Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 1161–1172
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Image registration by local histogram matching

Dinggang Shen*

Department of Radiology, Section of Biomedical Image Analysis, University of Pennsylvania, Philadelphia, PA, USA

Received 9 November 2005; received in revised form 28 June 2006; accepted 10 August 2006

Abstract

We previously presented an image registration method, referred to hierarchical attribute matching mechanism for elastic registration (HAMMER), which demonstrated relatively high accuracy in inter-subject registration of MR brain images. However, the HAMMER algorithm requires the pre-segmentation of brain tissues, since the attribute vectors used to hierarchically match the corresponding pairs of points are defined from the segmented image. 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 have used 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 importantly it also captures spatial information by integrating a number of local intensity histograms from multi-resolution images of original intensity image. The new attribute vectors are able to determine the corresponding points across individual images. Therefore, by hierarchically matching new attribute vectors, the proposed method can perform as successfully as the previous HAMMER algorithm did in registering MR brain images, while providing more generalized applications in registering images of various organs. Experimental results show good performance of the proposed method in registering MR brain images, DTI brain images, CT pelvis images, and MR mouse images.

Keywords: Deformable registration; Image warping; Non-rigid registration; Attribute vector; Invariants; Spatial histogram; Brain atlas; Atlas-based segmentation and labeling

1. Introduction

Deformable registration of medical images has been an active topic of research for over a decade. Many registration methods have been developed and they can be generally classified into three categories. The first category of methods is based on feature matching, where transformations are calculated based on a number of anatomical correspondences established on distinct anatomical features [1–10]. Such anatomical features were manually, semi-automatically, or automatically extracted from the landmark points, curves, or surfaces. The second category of methods is based on volumetric deformations [11–20], such as intensity-based image matching and registration algorithms. The last category of methods is based on matching of anatomical features and as well as deformation of volumetric images. The hierarchical attribute matching mechanism for elastic registration algorithm (HAMMER) [21,22] is one example of this category of methods, where each image point is characterized as a feature point and the distinctive points are designed to drive the initial deformation and registration of the images. This method has been shown to have relatively high accuracy in registering MR images of human brains.

However, the HAMMER algorithm requires that the images be pre-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 images. For the 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

* Tel.: +1 215 349 8588.
E-mail address: dgshen@rad.upenn.edu.

0031-3203/$30.00
doi:10.1016/j.patcog.2006.08.012

register two intensity images, yet without pre-segmenting those images before registering them. In HAMMER registration algorithm, geometric moments, used to characterize the underlying structure around each image point, were computed from the segmented images. To avoid the step of image segmentation, wavelets can be used to directly extract the detailed image features from the original intensity images [23]. However, wavelet features are computationally very expensive to extract, and actually not invariant to rotations of the image. Importantly, for image matching and classification, more features will not always produce better results, if some additional features are ambiguous. 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 confounding the image matching procedure. Accordingly, features used for image matching are not necessarily very detailed, but they must be robust to structural variations across individuals, and also invariant to image rotations.

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 can perform 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.

Actually, the histogram has been widely used to recognize and retrieve objects [24,25], because it can be computed and matched efficiently. However, a histogram of a single resolution image suffers from the inability to encode spatial image information. This can be overcome by simply combining histograms of an image at multiple resolutions to form a set of spatial histograms that incorporate spatial information [26]. With spatial information incorporated, spatial histograms are able to discriminate between different images even if those images have identical global histograms [26]. Notably, the formulation of spatial histograms is different from multi-resolution representation of a histogram where multi-resolutions are applied directly to the histogram [27]. Alternatively, spatial information can also be incorporated by combining the histograms of other spatial features such as edges, corners, textures, and regions [28].

The major contribution of this paper is the use of histogram-based features to directly register the original intensity images. To make the registration algorithm robust to the intensity-related problems, such as lighting and contrast, histograms are normalized globally and locally. Moreover, the deformation strategies developed in the HAMMER algorithm [21,22] are employed in the proposed method, and thus will be briefly explained here for the completeness of the paper.

This paper is organized as follows. The proposed method is provided in Section 2, with the focus of describing the definition of histogram-based attribute vector, and its ability in distinguishing points across images of the same subject and even of different subjects. The deformation strategies, similar to those in the HAMMER algorithm, are briefly mentioned in the end of Section 2. The performance of the proposed method has been evaluated by the images of different organs at different modalities, such as MR brain images, DTI brain images, CT pelvis images, and MR mouse images. The results are described and summarized in Section 3, and this paper concludes in Section 4.

2. Method

2.1. Histogram-based attribute vector

Definition: An attribute vector is defined for each point in the image, and used to characterize the geometric features around that point at different resolutions, thereby reducing 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 for image matching and correspondence detection. Although other detailed features such as wavelet coefficients of local images [23] can be used to characterize local image, the wavelet coefficients are computationally expensive to compute, and also they are not invariant to image rotations. Particularly for brain image registration, the detailed features are not necessarily good features for brain warping, since they are sensitive to structural variations across individual brain images.

Fig. 1 schematically demonstrates the procedure of calculating local intensity histograms at multi-resolution images. In addition to using the intensity histograms as attributes, boundary information, i.e., boundary strength, is also extracted from each resolution image and used as a spatial attribute, in order to discriminate boundary points from others. It is worth noting that features captured from the spatial histograms are coarse features, thus the use of these coarse features might affect the accuracy of correspondence detection. In order to identify the correspondences accurately as well, a boundary attribute is introduced to facilitate the boundary matching during the image registration procedure, thus increasing the accuracy in correspondence detection. Accordingly, Canny edge detector [29] is used here to quantify the strength of boundary on each point. 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.

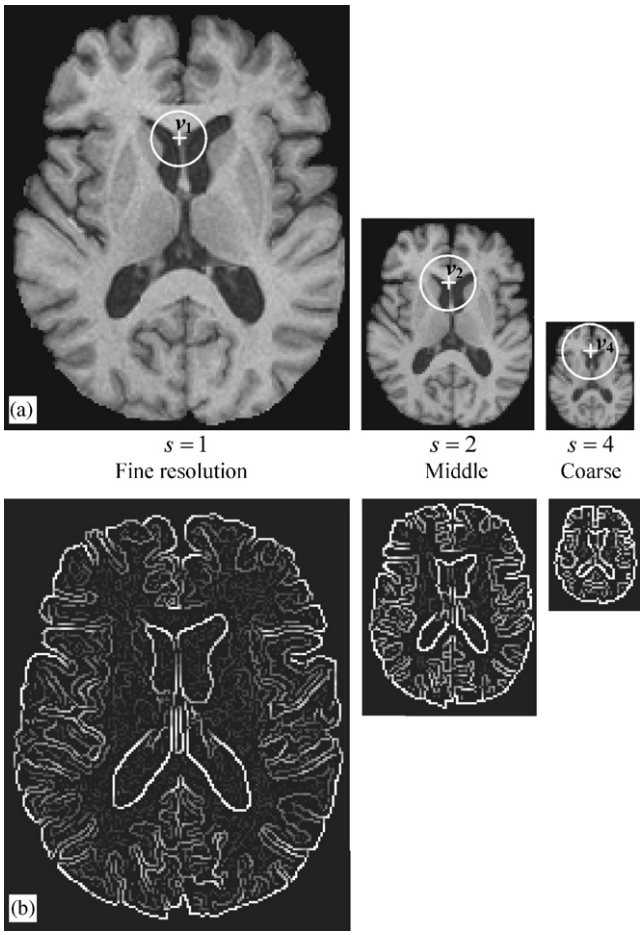

Fig. 1. Schematic demonstration of how the spatial histograms of multi-resolution images are calculated, for example, for a point v in the original image (a). Here, three levels of resolutions are used, and the histograms are computed within the circled regions of the point v_s, where $v_s = [v/s]$, $s = 1, 2, 4$, and v_s are the corresponding points of v in the different resolutions. The sizes of the spherical neighborhood are selected to be identical in voxel space, across different resolutions. Therefore, the histogram-based attributes, used to hierarchically characterize the local geometric structure, include hierarchical spatial information. Also, typical Canny boundary maps are provided in (b).

Histogram-based attributes are computed 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 at different levels, $f_s(v_s)$, where $v_s = [v/s]$ and thus $f_s(v_s) = f(v)$ when $s = 1$. Notably, for a point v in the original intensity image $f(v)$, its correspondence in the down-sampled image $f_s(v_s)$ is v_s. Gaussian filter is used in this study 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 is computed from a spherical region of point v_s. The radius of spherical region is set to be identical in voxel space across different resolutions, as demonstrated in Fig. 1. Therefore, for each point v in the original image $f(v)$, we can obtain several local histograms from the multi-resolution images, i.e., $\{\mathbf{h}_s(v_s) | s = 1, 2, 4\}$, which capture different levels of spatial image information around point v. *Thirdly*, the statistical features are, respectively, extracted from each histogram $\mathbf{h}_s(v_s)$ by calculating its regular geometric moments [30], i.e., $m(v_s, p) = \sum_i i^p \mathbf{h}_s(v_s, i)$, where $\mathbf{h}_s(v_s, i)$ is the frequency of intensity i in histogram $\mathbf{h}_s(v_s)$ and $m(v_s, p)$ is the pth order moment. By extracting these statistical features from a histogram, we can obtain a relatively short vector of attributes 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 $\mathbf{a}_s^{\mathbf{Hist}}(v)$ represent a vector of low-order geometric moments, i.e., $\{m(v_s, p),\ p=0, 1, 2\}$ obtained from a histogram $\mathbf{h}_s(v_s)$. Therefore, vector $\mathbf{a}_1^{\mathbf{Hist}}(v)$ captures relatively local features, while vector $\mathbf{a}_4^{\mathbf{Hist}}(v)$ captures relatively global features.

Boundary attributes $b_s(v_s)$, used to measure the boundary strength, are computed from each resolution image $f_s(v_s)$ by a Canny edge detector [29]. Notice that Canny edge detector outputs a point-wise boundary map, with zero as non-boundary and other values as strength of boundary. This point-wise boundary map provides a boundary attribute for each point in the image, as demonstrated in Fig. 1(b). For each point v in the original image $f(v)$, its corresponding boundary features in the three different 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 $b_s^{\mathbf{Bound}}(v) = b_s(v_s)$ to represent the boundary feature obtained at resolution s.

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

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

which includes three different levels of geometric features, with $[\mathbf{a}_1^{\mathbf{Hist}}(v)\ \ b_1^{\mathbf{Bound}}(v)]$ as local features, $[\mathbf{a}_2^{\mathbf{Hist}}(v)\ \ b_2^{\mathbf{Bound}}(v)]$ as middle-level features, and $[\mathbf{a}_4^{\mathbf{Hist}}(v)\ \ b_4^{\mathbf{Bound}}(v)]$ 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:

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

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

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 robust to intensity inhomogeneities in the image [31]. For example, by normalizing the global histogram, we can make any individual image have intensity distribution similar to that of a model

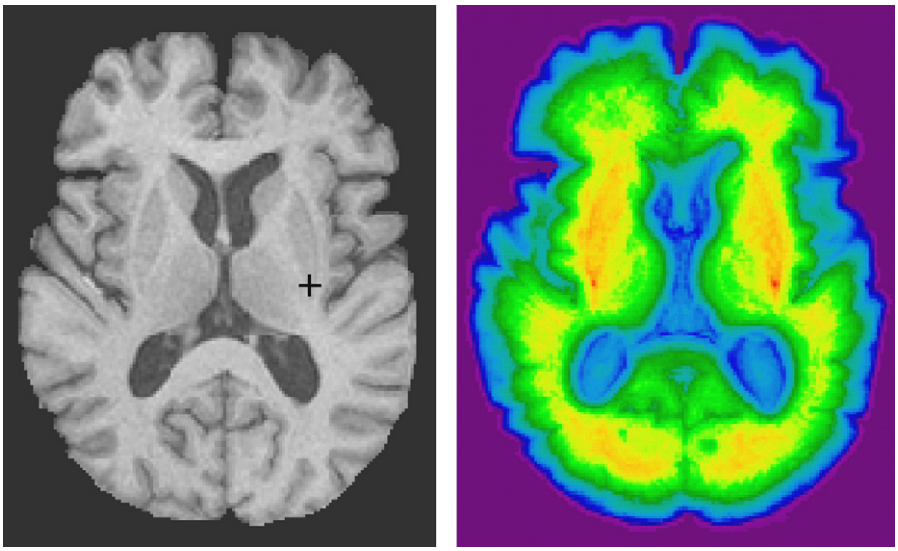

Fig. 2. Similarity of the points in the same individual. The attribute vector of a crossed point is compared with the attribute vectors of other points in the image, and the resulting map of similarities is color-coded and displayed in the right. Here, red is used to represent high similarity, and blue denotes low similarity. *This figure is best viewed with color.*

image. Currently, histogram normalization is implemented 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 nonlinearly 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: The ability of our histogram-based attribute vectors in discriminating points has been demonstrated by both Figs. 2 and 3. Fig. 2 shows a color-coded map of similarities between an attribute vector of a crossed point and the attribute vectors of all other points in the *same* individual. From the color-coded similarity map in the right, it is clear that the crossed point is similar only to the points in its small neighborhood and also in a small symmetric region. Here, the red color (or black in B/W print) represents the high similarity.

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, the attribute vector of a crossed point in Fig. 2 is compared with the attribute vectors of all points in the left image of Fig. 3. The two brain images in Figs. 2 and 3 are acquired from two *different* individuals. According to the color-coded similarity map in the right of Fig. 3, the crossed point of Fig. 2 is only similar to their corresponding points of Fig. 3, where the most similar point is labeled by a cross. Therefore, both Figs. 2 and 3 visually prove that our histogram-based attribute vectors are able to distinguish points in the same individual, and also 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, as visually indicated by an example in Fig. 4. If using only the part of local features ($[\mathbf{a}_1^{\mathbf{Hist}}(v)\ \ b_1^{\mathbf{Bound}}(v)]$), such as features obtained from the local histogram of fine-resolution image $f(v)$, the crossed point in Fig. 2 is similar to many points in its neighborhood, as shown in Fig. 4a. Similarly, if using only the part of middle-level features ($[\mathbf{a}_2^{\mathbf{Hist}}(v)\ \ b_2^{\mathbf{Bound}}(v)]$) or global features ($[\mathbf{a}_4^{\mathbf{Hist}}(v)\ \ b_4^{\mathbf{Bound}}(v)]$), it is still difficult to distinguish this crossed point from others. Only by using a complete vector of spatial attributes $\mathbf{a}(v)$, are we able to discriminate among different points, as we have seen in Fig. 2.

2.2. Energy function

The procedure of image registration is usually formulated as optimization of an energy function that evaluates the similarity of two images under deformable registration. Therefore, the actual definition of energy function is very important, since it directly determines the result of final image registration. Our energy function is designed to match the corresponding attribute vectors in the two images, according to the three criteria detailed below.

The first criterion is that the designed energy function should allow the registration algorithm to adaptively focus on different sets of image points in the different stages of image deformation. This is very important, since some image points have relatively distinctive attribute vectors, which

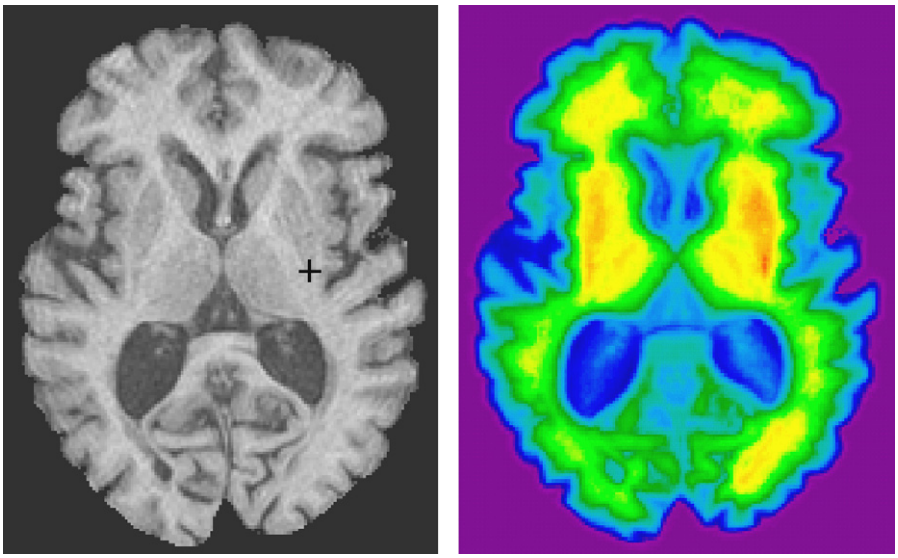

Fig. 3. Similarity of the points from the images of different individuals. The attribute vector of a crossed point in Fig. 2 is compared with the attribute vectors of all points in the left image of current figure. The resulting similarities are color-coded, as shown in the right image. The red represents high similarity, while the blue denotes low similarity. The crossed point in the left image is a detected point with the highest degree of similarity, compared to a crossed point in Fig. 2. *This figure is best viewed with color.*

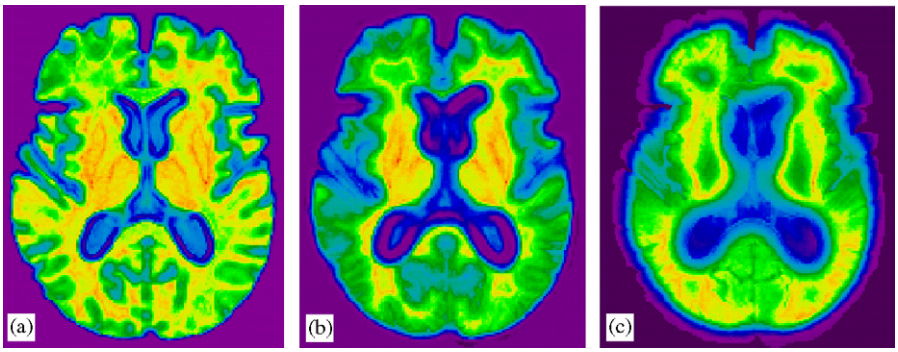

Fig. 4. Importance of integrating different levels of spatial geometric features for image comparison. (a) shows a color-coded map of similarities between the crossed point of Fig. 2 and all other points in Fig. 2, using only the part of local features in the attribute vector. Similarly, (b) and (c) correspond to the results where the parts of middle-level features and global features are, respectively, used. Clearly, those points cannot be discriminated among when only a part of features are used for image comparison. However, they become distinguishable when sufficient levels of spatial geometric features are integrated and used together for image comparison, as demonstrated by the color-coded similarity map in Fig. 2. *This figure is best viewed with color.*

can be identified relatively more reliably based on their morphological signatures, compared to other more ambiguous points. Therefore, these image points, which we call active points, should be allowed to drive the initial registration of two images. As the image registration progresses, a greater and greater number of image points become spatially closer to their counterparts, and thus become reliable to drive the image registration. Therefore, the number of the active points should be progressively increased during the deformable registration procedure. In this paper, each point is designed to have its own energy term, and the whole energy function is a weighted summation of all points' energy terms. By hierarchically assigning those weights according to the distinctiveness of attribute vectors, that is, assigning large weights for the energy terms of the points with distinctive attribute vectors and zero weights for the energy terms of other points, we can hierarchically focus on the most suitable points to actively drive the image registration. Effectively, this procedure approximates what would be 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 few local minima, because it is a function of the coordinates of active points, for which relatively unambiguous matches can be found. Therefore, using this strategy, we can speed up the performance of image registration and also reduce the chances of local minima, which in part result from ambiguities in determining the matching pairs of points.

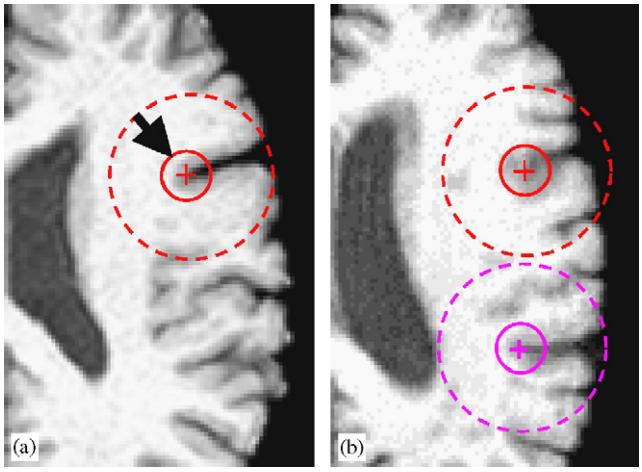

Fig. 5. Demonstrating the importance of using a hierarchical neighborhood matching mechanism for determining the corresponding structures in the model (a) and the subject (b). In the initial registration procedure, the large neighborhood will be used for image matching and correspondence detection, i.e., dotted circle; while in the later stages, the relative small neighborhood will be used, i.e., solid circle, in order to improve both the accuracy and the speed of registration algorithm. If directly using a small neighborhood, ambiguous correspondences, i.e., two candidates in (b), can be produced. The black arrow in (a) shows the decrease of neighborhood size with the progress of image registration, for achieving both robust and accurate correspondence detection.

The second criterion is that the energy function should be symmetrically designed for two images under registration, in order to construct consistent transformations that give identical mapping between two images, regardless of which of the two images is treated as the template [21,22,32,33]. That means, both forward and backward transformations should be evaluated in a single energy function, and forced to be consistent with each other. A classical approach for enforcing consistent transformations in a registration process is computationally very demanding. We instead enforce inverse consistency only on the active points that currently drive the registration of two images, since the transformations elsewhere are determined via interpolation from those active points. The detailed method for obtaining inverse consistent matches was described in Ref. [21].

The third criterion is that the similarity of attribute vectors of corresponding points should be evaluated along with points in the neighborhood during the energy function minimization procedure, since the matching of neighborhoods has been proved robust to false matches, as also demonstrated in Fig. 5. Consequently, the energy terms of neighboring points are better grouped in the energy function, according to the neighborhood relationship of points in the image space. The use of this neighborhood matching strategy provides at least two advantages as listed below:

- By adaptively changing the size of neighborhood, i.e., a spherical neighborhood, the correspondences in the two images can be robustly and also accurately established. For example, by using a spherical neighborhood with large radius r in the initial registration stages, the correspondences in the two images can be robustly established. On the other hand, by reducing the size of neighborhood with the progress of image registration, the correspondences in the two images can be established accurately. In particular, the radius r of the spherical neighborhood is decreased from 8 to 1 voxels in each resolution image. Notice that 8 voxels in the lowest resolution image equal to 32 voxels in the fine-resolution image, which means a large neighborhood will be used for correspondence detection in the initial registration procedure.
- In the real application, the detection of correspondence for an active point can be completed by firstly searching for several candidate points with similar attribute vectors in its neighbor, and then verifying each candidate match by checking the similarity of attribute vectors in the respective neighborhood, as did in Ref. [21]. The candidate point with the highest neighborhood-derived attribute vector similarity is defined as a corresponding point, provided that the similarity is above a pre-specified threshold. Thus, the use of this neighborhood matching strategy helps eliminate the false correspondence detection.

Our energy function is designed to obey the above-mentioned criteria exactly. Let us assume $T(u)$ to be a template image, and $S(v)$ to be a subject image. Correspondingly, $\mathbf{a}_T(u)$ denotes the attribute vector of a template point u, while $\mathbf{a}_S(v)$ represents the attribute vector of a subject point 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 $h(u)=u+d(u)$. Therefore, the backward transformation from the subject to the model is $h^{-1}(v)$, which is the inverse of the forward transformation $h(u)$. The following is the energy function that our image registration algorithm will minimize:

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

There are three energy terms in this energy function. The first energy term evaluates the match of model with subject, by using forward transformation $h(\cdot)$; while the second energy term evaluates the match of subject with model, by using backward transformation $h^{-1}(\cdot)$. Therefore, this energy function is symmetric, satisfying the second criterion listed above.

The first energy term is defined as the weighted summation of neighborhood matching degrees of all points u in the model image. $\omega_T(u)$ is used as a weight for a point u in the template image, representing the importance of this point in the image matching. The weight $\omega_T(u)$ can be adaptively

adjusted by the boundary attribute of the point u, during the image registration procedure. That is, points with the large boundary attributes will be assigned with large weights, while points with zero boundary attributes will be assigned with zero weights. This effectively excludes the energy terms corresponding to the non-boundary points, thus making the whole energy function simple and mainly focusing on critical points during the image registration procedure. Notably, this design is exactly required by the first criterion listed above.

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 match during the image registration, exactly satisfying the third criterion listed above. z is a neighboring point of u; its attribute vector $\mathbf{a}_T(z)$ is compared with the attribute vector $\mathbf{a}_S(h(z))$ of corresponding point $h(z)$ in the subject. The similarity is defined as $m(\cdot,\cdot)$, thereby the difference is $1-m(\cdot,\cdot)$. $\varepsilon(z)$ is a weight that is also adaptively determined by the boundary attribute of the point z, which is very similar to the determination of weight $\omega_T(u)$ as discussed above. The term $\sum_{\mathbf{z}\in n(u)}\varepsilon(z)$ is used as normalization. The design of the second energy term is the same as the first, in particular, the weight $\omega_S(v)$ is also determined by the boundary attribute of point v in the subject image.

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. For our applications, we use 0.5 for β.

3. Experimental results

This section provides a set of experiments to evaluate the performances of the proposed method in registering images of different individuals at different modalities, such as MRI, DTI, and CT. Both brain and pelvis images of humans are used as testing samples. The images of whole-body mice are also included as testing samples. Notably, all experiments are performed on the volumetric images. However, only the cross-sectional images are provided for a more readily accessible illustration of the work on paper.

3.1. MR brain images

The brains images used in this study are obtained from our project, Baltimore Longitudinal Study of Aging [34]. These images of elderly subjects pose several difficulties in image matching, including reduced tissue contrast, significant atrophy, and motion artifacts. In the following, we provide the results of our method in registering two individual brains, and in normalizing a group of individual brains to a standard space. Also, the validation results on simulated data are provided in the end of this subsection.

Registration of two individual brains: This experiment demonstrates the performance of the proposed method in registering images of two individual brains; one is used as the model, while the other is used as the subject. The images of these two individuals are quite different, as shown in Fig. 6a and c. For example, the one in Fig. 6c has quite large ventricles. After image warping, the two individual brains become very similar, as can be seen by comparing the warped image in Fig. 6b with the subject in Fig. 6c. The similarity of these two images can be further verified by overlapping a set of identical landmarks on them, as displayed in Fig. 6d and e. The underlying gray-level images in Fig. 6d and e are identical to those in Fig. 6b and c, respectively.

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 [21]. Notably, the ventricles and also other structures in these 18 brains are of various shapes and sizes [21]. By normalizing these 18 brains to the space of a randomly selected model, we can obtain an average image of these normalized 18 brains, as shown in Fig. 7. By comparing this average image with the model as in Fig. 7, 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. 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. 7, the two average images possess almost the same level of sharpness, indicating the comparable accuracy of the proposed method to that of HAMMER method. However, the registration method proposed in this paper does not require tissue segmentation, therefore making our registration method independent of the tissue-segmentation methods that may produce segmentation errors. Importantly, for certain images in certain modalities, the segmentation of tissues may be difficult, unreliable or even impossible 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.

Validation on simulated data: The performance of the proposed method has been tested on a simulated data set, used to validate HAMMER registration algorithm in Ref. [21]. The simulated data were obtained by an elastic warping algorithm called STAR [35]. In particular, we manually painted major sulci of the model and individuals, and used them as constraints to warp the model into individuals using the STAR algorithm, thereby obtaining five simulated brain images in our study. Notice that, for these simulated brain images, we exactly know their deformation fields relative to the model. Therefore, we can compare the deforma-

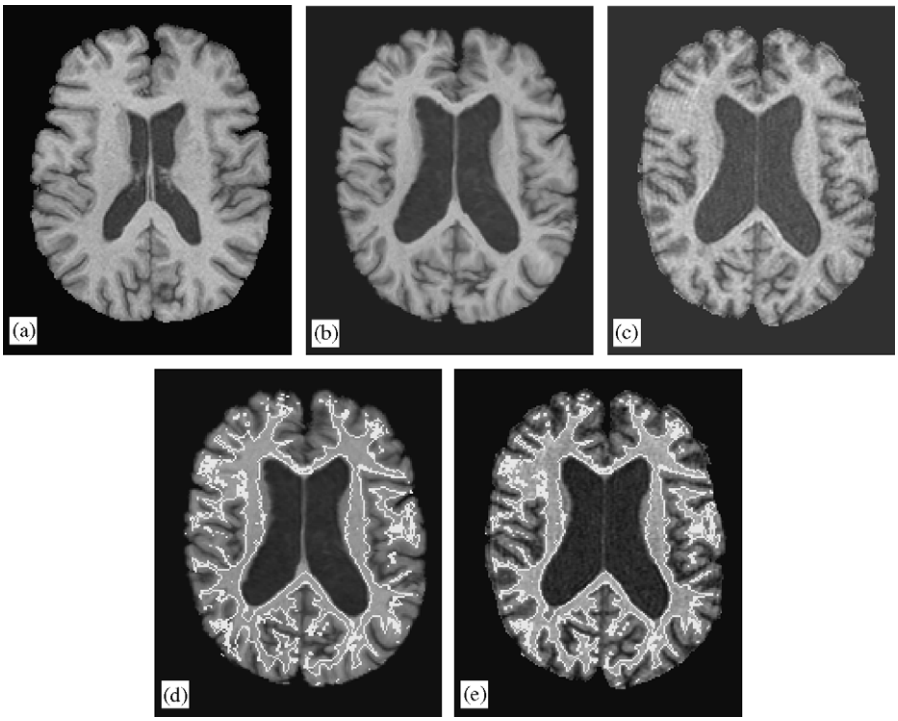

Fig. 6. The performance of warping a model brain in (a) to a subject brain in (c). The warped model brain in (b) becomes very similar to the subject in (c). This can be further confirmed by (d) and (e), where a set of identical landmarks are overlapped onto (b) and (c), respectively.

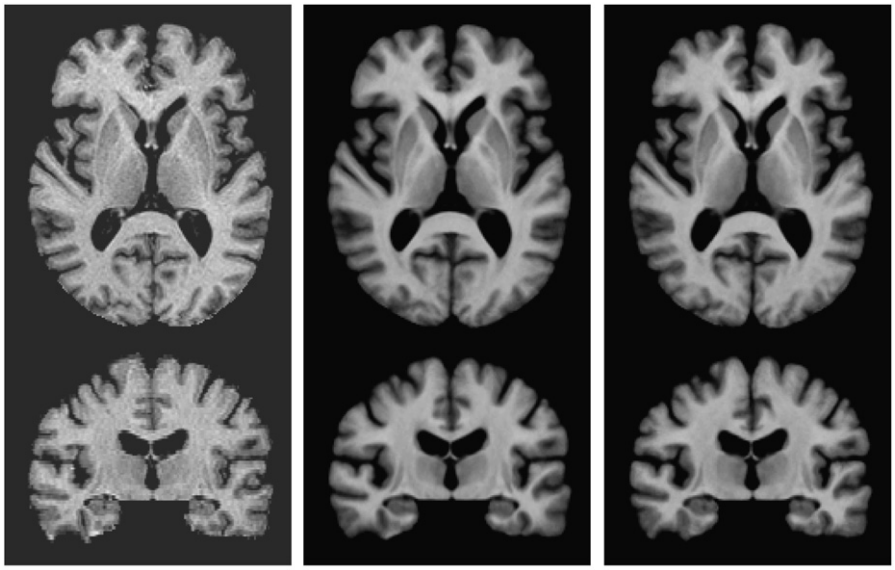

Fig. 7. Demonstration of the accuracy of the proposed method in averaging 18 individual brains. The model that we used to normalize the individual brains is displayed in the left, and the resulting average image of 18 normalized brains is given in the middle. Also, the average obtained by the proposed method is compared with that obtained by our previous HAMMER registration algorithm (in the right), which indicates almost the same level of registration accuracy for the two methods.

tion fields that our algorithm estimated with the ground-truth deformation fields, therefore providing a quantitative measurement on the accuracy of our registration algorithm. Our average registration error is ~ 1.0 mm, which is very similar to the one obtained by HAMMER registration algorithm as we reported in Ref. [21]. Notice that, although the proposed registration algorithm has similar accuracy compared to the HAMMER registration algorithm, the advantage of using

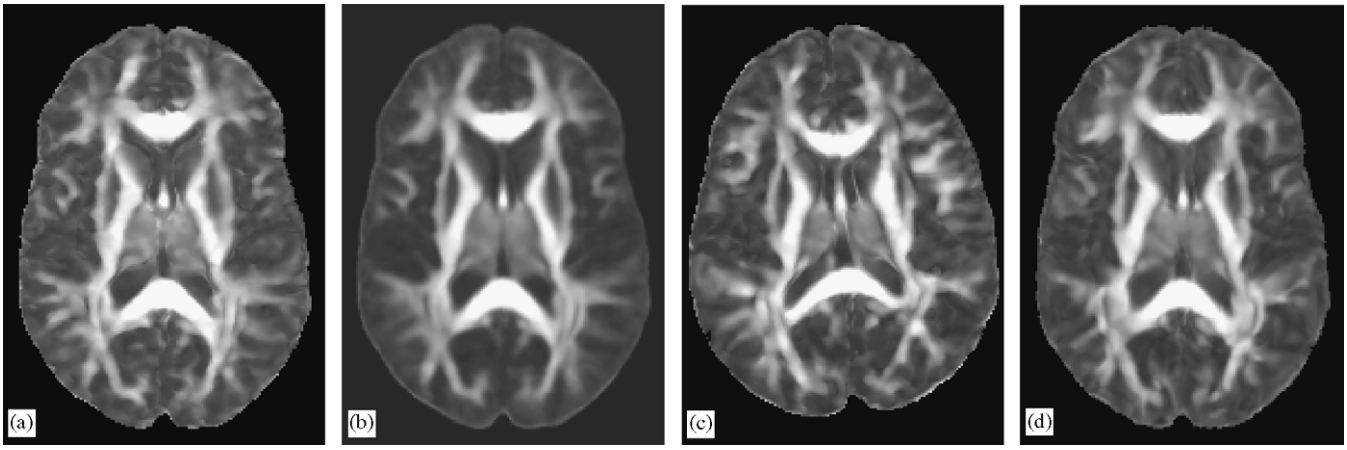

Fig. 8. The performance of the proposed image registration method in normalizing the FA maps of DTI brain images. The model used for normalization is shown in (a), and the average of 10 normalized FA maps is displayed in (b). Additionally, (c) and (d) demonstrate a subject before normalization and after normalization, using the proposed image registration method.

this new registration algorithm is that we can directly register two brain images without any tissue segmentation. This is particularly important for some images, as demonstrated next, where the tissue segmentation is difficult to perform.

3.2. DTI brain images

We also test the performance of our histogram-based registration method in warping fractional anisotropy (FA) maps of DTI brain images. It can be observed from Fig. 8 that FA maps are usually very noisy, which thus requires the registration algorithms be robust to noise. We test the performance of our registration method in registering individual FA maps, and also in averaging FA maps of 10 individuals. Fig. 8c and d show FA maps of a subject, before normalization and after normalization to the model (Fig. 8a). We can see that the normalized image of this subject is very similar to the model. The accuracy of registration can be further evaluated by checking the sharpness of the average FA map of 10 individuals, after spatial normalization to the model space. We can see a relatively clear average in Fig. 8b, partially indicating the accuracy of the proposed registration method. Actually, the average FA map is also relatively clear in the cortex, as shown by its 3D rendering in Fig. 9b. Compared to the 3D renderings of the model in Fig. 9a, the average FA map seems much clearer than the model. This is because a part of noise has been removed by averaging across accurately co-registered individuals.

3.3. CT pelvis images

The proposed registration method is also applicable in registering CT images of human pelvises, as shown in Fig. 10. Before registration, the shapes of two individual pelvises and also their internal structures are very different, according to the cross-sectional images and 3D renderings in Fig. 10a and c. After image registration, the two individual images become very similar, not only in their global shapes but their representation of internal tissues, as shown in Fig. 10b.

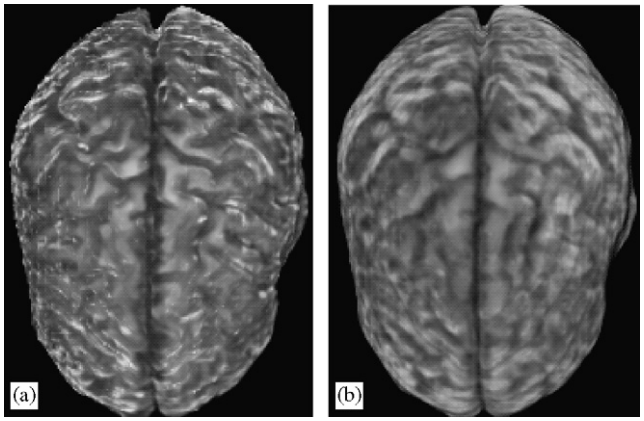

Fig. 9. Comparing the similarity of the model (a) and the average image of 10 normalized (b) FA maps, by using their 3D renderings. The cortex in the average image remains relatively clear.

3.4. MR mouse images

Our registration method is used to register the MR images of whole-body mice, as shown in Fig. 11. Before registration, the images of the two mice show differences in global shapes, such as the back of one mouse appearing bent, and the images also show differences in the shapes of internal organs, such as ventricles and bones. These differences are indicated by cross-sectional images, in the left and the right images of Fig. 11. After image-based registration, the images of the right mouse are deformed into the warped images as shown in the middle panel, which is very similar to the image of the mouse in the left that is used as the model. The registration algorithm did good job of registering the global

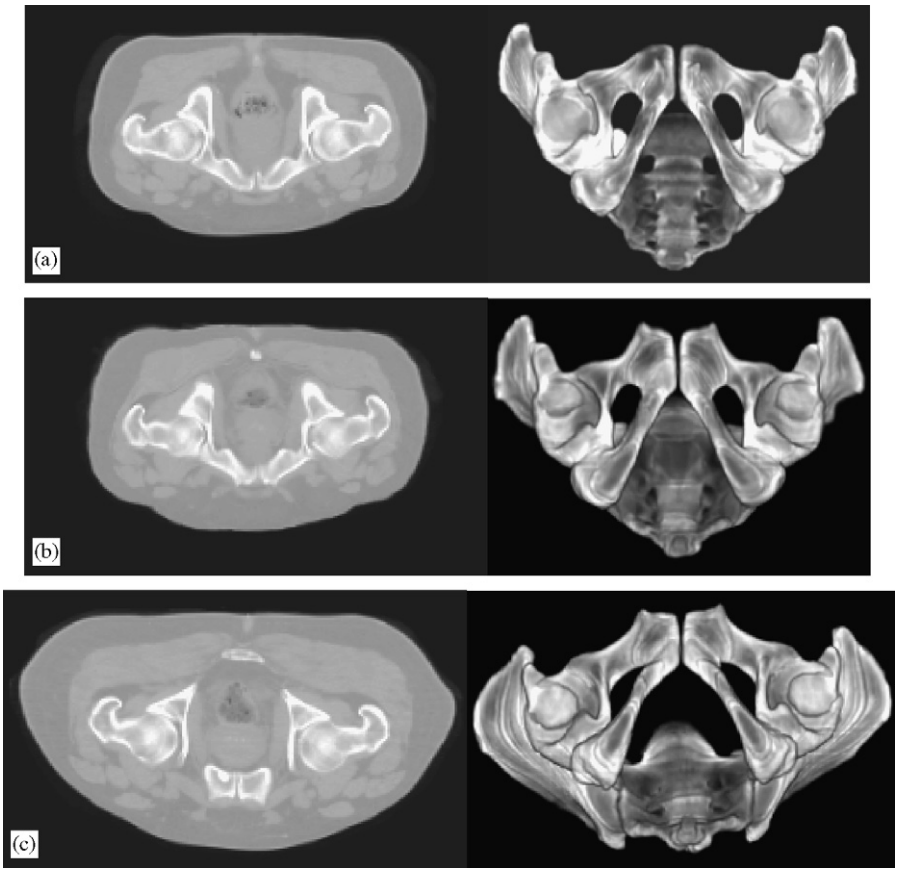

Fig. 10. Demonstration of the proposed method in registering CT images of pelvises. A subject in (c) is normalized to the space of a model in (a), resulting in a spatially normalized image in (b), which is very similar to the model. Both cross-sectional images and 3D renderings are provided for comparison.

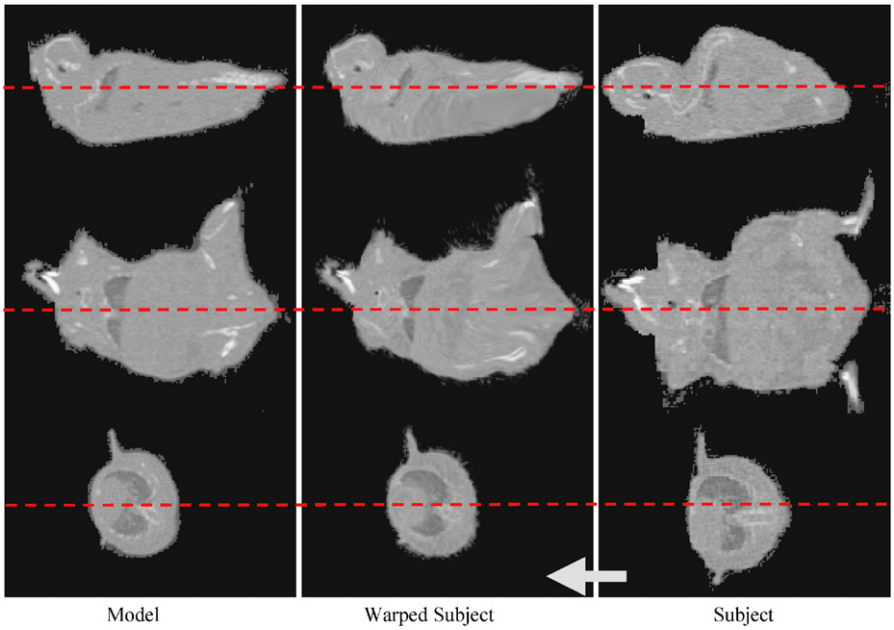

Fig. 11. The performance of the proposed method in registering MR images of whole-body mice. A subject in the right is warped to a model in the left, resulting in a warped subject as shown in the middle panel, which is very similar to the model.

shape of mice, as well as in registering internal structures such as ventricles. The dashed lines in Fig. 11 are used as landmarks for easy comparison of the shapes of the mice, before warping and after warping.

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 HAMMER was defined from tissue-segmented images. In this paper, the attribute vector 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 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 intensity histograms of a sufficient number of multi-resolution images.

The performance of the proposed method has been tested on image registration of different organs at different modalities, such as MR brain images, DTI brain images, CT pelvis images, and MR mouse images. In the future, we plan to evaluate our method on new testing data, i.e., additional CT pelvis images and MR mouse images. In particular, we will adapt the proposed method for the registration of images of developing mice, thereby establishing the possible relationship between structural changes and genetic effects.

Acknowledgment

The author would like to thank Dr. Susan Resnick and the BLSA for providing the brain data sets, Dr. Russell Taylor for providing pelvis data sets, and Dr. Susumu Mori for providing DTIs.

References

[1] C. Davatzikos, R.N. Bryan, Using a deformable surface model to obtain a shape representation of the cortex, IEEE Trans. Med. Imaging 15 (1996) 785–795.

[2] F.L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, IEEE Trans. Pattern Anal. Mach. Intell. 11 (6) (1989) 567–585.

[3] H. Chui, L. Win, R. Schultz, J. Duncan, A. Rangarajan, A unified feature registration method for brain mapping, Inf. Process. Med. Imaging 2001, pp. 300–314.

[4] J.P. Thirion, O. Monga, S. Benayoun, A. Gueziec, N. Ayache, Automatic registration of 3-D images using surface curvature, SPIE Proc. Math. Methods Med. Imaging 1768 (1992) 206–216.

[5] K. Rohr, Image registration based on thin plate splines and local estimates of anisotropic landmark localization uncertainties, Lecture Notes in Computer Science: MICCAI'98, vol. 1496, 1999, pp. 1174–1183.

[6] M. Breijl, M. Sonka, Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples, IEEE Trans. Med. Imaging 19 (2000) 973–985.

[7] M. Vaillant, C. Davatzikos, Hierarchical matching of cortical features for deformable brain image registration, Lecture Notes in Computer Science: Information Processing in Medical Imaging, vol. 1613, June 1999, pp. 182–195.

[8] M.I. Miller, G.E. Christensen, Y. Amit, U. Grenander, Mathematical textbook of deformable neuroanatomies, Proc. Nat. Acad. Sci. USA 90 (1993) 11944–11948.

[9] P. Thompson, A.W. Toga, A surface-based technique for warping three-dimensional images of the brain, IEEE Trans. Med. Imaging 15 (1996) 402–417.

[10] S.C. Joshi, M.I. Miller, G.E. Christensen, A. Banerjee, T. Coogan, U. Grenander, Hierarchical brain mapping via a generalized Dirichlet solution for mapping brain manifolds, in: Proceedings of the SPIE Conference on Geometry Methods in Applied Imaging, vol. 2573, July 1995, pp. 278–289.

[11] A.C. Evans, W. Dai, L. Collins, P. Neeling, S. Marett, Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis, SPIE Proc. Image Process. 1445 (1991) 236–246.

[12] G.E. Christensen, R.D. Rabbitt, M.I. Miller, Deformable templates using large deformation kinematics, IEEE Trans. Image Process. 5 (9) (1996).

[13] J.C. Gee, C. Barillot, L.L. Briquer, D.R. Haynor, R. Bajcsy, Matching structural images of the human brain using statistical and geometrical image features, Proc. SPIE Visualization Biomed. Comput. 2359 (1994) 191–204.

[14] J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons, Med. Image Anal. 2 (3) (1998) 243–260.

[15] K.J. Friston, A.P. Holmes, K. J Worsley, J.P. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Human Brain Mapp. (1995) 189–210.

[16] P.E. Roland, C.J. Graufelds, J. Wahlin, et al., Human brain atlas: for high-resolution functional and anatomical mapping, Human Brain Mapp. 1 (1994) 173–184.

[17] P.T. Fox, M.A. Mintum, E.M. Reiman, M.E. Reichle, Enhanced detection of focal brain responses using inter-subject averaging and distribution analysis of subtracted PET images, J. Cerebral Flow Metab. 8 (1988) 642–653.

[18] R. Bajcsy, R. Lieberson, M. Reivich, A computerized system for the elastic matching of deformed radiographic images to idealized atlas images, J. Comput. Assisted Tomography 7 (4) (1983) 618–625.

[19] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, D.J. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, IEEE Trans. Med. Imaging 18 (8) (1999) 712–721.

[20] R. Gan, A.C.S. Chung, Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude, IPMI'05, Glenwood Springs, Colorado, USA, July 10–15, 2005, Lecture Notes in Computer Science, vol. 3565, pp. 210–221.

[21] D. Shen, C. Davatzikos, HAMMER: hierarchical attribute matching mechanism for elastic registration, IEEE Trans. Med. Imaging 21 (11) (2002) 1421–1439.

[22] D. Shen, C. Davatzikos, Very high resolution morphometry using mass-preserving deformations and HAMMER elastic registration, NeuroImage 18 (1) (2003) 28–41.

[23] Z. Xue, D. Shen, C. Davatzikos, Determining correspondence in 3D MR brain images using attribute vectors as morphological signatures of voxels, IEEE Trans. Med. Imaging 23 (10) (2004) 1276–1291.
[24] M. Stricker, M. Orengo, Similarity of color images, in: Proceedings of the Storage and Retrieval for Image and Video Databases III, SPIE Conference Proceedings, vol. 2420, San Jose, CA, USA, February, 1995, pp. 381–392.
[25] H. Zhang, C. Low, W. Smoliar, J. Wu, Video parsing, retrieval and browsing: an integrated and content-based solution, ACM Multimedia (1995) 15–24.
[26] E. Hadjidemetriou, M. Grossberg, S.K. Nayar, Spatial information in multiresolution histograms, in: Proceedings of IEEE 2001 Conference on Computer Vision and Pattern Recognition, 2001.
[27] M. Carlotto, Histogram analysis using a scale-space approach, IEEE Trans. PAMI 9 (1) (1987) 121–129.
[28] B. Mel, Combining color shape and texture histogramming in a neurally inspired approach to visual object recognition, Neural Comput. 9 (1997) 777–804.
[29] J. Canny, A computational approach to edge detection, IEEE Trans. PAMI 8 (6) (1986).
[30] D. Shen, H.H.S. Ip, Generalized affine invariant image normalization, IEEE Trans. Pattern Anal. Mach. Intell. 19 (5) (1997) 431–440.
[31] L.G. Nyul, J.K. Udupa, X. Zhang, New variants of a method of MRI scale standardization, IEEE TMI 19 (2) (2000) 143–150.
[32] G.E. Christensen, H.J. Johnson, Consistent image registration, IEEE TMI 20 (7) (2001) 568–582.
[33] H.J. Johnson, G.E. Christensen, Consistent landmark and intensity-based image registration, IEEE TMI 21 (5) (2002) 450–461.
[34] S.M. Resnick, A.F. Goldszal, C. Davatzikos, S. Golski, M.A. Kraut, E.J. Metter, R.N. Bryan, A.B. Zonderman, One-year age changes in MRI brain volumes in older adults, Cereb. Cortex 10 (2000) 464–472.
[35] A.F. Goldszal, C. Davatzikos, D.L. Pham, M.X.H. Yan, R.N. Bryan, S.M. Resnick, An image-processing system for qualitative and quantitative volumetric analysis of brain images, J. Comput. Assisted Tomography 22 (5) (1998) 827–837.

About the Author—DINGGANG SHEN received all of his degrees from Shanghai JiaoTong University. He is an assistant professor (tenure-track) in the Department of Radiology at University of Pennsylvania (Upenn) since July 2002. Before moving to Upenn, he was a tenure-track faculty member in Johns Hopkins University. Dr. Shen is on the Editorial Board of Pattern Recognition, and served as a reviewer for numerous international journals and conferences. He has published over 140 articles in journals and proceedings of international conferences. His research interests include medical image analysis, pattern recognition, and computer vision.