

Fast Tensor Image Morphing for Elastic Registration

Pew-Thian Yap^{1,3}, Guorong Wu^{1,3}, Hongtu Zhu^{2,3}, Weili Lin^{1,3},
and Dinggang Shen^{1,3}

¹ Department of Radiology, ² Department of Biostatistics and ³ BRIC
University of North Carolina at Chapel Hill, NC

Abstract. We propose a novel algorithm, called *Fast Tensor Image Morphing for Elastic Registration* or F-TIMER. F-TIMER leverages multiscale tensor regional distributions and local boundaries for hierarchically driving deformable matching of tensor image volumes. Registration is achieved by aligning a set of automatically determined structural landmarks, via solving a soft correspondence problem. Based on the estimated correspondences, thin-plate splines are employed to generate a smooth, topology preserving, and dense transformation, and to avoid arbitrary mapping of non-landmark voxels. To mitigate the problem of local minima, which is common in the estimation of high dimensional transformations, we employ a hierarchical strategy where a small subset of voxels with more distinctive attribute vectors are first deployed as landmarks to estimate a relatively robust low-degrees-of-freedom transformation. As the registration progresses, an increasing number of voxels are permitted to participate in refining the correspondence matching. A scheme as such allows less conservative progression of the correspondence matching towards the optimal solution, and hence results in a faster matching speed. Results indicate that better accuracy can be achieved by F-TIMER, compared with other deformable registration algorithms [1, 2], with significantly reduced computation time cost of 4–14 folds.

1 Introduction

Diffusion tensor imaging (DTI) is capable of measuring water diffusion in vivo non-invasively, and has been widely employed to delineate potential neurological disease related white matter abnormalities. To this end, accurate registration of diffusion tensor images across different subjects is a critical prerequisite for detailed voxel-by-voxel statistical analysis. Spatial normalization of diffusion tensor images is, however, challenging both technically and computationally given that tensor data representation is inherently high dimensional, and the anisotropic nature of cellular water diffusivity calls for proper reorientation of the tensors on top of their spatial alignment, which adds another dimension of difficulty to the problem.

Conventional methods generally extract tensor scalar features from each tensor individually, and by constructing scalar maps, regional integration and other operations such as edge detection can be performed to extract the final features

for correspondence matching. These approaches, however, do not exploit full tensor information, and are limited in the sense that the regional and edge information extracted might not reflect the true underlying tensor structure. While it is possible to use existing scalar image registration algorithms to establish correspondence matching, it is natural to expect that the inclusion of additional information gained from considering the full diffusion tensor will lead to more accurate matching, and hence more robust registration.

In this paper, we propose a fast and accurate diffusion tensor image registration algorithm, called F-TIMER. F-TIMER, in addition to scalar map based features, gathers directly from the tensors regional statistical information and boundary edge information in a multiscale fashion. All the collected information is then grouped, for each voxel, into an *attribute vector* [3], serving as its structural signature. Based on the attribute vectors, salient points, signifying important anatomical structures, are automatically selected, and are utilized as landmarks for correspondence matching. For robust correspondence matching and to avoid being bias towards the template or the subject, the deformation is estimated by jointly considering the forward and reverse transformations [4]. Upon establishing landmark correspondences, thin-plate splines (TPS) [5] are employed to generate a smooth, topology preserving, and dense transformation, and to avoid arbitrary mapping of non-landmark voxels. The whole framework can be summarized as an *Attributive Symmetric Soft Assignment Problem* (ASSAP), which is an extension of the original soft assignment problem [6] by taking into account attribute vector labelled landmarks [3] and also symmetric transformation [4]. To obviate local minima, which are prone to happen, and often inevitable in estimating a transformation with very high degrees of freedom, we adopt a hierarchical strategy in optimizing the ASSAP related energy function. We progressively build up our optimization solution starting with landmarks resulting from a small subset of the voxels, which exhibit more distinctive attribute vectors — equivalent to optimizing an approximate lower-degrees-of-freedom version of the energy function — and include more voxels to refine the solution as registration progresses. Finally, reorientation of the tensors is performed using the algorithm presented in [7]. Such scheme, as can be validated from the experiments, results in more robust and accurate matching in a lesser amount of time.

2 Methods

2.1 Attribute Vectors

In medical imaging, it is important to build deformable anatomical models that take into account the underlying anatomy, and not simply the similarity of image intensities. To this end, a multiscale attribute vector is attached to each voxel, reflecting its underlying anatomical structure in a local scale, and also its relationship to more distant voxels in a more global scale. A rich enough attribute vector can potentially differentiate different parts of the anatomy that would otherwise look similar. F-TIMER characterizes a voxel using three different types of features: 1) Regional features $\mathbf{a}_r^{\text{Reg}}(\mathbf{x})$ (means and variances),

2) Edge features $\mathbf{a}_r^{\text{Edg}}(\mathbf{x})$ (tensor edges and FA map edges), and 3) Geometrical features $\mathbf{a}_r^{\text{Geo}}(\mathbf{x})$ (FA values and principal diffusivities). These features are computed in three different scales (Fine, Middle, Coarse) and are grouped, for each voxel \mathbf{x} , into an attribute vector: $\mathbf{a}(\mathbf{x}) = [\mathbf{a}_{\text{Fine}}(\mathbf{x}), \mathbf{a}_{\text{Middle}}(\mathbf{x}), \mathbf{a}_{\text{Coarse}}(\mathbf{x})]$, where, for each scale $r \in \{\text{Fine}, \text{Middle}, \text{Coarse}\}$, there are three types of features: $\mathbf{a}_r(\mathbf{x}) = [\mathbf{a}_r^{\text{Reg}}(\mathbf{x}), \mathbf{a}_r^{\text{Edg}}(\mathbf{x}), \mathbf{a}_r^{\text{Geo}}(\mathbf{x})]$. In the following, we describe and define $\mathbf{a}_r^{\text{Reg}}(\mathbf{x})$, $\mathbf{a}_r^{\text{Edg}}(\mathbf{x})$, $\mathbf{a}_r^{\text{Geo}}(\mathbf{x})$, and also measures for gauging attribute vector similarity.

Regional Features. Utilizing Log-Euclidean metrics, we can define the regional mean in the neighborhood $\mathcal{N}(\mathbf{x})$ of voxel \mathbf{x} as: $\mathbf{M}(\mathbf{x}) = \exp \left[\frac{\sum_{\mathbf{z} \in \mathcal{N}(\mathbf{x})} \log(\mathbf{D}(\mathbf{z}))}{|\mathcal{N}(\mathbf{x})|} \right]$, where $\mathbf{D}(\mathbf{x})$ is the tensor pertaining to voxel \mathbf{x} . From the mean, the principal diffusivities, i.e., the eigenvalues, can be computed: $\lambda_1^{(\mathbf{M})}(\mathbf{x}) \geq \lambda_2^{(\mathbf{M})}(\mathbf{x}) \geq \lambda_3^{(\mathbf{M})}(\mathbf{x})$, where $\lambda_k^{(\mathbf{M})}(\mathbf{x})$ represents the k -th largest eigenvalue of matrix $\mathbf{M}(\mathbf{x})$. Similarly, we can define the regional variance as: $\mathbf{V}(\mathbf{x}) = \left[\frac{\sum_{\mathbf{z} \in \mathcal{N}(\mathbf{x})} [\log(\mathbf{D}(\mathbf{z})) - \log(\mathbf{M}(\mathbf{x}))]^2}{|\mathcal{N}(\mathbf{x})|} \right]$, and the principal variabilities as: $\lambda_1^{(\mathbf{V})}(\mathbf{x}) \geq \lambda_2^{(\mathbf{V})}(\mathbf{x}) \geq \lambda_3^{(\mathbf{V})}(\mathbf{x})$. These eigenvalues are scaled according to the following equation to yield their mean normalized values: $\tilde{\lambda}_i^{(\mathbf{V})}(\mathbf{x}) = \lambda_i^{(\mathbf{V})}(\mathbf{x}) / \sum_{k=1}^3 \lambda_k^{(\mathbf{M}^2)}(\mathbf{x})$ where we have used \mathbf{M}^2 instead of \mathbf{M} to match the dynamic range of \mathbf{V} . The notation $[\cdot]$ indicates that the quantity in the bracket is in the log space. The regional features are computed for each scale r , and are grouped into: $\mathbf{a}_r^{\text{Reg}}(\mathbf{x}) = [\lambda_{1,r}^{(\mathbf{M})}(\mathbf{x}), \lambda_{2,r}^{(\mathbf{M})}(\mathbf{x}), \lambda_{3,r}^{(\mathbf{M})}(\mathbf{x}), \tilde{\lambda}_{1,r}^{(\mathbf{V})}(\mathbf{x}), \tilde{\lambda}_{2,r}^{(\mathbf{V})}(\mathbf{x}), \tilde{\lambda}_{3,r}^{(\mathbf{V})}(\mathbf{x})]$.

Edge Features. To better extract tissue boundaries, the Canny edge detector is extended to cater for diffusion tensor images. For fast edge detection, 3D Gaussian-based image filtering is implemented using three subsequent steps of one-dimensional (1D) Gaussian filtering along the anterior-posterior, superior-inferior and left-right directions, which is then followed by gradient maps computation. Using these steps, edge detection can be accomplished rapidly and robustly. Note that tensor edge detection is performed in the logarithmic space [8]. For each voxel in the volume, a gradient $\mathbf{H}^{\text{Tensor}}(\mathbf{x})$ can be computed, and from which, after non-maximum suppression, a final edge magnitude $H^{\text{Tensor}}(\mathbf{x})$ can be obtained. Edges from tensors and edges from FA map are complementary to each other and, by using both, potentially all major kinds of tissue boundaries, that is, those formed between white matter (WM), gray matter (GM) and cerebro-spinal fluid (CSF), can be detected and aligned in the registration. We denote the edge magnitude returned by the FA map detection at point \mathbf{x} as $H^{\text{FA}}(\mathbf{x})$. For scale r , the edge features are grouped into: $\mathbf{a}_r^{\text{Edg}}(\mathbf{x}) = [H_r^{\text{Tensor}}(\mathbf{x}), H_r^{\text{FA}}(\mathbf{x})]$.

Geometrical Features. The rest of the features used in F-TIMER includes the fractional anisotropy value $\text{FA}(\mathbf{x})$ and the principal diffusivities $\lambda_1^{(\mathbf{D})}(\mathbf{x}) \geq \lambda_2^{(\mathbf{D})}(\mathbf{x}) \geq \lambda_3^{(\mathbf{D})}(\mathbf{x})$, which characterize the geometrical shape of the tensor ellipsoid. For scale r , the geometrical features are grouped into: $\mathbf{a}_r^{\text{Geo}}(\mathbf{x}) = [\text{FA}_r(\mathbf{x}), \lambda_{1,r}^{(\mathbf{D})}(\mathbf{x}), \lambda_{2,r}^{(\mathbf{D})}(\mathbf{x}), \lambda_{3,r}^{(\mathbf{D})}(\mathbf{x})]$.

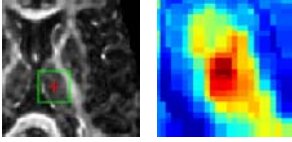


Fig. 1. Distinctiveness of attribute vector. The similarity map in the green box is magnified for a closer inspection. Dark red indicates high similar and dark blue otherwise.

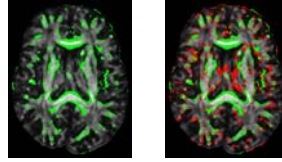


Fig. 2. Landmarks superimposed on the FA image. Shown on the left are the initial landmarks and on the right, in red, are the additional landmarks when the registration progresses to a later stage.

Similarity Measures. Prior to computing the attribute vector similarity, we normalize the elements of the attribute vectors to have a range of $[0, 1]$. For a template voxel \mathbf{x}_t with attribute vector $\mathbf{a}_T(\mathbf{x}_t)$ and a subject voxel \mathbf{x}_s with attribute vectors $\mathbf{a}_S(\mathbf{x}_s)$, their voxel similarity is defined as: $m(\mathbf{a}_T(\mathbf{x}_t), \mathbf{a}_S(\mathbf{x}_s)) = \prod_i (1 - |a_{T,i}(\mathbf{x}_t) - a_{S,i}(\mathbf{x}_s)|)$, where $a_{T,i}(\mathbf{x})$ and $a_{S,i}(\mathbf{x})$ are the i -th elements of $\mathbf{a}_T(\mathbf{x})$ and $\mathbf{a}_S(\mathbf{x})$, respectively. Fig. 1 illustrates that the attribute vectors are rich enough to warrant differentiation of different anatomical structures. For robust correspondence matching, rather than computing the similarity on a voxel-to-voxel basis, we compare the similarity of the voxels in the neighborhood of \mathbf{x}_t with that of \mathbf{x}_s , and vice-versa. For a voxel \mathbf{x} in the neighborhood of \mathbf{x}_t in the template space, the distance $\mathbf{x} - \mathbf{x}_t$ corresponds to $\Delta_s(\mathbf{x}, \mathbf{x}_t) = f(\mathbf{x}) - f(\mathbf{x}_t)$ in the subject space. Likewise, for voxels \mathbf{x} and \mathbf{x}_s in the subject space, their corresponding distance in the template space is $\Delta_t(\mathbf{x}, \mathbf{x}_s) = f^{-1}(\mathbf{x}) - f^{-1}(\mathbf{x}_s)$. We can hence define the regional similarity measures, in both forward and reverse directions, as:

$$\text{sim}_{T \rightarrow S}(\mathbf{x}_t, \mathbf{x}_s) = m(\mathbf{a}_T(\mathbf{x}_t), \mathbf{a}_S(\mathbf{x}_s)) \sum_{\mathbf{x} \in \mathcal{N}_1(\mathbf{x}_t)} \frac{m(\mathbf{a}_T(\mathbf{x}), \mathbf{a}_S(\mathbf{x}_s + \Delta_s(\mathbf{x}, \mathbf{x}_t)))}{|\mathcal{M}_1(\mathbf{x}_t)|},$$

$$\text{sim}_{S \rightarrow T}(\mathbf{x}_t, \mathbf{x}_s) = m(\mathbf{a}_T(\mathbf{x}_t), \mathbf{a}_S(\mathbf{x}_s)) \sum_{\mathbf{x} \in \mathcal{N}_2(\mathbf{x}_s)} \frac{m(\mathbf{a}_T(\mathbf{x}_t + \Delta_t(\mathbf{x}, \mathbf{x}_s)), \mathbf{a}_S(\mathbf{x}))}{|\mathcal{M}_2(\mathbf{x}_s)|},$$

where $\mathcal{N}_1(\cdot)$ and $\mathcal{N}_2(\cdot)$ denote the respective neighborhoods, and $|\mathcal{N}(\mathbf{x})|$ the cardinality of set $\mathcal{N}(\mathbf{x})$. These regional similarity measures are used to optimize the ASSAP energy function (1), which will be discussed in Section 2.3.

2.2 Landmark Selection

Brain medical images are inherently high dimensional and computations involved in registering these images can be prohibitive. In order to overcome this, out of all possible voxels $\mathbf{X}_T = \{\mathbf{x}_t : t = 1, \dots, M\}$ for the template, and $\mathbf{X}_S = \{\mathbf{x}_s : s = 1, \dots, N\}$ for the subject, ($\mathbf{X}_T, \mathbf{X}_S \in \mathcal{R}^3$), we select as landmarks a subset of voxels with more distinctive attribute vector $\mathbf{X}_S(k) = \{\mathbf{x}_t(k) :$

$t = 1, \dots, M(k)\} \subset \mathbf{X}_S$ and $\mathbf{X}_T(k) = \{\mathbf{x}_s(k) : s = 1, \dots, N(k)\} \subset \mathbf{X}_T$, at iteration k of the registration, for correspondence matching. As the registration progresses, an increasing number of landmarks are used to refine the registration. Besides the obvious benefit of making the correspondence matching problem more feasible, this approach also helps mitigate the problem of local minima. Selecting an initial smaller number of landmarks essentially means that we are now solving a lower-degrees-of-freedom approximation of the transformation, and is hence less prone to be trapped by local minima. As more landmarks start to participate, transformation of increasing complexity can be estimated. F-TIMER selects as landmarks a combination of voxels with the highest tensor edge magnitudes ($H_r^{\text{Tensor}}(\mathbf{x}) > \alpha_r^{\text{TsrEdge}}$), FA map edge magnitudes ($H_r^{\text{FA}}(\mathbf{x}) > \alpha_r^{\text{FAEdge}}$) and FA values ($\text{FA}_r(\mathbf{x}) > \alpha_r^{\text{FA}}$), since these voxels represent important anatomical structures, and are relatively easy to locate in images with sufficient contrast. The values of $\alpha_r^{\text{TsrEdge}}$, α_r^{FAEdge} and α_r^{FA} are initially high, but are progressively lowered to allow more voxels to participate in correspondence matching as registration progresses. Fig. 2 gives an illustration of the landmarks at two different stages of the registration.

2.3 Correspondence Matching and Transformation Estimation

The determination of the non-rigid spatial mapping can be cast into an *Attributive Symmetric Soft Assignment Problem* (ASSAP). We adopt a hierarchical strategy in minimizing the related energy function. At each iteration k , the active landmarks consist of a subset of voxels $\mathbf{X}_T(k) \subset \mathbf{X}_T$ from the template and in $\mathbf{X}_S(k) \subset \mathbf{X}_S$ from the subject. Let f denote a non-rigid spatial mapping. Our goal is to find the optimal correspondence matrices $\hat{\mathbf{P}}(k)$ and $\hat{\mathbf{Q}}(k)$ and an optimal spatial transformation \hat{f} by minimizing an energy function with constituent terms explained as follows:

- We would like to match the driving voxels in $\mathbf{X}_T(k)$ and $\mathbf{X}_S(k)$ as closely as possible in the subject space, but at the same time encourage matching of voxels with similar attribute vectors. And naturally, a voxel pair $\mathbf{x}_t(k)$ and $\mathbf{x}_s(k)$ satisfying these conditions will be deemed a more probable match and given a higher probability $p_{t,s}$ in the energy function. This can be realized with: $E(\mathbf{P}(k), f) = \sum_{t=1}^{M(k)} \sum_{s=1}^{N(k)} p_{t,s}(k) \left\{ \|f(\mathbf{x}_t(k)) - \mathbf{x}_s(k)\|^2 - \log [\text{sim}_{T \rightarrow S}(\mathbf{x}_t(k), \mathbf{x}_s(k))] \right\}$, where $\mathbf{P}(k) = \{p_{t,s}(k)\}$.
- We require mapping f to be consistent in that equal consideration is given to correspondence matching from the point of view of the template and the subject in determining f . This is to avoid bias towards the template or the subject, and to avoid local minima resulting from correspondence ambiguity. We thus incorporate a term in the energy function, which is symmetric to that defined above, by requiring similar matching in the template space: $E(\mathbf{Q}(k), f) = \sum_{t=1}^{M(k)} \sum_{s=1}^{N(k)} q_{t,s}(k) \left\{ \|\mathbf{x}_t(k) - f^{-1}(\mathbf{x}_s(k))\|^2 - \log [\text{sim}_{S \rightarrow T}(\mathbf{x}_t(k), \mathbf{x}_s(k))] \right\}$ where $\mathbf{Q}(k) = \{q_{t,s}(k)\}$.

- Soft correspondence are permitted in the initial stages of the registration so that the energy function is smooth and better behaved. Towards the end of the registration, more exact one-to-one correspondence is enforced. This is realized by energy terms: $E(\mathbf{P}(k)) = \alpha \sum_{t=1}^{M(k)} \sum_{s=1}^{N(k)} p_{t,s}(k) \log(p_{t,s}(k))$, $E(\mathbf{Q}(k)) = \alpha \sum_{t=1}^{M(k)} \sum_{s=1}^{N(k)} q_{t,s}(k) \log(q_{t,s}(k))$. Parameters α controls the degree of fuzziness of the matching. It has initially high values, encouraging fuzzy matching, and later progressively lower values, which enforce exact matching.
- Mapping f is required to be smooth so as to preserve the topology and to avoid arbitrary mapping of non-landmark voxels. This is enforced by energy term: $E(f) = \beta \|\mathcal{L}f\|^2$. \mathcal{L} is an operator which aids in measuring the bending energy. β is a weighting factor which is decreased throughout the registration to allow f to model deformation of increasing complexity.

The overall energy function that needs to be optimized is:

$$E(\mathbf{P}(k), \mathbf{Q}(k), f) = \{E(\mathbf{P}(k), f) + E(\mathbf{Q}(k), f) + E(\mathbf{P}(k)) + E(\mathbf{Q}(k)) + E(f)\}. \quad (1)$$

In the following, we describe an optimization strategy to minimize (1), which basically alternates between correspondence matching (**Step 1**) and dense transformation estimation (**Step 2**). We first fix f and solve for the \mathbf{P} and \mathbf{Q} by letting $\partial E(\mathbf{P}(k), \mathbf{Q}(k), f) / \partial p_{t,s}(k) = 0$, $\partial E(\mathbf{P}(k), \mathbf{Q}(k), f) / \partial q_{t,s}(k) = 0$. We then fix $\mathbf{P}(k)$ and $\mathbf{Q}(k)$, and solve for f using TPS.

Step 1: Correspondence Matching: The correspondence matrices $\mathbf{P}(k)$ and $\mathbf{Q}(k)$ can be updated as follows:

$$p_{t,s}(k) = \text{sim}_{T \rightarrow S}(\mathbf{x}_t(k), \mathbf{x}_s(k)) e^{-\|f(\mathbf{x}_t(k)) - \mathbf{x}_s(k)\|^2 / \alpha}, \quad p_{t,s}(k) \leftarrow p_{t,s}(k) / \sum_{s=1}^{N(k)} p_{t,s}(k)$$

$$q_{t,s}(k) = \text{sim}_{S \rightarrow T}(\mathbf{x}_t(k), \mathbf{x}_s(k)) e^{-\|\mathbf{x}_t(k) - f^{-1}(\mathbf{x}_s(k))\|^2 / \alpha}, \quad q_{t,s}(k) \leftarrow q_{t,s}(k) / \sum_{t=1}^{M(k)} q_{t,s}(k).$$

Overly weak matches are prevented by setting $p_{ts}(k)$ or $q_{ts}(k)$ falling below a predefined threshold to zero prior to normalization.

Step 2: Dense Transformation Estimation:

$$\sum_{t=1}^{M(k)} \sum_{s=1}^{N(k)} \left[p_{t,s}(k) \|f(\mathbf{x}_t(k)) - \mathbf{x}_s(k)\|^2 + q_{t,s}(k) \|\mathbf{x}_t(k) - f^{-1}(\mathbf{x}_s(k))\|^2 \right] + \beta \|\mathcal{L}f\|^2$$

which, as it stands, is very cumbersome. A slightly different form is implemented:

$$\min_f E(f) = \min_f \left\{ \sum_{t=1}^{M(k)} \|f(\mathbf{x}_t(k)) - \hat{\mathbf{x}}_t(k)\|^2 + \sum_{s=1}^{N(k)} \|f(\hat{\mathbf{x}}_s(k)) - \mathbf{x}_s(k)\|^2 + \beta \|\mathcal{L}f\|^2 \right\}$$

where $\hat{\mathbf{x}}_t(k) = \sum_{s=1}^{N(k)} p_{t,s}(k) \mathbf{x}_s(k)$, $\hat{\mathbf{x}}_s(k) = \sum_{t=1}^{M(k)} q_{t,s}(k) \mathbf{x}_t(k)$. The variables $\hat{\mathbf{x}}_t(k)$ and $\hat{\mathbf{x}}_s(k)$ can be seen as the newly estimated locations in the subject and

template spaces, which correspond to $\mathbf{x}_t(k)$ and $\mathbf{x}_s(k)$, respectively. The above equation can be solved using TPS fitting [5], which essentially minimizes the geometrical distance of the driving voxels and, at the same time, minimizes the bending energy $\|\mathcal{L}f\|^2$.

3 Experimental Results

The dataset consisted of diffusion tensor images of 22 subjects, acquired using a 1.5T scanner. The imaging dimension and resolution were 256×256 and $0.9375 \times 0.9375 \times 2.5\text{mm}^3$, respectively. Whenever appropriate, results yielded by the deformable registration algorithms, presented by Yang et al. [1] and Zhang et al. [2], will be included for comparison. They will be referred to as YANG and ZHANG, respectively, in the rest of the paper. It is worth noting here that registration using F-TIMER takes approximately 15 min on a 2.66GHz Linux machine, compared to 210 min using YANG, and 60 min using ZHANG.

3.1 Real Subjects

One subject was selected from the dataset and taken as the template. 21 subjects were then registered onto this template. By averaging all the registered images, we could visually inspect the accuracy of the registration. The result is shown in Fig. 3. It can be observed that for the FA map based affine registered images, their average image, shown in Fig. 3(b), is fuzzy especially in areas near the cortical region. In comparison, after registration with F-TIMER, the average image, Fig. 3(c), shows much improved sharpness. It is difficult to gauge the performance of F-TIMER by visual comparison with the average images yielded by YANG and ZHANG, shown in Fig. 3(c) & (d). We describe a quantitative approach in the upcoming section.

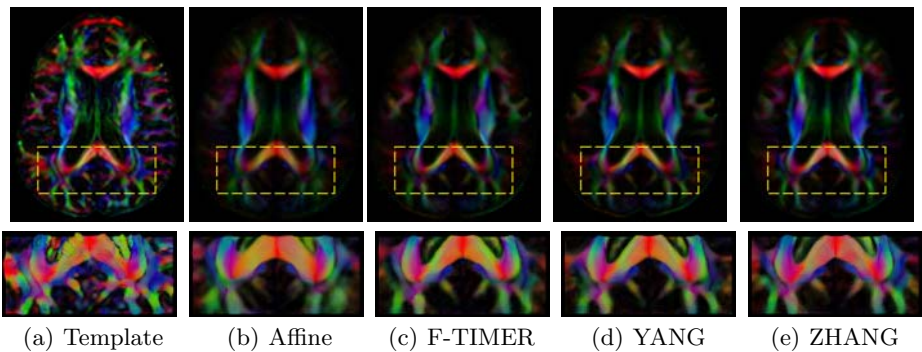


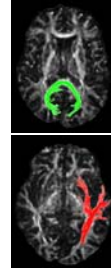
Fig. 3. Group-averaged images resulting from the registration of the 21 subjects. The FA weighted first principal directions are shown in their color coded representations: green for the anterior-posterior direction, blue for the superior-inferior direction, and red for the left-right direction. The tensors in the yellow boxes are shown in their FA weighted ellipsoidal representations in the bottom panels.

Table 1. Deformation estimation errors (*mm*)

	Whole Brain			Cortical Region		
	F-TIMER	YANG	ZHANG	F-TIMER	YANG	ZHANG
Mean	1.36	1.87	1.44	1.43	2.03	1.68
S. D.	1.14	1.48	1.75	1.17	1.59	1.96
<i>p</i> -value		3.50×10^{-8}	2.80×10^{-1}		7.71×10^{-8}	1.28×10^{-2}

Table 2. Fiber bundle distances (*mm*)

	F-TIMER	YANG	ZHANG
Splenium	0.87	0.91	0.98
S. D.	0.015	0.016	0.044
<i>p</i> -value		3.51×10^{-1}	5.06×10^{-2}
Cortical	1.34	1.49	1.91
S. D.	0.150	0.115	0.208
<i>p</i> -value		2.08×10^{-1}	1.30×10^{-4}



3.2 Simulated Subjects

To further evaluate the accuracy of F-TIMER, we generated 20 simulated deformation fields using the statistical model of deformation (SMD) proposed in [9]. One human brain was chosen from the dataset as the template and the 20 simulated deformation fields, which also served as the ground truths, were applied to the template, resulting in 20 simulated human brains. These 20 simulated brains were then registered back onto the template using F-TIMER and the deformation fields estimated by the registration were compared with the ground truths, using Euclidean distance. Results shown in Table 1 indicate that F-TIMER yields higher accuracy. F-TIMER is also more consistent as indicated by the smaller deviation values. Also shown in Table 1 are the results for the cortical region and similar conclusion can be drawn. The significance of the improvement of F-TIMER over YANG and ZHANG is indicated by the paired *t*-test *p*-values in the same table.

3.3 Fiber Tracking

Using a tractography method know as FACT [10], fiber bundles passing through some regions of interest (ROIs) were tracked, extracted, and compared for quantifying the registration accuracy of F-TIMER. Based on the simulated data generated above, we present here two sets of results. One ROI was selected so a fiber bundle residing in the splenium of the corpus callosum (CC) could be extracted for comparison. The second ROI was selected so that we could evaluate F-TIMER in a more difficult situation where a fiber bundle near the cortical surface was extracted. The fiber bundles are shown in Table 2. The distance of two fiber bundles was then measured using the *mean of the closest distances*, in

a way similar to that used in [2]. A summary of the results of all fiber bundles is shown in Table 2. The results indicate that F-TIMER yields better performance when compared to YANG and ZHANG.

4 Conclusion

F-TIMER extracts distinctive features from a diffusion tensor image, drawing on multiscale tensor regional and boundary information, as automated structural landmarks. Employing these landmarks to minimize an *Attributive Symmetric Soft Assignment Problem* (ASSAP) related energy function, a smooth, topology preserving, and consistent transformation can be derived. Experimental results show that F-TIMER can achieve sufficiently good accuracy with relatively low computation cost.

Acknowledgments. This work was supported in part by grant 1R01EB006733.

References

1. Yang, J., Shen, D., Davatzikos, C., Verma, R.: Diffusion tensor image registration using tensor geometry and orientation features. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 905–913. Springer, Heidelberg (2008)
2. Zhang, H., Yushkevich, P.A., Alexander, D.C., Gee, J.C.: Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Med. Image Anal.* 10(5), 764–785 (2006)
3. Shen, D., Davatzikos, C.: HAMMER: Heirarchical attribute matching mechanism for elastic registration. *IEEE TMI* 21(11), 1421–1439 (2002)
4. Christensen, G.E., Johnson, H.J.: Consistent image registration. *IEEE TMI* 20(7), 568–582 (2001)
5. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE TPAMI* 11(6), 567–585 (1989)
6. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. *CVIU* 89(2-3), 114–141 (2003)
7. Xu, D., Mori, S., Shen, D., van Zijl, P.C.M., Davatzikos, C.: Spatial normalization of diffusion tensor fields. *Mag. Res. in Med.* 50(1), 175–182 (2003)
8. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Mag. Res. in Med.* 56(2), 411–421 (2006)
9. Xue, Z., Shen, D., Karacali, B., Stern, J., Rottenberg, D., Davatzikos, C.: Simulating deformations of MR brain images for validation of atlas-based segmentation and registration algorithms. *NeuroImage* 33(3), 855–866 (2006)
10. Jiang, H., van Zijl, P.C., Kim, J., Pearlson, G.D., Mori, S.: DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking. *Computer Methods and Programs in Biomedicine* 81(2), 106–116 (2006)