

Groupwise Registration for Correcting Subject Motion and Eddy Current Distortions in Diffusion MRI Using a PCA Based Dissimilarity Metric

W. Huizinga, C.T. Metz, D.H.J. Poot, M. de Groot, W.J. Niessen, A. Leemans, and S. Klein

Abstract Before starting a diffusion-weighted MRI analysis, it is important to correct any misalignment between the diffusion-weighted images (DWIs) that was caused by subject motion and eddy current induced geometric distortions. Conventional methods adopt a pairwise registration approach, in which the non-DWI, a.k.a. the $b = 0$ image, is used as a reference. In this work, a groupwise affine registration framework, using a global dissimilarity metric, is proposed, which eliminates the need for selecting a reference image and which does not rely on a specific method that models the diffusion characteristics. The dissimilarity metric is based on principal component analysis (PCA) and is ideally suited for DWIs, in which the signal contrast varies drastically as a function of the applied gradient orientation. The proposed method is tested on synthetic data, with known ground-truth transformation parameters, and real diffusion MRI data, resulting in successful alignment.

Keywords Image registration • Diffusion-weighted MRI • Subject motion correction • Principal component analysis

W.J. Niessen

Department of Applied Sciences, Imaging Science & Technology, Delft University of Technology, Delft, The Netherlands

Department of Radiology & Medical Informatics, Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, The Netherlands

W. Huizinga (✉) · C.T. Metz · D.H.J. Poot · S. Klein

Department of Radiology & Medical Informatics, Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, The Netherlands

e-mail: w.huizinga@erasmusmc.nl

M. de Groot

Department of Radiology & Medical Informatics, Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, The Netherlands

Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands

A. Leemans

Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands

1 Introduction

Diffusion MRI of the brain is based on the anisotropic diffusion of water due to the presence of nerve fibers. For diffusion MRI models, such as the diffusion tensor model, it is crucial for reliable analysis that all acquired images are spatially aligned. However, during the acquisition, subject motion and eddy currents may cause misalignment, which are typically characterized with global 3D affine transformation models, [11, 17]. Alignment of the diffusion-weighted images (DWIs) is commonly achieved by a pairwise affine registration of every single DWI to the non-DWI (i.e., $b = 0$) image [6, 17]. Recently, several groupwise registration approaches were proposed, e.g. [1, 13, 15, 23], in which a global dissimilarity metric is minimized during a simultaneous optimization of the transform parameters for all images. In this way, a bias towards the reference image is avoided and the intensity information of all images is taken into account simultaneously, leading to more consistent registration results [13, 15]. Especially for data with a low signal-to-noise ratio (SNR) it is preferable to use the intensity information of all images simultaneously.

In this work, we propose a novel groupwise registration method for diffusion MRI data, using a dissimilarity metric based on principal component analysis (PCA). Our method is based on the assumption that aligned data can be modeled by a limited number of principal components with high eigenvalues, while unaligned data will need more principal components with high eigenvalues. Rohde et al. have previously used this principle as a post-hoc evaluation method [17]. Melbourne et al. [14] also use PCA in a registration. They use a progressive principal component registration (PPCR), iteratively subtracting the principal components from the original image data.

In our work we use the PCA during registration by deriving a dissimilarity metric that explicitly maximizes the selected eigenvalues. This approach does not make assumptions about the diffusion, such as the diffusion tensor model, which means that it can potentially be used for any other diffusion approach as well, such as diffusion kurtosis imaging, [5], or diffusion spectrum imaging [25].

2 Method

2.1 Groupwise Registration Framework

Let $M_g(\mathbf{x})$ be the series of $g \in \{1 \dots G\}$ images to be registered, with \mathbf{x} a 3D voxel position. During registration, a transformation $T_g(\mathbf{x}; \mu_g)$ is applied to each image $M_g(T_g(\mathbf{x}; \mu_g))$. The transformation $T_g(\mathbf{x}; \mu_g)$ is parameterized by μ_g and the goal is to find μ_g . For the groupwise registration approach the transform parameters for each separate volume are concatenated into one parameter vector

$\boldsymbol{\mu} = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, \dots, \boldsymbol{\mu}_G^T)^T$. Transform parameters $\hat{\boldsymbol{\mu}}$ are estimated by minimizing a dissimilarity metric $\mathcal{D} : \hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} \mathcal{D}(\boldsymbol{\mu})$.

2.2 Dissimilarity Metric

Given sample locations \mathbf{x}_i with $i \in \{1 \dots N\}$, we can define $\mathbf{M}(\boldsymbol{\mu})$:

$$\mathbf{M}(\boldsymbol{\mu}) = \begin{pmatrix} M_1(\mathbf{T}_1(\mathbf{x}_1; \boldsymbol{\mu}_1)) & \dots & M_G(\mathbf{T}_G(\mathbf{x}_1; \boldsymbol{\mu}_G)) \\ \vdots & \ddots & \vdots \\ M_1(\mathbf{T}_1(\mathbf{x}_N; \boldsymbol{\mu}_1)) & \dots & M_G(\mathbf{T}_G(\mathbf{x}_N; \boldsymbol{\mu}_G)) \end{pmatrix} \quad (1)$$

The dissimilarity metric is based on PCA performed on the measurements in $\mathbf{M}(\boldsymbol{\mu})$. Define the $G \times G$ correlation matrix \mathbf{K} associated with \mathbf{M} :

$$\mathbf{K} = \boldsymbol{\Sigma}^{-1} \mathbf{C} \boldsymbol{\Sigma}^{-1} = \frac{1}{N-1} \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \overline{\mathbf{M}})^T (\mathbf{M} - \overline{\mathbf{M}}) \boldsymbol{\Sigma}^{-1}, \quad (2)$$

$\overline{\mathbf{M}}$ is a matrix within each column the column-wise average of \mathbf{M} . $\boldsymbol{\Sigma}$ is a diagonal matrix that equals the square root of the diagonal of the covariance matrix \mathbf{C} : $\boldsymbol{\Sigma} = \text{diag}[\sqrt{C_{11}}, \dots, \sqrt{C_{gg}}, \dots, \sqrt{C_{GG}}]$. Element i, j of \mathbf{K} describes the correlation between the images M_i and M_j . The correlation between an image with itself is one by definition. On the diagonal of \mathbf{K} we therefore find only ones. The trace of \mathbf{K} is then equal to the number of images in the series, G . Our metric is then defined as

$$\mathcal{D}_{\text{PCA}} = \sum_{g=1}^G K_{gg} - \sum_{j=1}^L \lambda_j = G - \sum_{j=1}^L \lambda_j, \quad (3)$$

where λ_j is the j -highest eigenvalue of \mathbf{K} and L is a user-defined number. It is expected that for diffusion weighted images that follow the diffusion tensor model, the optimal value for L agrees with the number of free parameters in the tensor model. The diffusion tensor model has seven free parameters: the six independent diffusion tensor elements and the intensity for $b = 0$. Usually more images are acquired to obtain a better estimation of the model parameters. The redundancy in the number of measured images becomes visible when PCA is applied: a limited number (L) of eigenvectors describes most of the variance in the aligned data. However, when the images are misaligned, the number of eigenvalues that describe most of the data variance is higher. During optimization of \mathcal{D}_{PCA} , the total variance minus the L highest eigenvalues is minimized. The image data is registered such that it can be described best by the eigenvectors belonging to the L highest eigenvalues. Real data does not follow the diffusion tensor model everywhere, [2]. The real data

is more complicated than the model fitted so it is expected that for real data the optimal value for $L \geq 7$.

2.3 Metric Derivative

For minimization with gradient based optimizers the metric derivative with respect to $\boldsymbol{\mu}$ must be known. Differentiating Eq. (3) with respect to $\boldsymbol{\mu}$ and using the approach of van der Aa et al. [21] for calculating eigenvalue derivatives we get:

$$\frac{\partial \mathcal{D}}{\partial \boldsymbol{\mu}} = - \sum_{j=1}^L \frac{\partial \lambda_j}{\partial \boldsymbol{\mu}} = - \sum_{j=1}^L \mathbf{v}_j^T \frac{\partial \mathbf{K}}{\partial \boldsymbol{\mu}} \mathbf{v}_j, \quad (4)$$

where \mathbf{v}_j^T is the j th eigenvector of the matrix of \mathbf{K} . The unlikely repetition of eigenvalues which leads to linear combination of eigenvectors being also an eigenvector, invalidating the above expression, is ignored [21]. Schultz and Seidel [18] also use this approach, for calculating eigenvalue derivatives of the diffusion tensor in DW-MRI data. Using Eqs. (2) and (4) we get for the derivative of \mathcal{D} with respect to an element μ_p :

$$\begin{aligned} \frac{\partial \mathcal{D}}{\partial \mu_p} = & - \frac{2}{N-1} \sum_{i=1}^L \left[\mathbf{v}_i^T \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \bar{\mathbf{M}})^T \left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \bar{\mathbf{M}}}{\partial \mu_p} \right) \boldsymbol{\Sigma}^{-1} \mathbf{v}_i \right. \\ & \left. + \mathbf{v}_i^T \boldsymbol{\Sigma}^{-1} (\mathbf{M} - \bar{\mathbf{M}})^T (\mathbf{M} - \bar{\mathbf{M}}) \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p} \mathbf{v}_i \right]. \end{aligned} \quad (5)$$

The expression above is obtained after some simplifications and using the fact that

$$\mathbf{v}_i^T \mathbf{B}^T \mathbf{E} \mathbf{v}_i = \mathbf{v}_i^T \mathbf{E}^T \mathbf{B} \mathbf{v}_i. \quad (6)$$

for two matrices \mathbf{B} and \mathbf{E} and vector \mathbf{v}_i .

The derivative of $\boldsymbol{\Sigma}^{-1}$ with respect to μ_p is equal to

$$\frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p} = - \frac{\boldsymbol{\Sigma}^{-3}}{N-1} \text{diag} \left[(\mathbf{M} - \bar{\mathbf{M}})^T \left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \bar{\mathbf{M}}}{\partial \mu_p} \right) \right] \quad (7)$$

and $\partial \mathbf{M} / \partial \mu_p$ and $\partial \bar{\mathbf{M}} / \partial \mu_p$ are computed using

$$\frac{\partial \mathbf{M}_g(\mathbf{T}_g(\mathbf{x}_i; \boldsymbol{\mu}_g))}{\partial \mu_p} = \left(\frac{\partial \mathbf{M}_g}{\partial \mathbf{x}} \right)^T \left(\frac{\partial \mathbf{T}_g}{\partial \mu_p} \right) (\mathbf{x}_i; \boldsymbol{\mu}_g). \quad (8)$$

2.4 Transformation Model

The applied affine transformation is defined as $\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_g) = \mathbf{A}(\mathbf{x} - \mathbf{c}) + \mathbf{t} + \mathbf{c}$, where \mathbf{A} is a matrix without restrictions and \mathbf{c} is the center of rotation. For the parameterization we use an exponential mapping of \mathbf{A} , similar to [23]:

$$\mathbf{A} = \exp(\boldsymbol{\Gamma}) = \exp \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 \\ \mu_4 & \mu_5 & \mu_6 \\ \mu_7 & \mu_8 & \mu_9 \end{pmatrix} \quad \text{and} \quad \mathbf{t} = (\mu_{10}, \mu_{11}, \mu_{12})^T, \quad (9)$$

where $\exp(\cdot)$ is the matrix exponential and omitting subscript g for clarity. For the calculation of the metric derivative, $\partial \mathbf{T}_g / \partial \boldsymbol{\mu}_g$ is required. This derivative is trivial for the translation part of the transform. For the linear part the approach of Fung is applied [4]. Consider the following system of differential equations:

$$\frac{d}{dt} \mathbf{y} = \boldsymbol{\Gamma} \mathbf{y} \quad \text{with solution at } t = 1 : \quad \mathbf{y}(1) = \exp(\boldsymbol{\Gamma}) \mathbf{y}(0). \quad (10)$$

Now differentiate Eq. (10) with respect to μ_p :

$$\frac{d}{dt} \left(\frac{\partial}{\partial \mu_p} \mathbf{y} \right) = \frac{\partial \boldsymbol{\Gamma}}{\partial \mu_p} \mathbf{y} + \boldsymbol{\Gamma} \frac{\partial}{\partial \mu_p} \mathbf{y} \quad (11a)$$

$$\frac{\partial}{\partial \mu_p} \mathbf{y}(1) = \frac{\partial \exp(\boldsymbol{\Gamma})}{\partial \mu_p} \mathbf{y}(0) + \exp(\boldsymbol{\Gamma}) \frac{\partial}{\partial \mu_p} \mathbf{y}(0) \quad (11b)$$

and define:

$$\mathbf{z} = \begin{pmatrix} \frac{\partial}{\partial \mu_p} \mathbf{y} \\ \mathbf{y} \end{pmatrix} \quad \text{and} \quad \tilde{\boldsymbol{\Gamma}} = \begin{pmatrix} \boldsymbol{\Gamma} & \frac{\partial \boldsymbol{\Gamma}}{\partial \mu_p} \\ 0 & \boldsymbol{\Gamma} \end{pmatrix}. \quad (12)$$

Then Eqs. (11a) and (11b) can be written as:

$$\frac{d}{dt} \mathbf{z} = \tilde{\boldsymbol{\Gamma}} \mathbf{z}, \quad \text{with solution at } t = 1 : \quad \mathbf{z}(1) = \exp(\tilde{\boldsymbol{\Gamma}}) \mathbf{z}(0) \quad (13a)$$

$$\mathbf{z}(1) = \begin{pmatrix} \exp(\boldsymbol{\Gamma}) & \frac{\partial}{\partial \mu_p} \exp(\boldsymbol{\Gamma}) \\ 0 & \exp(\boldsymbol{\Gamma}) \end{pmatrix} \mathbf{z}(0). \quad (13b)$$

Combining Eqs. (13a) and (13b) it follows that $\frac{\partial}{\partial \mu_p} \exp(\boldsymbol{\Gamma})$ can be extracted from $\exp(\tilde{\boldsymbol{\Gamma}})$.

2.5 Optimization

The adaptive stochastic gradient descent (ASGD) [7] is used for optimization. This optimizer randomly samples the image in order to reduce computation time. A conventional multi-resolution strategy is used to avoid convergence to local minima. The number of random samples, the number of resolution levels and, the number of iterations per resolution level are user-defined parameters. The average deformation of the DWIs is constrained to zero. To guarantee this the approach of Balci et al. is applied [1]:

$$\frac{\partial \mathcal{D}^*}{\partial \boldsymbol{\mu}_g} = \frac{\partial \mathcal{D}}{\partial \boldsymbol{\mu}_g} - \frac{1}{G} \sum_{g'} \frac{\partial \mathcal{D}}{\partial \boldsymbol{\mu}'_{g'}}. \quad (14)$$

A scaling between the matrix elements of the transform and the translations is necessary, due to the different ranges in voxel displacement that they cause. The scaling is done according to Klein et al. [7].

2.6 Groupwise Approaches Proposed by Others

Wachinger et al. [23] proposed accumulated pair-wise estimates (APE) as a family of metrics. One of the metrics they propose is the sum of squared normalized correlation coefficients. This can be written as the sum of the squared elements of the correlation matrix \mathbf{K} . We implemented this metric as follows:

$$\mathcal{D}_{\text{APE}} = 1 - \frac{1}{G} \sqrt{\sum_j \sum_j K_{ij}^2}. \quad (15)$$

Metz et al. [15] proposed the sum of the variances, assuming no intensity changes between images. The metric is defined as:

$$\mathcal{D}_{\text{VAR}} = \frac{1}{\text{NG}} \sum_{i=1}^N \sum_{g=1}^G \left[M_g(\mathbf{T}_g(\mathbf{x}_i, \boldsymbol{\mu}_g)) - \frac{1}{G} \sum_g M_g(\mathbf{T}_g(\mathbf{x}_i, \boldsymbol{\mu}_g)) \right]^2. \quad (16)$$

Both metrics are compared with our method \mathcal{D}_{PCA} .

2.7 Implementation

The method is implemented in the publicly available registration package `elastix` [8]. In all experiments we used 10,000 samples, 2 resolutions, and 1,000 iterations

per resolution. The sampling was performed off the voxel grid, to reduce interpolation artifacts [8]. Masks are used in all experiments. For the synthetic data a mask is used to exclude the background, in order to only sample in the region of the brain. This is done to make sure that the zero voxel values in the background are not helping the registration. For both the synthetic data and the real data, a mask is used to exclude the high voxel values of the $b = 0$ image (in cerebrospinal fluid). When excluding the high voxel values the correlation between the $b = 0$ image and the DWIs is increased. In preliminary experiments it was observed that this mask is necessary for good alignment.

3 Experiments and Results

3.1 Synthetic Data

A noiseless, initially perfectly aligned, synthetic DWI set [9], with size $107 \times 79 \times 60$, $1.8 \times 1.8 \times 2.4 \text{ mm}^3$ voxel size and $G = 61$, was transformed with five random parameterized affine transformations. The matrix elements of Γ were drawn from a normal distribution $0.01 \mathcal{N}(0, 1)$ and t_x, t_y and t_z were drawn from $\mathcal{N}(0, 1)$ (mm). The five transformed image sets are registered with the three different metrics \mathcal{D}_{PCA} , \mathcal{D}_{APE} and \mathcal{D}_{VAR} , where for metric \mathcal{D}_{PCA} different values for L are investigated: $L \in \{1 \dots 10\}$. The synthetic data is simulated using the diffusion tensor model. For $L = G$, $\mathcal{D}_{\text{PCA}} = 0$ and the expected optimal value for L is 7, since this is the number of free parameters in the diffusion tensor model. The range chosen for investigating the impact of L is therefore expected to be sufficiently broad.

In the next experiment Gaussian noise¹ was added to the synthetic DWI set, which gives a resulting SNR of 8.65. Our metric, using $L = 6$, was used to register the noised DWI set with the same initial five transforms.

Evaluation measure Let $\mathbf{T}_g(\mathbf{x}; \hat{\boldsymbol{\mu}}_g)$ be the transformation that was computed by the registration. The deformation field of the initial aligned data, after application of the composition of $\mathbf{T}_g(\mathbf{x}; \hat{\boldsymbol{\mu}}_g)$ and $\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g})$ should be zero: $\mathbf{d}_g(\mathbf{x}) = \mathbf{T}_g(\mathbf{T}_g(\mathbf{x}; \boldsymbol{\mu}_{\text{init},g}); \hat{\boldsymbol{\mu}}_g) - \mathbf{x} = \mathbf{0}$. However, the constraint, Eq. (14), was not applied to the initial transformation, so we subtract the mean of the deformation field: $\mathbf{d}_g^*(\mathbf{x}) = \mathbf{d}_g(\mathbf{x}) - \frac{1}{G} \sum_{g=1}^G \mathbf{d}_g(\mathbf{x})$. Our measure for the registration error was then defined as the mean and standard deviation of $\|\mathbf{d}_g^*(\mathbf{x})\|$ for all \mathbf{x} and g .

Results Table 1 shows the mean and standard deviation of $\|\mathbf{d}^*(\mathbf{x})\|$ for the experiments with the synthetic DWI set, for each different metric and before registration. For $L = 6$ our method performs best. But for $L > 3$ our method is

¹Although noise in MR images is Rician distributed, for $\text{SNR} > 5$ it can be approximated with a Gaussian distribution [19].

Table 1 Mean and standard deviation of $\|\mathbf{d}^*(\mathbf{x})\|$ for the noiseless synthetic DWI set without registration, for registration with metric \mathcal{D}_{APE} , with metric \mathcal{D}_{VAR} and with metric \mathcal{D}_{PCA} for $L = 1 \dots 10$

T_{init}	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)
#	No reg.	\mathcal{D}_{APE}	\mathcal{D}_{VAR}	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$L = 6$	$L = 7$	$L = 8$	$L = 9$	$L = 10$
1	7.39 ± 4.06	0.70 ± 0.57	4.16 ± 6.42	1.25 ± 0.93	30.91 ± 13.05	2.04 ± 1.37	0.48 ± 0.30	0.31 ± 0.22	0.09 ± 0.20	0.16 ± 0.65	0.15 ± 0.61	0.17 ± 0.66	0.15 ± 0.58
2	7.29 ± 3.90	0.70 ± 0.56	4.38 ± 5.75	1.34 ± 1.05	33.63 ± 10.18	2.08 ± 1.32	0.48 ± 0.30	0.31 ± 0.22	0.09 ± 0.21	0.13 ± 0.48	0.13 ± 0.50	0.13 ± 0.50	0.13 ± 0.49
3	7.52 ± 4.00	0.68 ± 0.54	4.38 ± 5.23	1.21 ± 0.90	26.44 ± 14.45	10.00 ± 6.41	0.48 ± 0.30	0.30 ± 0.22	0.09 ± 0.21	0.14 ± 0.54	0.15 ± 0.56	0.14 ± 0.54	0.15 ± 0.60
4	7.55 ± 4.13	0.68 ± 0.55	3.54 ± 5.74	1.32 ± 1.02	21.06 ± 11.80	1.39 ± 0.87	0.46 ± 0.29	0.30 ± 0.22	0.09 ± 0.21	0.15 ± 0.56	0.15 ± 0.59	0.18 ± 0.70	0.21 ± 0.83
5	7.29 ± 3.86	0.69 ± 0.55	3.96 ± 6.06	1.25 ± 0.92	24.98 ± 13.68	3.21 ± 2.31	0.47 ± 0.29	0.30 ± 0.22	0.09 ± 0.21	0.23 ± 0.92	0.25 ± 0.98	0.24 ± 0.96	0.24 ± 0.90

Table 2 Results for synthetic DWI set with noise

\mathbf{T}_{init}	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)	$m \pm \sigma$ (mm)
#	\mathcal{D}_{PCA}	\mathcal{D}_{APE}	\mathcal{D}_{VAR}
1	0.11 ± 0.20	1.30 ± 0.85	4.51 ± 6.22
2	0.11 ± 0.20	1.37 ± 1.03	3.70 ± 6.25
3	0.11 ± 0.20	1.42 ± 1.00	3.60 ± 6.51
4	0.11 ± 0.20	1.48 ± 0.98	3.37 ± 5.82
5	0.11 ± 0.20	1.47 ± 1.04	3.98 ± 6.02

already better than \mathcal{D}_{APE} . For $L = 2$, the error is particularly high. Visual inspection revealed that in this case the registration resulted in two completely misaligned groups of images M_g , although the images within each group were properly aligned with each other. This apparently leads to a correlation matrix \mathbf{K} with two relatively high eigenvalues. Table 2 shows the results of the experiments with noisy synthetic DWI set, for all metrics. For the noisy dataset our method is the best of the three.

3.2 Real Diffusion Weighted Data

Five diffusion MRI data sets were obtained from different previous studies to evaluate our new approach. Details of these studies:

1. $10 \text{ b} = 0 \text{ s/mm}^2$; $60 \text{ b} = 700 \text{ s/mm}^2$; $2.0 \times 2.0 \times 2.0 \text{ mm}^3$ voxel size; 1.5T; [10]
2. $1 \text{ b} = 0 \text{ s/mm}^2$; $60 \text{ b} = 1,200 \text{ s/mm}^2$; $1.75 \times 1.75 \times 2.0 \text{ mm}^3$ voxel size; 3.0T; [3]
3. $1 \text{ b} = 0 \text{ s/mm}^2$; $32 \text{ b} = 800 \text{ s/mm}^2$; $1.75 \times 1.75 \times 2.0 \text{ mm}^3$ voxel size; 3.0T; [24]
4. $1 \text{ b} = 0 \text{ s/mm}^2$; $32 \text{ b} = 800 \text{ s/mm}^2$; $1.75 \times 1.75 \times 2.0 \text{ mm}^3$ voxel size; 1.5T; [22]
5. $1 \text{ b} = 0 \text{ s/mm}^2$; $45 \text{ b} = 1,200 \text{ s/mm}^2$; $1.72 \times 1.72 \times 2.0 \text{ mm}^3$ voxel size; 3.0T; [16]

For all data sets we choose $L = 6$.

To evaluate the registration, ExploreDTI [12] is used. B-matrix rotation is applied [3]. Directional Encoded Colormaps (DEC) and the standard deviation (STD) across the DWIs show a bright rim at the edge of the brain when the images are not aligned [20]. DEC maps and STDs of the DWIs of the five datasets before and after alignment are shown in Fig. 1. The bright rim is visible in the data pre alignment and is not visible in the data post alignment. The method is therefore successful in aligning the data. Metrics \mathcal{D}_{VAR} and \mathcal{D}_{APE} were also tested. In agreement with the results on synthetic data, the proposed metric \mathcal{D}_{PCA} outperformed \mathcal{D}_{VAR} and gave a slight improvement over \mathcal{D}_{APE} , based on inspection of the DEC and STD maps.

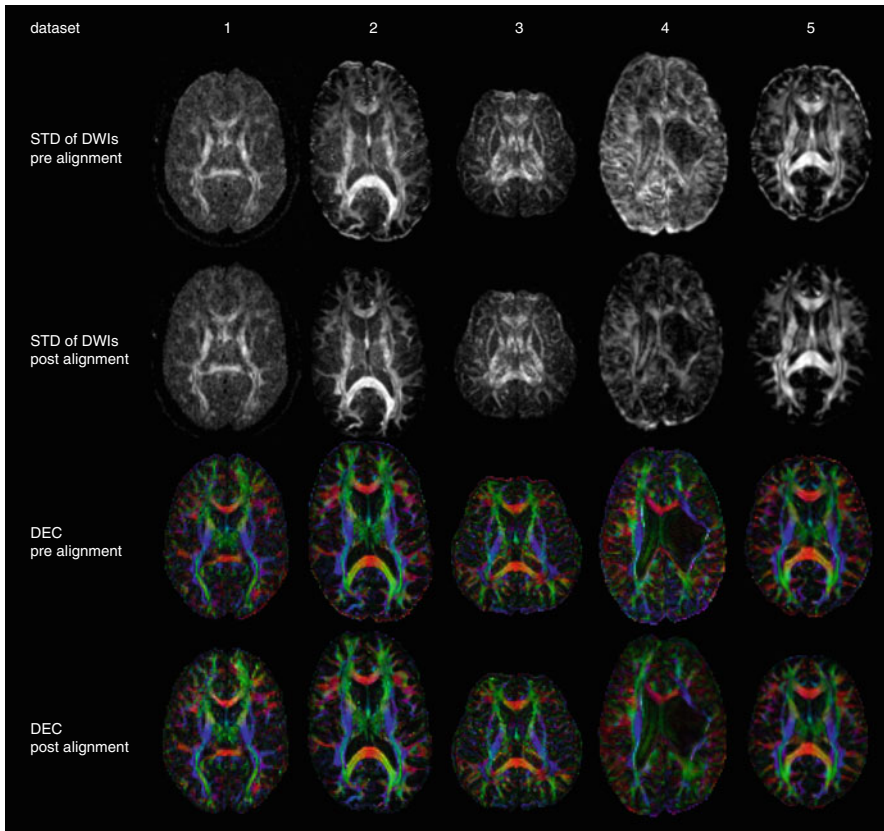


Fig. 1 STD (equal intensity range for pre and post registration) and DEC maps of datasets 1–5, pre and post alignment

4 Conclusions

With regard to the synthetic DWI set our method outperforms the two existing groupwise methods with which it was compared.

The use of $L = 6$ eigenvalues performed best. This is due to the underlying structure of the synthetic DWI data: the degrees of freedom are determined by the directions of the diffusion in the brain. The diffusion tensor model has 7 free model parameters, so the optimal value for L was expected to be 7. This optimal value is 6 due to the use of the mask: The eigenvalue spectrum of the aligned synthetic DWI shows that 99% of the data variance is described by 6 eigenvectors when the high values are masked, but without masking the high values, this number is 7. Masking out these high values decreases the data variance that is described by the $b = 0$ image, leading to a decrease in the number of eigenvectors describing most of the data variance.

The number of eigenvalues is a parameter that must be set correctly to obtain the best results, however the synthetic data experiments suggest that large values of L are preferable to small values of L . The optimal value for L for real data may not be 6. Further research qualifying the performance of the registration for real data is necessary to see if the optimal value should be >6 .

It would be an improvement if the method could also work without the necessity of the mask. Furthermore it would be interesting to investigate if the proposed method also works on other applications that involve intensity changes over time, such as perfusion-imaging or T1 mapping data.

Altogether the proposed method offers potential improvements over the current standard to align diffusion weighted data, due to the general benefits of groupwise registration, the fact that the method is not parameterized with a diffusion model, and the good results obtained on the synthetic and real data.

References

1. Balci, S. et al.: Free-form B-Spline deformation model for groupwise registration. Proc. Stat. Regis. Workshop (MICCAI). 23–30 (2007)
2. Behrens, T.E.J., et al.: Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *NeuroImage* **34**, 144–155 (2007)
3. de Geeter, N., et al.: A DTI-based model for TMS using the independent impedance method with frequency-dependent tissue parameters. *Phys. Med. Biol.* **57**, 2169–2188 (2012)
4. Fung, T.C.: Computation of the matrix exponential and its derivatives by scaling and squaring. *Int. J. Numer. Methods Eng.* **59**, 1273–1286 (2004)
5. Jensen, H.J., et al.: Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1340 (2005)
6. Jones, D.K., Leemans, A.: Diffusion tensor imaging. *Methods Mol. Biol.* **711**, 127–144 (2011)
7. Klein, S., et al.: Adaptive stochastic gradient descent optimization for image registration. *Int. J. Comput. Vis.* **81**, 227–239 (2009)
8. Klein, S., et al.: *elastix*: a toolbox for intensity based medical image registration. *IEEE Trans. Med. imaging* **29**, 196–205 (2010)
9. Leemans, A., et al.: Mathematical framework for simulating diffusion tensor MR neural fiber bundles. *Magn. Reson. Med.* **53**, 944–953 (2005)
10. Leemans, A., et al.: Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data. *Magn. Reson. Med.* **55**, 1414–1423 (2006)
11. Leemans, A., et al.: The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* **61**, 1336–1349 (2009)
12. Leemans, A., et al.: ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. In: 17th Annual Meeting of International Society for Magnetic Resonance in Medicine, Hawaii, 2009, p. 3537
13. Marsland, S., et al.: A minimum description length objective function for groupwise non-rigid image registration. *Image Vis. Comput.* **26**, 333–346 (2008)
14. Melbourne, A., et al.: Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). *Phys. Med. Biol.* **52**, 5147–5156 (2007)
15. Metz, C.T., et al.: Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. *Med. Image Anal.* **15**, 238–249 (2010)
16. Reijmer, Y.D., et al.: Improved sensitivity to cerebral white matter abnormalities in Alzheimer's disease with spherical deconvolution based tractography. *PLoS one.* **7**, 1371 (2012)

17. Rohde, G.K., et al.: Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magn. Reson. Med.* **51**, 103–114 (2004)
18. Schultz, T., Seidel, H.: Using eigenvalue derivatives for edge detection in DT-MRI data. In: Rigoll, G. (ed.) *Pattern Recognition*, vol. 5096, pp. 193–202. Springer, Berlin (2008)
19. Sijbers, J., et al.: Parameter estimation from magnitude MR images. *Int. J. Imaging Syst. Technol.* **10**, 109–114 (1999)
20. Tournier, J.D., et al.: Diffusion tensor imaging and beyond. *Magn. Reson. Med.* **65**, 1532–1556 (2011)
21. van der Aa, N.P., et al.: Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem. *Electron. J. Linear Algebra* **16**, 300–314 (2007)
22. van der Aa, N.E., et al.: Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke? *Stroke* **42**, 3410–3414 (2011)
23. Wachinger, C., et al.: Simultaneous registration of multiple images: similarity metrics and efficient optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **7**, 667–674 (2012)
24. Wang, H.C., et al.: Diffusion tensor imaging of vascular parkinsonism: structural changes in cerebral white matter and the association with clinical severity. *Arch. Neurol.* **69**, 1340–1348 (2012)
25. Wedeen, V.J., et al.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.* **54**, 1377–1386 (2005)
26. Woods, P.: Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. *NeuroImage*. **18**, 769–788 (2003)