Multichannel Image Registration by Feature-Based Information Fusion

Yang Li and Ragini Verma*, *Member, IEEE*

***Abstract*—This paper proposes a novel nonrigid inter-subject multichannel image registration method which combines information from different modalities/channels to produce a unified joint registration. Multichannel images are created using co-registered multimodality images of the same subject to utilize information across modalities comprehensively. Contrary to the existing methods which combine the information at the image/intensity level, the proposed method uses feature-level information fusion method to spatio-adaptively combine the complementary information from different modalities that characterize different tissue types, through Gabor wavelets transformation and Independent Component Analysis (ICA), to produce a robust inter-subject registration. Experiments on both simulated and real multichannel images illustrate the applicability and robustness of the proposed registration method that combines information across modalities. This inter-subject registration is expected to pave the way for subsequent unified population-based multichannel studies.**

***Index Terms*—Deformable registration, diffusion tensor imaging (DTI), independent component analysis (ICA), information fusion, Gabor filter, multichannel image registration.**

I. Introduction

With advances in technology, more and more imaging modalities have become available for clinical/research studies, for instance, X-ray computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI). In MRI, different protocols (T1, T2, FLAIR, mPRAGE and diffusion tensor imaging (DTI), etc.) can also be viewed as different modalities. Each of these modalities provides some unique and often complementary characterization of the underlying anatomy and tissue microstructure. For instance, structural T1 images provide contrast between the gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). On the other hand, DTI captures the directional micro-structural information within WM as a tensor, which complements the missing orientation information in structural images. This calls for a joint analysis of these multimodality images. We define a multichannel image for each subject as the co-registered collection of all single modality images that represent the same anatomy. Analysis of such images will be referred to as multichannel image analysis (MIA), which has become an active area of research with the growth in the number of imaging techniques, that provide different characterization for the same anatomy.

As in the analysis of single modality images, inter-subject spatial normalization/registration is the crucial first step for any population-based MIA. In contrast to the population-based image analysis using single modality, the population-based MIA utilizes the richer information carried by the multichannel image set. Although it is possible to use the deformation field from the registration of one modality to spatially normalize all other modalities in the multichannel image, this method has acceptable results only when there is a "dominating" modality that can give almost all anatomical information to drive the registration. For e.g., T1 structural images cannot do as well on white matter in registration, as compared to DTI, as DTI is a WM specific modality. Usually, since different modalities characterize different tissue types, using the information from only one channel will result in reduced accuracy in the areas where the driving modality does not characterize the underlying tissue type well. And this problem is more challenging when anatomical information is relatively evenly distributed across modalities without a "dominating" modality that gives enough anatomical information to drive the registration. Specifically, using a true inter-subject multichannel registration algorithm has several benefits. 1) By combining complementary information from different modalities which characterize different aspects of the anatomy, the multichannel image registration is expected to be more robust and accurate. 2) The deformation field obtained is common to all the modalities, therefore the spatially normalized images of all modalities can be jointly statistically analyzed, in contrast to using separate registrations for each of the modalities which usually generates inconsistent deformation fields. 3) Although different modalities can provide complementary information, a considerable portion of the information could be overlapping and hence "redundant" or less reliable (such as the WM information from T1 in comparison to that from DTI). Therefore, by controlling the redundancy (as explained above) appropriately, the multichannel image registration can be more efficient and reliable. Multichannel inter-subject registration is rendered challenging because there could be competing information from different modalities. We propose a multichannel inter subject registration that combines information across modalities in such a way as to alleviate these challenges and provide an effective joint registration.

More formally, we define a multichannel image as a set of images from different modalities or scanning parameters, which depict the same subject: $\mathbf{I} = [I_1, \ldots, I_n]$, where n is

Manuscript received August 15, 2010; revised October 22, 2010; accepted November 04, 2010. Date of publication November 22, 2010; date of current version March 02, 2011. This work was supported by the National Institute of Health under Grant R01-MH-079938. *Asterisk indicates corresponding author.*

Y. Li is with the Section of Biomedical Image Analysis (SBIA), Department of Radiology, University of Pennsylvania, Philadelphia, PA, 19104 USA

*R. Verma is with the Section of Biomedical Image Analysis (SBIA), Department of Radiology, University of Pennsylvania, Philadelphia, PA, 19104 USA (e-mail: ragini.verma@uphs.upenn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2010.2093908

the total number of modalities in the multichannel image. For any spatial location $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{I}(\mathbf{x}) \in \mathbb{R}^{\mathbf{n}}$. Throughout this paper, we assume that, all the image channels $I_1, \ldots, I_n$ have been co-registered to each other through mutual information (MI)-based intra-subject multimodal registration method, such as [1]. Although there are several established methods for single modality deformable registration, to the best of our knowledge, methods for multichannel image registration remain scarce in the literature, mainly owing to the challenge of combining information effectively. In [2], Park *et al.* proposed a deformable multichannel image registration method using multichannel demons [3], in which, the multichannel image consists of a T2-weighted image and other scalar images derived from DTI. After these scalar images were co-registered within each subject, two multichannel images $\mathbf{I}(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ were registered as multidimensional vector images using the metric $D(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2})) = \sqrt{\sum_{i=1}^{n} d(I_i(\mathbf{x_1}) - J_i(\mathbf{x_2}))^2}$, where d is the scalar image metric. In this method, every image channel is assigned equal importance in the image matching at every voxel. More recently, in [4], multiple channels consisting of T1, DTI and cerebral blood flow (CBF) image were registered by defining the multichannel similarity metric as a weighted summation of each modality's similarity metric $D(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2})) = \sqrt{\sum_{i=1}^{n} w_i \cdot d_i(I_i(\mathbf{x_1}) - J_i(\mathbf{x_2}))^2}$, where w_i and d_i are the weights and similarity metric for the ith image modality. Although, in the paper, spatially-consistent weight $w_i = 0.5$ were used in the experiments, the authors also suggested that a spatially-adaptive weighting function $\omega_i(\mathbf{x})$ properly-defined on the image domain could generate better results, due to the fact that different tissue types need different modalities to characterize. MI ([5], [6]) has been proven to be a very robust and reliable similarity metric for multimodal image registration [7]. In [8], multichannel MI was used to match the multichannel images, in which DTI was incorporated as a constraint into the registration of T1 images. Due to the impractical number of histogram bins required by the multichannel MI of all seven image channels (T1 and six tensor components), a simplification of only using two tensor components, D_{xx} and D_{yy}, was adopted based on the fact that the different diffusion directions are relatively uncorrelated. For other multichannel images with large number of uncorrelated channels, multichannel MI remains computationally expensive. To avoid the extremely high number of histogram bins required by the multichannel formulation, a Bayesian framework was proposed in [9] to register the multichannel image. Based on the assumption that the underlying anatomy consists of finite number of structures, the varying radiometric intensity values across different image modalities is estimated as the posterior mass functions associated with each of the structures. Then, the Kullback–Leibler divergence in the space of probability mass functions (acquired through a training process) is used as a similarity metric to drive the registration.

The main challenge associated with multichannel image registration is how to best fuse the multimodal information to guide the registration, so that competing and redundant information is minimized, and complementary information is emphasized. The meaning of "fuse" is two-fold here. Firstly, on different tissue types, it means to adaptively select the features that best characterize the underlying anatomical structure. In this way, the complementary information from different modalities are merged and, thereby, making the registration more accurate. Secondly, it aims at removing the redundant/less-reliable information. For a particular tissue type, usually, there is one or several modalities that give relatively more reliable information. For the most part, information from other suboptimal modalities could be low on reliability and somewhat redundant, due to the noisier signal with lower contrast. From another point of view, removing redundancy may also be understood as alleviating the possible "contradictoriness" between different modalities. That is, the same point could lead to different correspondences by using information from different modalities. In such a scenario, the most reliable modality is selected, omitting the less reliable. For example, any GM voxel will be better characterized by T1 than by fractional anisotropy (FA) measures obtained from DTI. Adding all the FA information to characterize GM may confuse the registration with CSF or fiber crossings. So losing the redundant part of FA information for this voxel may actually enhance the matching ability of this voxel. Likewise, two modalities such as T1 and gadolinium-enhanced T1 may convey the same information in all nonenhanced tissue and hence be overlapping and redundant. Therefore, to improve the efficiency and reliability of the multichannel image registration, for each tissue type, it is beneficial to only keep the most pertinent information from all the modalities, and discard the part that overlaps or less reliable. Such information will be referred to as "redundant" in the paper.

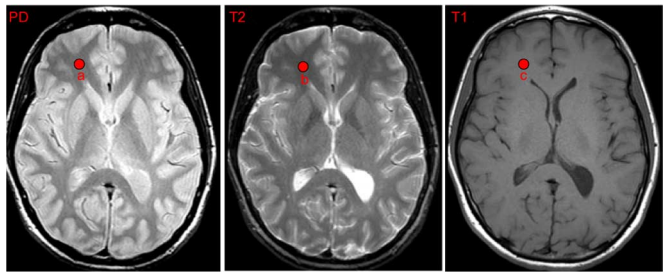


Fig. 1. Characterization ability of modalities. Images from three different modalities of the same subject are co-registered, in which, (a), (b), and (c) show the corresponding location on WM. The intensity values of the three points are related as: $b < a < c$. However, the PD image gives better characterization of the underlying WM structure, because it gives better contrast between WM and the surrounding tissues, although a has neither the lowest, nor the highest intensity value.

In contrast to the existing methods that are trying to solve the challenging information fusion problem in multichannel image registration at image/intensity level, in this paper, we propose to fuse the multimodal information using Gabor Wavelets transformation. The main drawback of dealing with the information fusion problem at image/intensity level is that, it is difficult to extract and distinguish between complementary and redundant information, since for a specific tissue type, a higher intensity value does not necessarily indicate a stronger tissue characterization ability which reflects how a modality can discriminate one tissue type from another. Here, we use an example to explain this. Fig. 1 shows the co-registered PD, T2 and T1 images of a single subject and the intensities have been normalized into the same range. a, b, and c are three corresponding points on WM from the three different modalities, respectively. For the

intensity values of three points, we can find that $b < a < c$. However, if it is asked which modality best characterizes the WM and discriminates it from other tissue types, we may say that, although a is neither the lowest, nor the highest intensity value, it is the PD image that gives the best characterization of the underlying WM structure. This is because the proton-density (PD) image provides the highest contrast between WM and other tissues. Therefore, as this example shows, it is difficult to compare the characterization ability of different modalities by directly comparing the intensity values. Instead, this characterization ability is proportional to the contrast level the modality can provide.

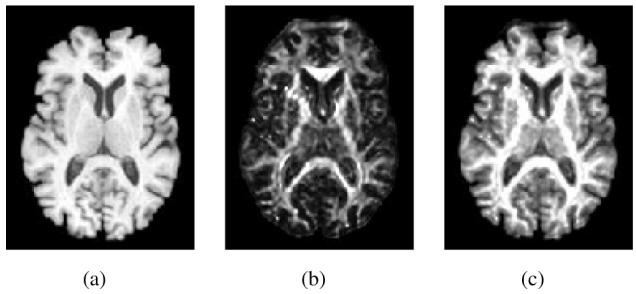
To measure this characterization ability, multiresolution analysis methods, such as, discrete wavelets transform (DWT) or Gabor wavelets transform have been shown to be appropriate tools [10], as the strength of the output signal from the high-pass (wavelets) or band-pass (Gabor) filter, which reflects the contrast level, is a reliable indicator of the characterization ability at different scales and locations. Based on this idea of measuring the characterization ability, multiscale decomposition (MSD)-based information fusion scheme was proposed and has been successfully used in image fusion [11]. In this paper, by adopting a similar information fusion scheme based on Gabor wavelets transform, we propose a novel multichannel image registration algorithm. The remaining part of this paper is organized as the follows. The basic idea of MSD-based information fusion scheme is introduced in Section II-A. After giving an overview of the proposed information fusion scheme and the dissimilarity metric for multichannel image registration in Section II-B, details of the proposed algorithm are described in Section II-C. Then in Section II-C6, all the components are incorporated into a registration framework for inter-subject multichannel image registration. After that, in Section III, we apply our new registration algorithm to both simulated and real multichannel image sets to validate the proposed method. The results show that the proposed multichannel image registration method can integrate and enhance complementary information, while eliminating the less reliable/redundant information from different channels and leads to a more accurate and robust inter-subject registration at low computational cost.

II. Methods

As the pivotal problem for information fusion, the method of measuring the characterization ability using discrete wavelets or Gabor wavelets transformation is first illustrated through an example of image fusion. It is worthwhile to mention that, although a similar information fusion scheme is used in our registration method, we do not explicitly fuse the images from different modalities and then perform the registration on the fused image. Instead, we fuse the information in the process of finding the correspondences between two images (see more details in Section II-C3).

A. Measuring the Characterization Ability, an Example of Image Fusion

Suppose we have two co-registered images of the same subject from different modalities, such as the T1 and FA image (computed from DTI data, depicting the underlying anisotropy of the tissue [12]) shown in Fig. 2. Both the images provide a lot of complementary information. The purpose of image fusion is to combine the information from the two modalities into a single image such that the resulting image is more informative than any of the single input images. To achieve this, in MSD (multiscale decomposition)-based information fusion, after being normalized into the same intensity range, an image $\mathbf{I}$ is decomposed into a series of wavelets coefficients $c^{\mathbf{I}}_{k,l,m,n}$, where k, l, and (m, n) indicate the decomposition level, the frequency band and the position in a frequency band, respectively. As $\|c^{\mathbf{I}}_{k,l,m,n}\|$ reflects the local energy of the output of the high-pass filter banks and is a good indicator of the characterization ability of each modality at that specific scale, frequency band and position, the simplest but efficient information fusion rule is the "*Choose-Max*" scheme [11]. That is, to fuse two images $\mathbf{X}$ and $\mathbf{Y}$, we can get the wavelets-based MSD representation of the fused image $\mathbf{Z}$ by only keeping the coefficient with the largest absolute value (magnitude). After that, an inverse DWT can be performed to finally get the fused image $\mathbf{Z}$. As an example, by using this method, the T1 and FA image are fused as shown in Fig. 2(c). As we can see, the fused image gives a good characterization of both the WM and cortical GM which cannot be observed simultaneously in the input T1 and FA image individually.



(a) (b) (c)

Fig. 2. An example of image fusion by the choose-max scheme. (a) T1 image. (b) FA image. (c) Fused image. As can be seen, the fused image displays the WM detail from the FA image and the GM detail from the T1 image.

Through this example, we show the effectiveness of the characterization ability measuring method using multi-scale decompositions, such as DWT and Gabor wavelets transform. Please note that, to make these coefficients from different image channels comparable, it is very important to make sure that the filter banks of wavelets are applied on images with same intensity range. More specifically, all image channels are not only intensity normalized to each other, but to corresponding channels in the other multiparametric image. That is, for instance, the structural T1 component of image A is intensity normalized to the structural T1 component of image B, where A and B are described by several image channels. By intensity normalization, we mean that they have the same range of intensities obtained by histogram equalization. In the next section, we propose our multichannel registration method.

B. Overview of the Proposed Multichannel Image Registration Method

Every feature-based registration method requires feature extraction and a definition of a metric that will help match points

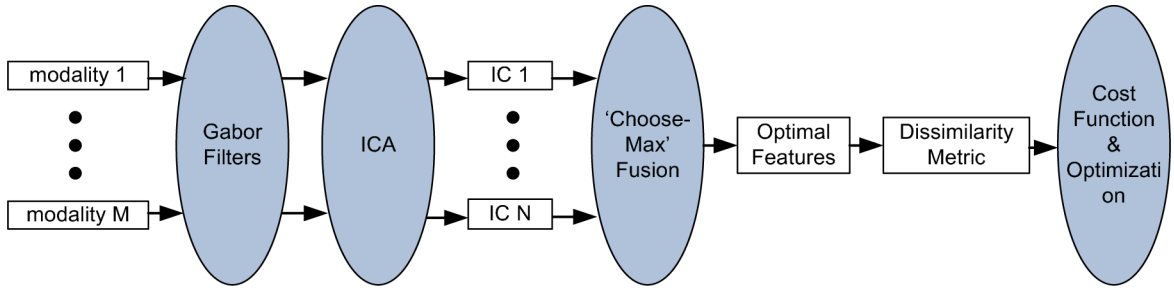


Fig. 3. Flowchart of the proposed multichannel information fusion scheme and the registration framework.

and regions between the two images, based on these features. This is then incorporated into an optimization framework with appropriate smoothing terms. Before going into the details, a schematic diagram of the entire framework is provided in Fig. 3 and summarized as the following:

1) Feature Extraction and the Measure of Tissue Characterization Ability: For the multiscale decomposition method, we use Gabor wavelets [13], as Gabor wavelet transformation has been shown to be optimal in the sense of minimizing the joint uncertainty in space and frequency, and has been widely used for feature extraction [14], and hence, more appropriate for the purpose of matching/registration. At the same time, just like coefficients of DWT, the magnitude of the coefficients of Gabor wavelets is also a reliable measure of the characterization ability of each modality. Thus Gabor wavelets are a better choice for both feature extraction/matching and tissue characterization ability measurement, for our proposed information fusion-based multichannel image registration.

2) Tissue-Pertinent Information Extraction Through Independent Component Analysis: As mentioned previously, for a particular tissue type, usually, there is one or a combination of several modalities that give relatively more reliable information. For other modalities, most part of the information could be less reliable and hence redundant, although they may still contain some useful information about the underlying tissue. Therefore, if the "Choose-Max" criteria is directly applied between different modalities, the residual useful information (usually a small amount) from the other "non-max" modality will be discarded. In some cases, if information about certain tissue type is quite evenly distributed among different modalities, the cost of this simple "Choose-Max" scheme would be high. To solve this problem and facilitate the subsequent information fusion, for each tissue type, it would be very beneficial to merge all the pertinent information of this tissue type from all modalities. In our proposed framework, independent component analysis (ICA) is adopted to achieve this goal. By applying ICA to the Gabor features extracted from different modalities, each of the resultant independent components (ICs) combines all the information regarding a particular tissue type in the feature space. In other words, after ICA, the information pertaining to different tissue types is "concentrated" into different ICs, and the input information is no longer separated by their sources (modalities), but by their targets (different tissue types). This information "repartitioning" is important to the subsequent "Choose-Max" step.

3) Information Fusion: As the underlying tissue type could be different for different voxels, different independent components (IC) are needed to acquire the best characterization. Therefore, "Choose-Max" scheme is adopted to select the optimal IC according to the underlying tissue type. In this spatially-adaptive manner, the complementary information is kept and enhanced while less reliable/redundancy is reduced, via information fusion.

4) Dis-Similarity Metric: For each voxel, based on the optimal IC obtained through ICA and "Choose-Max" scheme, a dissimilarity metric is defined to find the correspondence between two multichannel images.

5) Cost Function and Optimization: Finally, based on the dissimilarity metric, the image matching problem is formulated as a cost function. By optimizing it, the deformation field of the registration is obtained.

We now describe each of these steps in details. We will also illustrate each step with an example of registering T1 + DTI multichannel images.

C. Multichannel Image Registration by Information Fusion

1) Feature Extraction (Gabor Features): Although the registration is for 3D images, to alleviate the computational cost, we use 3 perpendicular (axial, coronal, and sagittal) 2D Gabor filter banks to extract the features. A 2D Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope: $G(x,y) = s(x,y)g(x,y)$, where $s(x,y)$ is a complex sinusoid: $s(x,y) = \exp[-j2\pi(u_0x+v_0y)]$ and $g(x,y)$ is a 2D Gaussian envelope

$$g(x,y) = \frac{1}{\sqrt{2\pi}\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right] \quad (1)$$

σ_x and σ_y characterize the spatial extent and bandwidth of g along the respective axes, u_0 and v_0 are the shifting frequency parameters in the frequency domain. Using $G(x,y)$ as the mother wavelet, the Gabor wavelets, a class of self-similar functions can be obtained by appropriate dilations and rotations of $G(x,y)$ through: $G_{m,n}(x,y) = a^{-m}G(x',y')$, where $x' = a^{-m}(x\cos\theta + y\sin\theta)$, $y' = a^{-m}(x\sin\theta + y\cos\theta)$, $a > 1, \theta = n\pi/O$, $m = 1 \ldots S$ $n = 1 \ldots O$. O indicates the number of orientations, S the number of scales in the multiresolution decomposition and a is the scaling factor between different scales. These parameters can be set according

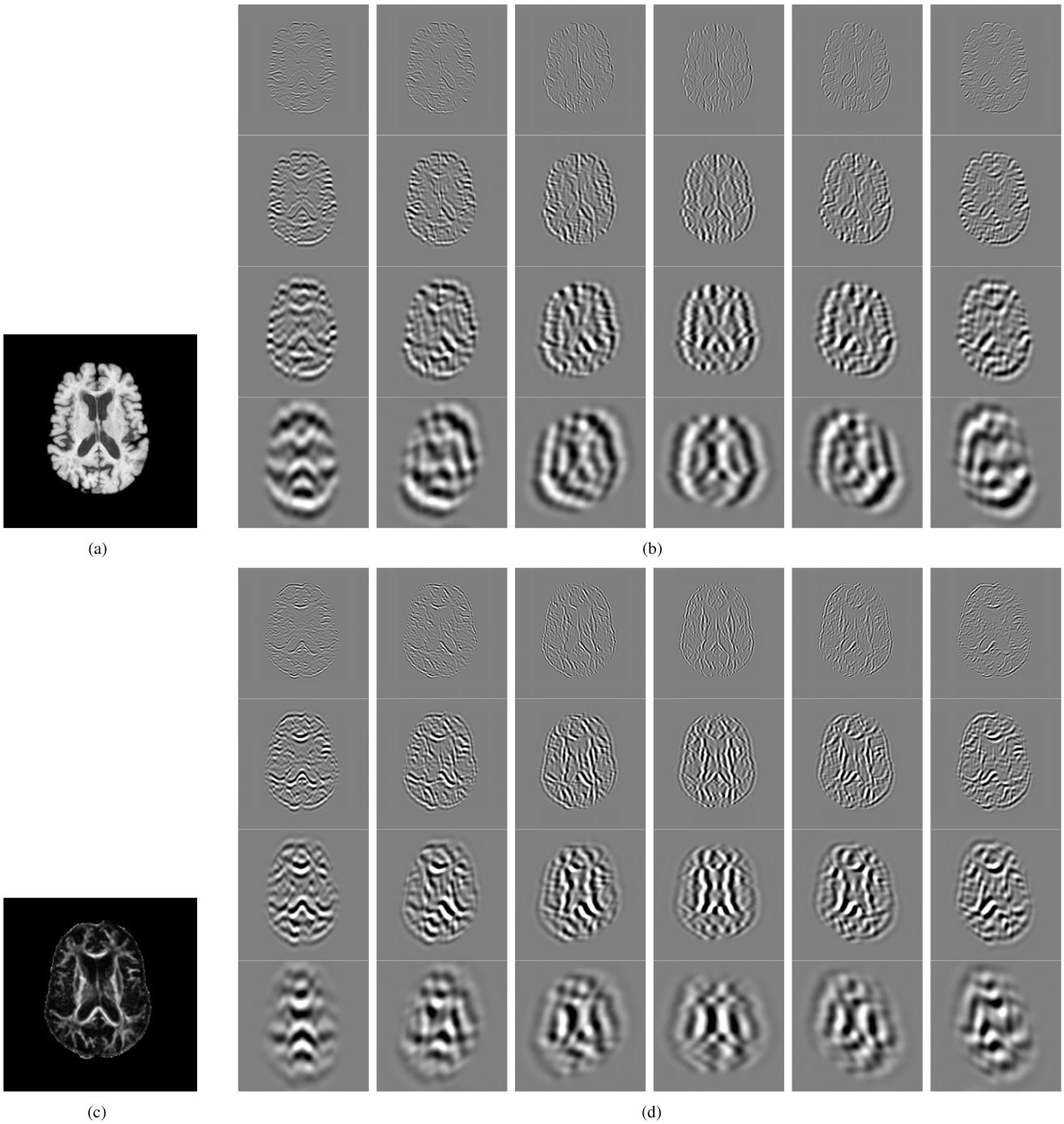


Fig. 4. The extracted Gabor features (imaginary part) on T1 (a) and FA (c) channels. (b) and (d) are the extracted Gabor features in four scales and six orientations, respectively. From left to right: orientation 0–5. From top to bottom: scale 0–3. It is shown that, on different location, scale and orientation, Gabor features from different modalities are needed to best delineate the underlying structure. For instance, on the corpus callosum, since it is mainly white matter, Gabor features from FA image give stronger response. (a) T1. (b) Gabor features of T1. (c) FA.(d) Gabor features of FA.

to [14] to reduce the redundant information (caused by the nonorthogonality of the Gabor wavelets) in the filtered images. Given an image $I(x, y)$, the Gabor transform with orientation n and scale m can be computed as

$$F_{m,n}(x,y) = \int I(x_1, y_1) G^*_{m,n}(x - x_1, y - y_1) dx_1 dy_1 \quad (2)$$

where $*$ indicates the complex conjugate. In our work, we set the Gabor filter to have $S = 4$ scale levels and $O = 6$ orientations. Fig. 4 gives the examples of the extracted Gabor features using the designed filter bank on T1 and FA images, respectively. As we can see, on different locations, scales and orientations, we need Gabor features from different modalities to best delineate the underlying structure. For instance, in the corpus callosum, since it is mainly white matter, Gabor features from FA image (computed from DTI) is expected to give stronger response.

2) Independent Components of the Extracted Gabor Features: To incorporate all the relevant information regarding each

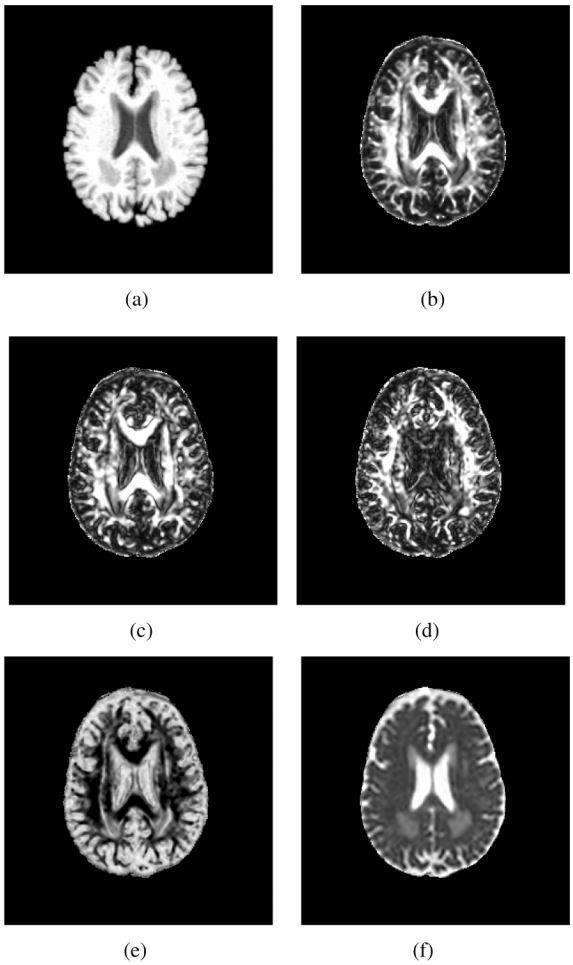


Fig. 5. The six channels from T1 and DTI-derived scalars that are used to create a multichannel image for the purposes of illustration. (a) T1. (b) FA. (c) C_l. (d) C_p. (e) C_s. (f) ADC.

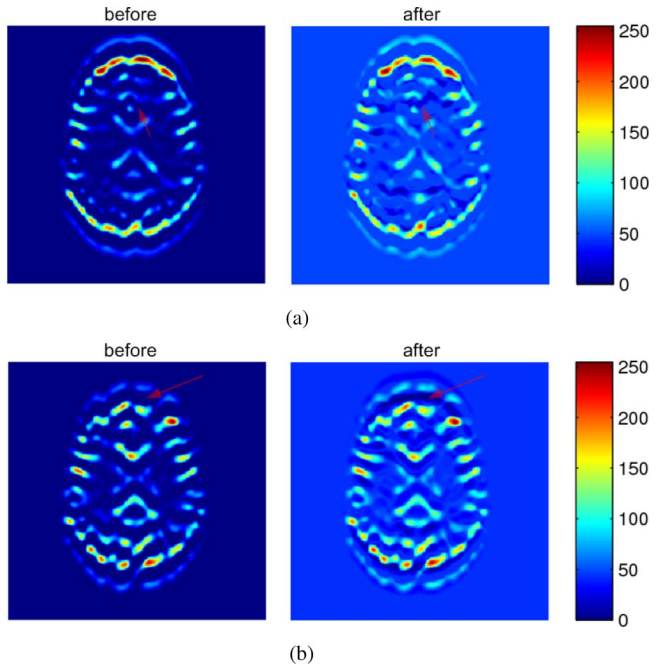
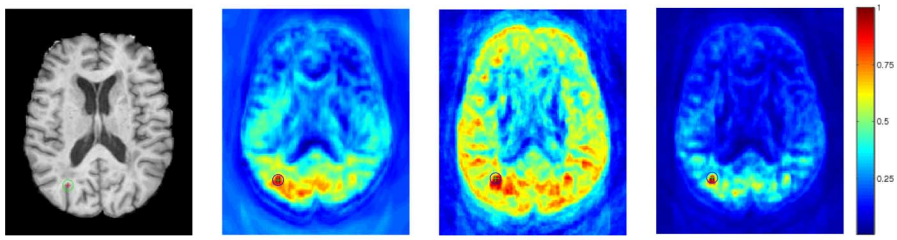


Fig. 6. The ICA-enhanced Gabor features. (a) and (b) Two examples of the Gabor feature images before-and-after ICA enhancement. All the feature images have been normalized to the same intensity range [0 255] and are visualized with the same colormap. More texture in the image is revealed after ICA. Red arrows indicate such an example, where new texture can be observed after ICA, where, at the same place, no texture is observable (same intensity as the background) prior to ICA. (a) Example A. (b) Example B.

particular tissue type from different modalities, and thus facilitate the subsequent "Choose-Max" information fusion scheme, we apply the independent component analysis (ICA) [15] on the Gabor features extracted. ICA has been successfully applied in MRI enhancement [16], [17], functional MRI (fMRI) [18] analysis and blind source separation [19]. The basic theory of applying ICA in multichannel image analysis can be briefly described as follows.

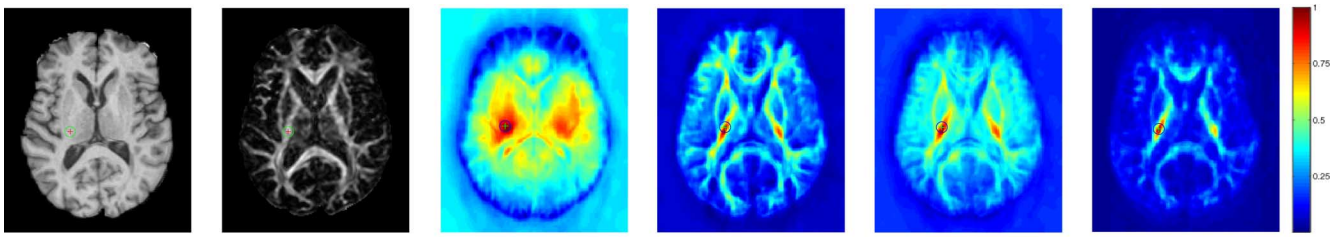
Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^T$ denote the multichannel image set generated by using different imaging modalities or scanning parameters. n is the number of modalities/channels. $\mathbf{X}$ is the signal we observed and it can be considered as the linear mixture ($\mathbf{X} = \mathbf{AS}$) of many independent sources ($\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_m]^T$), for instance water, blood, fat, GM, WM, CSF, and muscle, etc. The statistical independence of S is due to the fact that the underlying voxels are heterogeneous and consists of physically independent tissue components, e.g., water, blood, fat, GM, WM, CSF, and muscle, etc. Therefore, they are assumed to have independent spatial distribution. [16], [17]. $\mathbf{A}$ is the $n \times m$ mixing matrix which is decided by the imaging properties of different modalities. The objective of ICA is to find the matrix $\mathbf{A}^{-1}$ which can restore the independent source signals from the observed signals $\mathbf{X}$ by $\mathbf{S} = \mathbf{A}^{-1}\mathbf{X}$. As $\mathbf{S}$ and $\mathbf{A}$ are both unknown, it is impossible to estimate them simultaneously. However, it is known that the source signals in $\mathbf{S}$ are statistically independent. Therefore, in practice, an approximation of $\mathbf{S}$ can be estimated by iteratively modifying $\mathbf{W}$ in $\mathbf{S}' = \mathbf{WX}$ such that each element of $\mathbf{S}'$ becomes as independent as possible. If the iteration converges, the $m \times n$ matrix $\mathbf{W}$ is found to be an approximation of $\mathbf{A}^{-1}$ and $\mathbf{S}'$ is considered to be equivalent to $\mathbf{S}$, apart from the scale and the permutation. For details of ICA algorithm, we refer the readers to [20] and [21].

Specifically, to apply ICA to the analysis of multichannel images, there are two steps: training and decomposition. In the training step, a number of voxels are randomly selected from a few sample multichannel images and input to the ICA algorithm. After the above iterative optimization, when the algorithm converges and the error is below a specified threshold, the obtained decomposition matrix $\mathbf{W}$ can be applied to decompose the new multichannel image into independent component images, each of which highlights one of the tissue types. This decomposition process can be viewed as an information repartitioning process. Before decomposition, each image channel contains information about a combination of many tissue types. However, after decomposition, each independent component image is "specialized" for capturing one tissue type and contains all the relevant information (taken from all the modalities) pertaining to that tissue type. Because of this "purity," the decomposed independent component image provides higher contrast and better characterization than the original image (examples can be found in [16] and [17]). It is worthwhile to clarify that although each tissue type will be highlited in one IC,



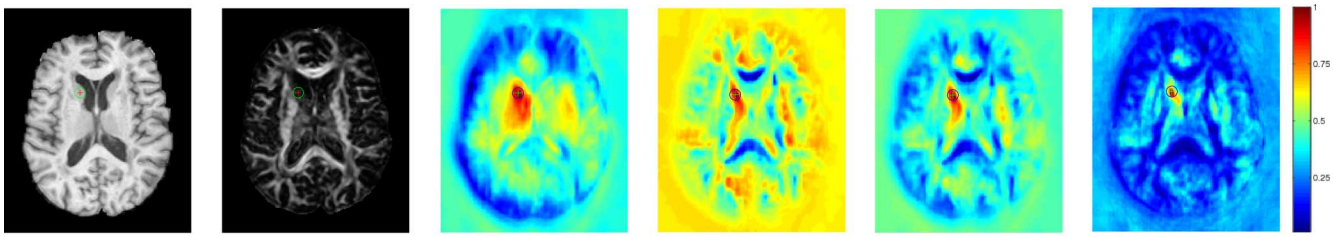
(a) (b) (c) (d)

Fig. 7. Similarity map generated by different metrics. (a) shows the reference point (denoted by a red cross) on T1. (b) is the similarity map generated by D_1. (c) is the similarity map generated by metric L (normalized to [0 1]). (d) is the similarity map generated by metric D. Red indicates high similarity. Metric D_1 gives more specific similarity map than the coarse metric L. By combining the two metrics in a coarse-to-fine manner, metric D gives the most specific similarity map which indicates the highest discriminative ability.



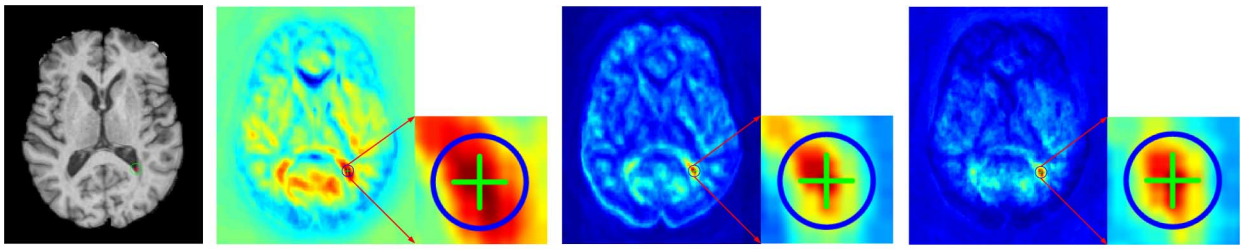
(a) (b) (c) (d) (e) (f)

Fig. 8. Comparison of the similarity maps of a reference point on WM. (a) and (b) show the reference point (denoted by a red cross) on T1 and DTI slices, respectively; (c) and (d) are the similarity map only using T1 and DTI, respectively; (e) is the similarity map using both T1 and DTI without fusing; (f) is the similarity map given by the new metric defined in (5). The color bar on the right shows that red is of highest similarity and blue is the lowest. The fused metric using both T1 and DTI information gives the most specific similarity map which indicates the highest discriminatory ability.



(a) (b) (c) (d) (e) (f)

Fig. 9. Comparison of the similarity maps of a reference point on GM. (a) and (b) show the reference point (denoted by a red cross) on T1 and DTI slices, respectively; (c) and (d) are similarity map only using T1 and DTI, respectively; (e) is the similarity map using both T1 and DTI without fusing; (f) is the similarity map given by the new metric defined in (5). The color bar on the right shows that red is of highest similarity and blue is the lowest. The fused metric using both T1 and DTI information gives the most specific similarity map which indicates the highest discriminatory ability.



(a) (b) (c) (d)

Fig. 10. The discriminatory ability of the ICA-enhanced Gabor features. (a) shows the reference point. (b) is the similarity map generated by the unfused Gabor features. (c) is the similarity map generated by the fused Gabor features without ICA-enhancement. (d) is the similarity map generated by the ICA-enhanced Gabor features after fusion. The ICA-enhanced Gabor features give the most specific similarity map which indicating the highest discriminatory ability among three types of features.

using ICA here is not expected to produce tissue segmentation, but to merge all the pertinent information regarding one tissue type into an IC. Therefore, the correspondence between each IC and the segmentation label, such as WM, GM, or CSF is not needed in our work, as the subsqence "choose-max" algorithm dose not needs this information.

In our work, instead of applying ICA to the multichannel image, we apply ICA to the multivariate Gabor features, since

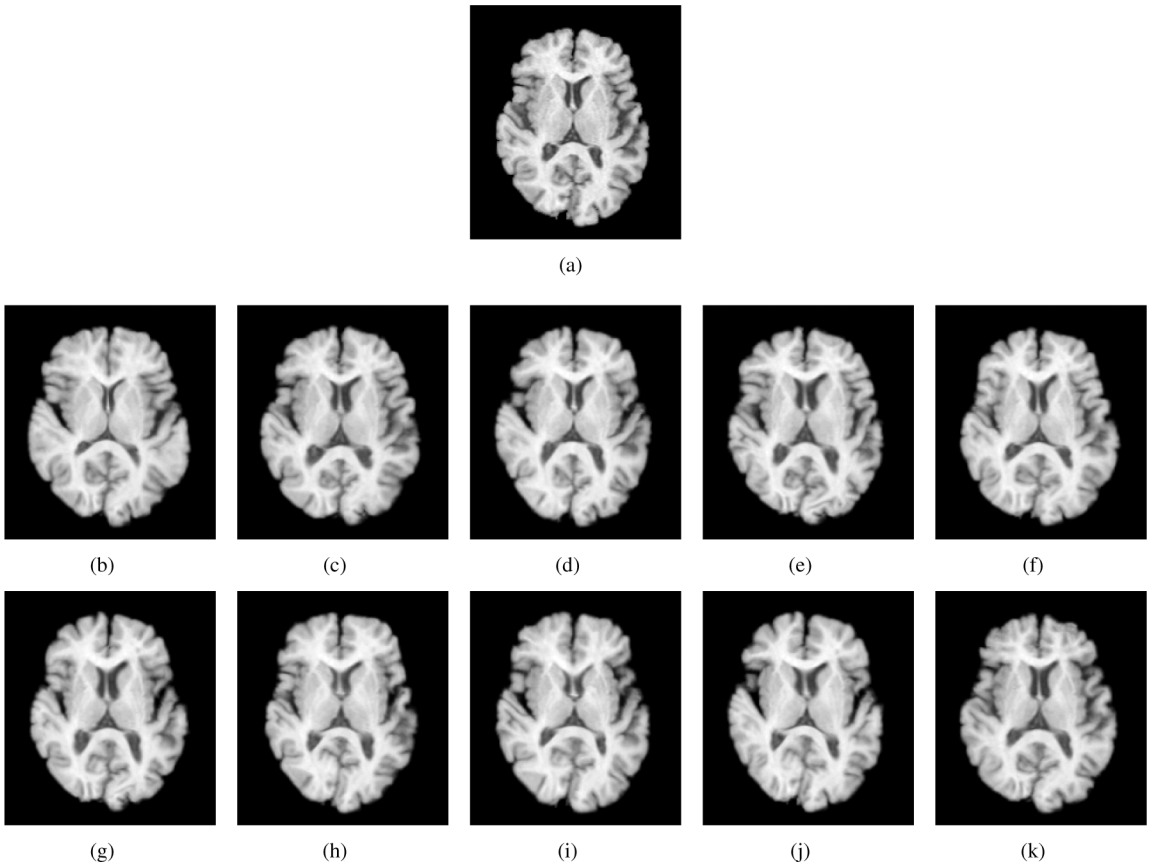


Fig. 11. The template [(a)] and the representative slices from 10 simulated subjects [(b)–(k)] used in the experiment of Section III-C.

our goal is to separate the information in Gabor feature space according to different tissue types. For each scale and orientation, we perform an ICA (using FastICA algorithm [21]) on all the Gabor features obtained from different modalities. We have proposed a general framework for multichannel image registration. Here we demonstrate its applicability on the combination of a T1 and DTI image as an example, as these images are routinely acquired in a clinical study. Specifically, we use a multichannel image created by combining T1 and five DTI-derived scalar images to illustrate and test the proposed algorithm. Besides T1, the five scalar images computed from DTI are FA, apparent diffusion coefficient (ADC) [12], C_s (sphericity), C_p (oblateness), and C_l (prolateness) [22], respectively. Examples of the six image channels of a subject are shown in Fig. 5. As each of them gives a somewhat different characterization of the underlying tissue, they can be viewed as different imaging modalities. Therefore, together they provide a good example of a multichannel image with complementary information between different channels. Contrast enhancement similar to [16] was also observed in our experiments. Such examples are shown in Fig. 6. By comparing Gabor features before and after ICA, it is shown that more texture of the image is revealed (indicated by red arrow in Fig. 6) that is not observable or salient prior to ICA. Note that in Fig. 6, both image intensities (before and after) have been normalized to the same range. The newly revealed/enhanced texture is attributed to ICA.

3) The "Choose-Max" Matching Metric: After the ICA step, each IC of Gabor features is "specialized" in depicting one particular tissue type. Therefore, by using the Choose-Max scheme on each voxel, we can choose the optimal Gabor features from the corresponding IC to characterize the underlying structure. Suppose we have two multichannel images $\mathbf{I} = [I_1, \ldots, I_M]$ and $\mathbf{J} = [J_1, \ldots, J_M]$, in which $I_1, \ldots, I_M$ and $J_1, \ldots, J_M$ are two collections of co-registered images generated by M different modalities. For each of the image channels, the raw Gabor features are extracted on S scales and O orientations. For orientation n and scale m, after ICA, the ith IC of the Gabor features of I_i and J_i are denoted as $F^{I_i}_{mn}$ and $F^{J_i}_{mn}$, respectively. Using the choose-max scheme, the dissimilarity between two voxels $\mathbf{I}(\mathbf{x_1})$ and $\mathbf{J}(\mathbf{x_2})$ can be measured as

$$D_1(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2})) = \sqrt{\sum_{m=1}^{S}\sum_{n=1}^{O}\sum_{i=1}^{M} w_{m,n,i}\left(\left\|F^{I_i}_{m,n}(\mathbf{x_1})\| - \|F^{J_i}_{m,n}(\mathbf{x_2})\right\|\right)^2} \quad (3)$$

where

$$w_{m,n,i} = \begin{cases} 1, & \text{if } i = \text{argmax}_{j=1\ldots M}\left\|F^{I_j}_{m,n}(\mathbf{x_1})\right\| \\ 0, & \text{otherwise} \end{cases}. \quad (4)$$

Equations (3) and (4) can be understood as follows. For each voxel on the first image $\mathbf{I}$, the appropriate IC (contains all the relevant information) corresponding to the underlying tissue type

is chosen [by (4)] to find the correspondence point on image $\mathbf{J}$ [by (3)]. Therefore (3) and (4) together define the dissimilarity metric which is used for finding the corresponding points on two images. Here, $\mathbf{x_1}$ and $\mathbf{x_2}$ are not necessarily same spatial location, since before comparison, the correspondence is yet unknown. In practice, for a fixed $\mathbf{x_1}$, $\mathbf{x_2}$ is usually chosen from a neighborhood centered at the corresponding location in the second image. We note that although metric D_1 is not symmetric, due to the "Choose-Max" scheme, this asymmetry can be rectified by defining a symmetric cost function which will be discussed in Section II-C6. Using D_1 on the six-channel T1+DTI multichannel image (as shown in Fig. 5), the similarity between a reference point on the template and other points on a subject can be computed and shown in Fig. 7(b) as a similarity map. As we can see, it gives substantial characterization ability of the underlying tissue.

4) Max-Index Metric for Matching: To fully utilize the information in the Choose-Max scheme, we design another metric using the max indices that we already have. Let $p^{\mathbf{I}}_{m,n}(\mathbf{x})$ denote the label of the IC which generates the strongest Gabor feature at scale m and orientation n on $\mathbf{I}(\mathbf{x})$. These labels are then arranged into a vector form $\mathbf{p}^{\mathbf{I}}(\mathbf{x}) = [p^{\mathbf{I}}_{1,1}(\mathbf{x}), \ldots, p^{\mathbf{I}}_{S,O}(\mathbf{x})]$. Then $\mathbf{p}^{\mathbf{I}}(\mathbf{x})$ is a very informative fingerprint of voxel $\mathbf{I}(\mathbf{x})$. Let $L(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2}))$ be the number of equal labels between $\mathbf{p}^{\mathbf{I}}(\mathbf{x_1})$ and $\mathbf{p}^{\mathbf{J}}(\mathbf{x_2})$. Thus, L is a similarity metric with increasing resolution as S and O increase. Here, the three perpendicular 2D Gabor filter banks each have four scales and six orientations. Therefore, L has the ability to differentiate the brain tissue into $4 \times 6 \times 3 = 72$ classes. Fig. 7(c) shows the similarity map generated by metric L. As we can see, at this resolution level, only L can give considerable discriminatory power. More importantly, this also shows the effectiveness of the choose-max fusion rule. Without it, we would not have been able to get such a specific similarity map by using only L. By combining L with (3), our metric for multichannel image matching is finally defined as

$$D(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2})) = \left[1 - \frac{L(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2}))}{3SO+1}\right] \cdot D_1(\mathbf{I}(\mathbf{x_1}), \mathbf{J}(\mathbf{x_2})) \quad (5)$$

In (5), D is designed as a coarse-to-fine metric, in which L (first part) gives a coarse/discrete similarity measure and D_1 refines the measure. The similarity map generated by the proposed metric D is shown in Fig. 7(d). As we can see, in this coarse-to-fine manner, the metric D has the ability to give a very specific correspondence measurement indicating high discriminative power.

5) Discriminatory Ability of the Metric: Before using the proposed similarity metric for multichannel image registration, we need to test the discriminatory ability of the proposed metric by comparing the similarity maps generated by different metrics. Three other metrics compared are: metric using the Gabor features from T1 or DTI only and the metric using features from both T1 and DTI but without fusing (as in (3) with $w_{m,n,i} = 1, \forall m, n, i$). As expected, for a reference point on WM, the metric using only DTI features [Fig. 8(d)] is more discriminative

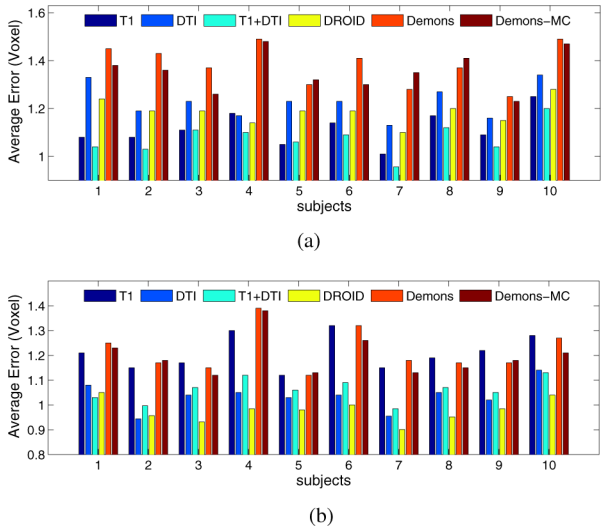


Fig. 12. Comparison of the registration accuracy on simulated images. T1 + DTI indicates the proposed multivariate registration method. T1 and DTI indicate the methods which use the same deformation model and features, but only use Gabor features from T1 or DTI, respectively. DROID and Demons are other two methods which have different deformation model and features. Multichannel Demons (Demons-MC) is another multichannel registration method, which is based on matching the image intensity. (a) Errors on GM. (b) Errors on WM.

than the metric using only T1 features [Fig. 8(c)]. And similarly, the metric only using T1 features [Fig. 9(c)] is more discriminative on GM than the metric only using DTI features [Fig. 9(d)]. Although for a reference point on GM, the metric using both T1 and DTI without fusing improves the discriminatory power [Fig. 9(e)], by comparing Fig. 8(e) with Fig. 8(c) and (d), we find that, without information fusion, Fig. 8(e) shows a less specific similarity map around the reference point. This indicates that the discriminatory ability of this unfused metric actually deteriorates on WM. This is a good example demonstrating that incorporating information from other suboptimal modalities directly without spatio-adaptive selection may even weaken the discriminative power. The explanation of this phenomenon is that the information given by some modalities is often noisy and less reliable, therefore incorporating it could "contaminate" the optimality of the information from the optimal modality or IC. This result indicates that in order to benefit from the complementary information from all the modalities without the above contamination, information fusion is a necessary step. For comparison, the proposed new metric [defined in (5)] correctly utilizes the information from both modalities and gives the most discriminative (shown by the most specific similarity map) result on both WM [Fig. 8(f)] and GM [Fig. 9(f)]. The gain of discriminatory ability through the ICA-enhancement is also shown in Fig. 10. By comparing Fig. 10(d) with (e) and (b), we can see that by using ICA on the Gabor features to keep more relevant information, we are able to get a more specific matching result and a smoother similarity map (which benefits the optimization step by producing less local minimum/maximum in the objective function).

6) Deformable Registration of Multichannel Image: With the dissimilarity metric defined in (5), the problem of registering

TABLE I
SUMMARIZATION OF THE EXPERIMENT RESULTS ON SIMULATED IMAGES

Methods	T1	DTI	T1+DTI	DROID	Demons	Demons-MC
mean $\pm$ std (voxel)	1.18 ± 0.07	1.09 ± 0.06	1.05 ± 0.05	1.05 ± 0.04	1.27 ± 0.08	1.20 ± 0.13
min (voxel)	1.09	1.01	0.97	0.97	1.16	1.10
max (voxel)	1.27	1.20	1.15	1.12	1.51	1.47

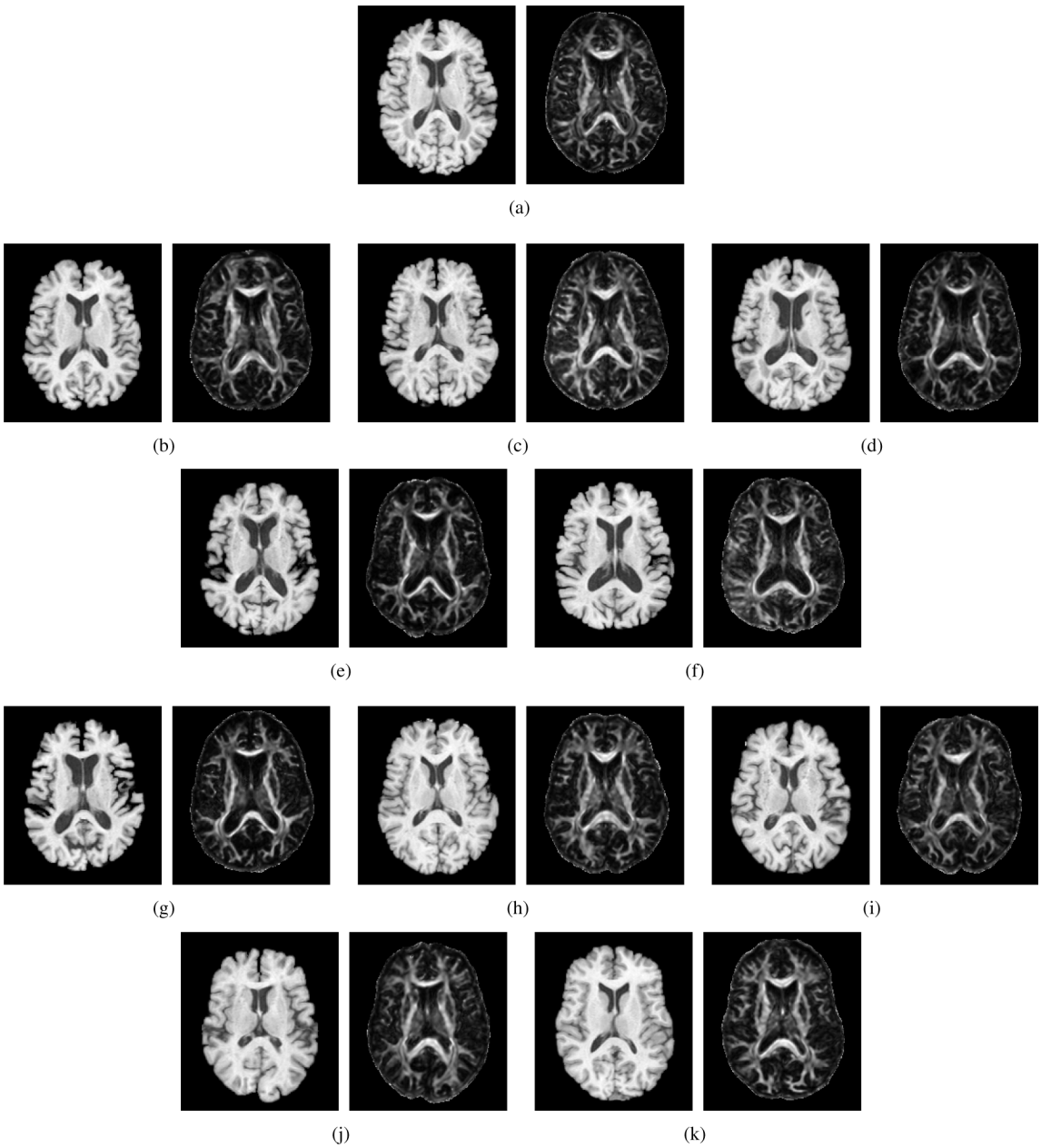


Fig. 13. The template (a) and representative slices from 10 randomly selected subjects (b)–(k) used the experiment on real image (Section III-D). In each subfigure, the left is T1 and the right is FA computed from DTI. (In the experiment, all the five scalar images from DTI are uses to extract the Gabor features, although here only FA is shown in this figure as a representative map computed from DTI.).

multichannel image $\mathbf{I}(\mathbf{x})$ to $\mathbf{J}(\mathbf{x})$, can be defined as finding an optimal transformation $h(\mathbf{x})$ which minimizes the cost function

$$\int_{\mathbf{x}\in\Phi_{\mathbf{I}}} D(\mathbf{I}(\mathbf{x}),\mathbf{J}(h(\mathbf{x})))d\mathbf{x} + \int_{\mathbf{x}\in\Phi_{\mathbf{J}}} D(\mathbf{I}(h^{-1}(\mathbf{x})),\mathbf{J}(\mathbf{x}))d\mathbf{x} + \alpha \int_{\mathbf{x}\in\Phi_{\mathbf{I}}} \|\nabla^2 h(\mathbf{x})\| d\mathbf{x}. \quad (6)$$

Here, the first term defines the cost of warping image $\mathbf{I}$ to image $\mathbf{J}$ and the second term defines the cost of warping image $\mathbf{J}$ to image $\mathbf{I}$. Adding both of these terms makes the cost function and, thus, the deformation symmetric. In addition, this symmetry also corrects the asymmetry in metric D_1 [defined in (3)], since swapping the moving and fixed images does not alternate the cost function. The third is the regularization term which makes the deformation field smooth. The above

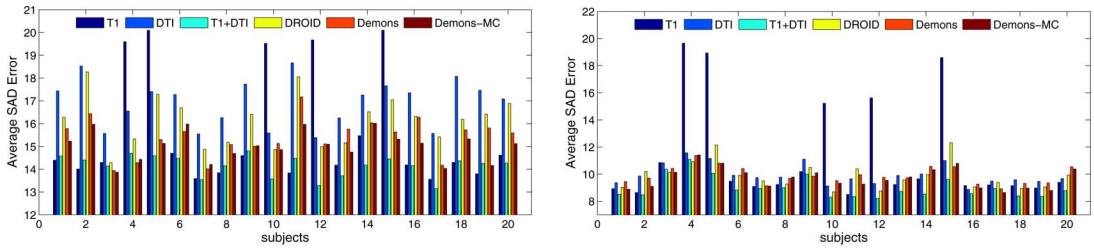


Fig. 14. Comparison of the registration accuracy on real T1 images. T1 + DTI indicates the proposed multichannel registration method. T1 and DTI indicate the methods which use the same deformation model and features, but only use information from T1 or DTI, respectively. DROID (for DTI) and Demons (for T1) are other two methods which have different deformation model and features. (a) Errors on GM. (b) Errors on WM.

TABLE II
SUMMARIZATION OF THE EXPERIMENT RESULTS ON REAL T1 IMAGES

Methods	T1	DTI	T1+DTI	DROID	Demons	Demons-MC
mean $\pm$ std (SAD)	15.37 ± 3.22	14.54 ± 0.97	11.89 ± 0.63	14.45 ± 1.18	14.36 ± 0.67	12.36 ± 0.21
min (SAD)	12.11	13.58	11.82	12.91	12.41	11.98
max (SAD)	19.57	15.93	13.48	15.86	14.36	13.57

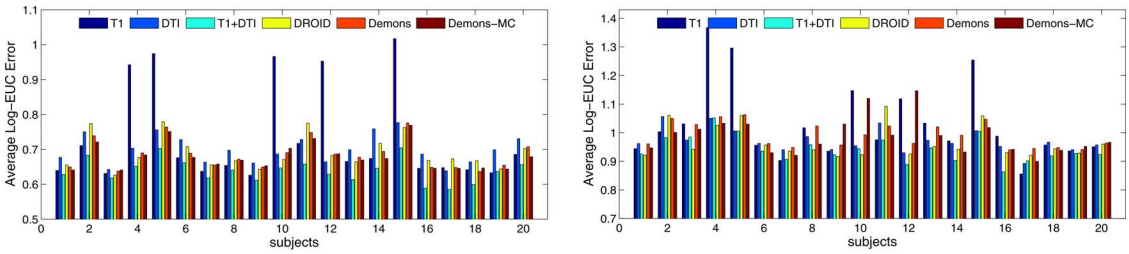


Fig. 15. Comparison of the registration accuracy on real DTI images. T1 + DTI refers the proposed multivariate registration method. T1 and DTI indicate the methods which use the same deformation model and features, but only use information from T1 or DTI, respectively. DROID (for DTI) and Demons (for T1) are other two methods used in the comparison but have different deformation model and features. Multichannel Demons (Demons-MC) is another multichannel registration method used for comparison, which is based on matching the image intensity. (a) Errors on GM. (b) Errors on WM.

TABLE III
SUMMARIZATION OF THE EXPERIMENT RESULTS ON REAL DTI IMAGES

Methods	T1	DTI	T1+DTI	DROID	Demons	Demons-MC
mean $\pm$ std (Log-EUC)	0.80 ± 0.15	0.70 ± 0.03	0.66 ± 0.03	0.71 ± 0.04	0.73 ± 0.06	0.71 ± 0.05
min (Log-EUC)	0.66	0.67	0.63	0.66	0.68	0.66
max (Log-EUC)	1.04	0.77	0.72	0.80	0.80	0.81

objective function is then minimized in a hierarchical manner [23] which can be briefly described as the following. In order to have a smoother energy function with fewer local minima, the images of different channels are firstly down-sampled to represent the anatomy in coarser levels. Then, starting from the coarsest level, in each iteration, landmark points such as edges or corners (similar geometric features are also used in [24] and [25]), are selected to find the correspondences on the other image and drive the deformation by using the metric defined in (5). The selected landmark points are first constrained to the most salient features at the beginning of the algorithm. While the ambiguity is reduced in the subsequent iterations, more and more points are added to guide the deformation. When such an iterative optimization process ends at the finest level, the final spatial transformation is generated by concatenating the hierarchical sequence of piecewise smooth transformations obtained at each stage.

III. EXPERIMENTS AND RESULTS

A. Experiment Design

This section reports the results of experiments performed on both simulated and real multichannel images to validate our proposed method. We use a six-channel T1 + DTI image (as shown in Fig. 5) in both the experiments. In each task, our proposed method is compared with other two single modality registration methods: Demons [3] for T1 and DROID [26] for DTI. In DROID, features consisting of the geometric moments invariants (GMI) from the 5 DTI-derived scalar images and the tensor orientation information are used as features to guide the transformation using the HAMMER [23] deformation model. Demons registration is a widely used fluid-like registration method based on image intensity. The basic Demons algorithm can be briefly described as the following. On each voxel, a velocity is defined using the intensity differences and gradient

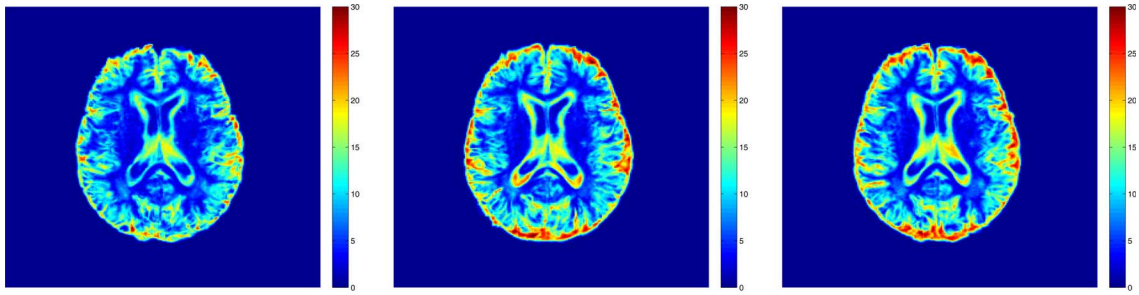


Fig. 16. The variance map of the registered T1. (a), (b) and (c) show the variance map of registered T1 using method T1 + DTI, DROID and DTI, respectively. The smaller variance value shows higher registration accuracy. (a) Method T1 + DTI. (b) Method DROID. (c) Method DTI.

information. Then this velocity field is smoothed and iteratively used to transform the moving image to register to the static image. Due to the fact that deformable registration is a very complex process, many factors can affect the final results, such as the deformation model and features. Therefore, to exclude these factors from the comparison, registration using single modality (either T1 or DTI) and with the same deformation model and features (Gabor features) are also performed. Finally, in order to compare with the multichannel image registration method that works on intensity information without fusion, the results from multichannel Demons [3] (Demons-MC) are also reported and compared. Since DTI is one of the modalities in our experiments, when warping the DTI images the deformation field obtained from each method is also used to determine the tensor reorientation by using a spatially adaptive procedure that estimates the underlying fiber orientation [27]. For simplicity, in the remaining part of the paper, we refer to our registration method as T1 + DTI. The other five methods we compared with are referred as: DROID, Demons, Demons-MC (multichannel Demons), T1 (only using Gabor features from T1), and DTI (only using Gabor features from DTI), respectively.

B. Image Acquisition and Preprocessing

In our experiments, 21 T1 + DTI datasets are used. The images were acquired using Philips 1.5 T REC Scanner. The datasets were acquired at 30 gradient directions with the diffusion weighting of $b = 700$ s/mm^2 (NEX $= 2$). The imaging dimension was 256×256 with a rectangular field of view (FOV) as 240×240 mm^2 and image resolution of $0.9375 \times 0.9375 \times 2.5$ mm. After the data was acquired from the scanner, we first use FSL's Diffusion Toolbox, FDT [28], for eddy current correction. Then, after the nonbrain tissue (skull, dura, and scalp) was removed (by masking), the diffusion tensors are reconstructed. Next, we align the $b = 0$ image first rigidly and then nonrigidly to T1, using FSL's linear image registration tool (FLIRT [29]) andnonlinear image registration tool (FNIRT [30]), respectively. The cost-function chosen is the normalized mutual information [5] in both two steps. The resultant deformation field is then applied to the reconstructed DT images to align the two modalities. As the last step, the five DTI-derived scalar images, FA, apparent diffusion coefficient (ADC) [12], C_s (sphericity), C_p (oblateness), and C_l (prolateness) [22], are calculated.

C. Registration of Simulated Images

For the simulated experiments, we choose one subject from the above dataset as the template [as shown in Fig. 11(a)]. After that, ten simulated DTI + T1 image sets are generated by applying ten simulated deformation fields to the template. The ten simulated images are shown in Fig. 11(b)–(k), respectively. For details of how to generate deformations that reflect realistic inter-subject variation, we refer the readers to [31]. Here we only give a brief description. In the first step, the inter-individual variability is captured by learning from the high dimensional deformation fields using a number of training samples. In this step, due to the spatial localization property of wavelet transforms, they are used to decompose the deformation fields to capture both the coarse and the fine characteristics. Then, principal component analysis (PCA) is performed on the corresponding wavelet coefficients at different scales and locations. The resultant eigenvectors thus represent the principal modes of variations. With these eigenvectors, any anatomical variation can be expressed as the mean variation (by averaging the deformation fields) plus a weighted linear combination of the principal variation modes (represented by the eigenvectors). Finally, in the second step, the simulated deformations can then be constructed by randomly sampling the weights in the above representation of variation.

After generating the simulated images, these ten simulated DTI + T1 image sets are then registered to the template using the proposed method and the five methods mentioned in Section III-A. To compare the performance of all the six methods, for each simulated subject, the deformation fields generated by the six methods are then compared with the simulated deformation fields (ground truth). The results are shown as averaged errors (in voxel) and summarized in Fig. 12. To have a clearer comparison, the errors on GM and WM are shown separately. As shown in Fig. 12(a), the proposed method (labeled as T1 + DTI) gives the smallest registration error on GM amongst all the 6 methods. This result shows that fusing information from DTI with T1 improves the registration even in the GM. This is consistent with the results reported in [8]. The result on WM is shown in Fig. 12(b). When compared with the method using Gabor features only from T1 or DTI, it is found that, by using the complementary information from the T1-weighted image, the registration accuracy is improved on WM as well. However, in contrast to the results on GM, DROID gives the best accuracy

on WM, as the features that DROID uses are WM specific and it utilizes the orientation information. To have an overall comparison of the 6 methods, the registration results on the whole brain are summarized in Table I. As we can see, the proposed method with the fused information gives better accuracy than most of the methods (except DROID which is slightly better) using single channel. Please note that the experiment with simulated datasets has the merit that ground truth is known, it is a far simpler registration task than the real inter-subject registration, because the simulated images are still very similar. Therefore, to evaluate the performance in practice, the results on real image are a more robust standard, in which our method consistently gives the best results.

D. Registration of Real Images

Next, we apply the proposed method to the 21 real T1 + DTI images. Amongst them, one subject is set as the template. The template and the randomly selected 10 subjects are shown in Fig. 13. As we can see, the anatomical variations are substantially large in this population and the subjects are quite different from each other. As in the experiment on real images the ground truth is not known, we use a different method to evaluate the registration performance. The deformation fields are computed from all the six methods and are applied to both the T1 and DTI images, to generate the registered T1 and DTI images, respectively. Then the registration accuracies are compared on both the registered T1 and the registered DTI based on the average sum absolute difference (SAD) and Log-Euclidean (Log-EUC) [32] metric, respectively.

Fig. 14 shows the comparison result on the registered T1 images. It can be seen that on both GM [Fig. 14(a)] and WM [Fig. 14(b)], the proposed method gives the best accuracy in most of the cases. There are a few cases, in which the proposed method is not best, for example, Subject 3 on WM and Subject 4 on GM. However, in these cases, the performance of the proposed method is very close to the method demonstrating the best accuracy. Table II summarizes the registration error of the registered T1 on the whole brain. As we can see, the proposed method statistically outperforms all the other method and has the smallest mean error. This suggests that fusing the information from different modalities can help in achieving robust and accurate registration. The comparison of the registered DTI gives the similar results as shown in Fig. 15 and Table III. In order to have more information of the spatial distribution of the registration error over the whole brain, we create the variance map of the registered T1 image as shown in Fig. 16. It is clearly shown that, by utilizing the information from the T1-weighted image, the proposed method has a more accurate registration on GM than by only using the information from DTI. Fig. 17 shows the mean registered T1 and DTI (depicted using a FA map) images. The sharpness and the preserved details of the mean image provide a visual validation of the accuracy of registration.

Based on the results on both the simulated and real images, it is shown that in general, our proposed method has a more stable performance on both tissue types (GW and WM). In contrast other methods using information from one modality only are found to perform very differently on different tissue types, although in some cases they can do slightly better than our method (depending on how much information can be extracted from a tissue type for that specific subject). Better performance of our method over the other multichannel image registration method, Demons-MC, also suggests the effectiveness of the proposed spatially adaptive information fusion method for multichannel inter-subject image registration.

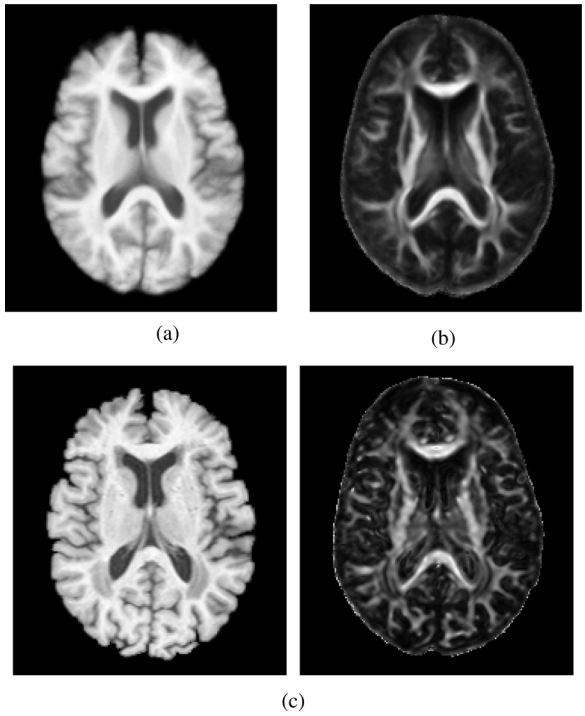


Fig. 17. The mean image of the registered T1 and DTI (shown as FA image). The sharpness and the preserved details of the mean image indicate the high accuracy of registration. (a) The mean T1. (b) The FA map of the mean DTI. (c) Template.

E. Computational Efficiency

For a pair of multichannel images with $256 \times 256 \times 128$ resolution, our proposed method takes about 100 min to get the results on a Linux server base on 2.8 GHz Opteron with 32 GB RAM. This time includes the cost on both Gabor feature extraction ($\approx$ 45 min) and the hierarchical registration process. Please note that the training of the ICA decomposition (introduced in Section II-C2) only needs to be performed once. For other images with the same parameters in imaging, the resultant separation matrix A^{-1} can be applied directly. In fact, recently, with the aid of parallel computation on the GP-GPU (General Purpose GPU), many complex computations in computer vision can be accelerated up to $\times 64$ times, including Gabor Filtering [33]. Therefore, together with an improved multithread implementation (current is single thread) of the registration part, the whole computation time is expected to be reduced further. In term of the computation complexity, the proposed method has another benefit that with the increase of the number of channels, the number of L^2 norm computations (which constitutes most computation in registration) will not increase in (4), since after "choose-max," there is always only one channel/IC left.

IV. Conclusion and Future Work

In this paper, we have proposed a multichannel inter-subject image registration framework that combines information from different modalities based on feature-level information fusion. The registration produces spatially normalized images of all the modalities acquired in the study. Thus the statistical analysis done on these jointly spatially normalized images is more comparable, as a unified registration scheme has been used to register them. We expect the proposed registration method to be very useful in large population clinical studies that acquire several modalities and a unified spatial normalization is needed for subsequent statistical analysis. Although the method is general and is applicable to any number of modalities on which features can be computed, we have applied it to the joint registration of T1 and DTI images, which are routinely acquired in all clinical studies. Experiments on both simulated and real multichannel images $(\mathrm{DTI} + \mathrm{T1})$ illustrate that the proposed method can effectively fuse the information from different modalities and result in a more accurate and robust registration. In the future, we plan to explore more advanced fusion schemes of the Gabor features. As an application, we also plan to apply this method on clinical studies for joint comparative statistics on T1 and DTI.

Acknowledgment

The authors would like to acknowledge the help from the work of Dr. D. Shen and Y. Ou.

References

[1] M. Jenkinson, P. Bannister, J. Brady, and S. Smith, "Improved optimisation for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.

[2] H.-J. Park, M. Kubicki, and M. Shenton, "Spatial normalization of diffusion tensor MRI using multiple channels," *Neuroimage*, vol. 20, pp. 1995–2009, 2003.

[3] J. P. Thirion, "Non-rigid matching using demons," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 1996, pp. 245–251.

[4] B. Avants, J. Duda, H. Zhang, and J. Gee, "Multivariate normalization with symmetric diffeomorphisms for multivariate studies," in *Med. Image Computing Computer-Assisted Intervent. (MICCAI)*, 2007, pp. 359–366.

[5] J. Pluim, J. Maintz, and M. Viergever, "Mutual information based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[6] G. Tourassi, E. Frederick, M. Markey, and C. Floyd, "Application of the mutual information criterion for feature selection in computer-aided diagnosis," *Med. Phys.*, vol. 28, no. 12, 2001.

[7] A. Andronache, M. von Siebenthal, G. Székely, and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Med. Image Anal.*, vol. 12, no. 1, pp. 3–15, Feb. 2008.

[8] C. Studholme, "Incorporating DTI data as a constraint in deformation tensor morphometry between T1 MR images," in *Inf. Process. Med. Imag. (IPMI)*, 2008, pp. 223–232.

[9] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi, "Multi-modal image set registration and atlas formation," *Med. Image Anal.*, vol. 10, pp. 440–451, 2006.

[10] B. Alfano, M. Ciampi, and G. De Pietro, "A wavelet-based algorithm for multimodal medical image fusion," in *Semantic Multimedia*, 2007, pp. 117–120.

[11] Z. Zhang and R. Blum, "A categorization of multiscale-decomposition-based image fusionschemes with a performance study for a digital camera application," *Proc. IEEE*, vol. 87, no. 8, pp. 1315–1326, 1999.

[12] P. J. Bassera and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *J. Magn. Reson.*, vol. 111, no. 3, pp. 209–219, 1996.

[13] J. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoustics*, vol. 36, no. 7, pp. 1169–1179, Jul. 1988.

[14] W. Ma and B. Manjunathi, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[15] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.

[16] S. Muraki, T. Nakai, Y. Kita, and K. Tsuda, "An attempt for coloring multichannel MR imaging data," *IEEE Trans. Vis. Comput. Graph.*, vol. 7, no. 3, pp. 265–273, 2001.

[17] T. Nakai, S. Muraki, and E. Bagarinao, "Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray and white matter," *NeuroImage*, vol. 21, no. 1, pp. 251–260, 2004.

[18] M. McKeown *et al.*, "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapp.*, vol. 6, no. 3, pp. 160–188, 1998.

[19] S. Makino, T.-W. Lee, and H. Sawada, *Blind Speech Separation*. New York: Springer, 2007.

[20] J. Friedman and J. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Trans. Computers*, vol. 23, no. 9, pp. 881–890, Sep. 1974.

[21] A. Hyvarinen, "Fast and robust fixed point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[22] C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis, "Processing and visualization of diffusion tensor MRI," *Med. Image Anal.*, vol. 6, no. 2, pp. 93–108, 2002.

[23] D. Shen, "Image registration by local histogram matching," *Pattern Recognit.*, vol. 40, pp. 1161–1171, 2007.

[24] A. Joshi, D. Shattuck, P. Thompson, and R. Leahy, "Surface-constrained volumetric brain registration using harmonic mappings," *IEEE Trans. Med. Imag.*, vol. 26, no. 12, pp. 1657–1669, Dec. 2007.

[25] G. Postelnicu, L. Zollei, and B. Fischl, "Combined volumetric and surface registration," *IEEE Trans. Med. Imag.*, vol. 28, no. 4, pp. 508–522, Apr. 2009.

[26] J. Yang, D. Shen, C. Davatzikos, and R. Verma, "Diffusion tensor image registration using tensor geometry and orientation features," in *Medical Image Computing Computer-Assisted Intervent. (MICCAI)*, 2008, pp. 905–913.

[27] D. Xu, S. Mori, D. Shen, P. C. M. V. Zijl, and C. Davatzikos, "Spatial normalization of diffusion tensor fields," *Magn. Reson. Med.*, vol. 50, no. 1, pp. 75–182, 2003.

[28] FDT—FMRIB's diffusion Toolbox. FMRIB Centre Dept. Clin. Neurol., Univ. Oxford [Online]. Available: http://poc.vl-e.nl/distribution/manual/fsl-3.2/fdt/index.html

[29] FLIRT—FMRIB's linear image registration tool. FMRIB Centre Dept. Clin. Neurol., Univ. Oxford [Online]. Available: http://www.fmrib.ox.ac.uk/fsl/flirt/index.html

[30] FNIRT—FMRIB's Non-Linear Image Registration Tool. FMRIB Centre Dept. Clin. Neurol., Univ. Oxford [Online]. Available: http://www.fmrib.ox.ac.uk/fsl/fnirt/index.html#fnirt_cost

[31] Z. Xue, D. Shen, B. Karacali, J. Stern, D. Rottenberg, and C. Davatzikos, "Simulating deformations of MR brain images for validation of atlas-based segmentation and registration algorithms," *NeuroImage*, vol. 33, no. 3, pp. 855–866, 2006.

[32] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magn. Reson. Med.*, vol. 56, pp. 411–421, 2006.

[33] I. Young, L. Van Vliet, and M. Van Ginkel, "Recursive Gabor filtering," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2798–2805, Nov. 2002.