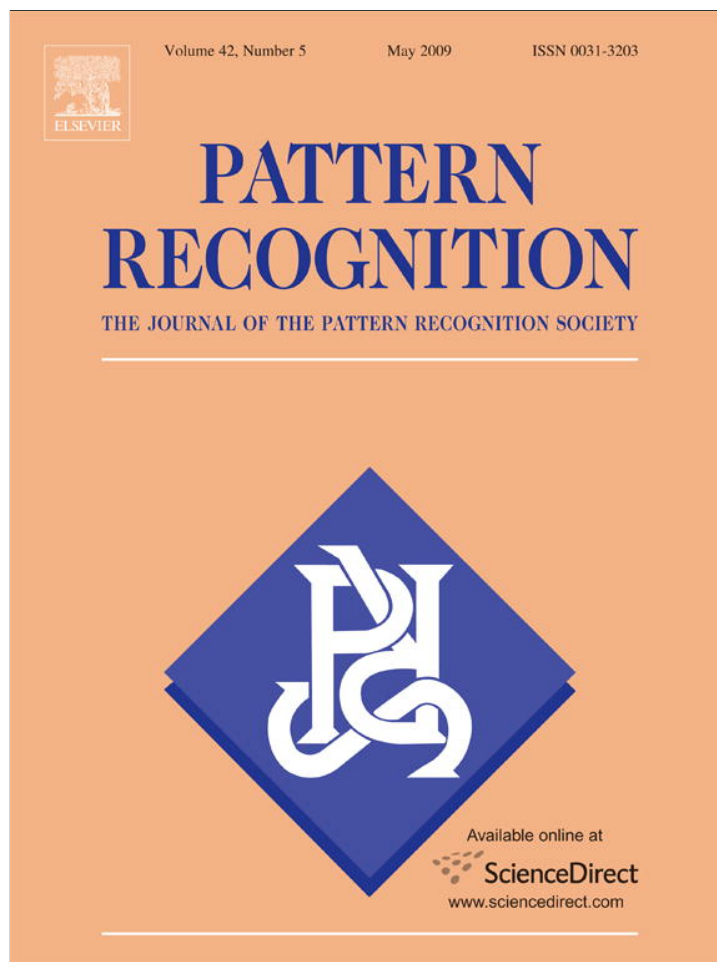


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/prFast image registration by hierarchical soft correspondence detection[☆]

Dinggang Shen

Department of Radiology and Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, NC 27599, USA

ARTICLE INFO

Article history:

Received 1 February 2008

Received in revised form 18 June 2008

Accepted 27 August 2008

Keywords:

Deformable registration

Non-rigid registration

Feature matching

Local descriptor

ABSTRACT

A new approach, based on the hierarchical soft correspondence detection, has been presented for significantly improving the speed of our previous HAMMER image registration algorithm. Currently, HAMMER takes a relative long time, e.g., up to 80 min, to register two regular sized images using Linux machine (with 2.40 GHz CPU and 2-Gbyte memory). This is because the results of correspondence detection, used to guide the image warping, can be ambiguous in complex structures and thus the image warping has to be conservative and accordingly takes long time to complete. In this paper, a hierarchical soft correspondence detection technique has been employed to detect correspondences more robustly, thereby allowing the image warping to be completed straightforwardly and fast. By incorporating this hierarchical soft correspondence detection technique into the HAMMER registration framework, both the robustness and the accuracy of registration (in terms of low average registration error) can be achieved. Experimental results on real and simulated data show that the new registration algorithm, based on the hierarchical soft correspondence detection, can run nine times faster than HAMMER while keeping the similar registration accuracy.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Deformable registration is important for comparing medical images across different subjects. So far, various image registration algorithms [1–19] have been proposed in the literature. Most registration methods are based on matching simple image information such as intensities, while other methods use features for image registration. Features can be local features [4,20] such as image boundary or curvature, or regional features such as geometric moment invariants (GMIs) [21,22], wavelet coefficients [23], multiple-scale histogram features [24], local descriptor features [25], or combinations of various local features [25]. In the image registration, besides maximizing the matching of image features, the estimated dense deformation field is required to be smooth as well. This can be achieved by using a Laplacian smoothness constraint [21,22], an elastic energy [8], or even the statistical deformation constraint learned from the training samples [26].

We have previously developed a hierarchical image registration algorithm, called HAMMER [21,22]. There are two novel ideas in this algorithm. *First*, an attribute vector is defined for each point in the image, in order to characterize and better distinguish this

point from its neighboring points; the image registration can be accordingly formulated as a problem of attribute vector matching or correspondence detection. *Second*, a strategy for hierarchical approximation of energy function is used to reduce the chance of local minima in the registration. In particular, the registration algorithm is designed to focus initially on the registration of only a small set of active points (with distinctive attribute vectors). This process implicitly selects only a small number of major energy terms in the whole energy function for optimization, thus potentially achieving global optimization for the new energy function. With the progress of registration, more and more points with less distinctive attribute vectors are added into the set of active points for actively driving the image registration, once these less distinctive points have been moved close to their targets and become stable for registration. In the final registration stage, all points in the images are considered as active points to drive image registration, which means that all energy terms in the whole energy function will be finally considered for optimization. By using this hierarchical registration strategy, we can improve the performance as well as the speed of registration. Also, since not all points are initially considered as active points to drive image registration, thus effectively reducing the scale of image registration problem and becoming fast and robust to complete.

HAMMER has been successfully applied to many large clinical research studies and clinical trials involving over 5000 MR brain images. These applications include one of the largest longitudinal

[☆] This work is partly supported by NIH grants 1R03EB008760, 7R03MH076970, and 1R01EB006733.

E-mail address: dgshen@med.unc.edu

studies of aging in the world to date—which involves an 18 year annual follow-up of 150 elderly individuals [27], a relatively large schizophrenia imaging study involving 148 participants [28], a study of cocaine-exposed children [29], a morphometric study of children with genetic abnormalities [30], a large study of the effects of organolead-exposure on the brain [31], and a study of effect of sustained, heavy drinking on the brain [32,33]. Many of these studies are the leading studies in their respective fields, and have received ample international attention.

To further improve the performance of HAMMER, various research works have been performed. In order to improve the performance of correspondence detection using attribute vectors, the best-scale GMIs have been learned for registration [34]. Also, the separate sets of best features have been specially learned for different parts of brain, since it is believed that different types of features have different functions in characterizing image points, and different parts of brain need different sets of best features to distinguish themselves. Moreover, by learning the normal brain variations from a set of brain image samples, we can use this learned information to constrain the image registration, thus improving the robustness of registration [25,26].

It is worth noting that all previous improvements were focusing on improving the robustness and accuracy of registration in HAMMER, e.g., using the average registration error. Less attention has been paid to increasing the registration speed. Currently, HAMMER takes up to 80 min for registering two MR brain images with the size of $256 \times 256 \times 124$, which might be too long for some applications. Therefore, it is important to design methods to improve the speed of registration algorithm.

One reason that HAMMER needs a relatively long time for registration is that the correspondence detection results, used to guide image registration, can be ambiguous in some complex structures, and image warping has to be performed conservatively and thus slow. Many advanced correspondence detection methods have been proposed in the literature [35,36]. Traditionally, one-to-one correspondence is established for each point in the image, which we can call as hard correspondence detection. Some advanced methods such as in [35,36] allow to identify correspondences in a soft way, e.g., each point can have multiple points as candidate correspondences, each assigned with a possibility. The total possibility is equal to 1. This is called soft correspondence detection. In the initial correspondence detection procedure, candidate correspondences can have similar amount of probability, indicating the uncertainty is high. With the progress of correspondence detection process, it will allow only a single correspondence or several candidate correspondences having the high probability and others close to zero, which increases the specificity of correspondence detection.

In this paper, we propose to detect soft correspondences [35,36] for active points selected during the registration procedure. Thus, correspondence detection can be robust and allows the image warping to perform fast. Moreover, soft correspondence detection can be implemented hierarchically to render the registration both robust and accurate. For example, in the initial registration procedure, the distant points with high attribute similarity can also contribute significantly to the soft correspondence detection, thus potentially increasing the robustness of correspondence detection. With the progress of registration, the distant points even with high attribute similarity decrease gradually their contribution to the soft correspondence detection, thus effectively increasing the specificity of final correspondence detection results. By mathematically formulating this soft correspondence detection idea into the HAMMER registration framework, the speed of the final registration algorithm is reduced to only the one-ninth of original computational cost required by HAMMER, without affecting much for registration performance. In particular, the proposed registration algorithm can complete the alignment of two $256 \times 256 \times 124$

images in about 9 min using Linux machine (with 2.40 GHz CPU and 2-Gbyte memory).

The structure of this paper is organized as follows. In Section 2, the definitions for attribute vector and similarity, a new energy function for registration, and details for implementation are all discussed one by one. In particular, the ideas of how to mathematically allow each active point to have multiple correspondences in the other image, and how to make soft correspondence detection strategy robust as well as accurate are also explained in detail. In Section 3, various experiments on simulated and real data are conducted to confirm that the performance of the proposed registration method is not decreased much even with the significant deduction of computation cost. This paper concludes in Section 4.

2. Method

Image registration problem is formulated here as an iterative process of detecting soft correspondences for the active points. The active points are selected according to the distinctiveness of their attribute vectors, which include image intensity, edge type, and geometric features from the neighborhood. The active points are also selected hierarchically during the registration procedure, e.g., initially a small set of active points with the most distinctive attribute vectors and gradually a larger and larger set of active points by including less distinctive points. The dense correspondences in the entire image domain are interpolated by the mean correspondences estimated in the active points.

The correspondence for each active point is softly determined. Instead of one-to-one correspondence detection, each active point is allowed to have multiple correspondences detected in the other image; each detected correspondence is assigned with a probability ranging from 0 to 1, and all the detected correspondences have a total probability of 1. The location of soft correspondence is determined as a statistical mean location of all detected correspondences. Note that, if a point \mathbf{v} is considered as a correspondence of an active point \mathbf{x} , the respective probability is determined by evaluating not only the similarity of attribute vectors of points \mathbf{v} and \mathbf{x} , but also the similarity of attribute vectors of all points in their neighborhoods. To increase the speed of detecting candidate correspondences and measuring their probabilities, the evaluation of the similarity of attribute vectors of the two points can be performed first, in order to select only a small set of candidate points $\{\mathbf{v}\}$ to perform the computation-consuming comparison of attribute vectors in the whole neighborhoods.

It is worth noting that the soft correspondence detection is more robust than the one-to-one correspondence detection, thus the robustness of registration using the soft correspondence detection can be improved. In the initial registration procedure, both local and distant points with similar attribute vectors can contribute to the correspondence detection. Gradually, in the later registration procedure, only the most similar points are allowed to contribute to the correspondence detection. By using these strategies, we can achieve both the robustness and the accuracy for the registration results.

In the following, we first briefly summarize the definitions for attribute vectors and their similarity, which are the same as in HAMMER [21,22]. Then, we give a definition of new energy function for the proposed registration algorithm. Finally, we provide the detailed descriptions of the soft correspondence matching within the original HAMMER registration framework.

2.1. Attribute vector and similarity

Definition: An attribute vector is defined for each point \mathbf{x} in the image, to better distinguish one point from its neighboring points and also to identify its corresponding points in the other image by matching attribute similarity. The attribute vector $a(\mathbf{x})$ includes three parts

of information, i.e., edge type $a_1(\mathbf{x})$, image intensity $a_2(\mathbf{x})$, and GMIs $a_3(\mathbf{x})$ that are calculated from the structure in the neighborhood. Hence, $\mathbf{a}(\mathbf{x}) = [a_1(\mathbf{x}) \ a_2(\mathbf{x}) \ a_3(\mathbf{x})]$; each of these three sub-vectors is explained one by one in the next three paragraphs.

For the *hard* tissue segmentation case [37] as employed in this paper, $a_1(\mathbf{x})$ is a 1×1 vector, which takes one of seven discrete values corresponding to seven edge types such as non-edge and six combinations of edges between grey matter (GM), white matter (GM), and CSF. In the *fuzzy* segmentation case [38], a 3×1 fuzzy vector $[C_{WM}, C_{GM}, C_{CSF}]^T$ is typically used to represent the probability of WM, GM, and CSF on an image point \mathbf{x} , where C_{WM} , C_{GM} , and C_{CSF} all take a value between 0 and 1 and $C_{WM} + C_{GM} + C_{CSF} = 1.0$. Thus, in the fuzzy segmentation case, $a_1(\mathbf{x})$ is designed as a 3×1 vector to measure the maximal direct difference (3×1 vector) of the tissue segmentation results between the image point \mathbf{x} and any neighboring points.

$a_2(\mathbf{x})$ represents direct information on the image point \mathbf{x} , which can be image intensity, tissue segmentation result, or both. In this paper, $a_2(\mathbf{x})$ is selected as a 1×1 vector to represent the intensity on the image point \mathbf{x} , which has been normalized between 0 and 1.

For the brain images, the vector $a_3(\mathbf{x})$ comprises the GMIs of the respective GM, WM, and CSF, calculated from the multiple-scale spherical neighborhoods around the image point \mathbf{x} . For a particular scale and tissue type, 13 rotation invariants can be calculated from the zero-order, second-order, and third-order 3D regular moments [39]. Therefore, $a_3(\mathbf{x})$ can be a vector with $K = 3 \times 13 \times L$, where L is the number of scales used. Four rotation invariants that are formulated from the zero-order and the second-order moments are listed next, while other nine rotation invariants, formulated from the third-order moments and both the second-order and the third-order moments, can be found in Ref. [39].

$$I_1 = M_{0,0,0},$$

$$I_2 = M_{2,0,0} + M_{0,2,0} + M_{0,0,2},$$

$$I_3 = M_{2,0,0}M_{0,2,0} + M_{2,0,0}M_{0,0,2} + M_{0,2,0}M_{0,0,2} - M_{1,0,1}^2 - M_{1,1,0}^2 - M_{0,1,1}^2,$$

$$I_4 = M_{2,0,0}M_{0,2,0}M_{0,0,2} - M_{0,0,2}M_{1,1,0}^2 + 2M_{1,1,0}M_{1,0,1}M_{0,1,1} - M_{0,2,0}M_{1,0,1}^2 - M_{2,0,0}M_{0,1,1}^2.$$

Note that the first GMI corresponds to the volume of a certain tissue.

Similarity: As mentioned above, there are three parts of attributes included in each attribute vector, with different parts of attributes having different contributions to the overall similarity of attribute vectors. For example, $a_1(\mathbf{x})$ represents the edge type, which provides important information for matching the same type of boundary points in the two individual brain images. $a_3(\mathbf{x})$ provides the geometric features, to allow the detailed comparison of geometric similarity between two images. Moreover, the sizes of these three different parts of attributes are different. Thus, we can first compare the similarity of short vectors of important attributes, such as $a_1(\mathbf{x})$, in the two attribute vectors, in order to avoid comparing the similarity of long vectors of attributes, such as $a_3(\mathbf{x})$, if the similarity on short vectors is low. In this way, we can significantly reduce the overall computation time for similarity measurement. In particular, the dissimilarity of two attribute vectors, $\mathbf{a}(\mathbf{x})$ and $\mathbf{a}(\mathbf{y})$, of two points \mathbf{x} and \mathbf{y} can be defined as

$$d(\mathbf{a}(\mathbf{x}), \mathbf{a}(\mathbf{y})) = \begin{cases} 1 & \text{if } |a_1(\mathbf{x}) - a_1(\mathbf{y})| > \varepsilon, \\ 1 - c([a_2(\mathbf{x})a_3(\mathbf{x})], [a_2(\mathbf{y})a_3(\mathbf{y})]) & \text{otherwise.} \end{cases}$$

If $|a_1(\mathbf{x}) - a_1(\mathbf{y})|$ is larger than ε , the overall dissimilarity $d(\mathbf{a}(\mathbf{x}), \mathbf{a}(\mathbf{y}))$ is one. In the hard tissue segmentation case, ε is equal to zero; in the

fuzzy segmentation case, ε is a certain value that should be determined in the experiments. $c([a_2(\mathbf{x}) \ a_3(\mathbf{x})], [a_2(\mathbf{y}) \ a_3(\mathbf{y})])$ is the similarity of the second and the third parts of the two attribute vectors in $\mathbf{a}(\mathbf{x})$ and $\mathbf{a}(\mathbf{y})$. To be strict on the similarity measure, each element in a_2 and a_3 is first normalized into a range from 0 to 1, and then compared and further integrated as a way described next:

$$c([a_2(\mathbf{x})a_3(\mathbf{x})], [a_2(\mathbf{y})a_3(\mathbf{y})]) = (1 - |a_2(\mathbf{x}) - a_2(\mathbf{y})|) \times \prod_{i=1}^K (1 - |a_3^i(\mathbf{x}) - a_3^i(\mathbf{y})|),$$

where $a_3^i(\mathbf{x})$ is the i -th element of $a_3(\mathbf{x})$ which has total K elements.

2.2. Energy function

The goal of image registration is to best align a template image $T(\mathbf{x})$ with a subject image $S(\mathbf{y})$ by maximizing the similarity of attribute vectors on the corresponding points. If a template point \mathbf{x} is corresponding to a subject point \mathbf{y} , their attribute vectors $a_T(\mathbf{x})$ and $a_S(\mathbf{y})$ should be as similar as possible, i.e., $d(a_T(\mathbf{x}), a_S(\mathbf{y}))$ should be as small as possible. On the other hand, the estimated deformation field $\{h(\mathbf{x}), \mathbf{x} \in \Omega_T\}$, where Ω_T denotes the template image domain, should be able to warp the template point \mathbf{x} very close to its corresponding subject point \mathbf{y} , i.e., $h(\mathbf{x}) = \mathbf{y}$.

It is important to allow each active point to have multiple correspondences detected in the other image, in order to increase the robustness of overall registration algorithm. This is significant particularly when the one-to-one correspondence detection is difficult. Accordingly, in this paper, each active point \mathbf{x} in the template is allowed to have multiple points $\{\mathbf{v}\}$ detected in the subject image as correspondences, with $\{p_{\mathbf{x},\mathbf{v}}\}$ representing the probability of a template active point \mathbf{x} and a subject point \mathbf{v} as a pair of corresponding points. The deformation field $h(\mathbf{x})$ in the template point \mathbf{x} can be thus estimated as a statistical mean location of all correspondences detected in the subject, i.e., $h(\mathbf{x}) = \sum_{\mathbf{v}} p_{\mathbf{x},\mathbf{v}} \cdot \mathbf{v}$. Note that the soft or fuzzy correspondence will be progressively forced to the one-to-one correspondence during the registration procedure, thus achieving not only the robustness but also the accuracy of registration. The robustness is achieved by *initially* allowing all points with *similar* attribute vectors to contribute to the correspondence detection, while the accuracy or the specificity of registration is achieved by *finally* allowing only the *most similar* points to contribute to the correspondence detection.

Our registration algorithm can be accordingly formulated as an iterative optimization process involving three steps. The *first* step is to select active points from both template and subject images, with the number of active points progressively increased during the registration. The *second* step is to detect multiple candidate correspondences and respective probabilities for each active point; candidate points can be selected initially from a large neighborhood and later from a small or even immediate neighborhood. The *third* step is to calculate the mean deformation for each template active point based on the detected correspondences, and also to interpolate the mean deformations in the other locations of the template image domain and make them smooth spatially. By iteratively performing these three steps, the registration algorithm can register images robustly and accurately.

The energy function is defined next to facilitate the iterative performance of these three steps in the registration procedure.

Energy terms: The first term in the whole energy function describes the overall cost of matching a template image with a subject image (see Fig. 1 for reference). For a template point \mathbf{x} , its mean deformation is defined as $h(\mathbf{x})$. For a subject point \mathbf{v} in the neighborhood of $h(\mathbf{x})$, $\mathbf{v} \in n_1(h(\mathbf{x}))$, its probability of being selected as a correspondence for the template point \mathbf{x} is $p_{\mathbf{x},\mathbf{v}}$. n_1 denotes

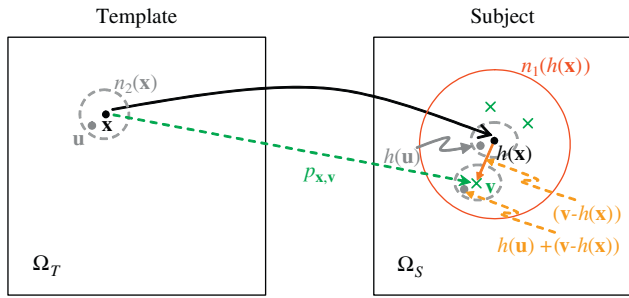


Fig. 1. Schematic explanation for correspondence detection, probability estimation, and neighborhood matching in our registration algorithm. A template point \mathbf{x} is tentatively warped to a location $h(\mathbf{x})$ in the subject. In the subject space, a neighboring point \mathbf{v} of $h(\mathbf{x})$ is considered as a candidate correspondence for the template point \mathbf{x} , with the probability $p_{\mathbf{x},\mathbf{v}}$ calculated from the similarity of attribute vectors on points \mathbf{x} and \mathbf{v} , as well as the similarity of attribute vectors in the respective neighborhoods (e.g., a dotted grey circle n_2 in the template and one deformed dotted grey circle around point \mathbf{v} in the subject). The deformed dotted circle around point \mathbf{v} is shifted by $\mathbf{v}-h(\mathbf{x})$ from the one around point $h(\mathbf{x})$, which is a warped version of an original circle in the template. In the searching neighborhood n_1 (red circle), only several points (green cross points) are detected as candidate correspondences for template point \mathbf{x} .

a searching neighborhood around the mean deformation $h(\mathbf{x})$; the size of n_1 is large initially, and decreases gradually with the progress of registration, in order to achieve the robustness as well as the accuracy of registration. Three types of costs need to be minimized once a subject point \mathbf{v} is selected as a correspondence of the template point \mathbf{x} . The *first* type of cost is the dissimilarity of attribute vectors of the template point \mathbf{x} and the subject point \mathbf{v} , i.e., $d(a_T(\mathbf{x}), a_S(\mathbf{v}))$. The *second* type of cost is the overall dissimilarity in the neighborhoods around the template point \mathbf{x} and the subject point \mathbf{v} , i.e.,

$$\frac{1}{\sum_{\mathbf{u} \in n_2(\mathbf{x})} 1} \sum_{\mathbf{u} \in n_2(\mathbf{x})} d(a_T(\mathbf{u}), a_S(h(\mathbf{u}) + (\mathbf{v} - h(\mathbf{x})))),$$

where all neighboring points $\{\mathbf{u}\}$ of the template point \mathbf{x} , $\mathbf{u} \in n_2(\mathbf{x})$, are compared with the neighboring points $h(\mathbf{u}) + (\mathbf{v} - h(\mathbf{x}))$ of the subject point \mathbf{v} , by using attribute vector dissimilarity. n_2 represents a matching neighborhood, whose size changes from an initial large size to a final small size with the progress of registration. The *third* type of cost is the distance of the subject point \mathbf{v} away from the mean deformation $h(\mathbf{x})$, i.e., $\|\mathbf{v} - h(\mathbf{x})\|^2$. By integrating all of these three types of cost from all template points, the total cost for aligning the template with the subject images can be obtained as the first energy term in whole energy function.

For increasing the consistency of registration [5,21,22,40,41], we also define a cost function for matching the subject image with the template image, as described by the second energy term in the whole energy function. $h^{-1}(\mathbf{y})$ is the backward transformation from the subject to the template, which is the inverse of the forward transformation $h(\mathbf{x})$. Cost definitions in the second energy term are similar to those in the first energy term.

As mentioned, by allowing each point to have multiple correspondences, the robustness of registration algorithm can be increased. Also, in order to increase the specificity or the accuracy of registration results, it is necessary to *finally determine only a single correspondence for each point*, or *require all the detected correspondences very close to the mean deformation $h(\mathbf{x})$* . The latter requirement can be implemented by increasing the parameter σ (in the energy function) with the progress of registration, to constrain the distance between each correspondence (\mathbf{v}) and the mean location of all correspondences ($h(\mathbf{x})$). Also, the progressive decrease of the size of searching neighborhood n_1 ensures the selection of candidate correspondences from the nearby points. The *former requirement* can be implemented by adding an entropy measurement on the

probabilities $\{p_{\mathbf{x},\mathbf{v}}\}$ and $\{p_{\mathbf{v},\mathbf{y}}\}$, which is included as the third energy term in the whole energy function. The parameter τ controls the significance of this energy term [42,43]. By adaptively adjusting parameter τ with the progress of registration, i.e., changing from extremely large to extremely small value, we can actively allow the registration algorithm to detect multiple correspondences from an initial large neighborhood to latter a small neighborhood, and also from initially even the low similarity points to latter only the high similarity points.

The last energy term evaluates the smoothness of the mean deformation field $\{h(\mathbf{x})\}$ estimated in the entire image domain, using a Laplacian operator. Based on the multiple correspondences and the respective probabilities detected for each active point, the mean deformations $\{h(\mathbf{x})\}$ in all active points $\{\mathbf{x}\}$ can be statistically determined and then used to refine the previous estimations in the same active points. The mean deformations in the other locations of the image can be interpolated with an interpolation kernel, e.g., a Gaussian kernel used in this paper. Finally, this updated deformation field can be made as regular as possible in the entire image domain by using a Laplacian smoothness constraint. A parameter β is used to control the importance of the Laplacian smoothness constraint in the whole energy function.

The following is the whole energy function that our image registration algorithm will minimize:

$$\begin{aligned} E = & \sum_{\mathbf{x} \in \Omega_T} \omega_T(\mathbf{x}) \sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} p_{\mathbf{x},\mathbf{v}} \left\{ d(a_T(\mathbf{x}), a_S(\mathbf{v})) \right. \\ & + \frac{1}{\sum_{\mathbf{u} \in n_2(\mathbf{x})} 1} \sum_{\mathbf{u} \in n_2(\mathbf{x})} d(a_T(\mathbf{u}), a_S(h(\mathbf{u}) + (\mathbf{v} - h(\mathbf{x})))) \\ & \left. + \sigma \cdot \|\mathbf{v} - h(\mathbf{x})\|^2 \right\} \\ & + \sum_{\mathbf{y} \in \Omega_S} \omega_S(\mathbf{y}) \sum_{\mathbf{v} \in n_1(h^{-1}(\mathbf{y}))} p_{\mathbf{v},\mathbf{y}} \left\{ d(a_T(\mathbf{v}), a_S(\mathbf{y})) \right. \\ & + \frac{1}{\sum_{\mathbf{u} \in n_2(\mathbf{y})} 1} \sum_{\mathbf{u} \in n_2(\mathbf{y})} d(a_T(h^{-1}(\mathbf{u}) + (\mathbf{v} - h^{-1}(\mathbf{y}))), a_S(\mathbf{u})) \\ & \left. + \sigma \cdot \|\mathbf{v} - h^{-1}(\mathbf{y})\|^2 \right\} + \tau \left\{ \sum_{\mathbf{x} \in \Omega_T} \sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} p_{\mathbf{x},\mathbf{v}} \log p_{\mathbf{x},\mathbf{v}} \right. \\ & \left. + \sum_{\mathbf{y} \in \Omega_S} \sum_{\mathbf{v} \in n_1(h^{-1}(\mathbf{y}))} p_{\mathbf{v},\mathbf{y}} \log p_{\mathbf{v},\mathbf{y}} \right\} + \beta \sum_{\mathbf{x}} \|\nabla^2 h(\mathbf{x})\|. \end{aligned}$$

Subject to the constraints:

$$0 \leq p_{\mathbf{x},\mathbf{v}} \leq 1, \quad \sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} p_{\mathbf{x},\mathbf{v}} = 1 \text{ for each template point } \mathbf{x};$$

$$0 \leq p_{\mathbf{v},\mathbf{y}} \leq 1, \quad \sum_{\mathbf{v} \in n_1(h^{-1}(\mathbf{y}))} p_{\mathbf{v},\mathbf{y}} = 1 \text{ for each subject point } \mathbf{y}.$$

Hierarchical optimization: As mentioned, our registration algorithm will hierarchically select a set of active points with distinctive attribute vectors to drive the image registration. The set of active points will be progressively increased with the progress of registration by adding less distinctive points when they are deformed close to their targets and become reliable to drive the image registration. This strategy can be implemented by hierarchically assigning different values to the parameters $\omega_T(\mathbf{x})$ and $\omega_S(\mathbf{y})$ in the energy function, thus effectively selecting the respective template and subject points (\mathbf{x} and \mathbf{y}) to be actively involved in the image registration.

We can initially assign large values to the points with distinctive attribute vectors, and zero to other points in the images. In this case, the whole energy function above describes a problem of determining correspondences for those selected active points and also interpolating a smooth dense deformation field according to the mean correspondence locations determined in those active points. After solving this relative small scale problem of correspondence detection and deformation field estimation, we can increase the set of active points by including other less distinctive points, and then start to solve a relatively large scale problem of correspondence detection on more active points. Since the sizes of the searching neighborhood n_1 and the matching neighborhood n_2 will be decreased with the progress of registration, the computational cost for detecting correspondences for more active points will not be necessarily increased as confirmed by our experiments.

In the final, by considering all image points into the energy minimization problem, we solve the exact energy function that our registration algorithm should originally minimize. Importantly, the advance of using this hierarchical optimization strategy is that we can have high chances to avoid local minima. The subsection next describes the details for implementing this hierarchical optimization strategy.

2.3. Implementation

The whole energy function above can be hierarchically minimized by iteratively estimating (1) the correspondence probabilities $\{p_{\mathbf{x},\mathbf{v}}\}$ and $\{p_{\mathbf{v},\mathbf{x}}\}$ for the current selected active points, (2) the mean deformations in the current selected active points, and (3) the dense mean deformation field in the entire image domain. The descriptions for each of these three iterative steps are provided next.

Estimating correspondence probabilities: The probability for each pair of correspondences in $\{p_{\mathbf{x},\mathbf{v}}\}$ can be determined by solving an equation of $\partial E/\partial p_{\mathbf{x},\mathbf{v}} = 0$, with the result given next. Note that the probabilities in $\{p_{\mathbf{v},\mathbf{x}}\}$ can be similarly determined.

$$p_{\mathbf{x},\mathbf{v}} = c \cdot m(\mathbf{x},\mathbf{v}) \cdot m_{\text{Sub}}(\mathbf{x},\mathbf{v}) \cdot m_{\text{Dist}}(\mathbf{x},\mathbf{v}),$$

where c is a constant, $m(\mathbf{x},\mathbf{v})$ measures the similarity of attribute vectors between a template point \mathbf{x} and a subject point \mathbf{v} , and $m_{\text{Sub}}(\mathbf{x},\mathbf{v})$ represents the similarity of whole neighborhoods around the template point \mathbf{x} and the subject point \mathbf{v} , and $m_{\text{Dist}}(\mathbf{x},\mathbf{v})$ measures how close two points \mathbf{x} and \mathbf{v} are. Expressions of these measures are given next:

$$m(\mathbf{x},\mathbf{v}) = \exp\left(-\frac{\omega_T(\mathbf{x}) \cdot d(a_T(\mathbf{x}), a_S(\mathbf{v}))}{\tau}\right),$$

$$m_{\text{Sub}}(\mathbf{x},\mathbf{v}) = \exp\left(\frac{\omega_T(\mathbf{x}) \cdot \frac{1}{\sum_{\mathbf{u} \in n_2(\mathbf{x})} 1} \sum_{\mathbf{u} \in n_2(\mathbf{x})} d(a_T(\mathbf{u}), a_S(h(\mathbf{u}) + (\mathbf{v} - h(\mathbf{x}))))}{\tau}\right),$$

$$m_{\text{Dist}}(\mathbf{x},\mathbf{v}) = \exp\left(-\frac{\omega_T(\mathbf{x}) \cdot \sigma \|\mathbf{v} - h(\mathbf{x})\|^2}{\tau}\right).$$

The estimated probabilities $\{p_{\mathbf{x},\mathbf{v}}\}$ are normalized separately for each template point \mathbf{x} as follows, in order to ensure that the total probability for detecting correspondences for the point \mathbf{x} is 1.0.

$$p_{\mathbf{x},\mathbf{v}} = \frac{p_{\mathbf{x},\mathbf{v}}}{\sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} p_{\mathbf{x},\mathbf{v}}} \quad \text{for every template point } \mathbf{x}.$$

It is worth noting that the similarity of neighborhoods $m_{\text{Sub}}(\mathbf{x},\mathbf{v})$ is generally expensive to compute, if the number of points $\{\mathbf{v}\}$ in the neighborhood n_2 is large. To avoid this computation problem, we

propose to *first* calculate the measurement $m(\mathbf{x},\mathbf{v})$ for all neighboring points $\{\mathbf{v}\}$. Then, we check whether any neighboring points have the measurement $m(\mathbf{x},\mathbf{v})$ over a certain threshold. (Note that this threshold can be gradually adjusted with the progress of registration, e.g., from initially a large value (such as 0.8 in our paper) to later a small value.) Only for the points with the expected measurement $m(\mathbf{x},\mathbf{v})$, their neighborhood similarity $m_{\text{Sub}}(\mathbf{x},\mathbf{v})$ will be consequently computed. By using this strategy, we can potentially assign zero probability for many neighboring points \mathbf{v} , thus significantly reducing the overall computational time.

Estimating the mean deformations in active points: Once the probabilities $\{p_{\mathbf{x},\mathbf{v}}\}$ and $\{p_{\mathbf{v},\mathbf{x}}\}$ have been determined, the mean deformation, or the location of soft correspondence, for each active point can be immediately determined. By taking active points in the template image as an example, we can determine the mean deformation for each template active point \mathbf{x} by optimizing the sub-energy term, $\sum_{\mathbf{x} \in \Omega_T} \omega_T(\mathbf{x}) \sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} p_{\mathbf{x},\mathbf{v}} \left\{ \sigma \cdot \|\mathbf{v} - h(\mathbf{x})\|^2 \right\}$. The solution is $h(\mathbf{x}) = \sum_{\mathbf{v} \in n_1(h(\mathbf{x}))} (p_{\mathbf{x},\mathbf{v}} \cdot \mathbf{v})$. This means that the new location of soft correspondence for the active point \mathbf{x} is the statistical mean location of all correspondences detected in the subject image.

By determining the new location for each of these active points in the template and the subject images, we can deform all points in the neighborhoods of these active points by using a Gaussian kernel as employed in HAMMER [21,22]. By selecting the standard deviation of the Gaussian kernel as small as the one-third of the size of searching neighborhood n_1 , we can smoothly update the deformation field without introducing possible self-intersection. Note that the entire deformation field might be not sufficiently smooth due to this independent, local updating of deformations. This necessitates the use of Laplacian smoothness constraint to regularize the entire deformation field, as briefly summarized next.

Regularizing the dense deformation field: Once deformations have been updated in active points and their neighborhoods, the last energy term, e.g., Laplacian smoothness constraint, can be used to ensure both global and local smoothness of the entire deformation field. In the discrete case, the Laplacian smoothness constraint actually requires the deformation of each point to be close to the average deformation in its immediate neighborhood. Therefore, in the implementation, we can gradually push the deformation of each point to the average deformation in the immediate neighborhood, e.g., a $3 \times 3 \times 3$ neighborhood. By repeating this step, we can finally obtain a smooth dense deformation field $\{h(\mathbf{x})\}$.

3. Experimental results

The performance of the proposed registration algorithm has been evaluated using both simulated and real data. The results on simulated data show that the proposed registration method can achieve a registration accuracy comparable to the one produced by HAMMER, while the results on the real longitudinal data indicate that the proposed registration algorithm can achieve better longitudinal consistency than HAMMER. Importantly, the *speed* of the proposed registration algorithm has been significantly improved, e.g., taking only about the one-ninth of computation time required by HAMMER. For example, it takes about 9 min for registration of two $256 \times 256 \times 124$ brain images, using Linux machine (with 2.40 GHz CPU and 2-Gbyte memory).

Three subsections next compare the performances of our proposed registration algorithm and HAMMER in registering simulated data, constructing average brains from 18 individual brains, and aligning longitudinal data of real patients, respectively. The registration is performed in the isotropic image space. Note that all of these experimental results are provided to confirm that the performance of the proposed registration algorithm is comparable to that of HAMMER, even with the significant improvement of registration speed.

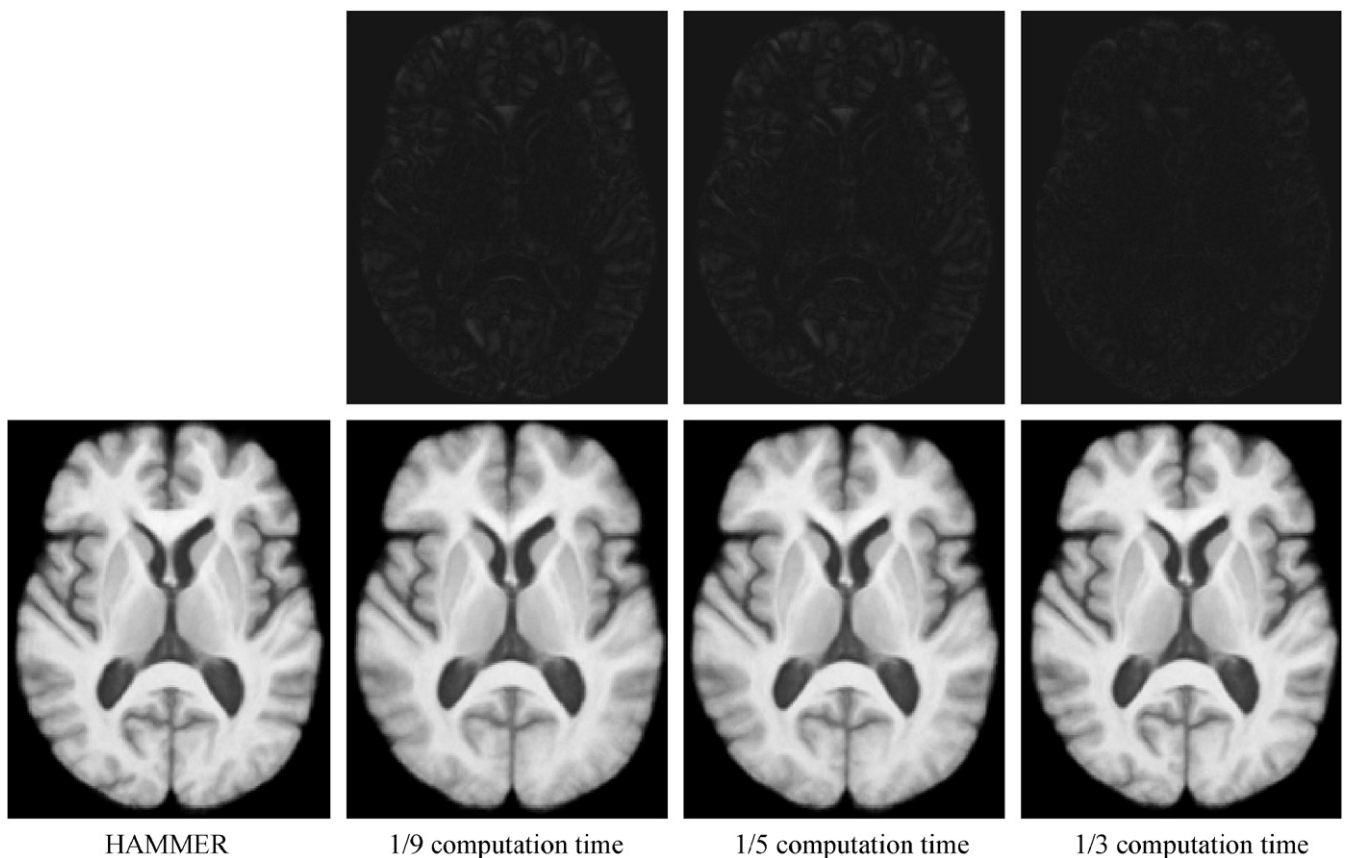


Fig. 2. Average brain images constructed from 18 normalized individual brain images by HAMMER, and by the proposed registration algorithm that took about $\frac{1}{9}$, $\frac{1}{5}$, or $\frac{1}{3}$ computation time of HAMMER. The difference image between the result by HAMMER and each result by the proposed method is provided in the top row. These different images have been enhanced for better visual inspection.

3.1. Simulated data

The proposed registration method has been applied to the simulated data for quantitatively measuring its registration performance. The same simulated data used to validate HAMMER in [21] are used here. The simulated data were created by *first* randomly selecting a brain, and *then* using the STAR elastic warping method [8] to warp this selected brain to five individual human brains by respectively placing forces on the ventricles and the cortex. *Finally*, five simulated brains were obtained, which are actually the different warped versions of the same selected brain. Note that the overall shapes of these five simulated brains are similar to those five individual brains used to guide the warping.

It is worth noting that the ground-truth deformations from the selected brain to each of these simulated brains are known. Thus, by applying the proposed registration method to estimating the deformations between the selected brain and each of these five simulated brains, we can quantitatively measure the accuracy of the proposed registration algorithm by comparing our estimated deformations with the ground-truth ones. The average registration error by the proposed algorithm is 1.13 mm, which is comparable to 0.97 mm error produced by HAMMER. Note that, as mentioned above, the proposed algorithm took only one-ninth of the computational cost required by HAMMER.

3.2. Average brain

The proposed registration algorithm has been applied to 18 real brain images for constructing an average brain from these 18 normalized brain images in the template space. The same 18 brain

images used in HAMMER [21] are used here for direct comparison of our constructed average brain with that by HAMMER.

Fig. 2 shows the average images constructed from 18 individual brain images using the proposed registration algorithm and HAMMER, respectively. The average brain image constructed by the proposed method (the 2nd from left in the bottom row) is relatively fuzzy compared to the one constructed by HAMMER (left). Considering that the proposed registration method takes only one-ninth of the computation time to align images, the quality of registration is still promising. If we spend a relatively long time for registration, i.e., about $\frac{1}{5}$ or $\frac{1}{3}$ of computation time used by HAMMER, we can obtain average brain images closer and closer to the one obtained by HAMMER.

3.3. Longitudinal data

The proposed registration method has been applied to the longitudinal data obtained from our Baltimore Longitudinal Study of Aging (BLSA) [27], to check the longitudinal stability of our registration algorithm in aligning images of the same subject acquired at different times. In the BLSA project, SPGR MR brain images have been acquired annually by a GE 1.5 T scanner for 150 elderly subjects over 13 years. We applied our registration algorithm to the first five images of the same patients (acquired in the first five years), and the performance is compared with that achieved by HAMMER.

Fig. 3 shows the results of five longitudinally acquired images collected once per year for five years of the same subject. These images were spatially normalized by HAMMER (1st row) and the proposed algorithm (2nd row), respectively. Since the normalization for each time image was performed independently, the normalization

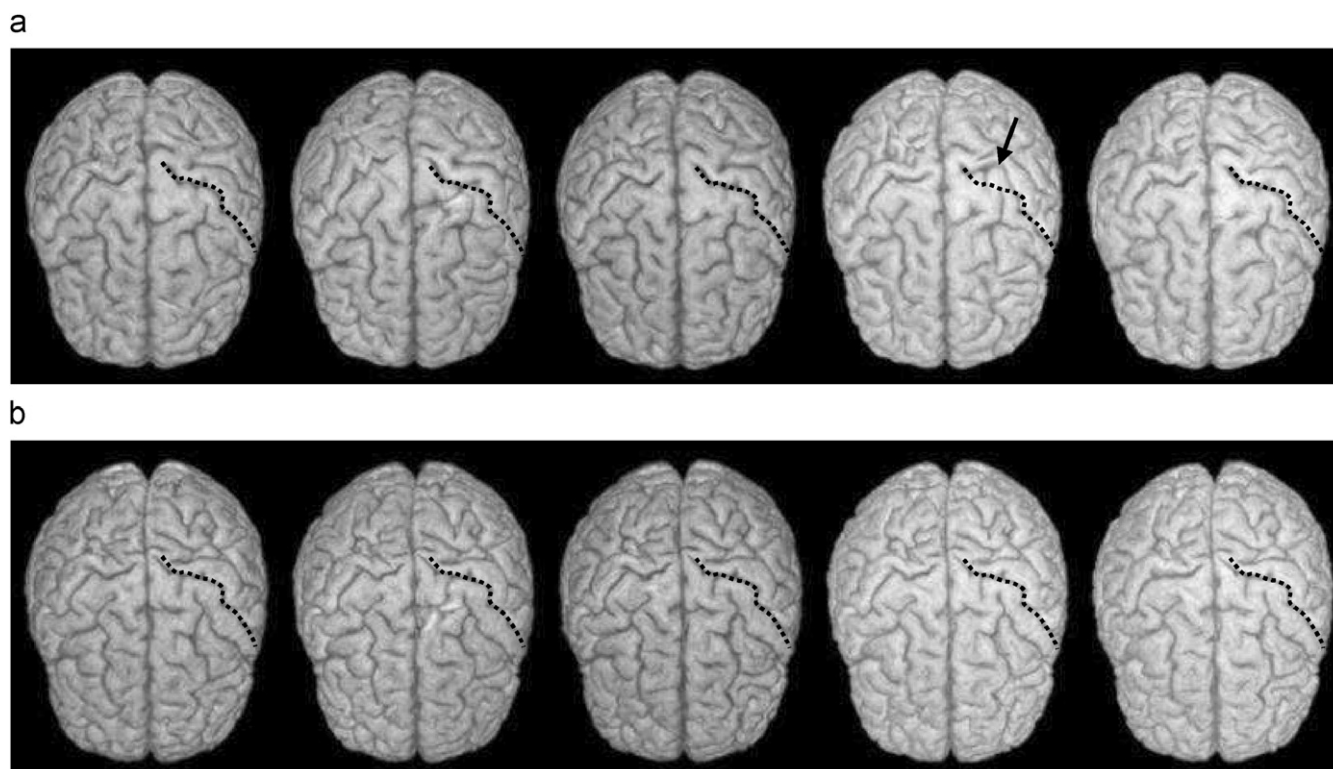


Fig. 3. Comparison on the longitudinal consistency of the registration on five different images of the same subject, using (a) HAMMER and (b) the proposed registration algorithm.

results might be not longitudinally consistent. It can be observed that the normalization results by HAMMER are not consistent in the fourth time image as indicated by the arrow. In contrast, the normalization results achieved by the proposed algorithm are more stable across different times, indicating the better performance of the proposed algorithm than HAMMER in aligning longitudinal data. On the other hand, it is worth noting that the normalization results from these two methods are somewhat different, particularly in the light hemisphere sulcal patterns (right in the image). More evaluations need to be done in the future for complete comparison of the two methods.

4. Conclusion

We have presented a new approach to significantly improve the speed of our previous HAMMER registration algorithm. This is achieved by using a hierarchical soft correspondence detection technique to replace the one-to-one correspondence detection strategy in HAMMER, thus the correspondence detection can be performed more robustly and the image warping can be conducted straightforwardly and fast. In particular, for each active point selected during the registration procedure, it is allowed *initially* to have multiple correspondences detected in the other image, and *gradually* forced to have few (or even one) correspondences detected in the other image. By using this hierarchical soft correspondence detection technique, we can achieve not only the robustness but also the accuracy of registration.

In summary, the image registration is here formulated as a process of hierarchically detecting a soft correspondence for each active point selected during the registration procedure, and also a process of interpolating and regularizing the dense smooth deformation field in the entire image domain. The experimental results indicate that,

compared to HAMMER, *first* the speed of the proposed registration algorithm has been significantly increased without sacrificing much for the registration performance, and *second* the longitudinal consistency has been considerably improved.

In future work, we will extensively evaluate and optimize all the parameters in our proposed registration algorithm, by using more real and simulated datasets with ground-truth information provided by manual raters or computer simulations. The final registration algorithm will be released publicly, as we did for the HAMMER registration algorithm, <http://www.nitrc.org/projects/hammer/>.

References

- [1] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, G. Marchal, Automated multi-modality image registration based on information theory, in: Y. Bizais, C. Barillot, R.D. Paola (Eds.), *Information Processing in Medical Imaging*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995, p. 263274.
- [2] D.C. Alexander, J.C. Gee, Elastic matching of diffusion tensor images, *Comp. Vision Image Understand.* 77 (1999) 233–250.
- [3] J. Ashburner, K.J. Friston, Voxel-based morphometry: the methods, *Neuroimage* 11 (2000) 805–821.
- [4] F.L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 567–585.
- [5] G.E. Christensen, H.J. Johnson, Consistent image registration, *IEEE Trans. Med. Imaging* 20 (2001) 568–582.
- [6] A. Collignon, Multi-modality medical image registration by maximization of mutual information, Catholic University Leuven, Leuven, Belgium, 1998.
- [7] D.L. Collins, A.C. Evans, Automatic 3D estimation of gross morphometric variability in the human brain, *Neuroimage* 3 (1996) S129.
- [8] C. Davatzikos, Spatial transformation and registration of brain images using elastically deformable models, *Comp. Vision Image Understand.* 66 (1997) 207–222.
- [9] B.M. Dawant, S.L. Hartmann, S. Gadamsetty, Brain atlas deformation in the presence of large space-occupying tumours, in: *Lecture Notes in Computer Science: MICCAI'99*, vol. 1679, 1999, pp. 589–596.
- [10] K.J. Friston, J. Ashburner, C.D. Frith, J.B. Poline, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, *Hum. Brain Mapp.* 2 (1995) 165–189.

- [11] J.C. Gee, M. Reivich, R. Bajcsy, Elastically deforming 3D atlas to match anatomical brain images, *J. Comput. Assisted Tomography* 17 (1993) 225–236.
- [12] G. Gerig, M. Styner, J. Lieberman, Shape versus size: improved understanding of the morphology of brain structures, presented at MICCAI 2001, Utrecht, The Netherlands, 2001.
- [13] S.C. Joshi, M.I. Miller, G. Christensen, A. Banerjee, T. Coogan, U. Grenander, Hierarchical brain mapping via a generalized Dirichlet solution for mapping brain manifolds, *Proceedings of the SPIE Conference on Geometrical Methods in Applied Imaging* 2573 (1995) 278–289.
- [14] A. Mohamed, D. Shen, C. Davatzikos, Deformable registration of brain tumor images via a statistical model of tumor-induced deformation, presented at MICCAI, Palm Springs, CA, 2005.
- [15] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, D.J. Hawkes, Non-rigid registration using free-form deformations: application to breast MR images, *IEEE Trans. Med. Imaging* 18 (1999) 712–721.
- [16] J.P. Thirion, Non-rigid matching using demons, presented at, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1996.
- [17] 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.
- [18] Y. Wang, B.S. Peterson, L.H. Staib, 3D brain surface matching based on geodesics and local geometry, *Comput. Vision Image Understand.* 89 (2003) 252–271.
- [19] I. Wells, M. William, P. Viola, H. Atsumi, S. Nakajima, R. Kikinis, Multi-modal volume registration by maximization of mutual information, *Med. Image Anal.* 1 (1996) 35–51.
- [20] K. Rohr, Image registration based on thin plate splines and local estimates of anisotropic landmark localization uncertainties, in: *Lecture Notes on Computer Science: MICCAI'98*, vol. 1496, 1999, pp. 1174–1183.
- [21] D. Shen, C. Davatzikos, HAMMER: hierarchical attribute matching mechanism for elastic registration, *IEEE Trans. Med. Imaging* 21 (2002) 1421–1439.
- [22] D.G. Shen, C. Davatzikos, Very high resolution morphometry using mass-preserving deformations and HAMMER elastic registration, *NeuroImage* 18 (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 (2004) 1276–1291.
- [24] D. Shen, Image registration by hierarchical matching of local spatial intensity histograms, presented at MICCAI'04, St. Malo, France, 2004.
- [25] G. Wu, F. Qi, D. Shen, Learning best features and deformation statistics for hierarchical registration of MR brain images, presented at IPMI, The Netherlands, 2007.
- [26] Z. Xue, D. Shen, C. Davatzikos, Statistically-constrained high-dimensional warping using wavelet-based priors, presented at *Proceedings of IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2006)*, New York City, NY, USA, 2006.
- [27] S.M. Resnick, A. 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.
- [28] C. Davatzikos, D.G. Shen, X. Wu, Z. Lao, P. Huggett, B.I. Turetsky, R.C. Gur, R.E. Gur, Whole-brain morphometric study of schizophrenia reveals a spatially complex set of focal abnormalities, *JAMA Arch. Gen. Psychiatry* 62 (2005) 1218–1227.
- [29] Y. Fan, H. Rao, J. Giannetta, H. Hurt, J. Wang, C. Davatzikos, D. Shen, Diagnosis of brain abnormality using both structural and functional MR images, *IEEE 2006 International Conference of the Engineering in Medicine and Biology Society (EMBS)*, NY, USA, 2006.
- [30] D. Shen, D. Liu, H. Liu, L. Clasen, J. Giedd, C. Davatzikos, Automated morphometric study of brain variation in XXY males, *NeuroImage* 23 (2004) 648–653.
- [31] W.F. Stewart, B.S. Schwartz, C. Davatzikos, D. Shen, D. Liu, X. Wu, A.C. Todd, W. Shi, S. Bassett, D. Youssef, Past adult lead exposure is linked to neurodegeneration measured by brain MRI, *Neurology* 66 (2006) 1476–1484.
- [32] T.P. Beresford, D.B. Arciniegas, J. Alfors, L. Clapp, B. Martin, Y. Du, D. Liu, D. Shen, C. Davatzikos, Hippocampus volume loss due to chronic heavy drinking, *Alcoholism: Clinical and Experimental Research* 30 (11) (2006) 1866–1870.
- [33] T.P. Beresford, D.B. Arciniegas, J. Alfors, L. Clapp, B. Martin, H.F. Beresford, Y. Du, D. Liu, D. Shen, C. Davatzikos, M.L. Laudenslager, Hypercortisolism in alcohol dependence and its relation to hippocampal volume loss, *Journal of Studies on Alcohol* 76 (6) (2006) 861–867.
- [34] G. Wu, F. Qi, D. Shen, Learning-based deformable registration of MR brain images, *IEEE Trans. Med. Imaging* 25 (2006) 1145–1157.
- [35] H. Chui, A. Rangarajan, A new point matching algorithm for non-rigid registration, *Comput. Vision Image Understanding* 89 (2003) 114–141.
- [36] J. Yang, J.P. Williams, Y. Sun, R.S. Blum, C. Xu, Non-rigid image registration using geometric features and local salient region features, presented at *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.
- [37] A.F. Goldszal, C. Davatzikos, D. Pham, M. Yan, R.N. Bryan, S.M. Resnick, An image processing protocol for the analysis of MR images from an elderly population, *J. Comput. Assisted Tomography* 22 (1998) 827–837.
- [38] D.L. Pham, J.L. Prince, Adaptive fuzzy segmentation of magnetic resonance images, *IEEE Trans. Med. Imaging* 18 (1999) 737–752.
- [39] C.H. Lo, H.S. Don, 3-D Moment forms: their construction and application to object identification and positioning, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 1053–1064.
- [40] Z. Xue, D. Shen, C. Davatzikos, CLASSIC: consistent longitudinal alignment and segmentation for serial image computing, presented at *Information Processing in Medical Imaging*, Glenwood Spring, Colorado, 2005.
- [41] H.J. Johnson, G.E. Christensen, Consistent landmark and intensity-based image registration, *IEEE Trans. Med. Imaging* 21 (2002) 450–461.
- [42] S. Geman, Seminar on MRFs, Gibbs distributions, and simulated annealing, presented at *Joint Centers for Intelligent Control Computer Vision Workshop*, Harvard University, 1987.
- [43] S. Kirkpatrick, C.D. Gellatt Jr., M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (1983) 671–680.

About the author—DINGGANG SHEN received all of his degrees from Shanghai JiaoTong University. He is an Associate Professor in the Department of Radiology at UNC-Chapel Hill since April 2008. He was an assistant professor (tenure-track) in the Department of Radiology at University of Pennsylvania (Upenn), from July 2002 to March 2008, and a faculty member in Johns Hopkins University from Jan 2001 and June 2002. Dr. Shen is on the Editorial Board of *Pattern Recognition*, *International Journal of Image and Graphics*, and *International Journal for Computation Vision and Biomechanics*. He also served as a reviewer for numerous international journals and conferences, as well as NIH, NSF and other grant foundations. He has published over 180 articles in journals and proceedings of international conferences. His research interests include medical image analysis, pattern recognition, and computer vision.