI brain volumes in older adults", *Cerebral Cortex*, 10:464-472, 2000.
10. H. Chui, L. Win, R. Schultz, J. Duncan, and A. Rangarajan. "A Unified Feature Registration Method for Brain Mapping". *IPMI*, p.300-314, Davis, CA, USA, June 18-22, 2001.
11. J.P. Thirion, O. Monga, S. Benayoun, A. Gueziec and N. Ayache. "Automatic registration of 3-D images using surface curvature". *SPIE Proc., Mathematical Methods in Medical Imaging*, volume 1768, p.206-216, 1992.

12. P. Thompson and A.W. Toga. "A surface-based technique for warping three-dimensional images of the brain". *IEEE Trans. on Med. Imaging*, volume 15, p.402-417, 1996.

13. J.C. Gee, C. Barillot, L.L. Briquer, D.R. Haynor and R. Bajcsy. "Matching structural images of the human brain using statistical and geometrical image features". *Proc. SPIE Visualization in Biomedical Computing*, vol. 2359, pp.191-204, 1994.

14. S.C. Joshi, M.I. Miller, G.E. Christensen, *et al*. "Hierarchical brain mapping via a generalized Dirichlet solution for mapping brain manifolds". *SPIE Conf. on Geom. Methods in Applied Imaging*, vol 2573, p.278-289, July 1995.

15. A.C. Evans, W. Dai , L. Collins, P. Neeling and S. Marett. "Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis". *SPIE Proc., Image Processing*, volume 1445, p.236-246, 1991.