Similarity Measures for Matching Diffusion Tensor Images

Daniel Alexander, James Gee, Ruzena Bajcsy

GRASP Lab. and Dept. Radiology

U. of Penn. 3401 Walnut St., Philadelphia, PA 19104, USA

`daniel2@grip.cis.upenn.edu`

Abstract

In this paper, we discuss matching of diffusion tensor (DT) MRIs of the human brain. Issues concerned with matching and transforming these complex images are discussed. A number of similarity measures are proposed, based on indices derived from the DT, the DT itself and the DT deviatoric. Each measure is used to drive an elastic matching algorithm applied to the task of registration of 3D images of the human brain. The performance of the various similarity measures is compared empirically by use of several quality of match measures computed over a pair of matched images. Results indicate that the best matches are obtained from a Euclidean difference measure using the full DT.

1 Introduction

Diffusion tensor (DT) imaging is a recent innovation in magnetic resonance imaging (MRI), [1]. In DT imaging, the measurement acquired at each voxel in an image volume is a symmetric second order tensor, which describes the local water diffusion properties of the material being imaged. The DT may be thought of as a trivariate Gaussian distribution on the position of a water molecule, initially at the centre of a voxel, after some fixed time. It is represented by a symmetric 3x3 matrix and thus has 6 independent components.

DT imaging of the brain has provoked particular interest because of the added insight it provides into the structure of white matter regions. Neuronal axons form the connections between different cells in the brain. These axons are fibrous structures in which water is free to diffuse along the fibre, but diffusion in perpendicular directions is restricted by the cell wall. DT measurements taken from white matter regions, where large numbers of axons are bundled together, thus tend to exhibit high anisotropy and the principal direction (PD) of the DT points along the axes of the bundled fibres. By associating neighbouring voxels according to the PD of their DTs, pathways within the brain can be traced and mapped, see for example [2], although the resolution at which these images can currently be obtained is such that only major pathways can be extracted reliably. There are an increasing number of clinical applications of DT imaging, for example, the analysis of stroke and multiple sclerosis, [3].

Visualisation of tensor fields, such as a 3D DT image, is an on going subject for research. In practice derived indices are used to express the information contained

within the DT. The trace, the anisotropy and the PD are the most important of these. The trace image is proportional to the trace of the DT matrix, which is a measure of the total amount of diffusion at a point. Diffusion anisotropy is a measure of directionality of the diffusion – it is high if the amount of diffusion is weighted in one particular direction. A review of measures of anisotropy can be found in [4], where one particular measure – the *lattice anisotropy* (see section 3) was found to be the most robust. Perhaps the most interesting index obtained from DT images is the PD, which is the major eigenvector of the DT at each point. The information provided by the PD is entirely complementary to standard MRI data and gives structure to regions that appear homogeneous in standard imagery. Figure 1 shows PD (left), anisotropy (top right) and trace (bottom right) images from one slice of a DT image of a human brain. The PD is shown only in anisotropic regions where it is meaningful; lighter lines indicate that the PD is more out of the plane of the image. Strong directional structure can be observed in regions that appear homogeneous in the anisotropy image.
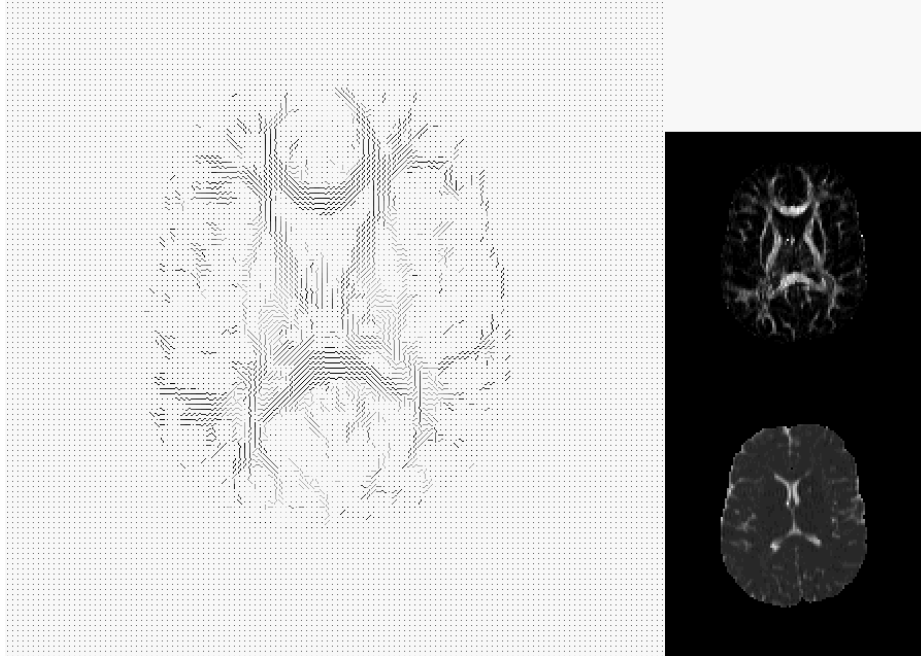

Figure 1 Images derived from a slice of a DT image. Left: principal DT eigenvector projected into the xy-plane - lighter lines indicate greater z-component. No line is drawn at points where the anisotropy falls below a certain threshold. Top right: lattice anisotropy image. Bottom right: DT trace image.

Here we are concerned with registration of DT images of the human brain. The multi-resolution elastic matching algorithm, described in detail elsewhere, [5-7], is adapted for use with this new data type. There are two issues of concern when matching DT images. Firstly, a pointwise similarity measure is required to drive the match. Secondly, the effect of image transformations on the DT values at each voxel must be considered. Each DT has intrinsic orientation with respect to the surrounding anatomical structure of the image. We must ensure that the orientation remains consistent with patient anatomy when transformations of the image are made. In this

paper, we are mainly concerned with measures of similarity. Reorientation of the DTs will be discussed briefly, but is addressed more fully elsewhere, [8].

Registration of 3D brain images can greatly simplify clinical studies into the variation of properties of the brain over large populations. Motivation for this work lies initially in assisting studies into the variation of anisotropy and other diffusion related properties, see for example [9]. However, we further hope that the added structural information in DT images can improve the general standard of matches that can be computed between the images of brains from different patients. A long term goal is to use high quality matches computed using DT data combined with more standard structural imagery, such as T1 or T2-weighted images, to generate statistical models of the variation in brain layout over population groups. Such a model could be used to constrain matches computed when only less descriptive data is available.

In the next section a brief description of the matching algorithm is given with some discussion of its application to this new data type. In section 3 a number of possible similarity measures for use with DT data are described. In order to decide which of these measures is more effective, we need to be able to assess the quality of the matches that they provide. Some discussion of this evaluation problem is given in section 4. Section 5 describes experiments and results, and conclusions are drawn in section 6.

2 Elastic Matching of DT Images

The elastic matching algorithm used here is that developed by Bajcsy, Gee, et al, [5-7]. Given a pair of 3D images, the algorithm computes a high-dimensional warp, described by a displacement vector field, from one fixed, reference image to the other, target image. An iterative finite element method is used at consecutive levels of a multi-resolution pyramid to find a displacement field that minimises an energy function of the form:

$$E = \lambda.\text{deformation} - \alpha.\text{similarity}. \qquad (1)$$

α and λ are weighting parameters. The similarity is a measure of the voxel to voxel correspondence of the image pair. Its form depends on the type of data in the images. For the deformation term we use the elastic body deformation described in [12], as in [6,7]. An initial global affine registration of the two images is performed to ensure a good starting point for the elastic matching. In the work reported here, we used Woods AIR (Automated Image Registration) algorithm, [10,11], applied to single valued DT trace images (see section 3), to compute an appropriate affine transformation.

The form of the similarity term in (1) appropriate for DT images is the subject of this paper and some possibilities are proposed in the next section. As mentioned in the introduction, DTs have intrinsic orientation and care must be taken when transformations are applied to DT images to ensure that these orientations remain consistent with surrounding anatomy. This requirement is illustrated in Figure 2. If the DTs in the transformed image are simply copied from the corresponding voxels in the untransformed image, without reorientation, their PDs would no longer point along the axis of the fibre, which would be anatomically incorrect. Here we estimate the appropriate reorientation at each point from the local displacement gradient. Several approaches to this estimation are discussed in a companion paper, [8], in which a method that applies the local displacement directly to the principal axes of the DT was found to be more effective to techniques derived from classical continuum mechanics,

[12]. We adopt that method in all the experiments reported here. The orientation of the DTs is updated throughout the matching process according to the current estimate of the displacement field. We note, however, that there is no term in the energy function that explicitly couples the displacement with the reorientation of the DT so we do not explicitly optimise over the DT orientation.
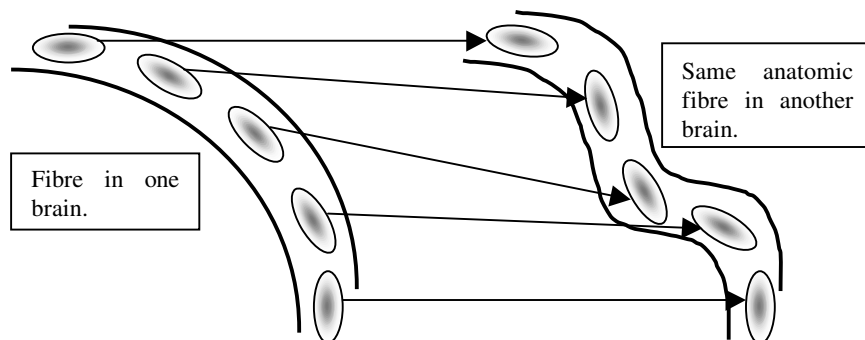

Figure 2 Depicts the same anatomic fibre in two different brains and the orientations of the DTs within the fibres. When the shape of one fibre is warped to that of the other, the DT orientations must change accordingly.

3 Similarity Measures for DTs

In this section, a number of similarity measures that might be used to match pairs of DT images are proposed.

3.1 DT Similarity Measures

A simple comparison between two tensor quantities, $\mathbf{D_1}$ and $\mathbf{D_2}$, is provided by the tensor scalar product, $\mathbf{D_1.D_2}$, given by, [12],

$$\mathbf{D_1.D_2} = \sum_{j=1}^{3}\sum_{i=1}^{3} D_{ij}^{1}.D_{ij}^{2}\,, \tag{2}$$

where D_{ij}^{n} is the ij-th entry of the matrix representing $\mathbf{D_n}$.

Another measure, $\mathbf{D_1:D_2}$, which is the sum of the squared scalar products between each pair of semi-major axes, weighted by the corresponding eigenvalues, was proposed in [4] as a measure of diffusive similarity between two media:

$$\mathbf{D_1 : D_2} = \sum_{j=1}^{3}\sum_{i=1}^{3} \lambda_i^1 \lambda_j^2 \left(\underline{e}_i^1.\underline{e}_j^2\right)^2\,, \tag{3}$$

where λ_k^n is the k-th eigenvalue of $\mathbf{D_n}$ and $\underline{e}_k^n$ is the corresponding unit eigenvector.

In preliminary work on this subject, [13], we used a Euclidean distance metric, $d(\mathbf{D_1},\mathbf{D_2})$, which can be negated to provide a DT similarity function:

$$d(\mathbf{D_1},\mathbf{D_2}) = \sqrt{\text{trace}\left[(\mathbf{D_1} - \mathbf{D_2})^2\right]}\,. \tag{4}$$

The expression in the square root is equivalent to the sum of square differences of the elements of the two DT matrices.

The total diffusivity at each point is closely linked to the concentration of water; for this reason, similar features are exposed by the trace image as standard structural MRIs. The benefit of DT imaging is the added structural information it provides in regions, such as areas of white matter, where the anisotropy is high. Furthermore, there is little evidence to suggest that the magnitude of the diffusivity is consistent in corresponding regions in different brains. The *deviatoric, D,* of a DT, **D**, given by,

$$D = \mathbf{D} - \frac{1}{3}\operatorname{trace}(\mathbf{D}).\mathbf{I} , \tag{5}$$

is the anisotropic part of the DT, which expresses just the shape and orientation of the DT, independent of its scale, [4, 12]. Each similarity measure above can equally be applied to deviatorics. Furthermore, each can be expressed simply in terms of the similarity of the original tensors, thus:

$$D_1.D_2 = \mathbf{D_1}.\mathbf{D_2} - \frac{1}{3}\operatorname{trace}(\mathbf{D_1})\operatorname{trace}(\mathbf{D_2}) , \tag{6}$$

$$D_1 : D_2 = \mathbf{D_1} : \mathbf{D_2} - \frac{1}{3}\operatorname{trace}(\mathbf{D_1})\operatorname{trace}(\mathbf{D_2}) , \text{ and} \tag{7}$$

$$[\operatorname{d}(D_1, D_2)]^2 = [\operatorname{d}(\mathbf{D_1}, \mathbf{D_2})]^2 - \frac{1}{3}[\operatorname{trace}(\mathbf{D_1}) - \operatorname{trace}(\mathbf{D_2})]^2 . \tag{8}$$

Since the PD is one of the primary indices of interest from DT images, a reasonable strategy might be to match purely on the principal diffusion direction. A natural way to represent the similarity of DTs in terms of their PD is by the cosine of the angle between them. In regions where the diffusion is isotropic, the PD is poorly defined. In order to avoid problems in these regions, the similarity is weighted by the geometric mean of the anisotropy of each DT to obtain the measure:

$$\sqrt{v(\mathbf{D_1}).v(\mathbf{D_2})}.\bigl(\operatorname{pd}(\mathbf{D_1}).\operatorname{pd}(\mathbf{D_2})\bigr), \tag{9}$$

where $v(\mathbf{D})$ represents the lattice anisotropy of **D**, defined in the next subsection, and $\operatorname{pd}(\mathbf{D})$ is the unit principal eigenvector.

In summary, we have defined three DT similarity measures, three deviatoric similarity measures and one PD similarity:

- **$\mathbf{D_1}.\mathbf{D_2}$**, of equation (2), and $D_1.D_2$, of equation (6),
- **$\mathbf{D_1}:\mathbf{D_2}$**, of equation (3), and $D_1:D_2$, of equation (7),
- -d(**$\mathbf{D_1}$, $\mathbf{D_2}$**), of equation (4) and -d(D_1, D_2), of equation (8),
- and the anisotropy weighted PD similarity defined in equation (9).

In order to improve the dynamic range of the similarity measures, the magnitudes of **$\mathbf{D_1}.\mathbf{D_2}$**, **$\mathbf{D_1}:\mathbf{D_2}$** and their deviatoric equivalents are square rooted in the energy function.

3.2 Intensity Correlation

Good results on single valued intensity images have been obtained in the past when the local intensity correlation has been the similarity measure driving the elastic matching algorithm, [5]. DT images can be matched by using this similarity measure on 1D images derived from the DT image, such as the trace and anisotropy images. The

correlation is a region based measure computed over a small neighbourhood of the current voxel and so may be more informative than the single voxel measures described in the previous subsection. For the images used in the experiments reported here, structural images were acquired at the same time as the DT images and so a match of the DT images can also be obtained by matching these structural images. Here we compute matches using the correlation of three separate single valued images:

- T1-weighted image,
- DT trace image,
- Lattice anisotropy image.

As mentioned in the introduction, we use the lattice anisotropy measure, which was shown in [4] to be less susceptible to bias, introduced by noise in the data, than other measures. The lattice anisotropy, v, is based on the *lattice index*, for a reference DT, $\mathbf{D_{ref}}$, and another DT, $\mathbf{D_N}$, which is defined in terms of the similarity in (2) and (6):

$$LI_N = \frac{\sqrt{3}}{\sqrt{8}} \frac{\sqrt{D_{ref} : D_N}}{\sqrt{\mathbf{D_{ref}} : \mathbf{D_N}}} + \frac{3}{4} \frac{D_{ref} : D_N}{\sqrt{\mathbf{D_{ref}} : \mathbf{D_{ref}}} \sqrt{\mathbf{D_N} : \mathbf{D_N}}} . \tag{10}$$

To obtain v at a point in a DT image, $\mathbf{D_{ref}}$ is set to the DT at that point and the average of LI_N weighted by the distance between the voxel centres, is computed, with N varying over the 8 in-plane neighbours.

4 Performance and Quality of Match

In order to compare the performance of the similarity measures suggested in the previous section, a measure of quality of the match obtained between two images is required. For the tasks of interest to us, i.e., inter-patient matching of 3D brain volumes, this is difficult, as ground truth is hard to obtain. The goal is to achieve an optimal *anatomical* match, where voxels representing the same anatomical features are associated. Perfect anatomical matches cannot be obtained, because features may be individual to particular brains and not present in others. However, we would like to ensure that anatomical features that do appear in both images are associated and that their internal structure is preserved as well as possible.

Any of the similarity measures described in the previous section can be used to provide a measure of match quality by computing the average value over the two images after matching. Such measures can be informative, but they are biased in general, and may not be a direct indication of the anatomical match quality. The best value for such a quality of match measure is most likely to be obtained by using the same measure to drive the matching process and there is no guarantee that maximal similarity corresponds to best anatomical association of the two images.

Another way to assess the performance is to compare the match with an association of anatomical features provided by hand. Such data could take the form of either a number of associated point landmarks or corresponding regions outlined in both images. Although measures generated in this way are more objective than those obtained by summing voxelwise similarity measures, there are still problems. Firstly, it can be difficult to place landmarks accurately by eye and localisation is particularly difficult in the out of plane axis. Also, landmarks tend to be defined in places in the image where the anatomical association is obvious from surrounding edges in single valued intensity images. These are the regions that tend to matched easily by the

algorithm and so closeness of landmarks may not imply a good match of structure within regions. Similar problems exist with measures generated from the overlap of hand outlined anatomical regions, although the edges of the region may be well matched, we also require a good voxel to voxel match within the region. This issue is particularly important for DT image matching as the reorientation of the DTs can only be computed correctly if the voxel to voxel association is good. If the voxel association is poor, much of the useful structure of the DT image may be lost.

Because of the problems associated with these methods for evaluating match quality, we use a consensus approach to compare the suggested similarity measures. Optimal algorithm parameter settings, in particular the balance of α and λ in (1), are selected by minimising a hand-defined landmark separation measure. Match quality is then assessed by averaging similarity measures over the matched pair of images.

5 Experiments and Results

In this section, we describe a set of experiments designed to compare the effectiveness of the similarity measures proposed in section 3. First we give details of our data sets. The experiments are then described and results presented.

5.1 Data

A single pair of brain volumes is used in the following experiments. Both are images of human brains taken from young female subjects. The image acquisition methodology is similar to that reported in [1]. Images were acquired using a 1.5T GE Signa Horizon EchoSpeed spectrometer. Each DT image consists of 33 contiguous axial slices, with slice thickness 3.5mm, 220mm field of view and 128x128 in-plane resolution. Six gradient directions were sampled and 4 images were acquired for each direction. 4 images with no diffusion weighting were also acquired and so a total of 28 T2-weighted acquisitions were made per slice of the DT image volume. Structural, T1-, T2- and PD-weighted, images matching the slices of the DT images were obtained at the same time, but with in-plane resolution of 256x256.

A slight misregistration between the DT images and the corresponding structural images can be observed and so a pre-alignment is performed on each DT image using Woods' AIR program. The brain volumes are then extracted from the images by hand.

By simultaneously viewing the trace, anisotropy and T1-weighted images of each volume, 237 associated landmarks were placed in the two volumes. Care was taken to ensure that the spatial distribution of these landmarks over the image volume was as uniform as possible. A globally registered version of one image was used in the display to assist the localisation of points in the out of plane axis.

5.2 Experiments

Each data set was interpolated, using trilinear interpolation separately for each of the six DT component images, to a uniform $1mm^3$ volume. After an initial global registration of one volume to the other, multi-resolution pyramids are computed from each data set, by sub-sampling at a factor of two down to a lowest resolution of one-sixteenth the size of the isotropic image in each dimension. Elastic matching is then

performed at four levels, from the lowest resolution up to half the size of the original image. λ in (1) is reduced by a factor of 2 at each level from 1.0 at the lowest resolution, down to 0.125 at the highest.

An optimal setting for α is found separately for each similarity measure by minimising the summed Euclidean distance between corresponding landmarks. The following similarity measures, averaged over the overlap of the within-brain regions of both images, are then computed to indicate the quality of each match:

- Squared DT trace difference (lower value implies better performance).
- Squared anisotropy difference (lower implies better).
- Euclidean DT difference given by (1) (lower implies better).
- Anisotropy weighted PD similarity given by (5) (higher implies better).

The first two measures provide an indication of how well the prominent features in the trace and anisotropy images are aligned. The latter two give further insight into the quality of the match that is obtained within homogenous regions. The Euclidean DT difference measure is chosen in preference to the other DT similarity measures, because the landmark separation measure suggests that it produces better results in general, as can be seen from the results below.

The numerical results are summarised in Table 1. A row is included with numerical results from the match obtained by the global transformation alone for comparison with those obtained with the additional elastic matching phase.

	Summed landmark separation	Av. Sq. DT trace diff.	Av. Sq. anisotropy diff.	Av. DT diff.	Av. Weighted PD sim.
Global only.	1139	69.2	73.4	650.3	0.743
T1-Correlation	**971**	**68.5**	**71.9**	**644.8**	**0.748**
Trace-Corr.	**980**	*65.6*	**69.3**	**627.2**	**0.761**
Anis.-Corr.	**980**	**69.1**	*66.4*	**639.1**	**0.765**
$\mathbf{D_1}.\mathbf{D_2}$ – eq. (1)	**1006**	70.1	74.4	659.0	0.739
$\mathbf{D_1}:\mathbf{D_2}$ – eq. (2)	**1007**	69.9	74.3	658.6	0.738
-d($\mathbf{D_1},\mathbf{D_2}$) – eq. (3)	**978**	**65.9**	**66.9**	*622.2*	*0.772*
$D_1.D_2$ – eq. (5)	**1005**	69.6	**72.3**	650.5	**0.752**
$D_1.D_2$ – eq. (6)	**1003**	69.7	**73.0**	653.5	**0.745**
-d(D_1,D_2) – eq. (7)	**982**	69.5	**67.3**	**633.6**	**0.778**
PD Sim. – eq. (8)	**1006**	70.1	74.3	659.0	0.740

Table 1 Numerical comparison of matches obtained using different similarity measures. Bold lettering indicates improvements over the global registration.

6 Conclusions

Good results are obtained from the correlation based matches and from the Euclidean distance measure applied to both the full DT and the deviatoric. The other tensor similarity measures, $\mathbf{D_1}.\mathbf{D_2}$ and $\mathbf{D_1}:\mathbf{D_2}$, are clearly inferior for this task. Some improvement is evident when they are applied to the deviatoric, rather than the full DT, but performance in terms of both landmark association and voxel similarity is still worse than for the Euclidean measure. The match obtained from the PD alone is also

poor. Much of the information contained in the DT is discarded when the PD alone is used and in many regions its value may be poorly defined. The resultant lack of features over large areas of the image means that the match is poorly constrained and, consequently, results are poor even in regions where the PD is strongly defined.

The match obtained using T1 correlation produces the best performance in terms of the landmark separation measure, but scores relatively poorly in terms of the similarity measures. Matches based on correlation of the trace and anisotropy images produce good similarity of the index used to compute the correlation, but relatively poor similarity of the other. We note that although it is not unreasonable to expect that better landmark association might be obtained by using the higher resolution T1 image, this measure may be biased in favour of the T1 correlation match, since it is also easier to place landmarks in this more detailed image. Residual mis-registration of the DT and T1 images may increase the separation of landmarks in the image in which the landmark was not placed directly.

Unlike any of the correlation measures the Euclidean distance, which takes into account all the information contained in the DT, scores well in terms of all the similarity measures. Use of the full DT appears to produce the most consistently good similarity, but it is not clear whether the best results are obtained from the full DT or from the deviatoric. Further tests over an extended image ensemble are required to establish more conclusively which approach is best.

These results are preliminary in the sense that they are computed over just a single pair of images. Further testing is planned as data from large-scale clinical studies becomes available. However, we can draw some conclusions from the results presented above. In summary,

- Correlation of single valued indices produces good association of landmarks, which tend to appear at edges in the image, and good voxel correspondence of the index used in the correlation. However, the correspondence of complementary indices tends to be poor.

- More reliable results can be obtained by exploiting all the available information in the DT. In particular, matches computed using the Euclidean distance measure between vectors of DT components produces similar levels of landmark correspondence to the correlation measures, but good levels of similarity of *all* derived indices are obtained.

As well as extending testing to larger a data set, there are some other areas on which attention will be focussed in future work on this subject:

- It is not clear that the Euclidean measure is ideal and it may be possible to construct more suitable alternatives, for example, by considering the physics and statistics of the imaging process.

- Strong performance of the correlation measures above suggests that the use of neighbourhood based similarity measures can enhance the performance. It may be possible to devise multivariate correlation measures that can be used with full DTs.

Acknowledgement

References

[1] C. Pierpaoli, P. Jezzard, P.J. Basser, A. Barnett, G. Di Chiro, *Diffusion Tensor MR Imaging of the Human Brain*, Radiology, Vol. 201, No. 3, pp. 637-648, 1996.

[2] D.K. Jones, A. Simmons, S.C.R. Williams, M.A. Horsfield, *Non-invasive assessment of axonal fibre connectivity in the human brain via diffusion tensor MRI*, in press, Magnetic Resonance in Medicine, 1999

[3] M.A. Horsfield, H.B.W. Larsson, D.K. Jones, A. Gass, *Diffusion magnetic resonance imaging in multiple sclerosis*, Journal of Neurology, Neurosurgery and Psychiatry, Vol. 64 (Supplement), pp S80-S84, 1998.

[4] C. Pierpaoli and P.J. Basser, *Toward a Quantitative Assessment of Diffusion Anisotropy*, Magnetic Resonance Medicine, Vol. 36, pp. 893-906, 1996.

[5] R. Bajcsy and S. Kovacic, *Multi-Resolution Elastic Matching*, Computer Vision, Graphics and Image Processing, Vol., 46, pp. 1-21, 1989.

[6] J.C. Gee and D.R. Haynor, *Numerical Methods for High Dimensional Warps*. Chapter in "*Brain Warping*", ed. A.W. Toga, Academic Press, 1998.

[7] J.C. Gee and R.K. Bajcsy, *Elastic Matching: Continuum Mechanical and Probabilistic Analysis*. Chapter in "*Brain Warping*", ed. A.W. Toga, Academic Press, 1998.

[8] D. Alexander, J.C. Gee and R.K. Bajcsy, *Strategies for Data Reorientation during Non-Rigid Warps of Diffusion Tensor Images* Submitted to MICCAI'99.

[9] A. Virta, A. Barnett, C. Pierpaoli, et al. *Projection Pathways at the Level of the Cerebral Peduncle, Proc.* ISMRM, Sydney, Australia, pp. 1347, 1998.

[10] R.P. Woods, S.T. Grafton, C.J. Holmes, S.R. Cherry and J.C. Mazziotta, *Automated Image Registration: I General Methods and intra-subject intra-modality validation*, Journal of Computer Assisted Tomography, Vol. 22 pp. 141-154, 1998.

[11] R.P. Woods, S.T. Grafton, J.D.G. Watson, N.L. Sicotte and J.C. Mazziotta, *Automated Image Registration: II Inter-subject validation of linear and non-linear models*, Journal of Computer Assisted Tomography, Vol. 22 pp. 155-165, 1998.

[12] L.E. Malvern, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Inc. Englewood Cliffs, N.J., 1969.