DTI registration in atlas based fiber analysis of infantile Krabbe disease

Yi Wang, Aditya Gupta, Zhexing Liu, Hui Zhang, Maria L. Escolar, John H. Gilmore, Sylvain Gouttard, Pierre Fillard, Eric Maltbie, Guido Gerig, Martin Styner

Abstract

In recent years, diffusion tensor imaging (DTI) has become the modality of choice to investigate white matter pathology in the developing brain. To study neonate Krabbe disease with DTI, we evaluate the performance of linear and non-linear DTI registration algorithms for atlas based fiber tract analysis. The DTI scans of 10 age-matched neonates with infantile Krabbe disease are mapped into an atlas for the analysis of major fiber tracts—the genu and splenium of the corpus callosum, the internal capsules tracts and the uncinate fasciculi. The neonate atlas is based on 377 healthy control subjects, generated using an unbiased diffeomorphic atlas building method. To evaluate the performance of one linear and seven nonlinear commonly used registration algorithms for DTI we propose the use of two novel evaluation metrics: a regional matching quality criterion incorporating the local tensor orientation similarity, and a fiber property profile based metric using normative correlation. Our experimental results indicate that the whole tensor based registration method within the DTI-ToolKit (DTI-TK) shows the best performance for our application.

Corresponding author. E-mail address: adigupta@med.unc.edu (A. Gupta).
et al., 2001) and along the DTI fiber bundle of internal capsules (IC) when compared with healthy age-matched controls (Escolar et al., 2009). Escolar et al. (2009) also showed a correlation of pretreatment FA measurements with post treatment gross motor function.

Based on the above research findings (Escolar et al., 2009; Goodlett et al., 2009), we use an atlas based fiber tract analysis for analyzing DTI images of Krabbe subjects. For an accurate analysis it is crucial to establish a registration based voxel-wise correspondence between a normal control neonate DTI atlas (with prior information of fiber tract locations) and the Krabbe subjects’ DTI images. The research presented in this paper highlights our work to determine the best state-of-the-art approach to individually register DTI images of Krabbe subjects into the atlas space.

**Challenges in DTI registration for Krabbe neonates**

The registration of diffusion tensor images is particularly challenging when compared to registering scalar images as DTI data is multi-dimensional and the tensor orientations after image transformations must remain consistent with the anatomy (Alexander et al., 2001; Gee and Alexander, 2005). The application of the registration methods on DTI of Krabbe neonates makes the problem even more challenging due to the following factors. Most of the registration methods discussed in this paper are based on the intensity of the fiber tracts in the fractional anisotropy maps and as discussed earlier, the Krabbe patients have lower FA values as compared to the control group. Lower FA values are due to the anisotropy caused by the demyelination of the nerves. Relatively rapid changes occur in white matter during the first year of life restricting the control provided age matched controls to a relatively narrow age range relative to the patient. Also regional variations between FA values in white matter sites could cause inaccurate comparisons and hence the analysis needs to be performed in specific well defined white matter structures (Provenzale et al., 2005). In addition to these points, the analysis in this paper is restricted to neonates and this adds to the complexity as DTI MRI of neonates have low signal-to-noise (SNR) and poorly developed white matter tracts.

**DTI registration algorithms**

DTI registration algorithms can be broadly categorized into two groups (Zhang et al., 2006). The first kind uses scalar images derived from DTI images and performs deformable registration with traditional image registration algorithms (Schnabel et al., 2001; Joshi et al., 2004; Andersson et al., 2007; Christensen et al., 1994; Christensen et al., 1997). Although this group discards the orientation component of the data, it is the most commonly used method because of the simplicity and the ease of implementation. The second group of DTI registration algorithms directly use higher order information of diffusion tensor images like the corresponding principal eigenvectors (Yap et al., 2009), or the full tensor information (Zhang et al., 2006; Yeo et al., 2008). Due to the complexity involved and the difficulty in realizing such algorithms, this group has not been explored extensively.

In this paper, we investigate eight DTI registration approaches from both groups, available either in-house or publicly:

1) Affine registration by Studholme et al. (1999) using normalized mutual information as a registration metric within the Image Registration Toolkit1 (referred to as Affine in this paper).

2) B-spline based registration by Schnabel et al. (2001) using normalized mutual information as a registration metric within the Image Registration Toolkit (referred to as B spline in this paper).

3) B-spline based registration by Andersson et al. (2007) using weighted sum of scaled sum-of-squared differences as a registration metric via the “fnirt” implementation within FSL2 (referred to as FSL in this paper).

4) Diffeomorphic demons3 by Vercauteren et al. (2009) using sum-of-squared differences as a registration metric (referred to as Demons in this paper).

5) Log demons3 by Vercauteren et al. (2008) using sum-of-squared differences as a registration metric (referred to as Demons-log in this paper).

6) Fluid registration by Joshi et al. (2004) using sum-of-squared differences as a registration metric (referred to as Fluid in this paper).

7) Tensor-based registration by Zhang et al. (2006) using explicit optimization of tensor reorientation in an analytic manner within DTI-ToolKit4 (referred to as DTI-TK in this paper).

8) Diffeomorphic tensor-based registration by Yeo et al. (2008) using the exact finite strain gradient within MedINRIA5 (referred to as MedINRIA in this paper).

The first six methods are based on normalized FA maps whereas the last two are whole tensor based registration methods. An evaluation of algorithms from both the groups will give an insight into the higher performance of one group over the other, particularly considering the complexities in registering Krabbe neonates. To evaluate the performance of the registration algorithms, we introduce two novel evaluation metrics. The first metric is based on the matching quality of the local tensor orientation and atlas anisotropy in each voxel. The voxel-wise metric values are averaged over predefined regions within the atlas (such as the genu, splenium, internal capsules and uncinates). The second evaluation metric employs a normative fiber tract profile based criterion, which computes the correlation of the FA profile along the major tracts in the registered dataset and the atlas.

**Materials and methods**

**Subjects**

The studies are approved by the institutional review board at the University of North Carolina. Due to the difficulty of Krabbe data acquisition, only ten neonates with Krabbe disease identified by family history or through the New York State screening program were used in this study. The ten Krabbe neonates are aged 8 to 67 days (mean: 22 days) at the time of scan. These subjects were referred to the Program for Neuro-developmental Function in Rare Disorders (NFRD) at the University of North Carolina at Chapel Hill for assessment of baseline neurologic function before receiving unrelated umbilical cord blood transplantation at Duke University Medical Center. The assessment included a detailed neurodevelopmental evaluation concurrent with a brain MR imaging within the first four weeks of life. 377 age-matched neonatal controls (aged 7 days to 92 days with a mean value of 23 days) were recruited in a separate, unrelated study of brain development in normal controls and high risk offspring as part of UNC’s Conte center (Knickmeyer et al., 2008).

**Scans**

All neonates (control and Krabbe subjects) were scanned without sedation on an Allegra 3T head-only MR scanner (Magnetom Allegra; Siemens, Erlangen, Germany). Two separate DTI protocols were employed due to the change in DTI acquisition methodology. Protocol 1, the protocol employed in scans — before July 2008, acquired seven references.

---

images, one without diffusion gradient ($b=0$) and six diffusion weighted images along unique gradient directions with $b=1000$ s/mm$^2$ (TR=4219 ms; TE=92.2 ms; in-plane resolution = $2 \times 2$ mm$^2$; slice thickness = 2 mm; five averages). Since July 2008, a newer protocol (protocol 2) was employed to improve SNR and the gradient direction acquisition scheme. Protocol 2, forty-nine images are acquired, seven without diffusion gradients ($b=0$) and 42 diffusion weighted images along unique gradient directions with $b=1000$ s/mm$^2$ (TR = 7680 ms; TE = 82 ms; in-plane resolution = $2 \times 2$ mm$^2$; slice thickness = 2 mm; one average). The first seven Krabbe neonates as well as all healthy control subjects were scanned with protocol 1 (K1 to K7). The three final Krabbe neonates were scanned with protocol 2 (K8 to K10). No sedation was used; all scans were performed with subjects fully asleep. Neonates were fed before scanning, then swaddled, put to sleep and were fitted with ear protection and had their heads secured in a vacuum-fixation device. A physician or nurse was present during each scan; a pulse oximeter was used to monitor heart rate and oxygen saturation. More details of the image acquisition and processing can be found in Gilmore et al. (2004).

**DTI atlas building**

In order to build the normative DTI atlas, we used a scalar, unbiased diffeomorphic atlas building method based on a nonlinear high-dimensional fluid deformation method (Joshi et al., 2004; Goodlett et al., 2006). The DTI derived intensity-histogram normalized FA is selected as the feature for atlas building. Nonlinear transformations are applied on the feature image to produce a deformation field for each image. All the tensor images are then reoriented into the unbiased space using the finite strain approximation proposed by Alexander et al. (2001). The atlas is then developed by averaging all the reoriented tensor images in log-Euclidean space (Arsigny et al., 2006). The selection of normalized FA image as the feature is based on the studies of Liu et al. (2010), wherein the authors show that this feature is the best scalar feature for DTI atlas building among all the other scalar measurements and their combinations. We did not investigate the use of alternative atlas building methods as part of this paper.

**Pre-processing of Krabbe datasets**

All the Krabbe datasets were subjected to a quality control (QC) using the DTIPrep tool to identify any artifacts in the diffusion weighted images (DWI), as well as to correct for motion and eddy current artifacts. The datasets were also cropped or embedded into consistent image dimensions. Diffusion tensors were then estimated for each dataset from the QCed DWIs using weighted least squares tensor estimation (Salvador et al., 2005). Skull stripping was performed semi-automatically for all Krabbe datasets by a trained expert.

**Registration methods**

In this section, we briefly present the working principle of the registration algorithms evaluated in this paper. The first five methods are based on intensity-histogram normalized FA images, while the last two are based on the whole tensor information. **Affine** registration is a linear transform method that is commonly used as an initialization step for most deformable registrations (Studholme et al., 1999). The **Affine** registration used in this paper optimizes fifteen linear parameters (three for rotation, translation and scaling and six for skewing—defining the skewing angles in different planes) by maximizing the normalized mutual information. This is accomplished in a multi-resolution framework using Gaussian smoothing to compute lower resolution steps.

**B-spline** is a parametric, non-rigid image registration method based on multi-resolution adaptable free-form deformations using B-splines (Schnabel et al., 2001; Rueckert et al., 1999). Similar to **Affine**, this method also maximizes normalized mutual information in a multi-resolution framework using Gaussian smoothing to compute lower resolution steps.

**FSL** (or rather “FSL-B-Spline”) is similar to the previous method in that it represents displacement fields as B-splines on a regular grid (Andersson et al., 2007). But in this method the regularization of the field is based on membrane energy and the registration criterion is based on the weighted sum-of-squared intensity differences and the membrane energy.

**Demons** is a non-parametric, diffeomorphic deformable image registration algorithm based on the Thirion (Thirion, 1998) Demons warp software in the Insight Toolkit (V ercauteren et al., 2007a,b and Vercauteren et al., 2009). The deformation model is based on optical flow and the registration criterion is based on the sum-of-squared intensity differences.

**Demons-log** is similar to the above Demons but works completely in the log-domain, i.e. it uses a stationary velocity field to encode the spatial transformation as its exponential (Vercauteren et al. 2008). **Fluid** is a non-parametric, diffeomorphic deformable image registration that employs a dilatational-viscous fluid flow formulation (Joshi et al., 2004) with sum-of-squared intensity differences as the registration criterion.

**DTI-TK** is a non-parametric, diffeomorphic deformable image registration (Zhang et al., 2007) that incrementally estimates its displacement field using a tensor-based registration formulation (Zhang et al., 2006). It is designed to take advantage of similarity measures comparing whole tensors via explicit optimization of tensor reorientation (Zhang et al., 2006).

**MedINRIA** is also a diffeomorphic deformable image registration (Yeo et al., 2008) that incorporates the exact finite strain gradient into a diffeomorphic DTI registration scheme.

In order to be consistent across methods, we adopted the deformation fields from each registration method and performed DTI reorientation and interpolation using the same software (ResampleDTIlogEuclidean$^2$) based on standard finite strain tensor realignment (Alexander et al., 2001).

For all the registration methods, the default parameters were used except in the case of Fluid, wherein the parameters were slightly modified for comparable performance.

**Evaluation of registration accuracy**

In central WM, FA values in DTI of neonates are considerably lower than those at older ages (Gilmore et al., 2003). The WM pathology causes FA values of Krabbe patients to be even lower than those of healthy age-matched controls. In addition, Krabbe subjects are likely to have regionally differing levels of white matter pathology. All of these observations indicate that the development of an evaluation criterion for the registration algorithms is a challenging task, but also that the results may not easily be generalized to other settings.

For our evaluation, we tested one linear and seven nonlinear algorithms to determine the most suitable method for our application. We mainly focused on the tracts of i) the genu of the corpus callosum ii) the splenium of the corpus callosum iii) the internal capsule of both the hemispheres (left and right) and iv) the uncinate tracts (left and right). The same tests can be further extended to a larger selection of tracts.

---


7. [https://www.ia.unc.edu/dev/](https://www.ia.unc.edu/dev/).
While there are several ongoing initiatives towards an unbiased evaluation of deformable registration algorithms (Studholme, 1997; Cao et al., 2005), there is currently no widely accepted metric standard for the evaluation of nonlinear registration algorithms, even more so for DTI registration. The following sections discuss our evaluation strategy.

**Visual assessment**

The first step of our evaluation strategy consists of a qualitative, visual quality control. To achieve this, we visualize FA and color-oriented FA images of all the registered datasets using a multi-dataset overview with MRIWatcher. While this kind of assessment is subjective, significant errors can be easily detected. The registration is judged to have failed on datasets that show large errors.

**Regional matching quality criterion**

For the second step in our evaluation framework, we propose a novel regional matching criterion that is tailored to atlas based analysis methods. In our specific setting, we have the following conditions: 1) the streamline fiber tractography employed in our fiber based analysis framework follows a concept developed by Mori et al. (1999) and Xu et al. (2002), which is based on the local principal eigenvectors \( e_p \) (i.e. the vector associated with the largest of the three principal directions of the diffusion tensor: \( \lambda_1 \geq \lambda_2 \geq \lambda_3 \)); 2) fiber tracts of the genu and splenium of the corpus callosum as well as both hemispheric internal capsule and uncinate tracts have higher intensity in FA images as compared to their neighboring tracts.

The orientation agreement between the principal eigenvectors of the individual subject (source) and the atlas (target) is the basis of this criterion. In order to enhance the specific regions associated with the selected fiber tracts and to render the method stable against small changes in the regional definition, we use the FA value of the atlas as a weight on the local orientation criterion. Thus, the proposed similarity value \( s_v \) is defined for each voxel as:

\[
    s_v = |e_p \cdot e_{pa}| FA_A
\]

where \( e_p \) is the subject’s principal eigenvector, \( e_{pa} \) is the atlas’ principal eigenvector, and \( FA_A \in (0, 1) \) is the atlas FA value. Notation \( | \cdot | \) in the above equation indicates the absolute value and \( \cdot \cdot \) indicates the dot product. In the particular case that the principal eigenvectors of the individual subject and the atlas are oriented in the same or fully opposite direction, the term \( |e_p \cdot e_{pa}| \) becomes \( ||e_{pa}||^2 \), which is 1, and \( s_v \) will be equal to \( FA_A \). Using this local criterion, we compute a scalar matching image representing the registration quality at each voxel.

Next, regions of interest (ROIs) on the atlas are defined representing WM sections the major fiber tracts. The average regional similarity value on these ROIs represents the regional matching quality criterion. Thus for region \( r \), the average similarity value \( \bar{s}_r \) is:

\[
    \bar{s}_r = \frac{1}{N_r} \sum_{i=1}^{N_r} s_{v,i}
\]

where \( N_r \) is the number of voxels in region \( r \), \( s_{v,i} \) is the similarity value at voxel \( i \). Larger values of \( \bar{s}_r \) represent better registration accuracy in our settings.

---

8 https://www.ia.unc.edu/dev/download/mriwatcher.
Fiber property profile based criterion

As a third step in our evaluation framework, we propose another novel matching criterion that evaluates the DTI property measurements along the fiber tracts, called tract profiles (Goodlett et al., 2009). The fiber tracts tracked in atlas space are resampled in each registered DTI dataset. Using a prior definition of a tract origin plane, which defines a curvilinear re-parameterization of the tracts, corresponding average tract property profiles (we focus on FA profiles) are extracted from each individual fiber tract. The average is thereby computed across the individual streamlines and the profile is plotted along the fiber tract.

For the evaluation, we calculated the normative correlation between each fiber tract profile in the registered subjects’ DTI datasets and the atlas. For this criterion also we expect larger values to represent better accuracy in the registration. It is further noteworthy that high degrees of white matter pathology are likely to decrease this evaluation metric, though that does not lessen its comparative merit in the presented work here.

Results

Visualization results

We present detailed results for two individual representative cases, K1 (with protocol 1) and K8 (with protocol 2), as well as the summary results across the whole Krabbe population of 10 subjects. As illustrated in Fig. 1, qualitative inspection of the registration results indicate that all deformable registration algorithms show satisfactory results. The linear Affine registration method fails to map the fiber tracts of the subjects into the atlas space, as clearly seen for the internal capsule tracts. Several qualitative differences can be
seen between the registration results, like the result from the B-spline algorithm captures the local anatomy and orientations poorly when compared to the other methods.

Regional matching quality criterion results

To test the regional matching quality criterion, we manually defined six regions on the atlas labeled 1 to 6 (Fig. 2). The regions defined represent the six tracts of interest — genu, splenium, internal capsules (left and right) and uncinate fasciculi (left and right). Fig. 2(a) shows the 3D visualization of ROIs' position inside the brain with Slicer.9 The relationship between the ROIs and the atlas FA image in 2D are shown in Figs. 2(b) and (c) using ITK-SNAP10 (Yushkevich et al., 2006). Fig. 2(d) shows the 3D color visualization of the target fiber tracts — genu (red), splenium (yellow), right hemisphere internal capsule (purple), left hemisphere internal capsule (blue) and uncinate (green) with Slicer (Catani and Thiebaut de Schotten, 2008; Wakana et al., 2004).

Table 1

<table>
<thead>
<tr>
<th>Fiber Tracts</th>
<th>Affine</th>
<th>B-spline</th>
<th>FSL</th>
<th>Demons</th>
<th>Demons-log</th>
<th>Fluid</th>
<th>DTI-TK</th>
<th>MedINRIA</th>
<th>Best Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genu</td>
<td>MEAN</td>
<td>0.1177</td>
<td>0.1098</td>
<td>0.1257</td>
<td>0.1247</td>
<td>0.1233</td>
<td>0.1237</td>
<td>0.1272</td>
<td>0.124 DTI-TK</td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0055</td>
<td>0.0063</td>
<td>0.0031</td>
<td>0.0028</td>
<td>0.0033</td>
<td>0.0025</td>
<td>0.0032</td>
<td>0.0042</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0007**</td>
<td>0.0001**</td>
<td>0.0174*</td>
<td>0.0924</td>
<td>0.0184</td>
<td>0.0101*</td>
<td>0.0538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Splenium</td>
<td>MEAN</td>
<td>0.1414</td>
<td>0.132</td>
<td>0.1506</td>
<td>0.1494</td>
<td>0.1473</td>
<td>0.1475</td>
<td>0.1519</td>
<td>0.149 DTI-TK</td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0062</td>
<td>0.0071</td>
<td>0.0029</td>
<td>0.0038</td>
<td>0.0025</td>
<td>0.0025</td>
<td>0.0037</td>
<td>0.0059</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0008**</td>
<td>0.0001**</td>
<td>0.3114</td>
<td>0.1179</td>
<td>0.0043</td>
<td>0.0089**</td>
<td>0.1711</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Left hemisphere internal capsule</td>
<td>MEAN</td>
<td>0.1325</td>
<td>0.1501</td>
<td>0.1908</td>
<td>0.1882</td>
<td>0.187</td>
<td>0.1843</td>
<td>0.1925</td>
<td>0.1895 DTI-TK</td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0091</td>
<td>0.0166</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0045</td>
<td>0.0023</td>
<td>0.0032</td>
<td>0.0018</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0001**</td>
<td>0.0001**</td>
<td>0.3014</td>
<td>0.0243</td>
<td>0.0186</td>
<td>0.0001**</td>
<td>0.0295*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Left hemisphere uncinate</td>
<td>MEAN</td>
<td>0.0746</td>
<td>0.0883</td>
<td>0.1357</td>
<td>0.1181</td>
<td>0.1158</td>
<td>0.0926</td>
<td>0.135</td>
<td>0.1317 FSL</td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0219</td>
<td>0.0156</td>
<td>0.0018</td>
<td>0.0215</td>
<td>0.0224</td>
<td>0.0203</td>
<td>0.0039</td>
<td>0.0027</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001**</td>
<td>0.0001**</td>
<td>0.0475</td>
<td>0.0314</td>
<td>0.0001**</td>
<td>0.6117</td>
<td>0.0039**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Average FA profiles of the splenium for the atlas and the ten subjects for the eight registration methods. The black profile indicates mean of the Krabbe subjects. X-axis: points along the fiber tracts.

9 http://www.slicer.org/.  
10 http://www.itksnap.org/.
The average similarity values for each registration method for genu, splenium, left hemisphere internal capsule and left hemisphere uncinate is shown in Table 1. The analysis on the right hemisphere internal capsule and uncinate show similar results to their corresponding left hemisphere tracts and hence their tables are not shown. The values in the tables indicate that the similarity values of subject K1 and subject K8 agree with the visualization results. This illustrates the effectiveness of our regional matching quality criterion as a potential for quality control of DTI registration, as well as a valid evaluation measure that highlights differences across methods.

Results from Table 1 show that DTI-TK gives the best results. This algorithm shows the best performance for the tracts of genu, splenium and both hemispheric internal capsules. For the uncinate fasciculi, the performance is second best to the FSL method. None of these seven registration methods can be said to give optimal results on each ROI for every subject.

We repeated the full evaluation with slightly modified regional definitions (the six regions were independently and manually redefined). The ranking of the methods was preserved in all regions, thus indicating that the computed regional matching criterion is stable and reliable.

**Fiber property profile based criterion results**

Fiber tractography of the genu, splenium, both hemispheric internal capsules and uncinctes is performed on the atlas. Fiber bundles of each subject are then created using the method in Goodlett et al. (2009). The FA profiles along the fiber tracts — genu, splenium, left internal capsule and left uncinate, generated from the seven registration methods are plotted for the atlas and the ten subjects (Figs. 3–6).

Analysis of the FA profiles gives further information on the performance of the registration methods. Considering all the fiber tracts, the Fluid registration results in slightly higher magnitudes of the subjects’ FA profiles compared to the other methods. The selected tracts are the tracts with the highest FA intensities and hence higher values of FA indicates better mapping of the subject into the atlas and hence better registration. So based on this criteria, Fluid has a slightly better performance. The other aspect that can be interpreted though only visually from the FA profiles is the visual match of the average Krabbe profile with the atlas profile. In that regard, Affine registration, followed by B-spline, clearly shows a poor mapping. FSL and both the tensor based registration methods — DTI-TK and MediNRIA, show a strong similarity of the FA profiles of the subjects to the atlas FA profile, with a slightly better matching for DTI-TK. The mean of the Krabbe subjects (black line) is very similar to the atlas (red line) both in terms of magnitude and shape for these three algorithms. In the case of Fluid, Demons and Demons-log the mean of the Krabbe subjects has a very similar shape profile as the atlas but has a different (higher) magnitude. Thus, considering the magnitude of the FA profiles, Fluid shows a higher performance than the other methods, whereas the shape of the FA profiles match visually best for the DTI-TK.

The FA profiles indicate the quasi-Euclidean distance of the tensors to a spherical shape and from the shape and magnitude of the profiles it appears that the tensor based methods are compensating the shape of the tensors (to make them more spherical — isotropic) while trying to map the tensors in to the atlas space. In a certain sense, it appears...
that these methods are slightly over-fitting the data. Correlation coefficients between the FA profiles in the registered subjects' dataset and the atlas for all the registration methods are shown in Table 2. No one algorithm shows the best performance for all the tracts. Demons-log shows the best performance for both the internal capsules and the right uncinate fasciculus. DTI-TK and FSL show the best result for the splenium, genu and the left uncinate fasciculus. Hence it is difficult to identify one best algorithm based on the normative tract profile correlation evaluation though summarized over all fiber tracts Demons-log followed by DTI-TK seems to do the best.

Based on the correlation coefficients, we used an additional evaluation criterion to determine the number of subjects' ROIs that the algorithms maps correctly into the atlas. We considered three different correlation values of 0.8, 0.85 and 0.9 as thresholds and correlation co-efficient below the threshold are marked as a failure for mapping the fiber tract into the atlas. Table 3 shows the number of

---

**Table 2**

Correlation coefficients between FA profiles for various fiber tracts in registered subjects and the atlas for the eight registration algorithms. Results significantly different from the best performance algorithm are marked with * (p<.05) and ** (p<.01).

<table>
<thead>
<tr>
<th>Fiber Tracts</th>
<th>Affine</th>
<th>B-spline</th>
<th>FSL</th>
<th>Demons</th>
<th>Demons-log</th>
<th>Fluid</th>
<th>DTI-TK</th>
<th>MedINRIA</th>
<th>Best Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genu</strong></td>
<td>0.5082</td>
<td>0.9057</td>
<td>0.9506</td>
<td>0.9039</td>
<td>0.9227</td>
<td>0.926</td>
<td>0.9487</td>
<td>0.9369</td>
<td><strong>FSL</strong></td>
</tr>
<tr>
<td>MEAN</td>
<td>0.2851</td>
<td>0.0691</td>
<td>0.0325</td>
<td>0.0955</td>
<td>0.0429</td>
<td>0.056</td>
<td>0.068</td>
<td>0.0815</td>
<td><strong>DTI-TK</strong></td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0007**</td>
<td>0.0721</td>
<td>0.2017</td>
<td>0.1262</td>
<td>0.2644</td>
<td>0.0315</td>
<td>0.2822</td>
<td>0.0976</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0007**</td>
<td>0.0721</td>
<td>0.2017</td>
<td>0.1262</td>
<td>0.2644</td>
<td>0.0315</td>
<td>0.2822</td>
<td>0.0976</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td><strong>Splenium</strong></td>
<td>0.6431</td>
<td>0.8594</td>
<td>0.8868</td>
<td>0.8173</td>
<td>0.8386</td>
<td>0.8648</td>
<td>0.8927</td>
<td>0.8368</td>
<td><strong>DTI-TK</strong></td>
</tr>
<tr>
<td>MEAN</td>
<td>0.2884</td>
<td>0.0709</td>
<td>0.0629</td>
<td>0.064</td>
<td>0.0559</td>
<td>0.0952</td>
<td>0.0479</td>
<td>0.074</td>
<td><strong>DTI-TK</strong></td>
</tr>
<tr>
<td>STDEV</td>
<td>0.0200*</td>
<td>0.1042</td>
<td>0.7485</td>
<td>0.0013</td>
<td>0.0114</td>
<td>0.2037</td>
<td>0.0585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0200*</td>
<td>0.1042</td>
<td>0.7485</td>
<td>0.0013</td>
<td>0.0114</td>
<td>0.2037</td>
<td>0.0585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td><strong>Left hemisphere internal capsule</strong></td>
<td>0.5712</td>
<td>0.8445</td>
<td>0.9453</td>
<td>0.9617</td>
<td>0.9693</td>
<td>0.9578</td>
<td>0.9499</td>
<td>0.9318</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td>MEAN</td>
<td>0.0047</td>
<td>0.1456</td>
<td>0.0223</td>
<td>0.0145</td>
<td>0.0104</td>
<td>0.0106</td>
<td>0.0118</td>
<td>0.0272</td>
<td><strong>Demons-log</strong></td>
</tr>
<tr>
<td>STDEV</td>
<td>&lt;0.0001**</td>
<td>0.233</td>
<td>0.0172</td>
<td>0.0546</td>
<td></td>
<td>0.0048</td>
<td>0.0052</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001**</td>
<td>0.233</td>
<td>0.0172</td>
<td>0.0546</td>
<td></td>
<td>0.0048</td>
<td>0.0052</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td><strong>Left hemisphere uncinate</strong></td>
<td>0.5568</td>
<td>0.8546</td>
<td>0.9146</td>
<td>0.8923</td>
<td>0.8988</td>
<td>0.7282</td>
<td>0.8995</td>
<td>0.8625</td>
<td><strong>FSL</strong></td>
</tr>
<tr>
<td>MEAN</td>
<td>0.133</td>
<td>0.051</td>
<td>0.0292</td>
<td>0.0808</td>
<td>0.0993</td>
<td>0.0838</td>
<td>0.0318</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>STDEV</td>
<td>&lt;0.0001**</td>
<td>0.0018</td>
<td>0.0479</td>
<td>0.0765</td>
<td></td>
<td>0.0001**</td>
<td>0.195</td>
<td>0.0046**</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001**</td>
<td>0.0018</td>
<td>0.0479</td>
<td>0.0765</td>
<td></td>
<td>0.0001**</td>
<td>0.195</td>
<td>0.0046**</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 6.** Average FA profiles of the left hemispheric uncinate for the atlas and the ten subjects for the eight registration methods. The black profile indicates mean of the Krabbe subjects. X-axis: points along the fiber tracts.
instances the algorithms fail to map the six DTI fiber bundles into the atlas for the ten subjects for a threshold of 0.85. Affine fails for almost all the cases even for the 0.8 threshold. DTI-TK results in minimum number of failures for all the three thresholds and can be considered as the best algorithm based on this criterion. The success of DTI-TK in correctly mapping the six fiber tracts for all the subject cases can be attributed to the fact that the algorithm exploits the whole tensor orientation information for registration compared to the scalar FA values. The Demons-log and the FSL algorithms show the next best performance. This can be attributed to a histogram based intensity normalization step of the subjects to the atlas prior to these two registration methods. The small variation in the intensities of the six selected tracts results in the FA based registration methods mapping certain regions of the subject to the atlas more accurately than the other regions. The tensor based methods use the orientation information and hence have an advantage in mapping fiber tract related information of the subjects to the atlas more accurately.

Discussion

In this paper, we evaluated one linear and seven nonlinear registration methods for use in an atlas based DTI fiber analysis framework on 10 neonates with infantile Krabbe disease. No difference was observed between the two different protocols in terms of their registration accuracy. We used visual evaluation, tensor orientation based criteria, FA profile based criteria, the correlation of the FA values and the number of failures to evaluate the performance of the registration methods. By visual evaluation, the linear Affine registration method and the B-spline method show a poor matching of the subject to the atlas. The regional matching quality criterion based on the local orientation of the tensors, which is highly relevant to fiber tract analysis, shows that on average, the whole tensor registration using explicit optimization of tensor reorientation — DTI-TK method performed better than the other methods. The criterion based on the correlation values of the subjects’ to the atlas shows that Demons-log, followed by DTI-TK and FSL show a better performance. Considering the number of failure rates in mapping the subjects’ fiber tracts to the atlas, the DTI-TK algorithm has the lowest failure rate.

From the results obtained based on the various criteria, it appears that all the deformable methods give a satisfactory performance. Depending on the selected criteria, different algorithms show slightly higher performance than the others. Of the above discussed criteria, for DTI fiber tract analysis, the local orientation of the tensors and fiber mapping minimal failure rate are the most crucial. Based on these two criteria, we recommend the DTI-TK registration method based on explicit optimization of tensor reorientation for fiber tract analysis.

As is the case with any evaluation metric, the question can be raised, whether some of these evaluation measures could be used for the purpose of registration itself. The answer to this question with respect to the regional matching criterion, which can be represented in a voxel-wise manner, is currently under investigation in our lab.

Acknowledgments

This work was supported by the Natural National Science Foundation of China (Grant No.60903127); the National Alliance for Medical Image Computing (NAMIC, NIH U54 EB005149); the National Institutes of Health (NIH) Roadmap for Medical Research (U54 EB005149–01); the Autism Centers of Excellence Network at UNC-CH (NIH R01 HD055741), Penn Image Computing & Science Laboratory (PICSL) and in part by the NIH Biomedical Imaging and Bioengineering (NIBIB) and the NIH Blueprint for Neuroscience (RO3-EB009321), the Neurodevelopmental Research Center at UNC-CH (NIH P30 HD03110); Ao-Xiang Star Project at Northwestern Polytechnical University, Xi’an, Shaanxi, China and the National Institute of Mental Health Conite Center at UNC-CH (MH064065).

We thank Daniel Rueckert and IXICO. The Image Registration Toolkit of Fluid was used under License from Ixico Ltd.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2011.01.038.

References


