Multimodal MRI Template Creation in the Ring-Tailed Lemur and Rhesus Macaque

Frederik J. Lange[1(✉)], Stephen M. Smith[1], Mads F. Bertelsen[2],
Alexandre A. Khrapitchev[3], Paul R. Manger[4], Rogier B. Mars[1,5],
and Jesper L. R. Andersson[1]

[1] Oxford Centre for Functional MRI of the Brain (FMRIB),
Wellcome Centre for Integrative Neuroimaging, Nuffield Department
of Clinical Neurosciences, University of Oxford, Oxford, UK
`frederik.lange@dtc.ox.ac.uk`
[2] Centre for Zoo and Wild Animal Health, Copenhagen Zoo, Frederiksberg, Denmark
[3] CRUK/MRC Oxford Institute for Radiation Oncology, Department of Oncology,
University of Oxford, Oxford, UK
[4] School of Anatomical Sciences, Faculty of Health Sciences,
University of the Witwatersrand, Johannesburg, South Africa
[5] Donders Institute for Brain, Cognition and Behaviour,
Radboud University Nijmegen, Nijmegen, The Netherlands

Abstract. We present a multimodal registration algorithm for simultaneous alignment of datasets with both scalar and tensor MRI images. We employ a volumetric, cubic B-spline parametrised transformation model. Regularisation is based on the logarithm of the singular values of the local Jacobian and ensures diffeomorphic warps. Tensor registration takes reorientation into account during optimisation, through a finite-strain approximation of rotation due to the warp. The combination of scalar, tensor and regularisation cost functions allows us to optimise the deformations in terms of tissue matching, orientation matching and distortion minimisation simultaneously. We apply our method to creating multimodal T2 and DTI MRI brain templates of two small primates (the ring-tailed lemur and rhesus macaque) from high-quality, *ex vivo*, 0.5/0.6 mm isotropic data. The resulting templates are of very high quality across both modalities and species. Tissue contrast in the T2 channel is high indicating excellent tissue-boundary alignment. The DTI channel displays strong anisotropy in white matter, as well as consistent left/right orientation information even in relatively isotropic grey matter regions. Finally, we demonstrate where the multimodal templating approach overcomes anatomical inconsistencies introduced by unimodal only methods.

Keywords: Registration · Multimodal · Lemur · Macaque · Template · MRI

1   Introduction

The most common approach to spatial normalisation of multimodal MRI datasets is to register a single, scalar modality to a template and then transform all modalities through the resulting warp. However, this approach is only valid when the within-tissue information is comparable across modalities (*e.g.*, T1 and T2-weighted scans). When a modality contains additional within-tissue information compared to the registered modality (*e.g.*, orientation in DTI), then there is no reason to believe that the resulting warps will modulate this information in a consistent manner across subjects.

One method of overcoming this issue is to register each modality independently. However, from a generative modelling perspective there is only a *single* true warp which maps each subject to the template. This true warp cannot be estimated by averaging the unimodal warps as the result would not guarantee the preservation of desirable properties such as diffeomorphism. A preferable approach is to simultaneously optimise over all modalities, thereby finding the solution which jointly maximises the similarity across all modalities.

To our knowledge, there is currently only one method (`DR-TAMAS`) capable of performing such joint optimisations [1]. This is an extension of the SyN method [2] to include a similarity term sensitive to local rotations due to warping. One limitation of `DR-TAMAS`'s *plastic* transformation model is that desirable measures of deformation, such as the local Jacobian determinant, are prohibitively difficult to regularise explicitly [1]. This necessitates using a simpler but less biologically meaningful regularisation of the velocity field (*e.g.*, Gaussian smoothing).

Our method overcomes this limitation by employing a cubic B-spline, *elastic* transformation which allows for direct regularisation of the Jacobian field, whilst simultaneously explicitly optimising for local rotations.

A prerequisite for making multimodal registration useful is the existence of multimodal templates. This work aims to present how our method can be used to generate such templates, and demonstrate some benefits they provide over their unimodal counterparts.

2   Registration Method

2.1   Framework

Transformation Model. Our registration framework utilises a 3D cubic B-spline parametrised transformation $\mathbf{t}(\mathbf{x})$. The finite spatial support of B-splines results in predictable sparseness of the Hessian of the cost function, which we leverage during optimisation. Additionally, this transformation has analytical derivatives, which allows for the formulation of explicit relationships between the transformation parameters and the local Jacobian field. We overcome historical concerns regarding diffeomorphism of elastic transformations through explicit regularisation of the Jacobian field as described in Sect. 2.4.

Optimisation Strategy. We employ the Levenberg-Marquardt variant of Gauss-Newton optimisation due to its robustness and rapid convergence properties [3]. Additionally, it does not require manually choosing a learning rate or performing a line search as is necessary when employing gradient only methods. This necessitates manipulating the cost function into the form $C(\mathbf{w}) = f^2(\mathbf{w})$ and iteratively updating the warp parameters $\mathbf{w}$ according to $\Delta\mathbf{w} = -\mathbf{H}^{-1}\mathbf{b}$, where $\mathbf{b}$ is the gradient and $\mathbf{H}$ is the Hessian of C. The Gauss-Newton Hessian is then $\mathbf{H}_{GN} = 2\frac{\delta f}{\delta \mathbf{w}}^{\mathsf{T}}\frac{\delta f}{\delta \mathbf{w}}$, and the Levenberg-Marquardt variant is $\mathbf{H}_{LM} = \mathbf{H}_{GN} + \lambda_{LM}\mathbf{I}$. Calculating $\mathbf{H}_{GN}$ is, in general, computationally expensive. However, this is aided by our choice of B-spline parametrisation as described in Sect. 2.1. At higher resolutions $\mathbf{H}_{GN}$ may become too large to store in memory (despite its sparseness). At this point we transition to using a *majorising* approximation $\mathbf{H}_{MM} = \texttt{diag}(\texttt{abs}(\mathbf{H}_{LM})\mathbf{1})$ [5], where $\mathbf{1}$ is a column vector of ones. $\mathbf{H}_{MM}$ is then a diagonal matrix where the diagonal entries are the sum of the absolute values across each row of $\mathbf{H}_{LM}$. This has the property that $\mathbf{H}_{MM} \succeq \mathbf{H}_{LM}$, and is therefore compatible with the Gauss-Newton family of algorithms. This greatly reduces memory requirements whilst allowing for a consistent optimisation strategy across all warp resolutions.

Cost Function Weighting. The relative weighting of each modality and the regularisation penalty can be set globally (*i.e.*, the importance of each modality to the overall cost function, controlled by a modality specific λ) as well as locally (*i.e.*, the importance of each area within a modality to that modality's cost) controlled by a spatially varying multiplicative mask.

2.2   Scalar Cost Function

The cost function for scalar modalities is the mean-squares dissimilarity, shown in Eq. (1) for a reference image f and moving image g defined over $N_{\mathbf{x}}$ voxels, where $\mathbf{x} \in \mathbb{R}^3$. This requires bias corrected images with the same nominal contrast. Residual differences in tissue intensities are modelled using quartic-polynomial intensity matching. This is re-evaluated at each iteration of the optimisation. The weight applied to the scalar cost can be modulated by the voxel-wise variance after each iteration to effectively up- or down-weight the regularisation prior.

$$C_S = \lambda_S \frac{1}{N_{\mathbf{x}}} \sum_{\mathbf{x} \in \mathbb{R}^3} \Big(f(\mathbf{x}) - g(\mathbf{t}(\mathbf{x})) \Big)^2 \qquad (1)$$

2.3   Tensor Cost Function

The tensor cost function can be chosen as either the Euclidean, equation (2), or log-Euclidean, equation (3), distance between two tensor volumes $\mathbf{F}$ and $\mathbf{G}$. These metrics are sensitive to both scalar and vector characteristics of the tensor, relatively insensitive to variations in tensor fitting, and efficient to implement [4,6]. The transformation affects these cost functions in two ways: through displacement of the tensor elements as if they were individual scalar modalities

($\mathbf{t}_{dis}$), and through the local rotation of tensors ($\mathbf{t}_{rot}$). The method proposed by Yeo [7] allows us to calculate analytical forms of the gradient and Hessian as a function of the transformation parameters, facilitating efficient implementation within our optimisation framework. The results in this work were generated using C_E.

$$C_E = \lambda_E \frac{1}{N_{\mathbf{x}}} \sum_{\mathbf{x}\in\mathbb{R}^3} \mathrm{tr}\left(\left(\mathbf{F}\big(\mathbf{t}_{rot}(\mathbf{x})\big) - \mathbf{G}\big(\mathbf{t}_{dis}(\mathbf{x})\big) \right)^2 \right) \tag{2}$$

$$C_{LE} = \lambda_{LE} \frac{1}{N_{\mathbf{x}}} \sum_{\mathbf{x}\in\mathbb{R}^3} \mathrm{tr}\left(\left(\log\big(\mathbf{F}(\mathbf{t}_{rot}(\mathbf{x}))\big) - \log\big(\mathbf{G}(\mathbf{t}_{dis}(\mathbf{x}))\big) \right)^2 \right) \tag{3}$$

2.4   Regularisation

The regularisation cost in Eq. (4) is based on work by Ashburner [8], and penalises the log singular values s_i of the local Jacobian $\mathbf{J}$. This imposes a log-normal prior on lineal, areal and volumetric changes, centred at a value of 1. Expansions and contractions are therefore penalised symmetrically. Additionally, the penalty tends to ∞ as $|\mathbf{J}|$ approaches both 0 and ∞, ensuring the warps remain diffeomorphic. By penalising s_i rather than simply $|\mathbf{J}|$ we ensure that the shape as well as the volume of the transformed images is kept within reasonable limits. Finally, the highly non-linear nature of the penalty allows for larger deformations than traditional linear elastic models, bringing our method in line with the capabilities of LDDMM-based transformation methods.

$$C_R = \lambda_R \frac{1}{N_{\mathbf{x}}} \sum_{\mathbf{x}\in\mathbb{R}} \left(1 + |\mathbf{J}(\mathbf{t}(\mathbf{x}))|\right) \sum_{i=1}^{3} \log^2 s_i\big(\mathbf{t}(\mathbf{x})\big) \tag{4}$$

3   Methods

3.1   Data Acquisition

Full extracted brains from three ring-tailed lemurs (Lemur catta) and three rhesus macaques (Macaca mulatta) were obtained. Brains were perfusion fixed using paraformaldehyde after euthanasia for causes unrelated to the current research. Data were acquired on a 7T magnet with an Agilent DirectDrive console[1]. A 2D diffusion-weighted spin-echo protocol with single line readout was used (DW-SEMS; TE/TR: 25 ms/10 s; matrix size: 128 × 128; number of slices: 128; resolution: 0.5 mm (lemur) or 0.6 mm (macaque) isotropic. Sixteen non-diffusion-weighted (b = 0 s/mm^2) and 128 diffusion-weighted (b = 4000 s/mm^2) volumes were acquired with diffusion directions distributed over the whole sphere. The brains were stored in PBS before scanning and placed in Fluorinert during the scan.

---

[1] Agilent Technologies, Santa Clara, CA, USA.

3.2   Data Preprocessing

Diffusion tensors were fit to the data using the FSL `FDT` toolbox [9]. Additionally, `FDT` was used to generate a "no-diffusion" b = 0 image with T2 contrast. The T2 images were then bias-field corrected using FSL `FAST` [10]. Note that T2 and DTI images *within* each subject are already co-registered.

3.3   Template Creation

A combined T2 and DTI template was created for each species individually, using three subjects per species. Template creation followed a multi-resolution iterative approach. An initial template space was chosen by affine alignment of the T2 images to one randomly chosen subject using FSL `FLIRT` [11]. T2 images were intensity normalised before being resampled into this space, followed by voxelwise averaging across subjects. DTI images were resampled with reorientation using FSL `vecreg`, followed by log-tensor averaging. Each subject was then non-linearly aligned to the initial template at a warp resolution of 16 mm isotropic, and averaged in the same way to create a new template. This process was repeated, doubling the warp resolution each time, to a final resolution of 0.5 mm isotropic. At each iteration all images were smoothed using an isotropic Gaussian kernel, with a full width at half maximum equal to a quarter of the current warp resolution. The amount of regularisation was empirically decreased as warp resolution increased.

Spatial unbiasing of the template was carried out after each iteration, such that the average displacement of every voxel in the template was approximately 0. The methodology followed is given in Eqs. (5) and (6), where W_m is the warp from the template $\tilde{f}$ to m^{th} subject.

$$\bar{W}(\mathbf{x}) = \frac{1}{M} \sum_{m=1}^{M} W_m(\mathbf{x}) \tag{5}$$

$$\tilde{f}_{unbiased}(\mathbf{x}) = \tilde{f}_{biased}\left(\bar{W}^{-1}(\mathbf{x})\right) \tag{6}$$

Although $\bar{W}^{-1}(\mathbf{x})$ is not guaranteed to be diffeomorphic, we enforced diffeomorphism by projecting this field onto its closest diffeomorphic representation using the method proposed by Karacali [12].

Relative modality weighting was determined in two steps. First, by setting $\lambda_S = 1$, $\lambda_E = 0$ and varying λ_R until the T2 template appeared visually acceptable. Second, by setting $\lambda_S = 0$, fixing λ_R to the final value from step 1, and varying λ_E until the range of $|\mathbf{J}|$ for the warps was similar to that in step 1. In this way, we aimed to weight the influence of the T2 and DTI modalities on the warp approximately equally.

The template creation process was carried out three times using different combinations of the modalities to drive the registration: T2 image only, DTI image only, and multimodal using both T2 and DTI. We will refer to these methods and resulting templates as T_{T2}, T_{DT} and T_{MM} respectively. The weighting of individual modalities and the amount of regularisation was constant across all methods.

3.4   Template Quality Assessment

The quality of the T2 and DTI templates were visually evaluated, and examined for obvious inconsistencies. The T2 template quality was then quantified using the average *mutual information* (MI) between each warped subject and the template as calculated by FSL `FLIRT`. MI ranges between 0 (worst) and ∞. The DTI template quality was evaluated using the average *overlap of eigenvalue-eigenvector pairs* (OVL) metric between each warped subject and the template [1,13]. OVL is defined in Eq. (7), where λ_i and $\mathbf{e}_i$ are the ith eigenvalue and eigenvector of the tensor respectively. This provides a voxelwise similarity measure of the complete eigenvalue-eigenvector tensor descriptor. It ranges between 0 and 1, with 0 representing complete dissimilarity and 1 representing identical tensors. OVL was evaluated in 3 regions, namely the entire brain, the entire brain weighted by FA, and within a mask where FA > 0.2.

$$OVL = \frac{1}{N_{\mathbf{x}}} \sum_{\mathbf{x} \in \mathbb{R}^3} \frac{\sum_{i=1}^{3} \lambda_i^F(\mathbf{x})\lambda_i^G(\mathbf{x}) \left(\mathbf{e}_i^F(\mathbf{x})^{\intercal}\mathbf{e}_i^G(\mathbf{x})\right)^2}{\sum_{i=1}^{3} \lambda_i^F(\mathbf{x})\lambda_i^G(\mathbf{x})} \tag{7}$$

4   Results and Discussion

4.1   Visual Evaluation

Figure 1 shows the DTI, FA and T2 volumes of the final T_{MM} template for both the lemur and macaque. We do not show the entire T_{T2} and T_{DT} templates as at this scale they are visually difficult to distinguish from one another. Instead, we first describe the common appearance of the templates, and then focus on select regions which demonstrate difference between the methods.

Visually the T2 volumes are sharp, showing good contrast between tissue types, even within fairly complex structures such as the cerebellum. This is particularly true for the lemur template where the relatively simpler gyrification leads to smaller differences between subjects making spatial normalisation somewhat easier. The posterior of the macaque brain showed the highest variability between subjects and thus is unsurprisingly the least sharp template region. In general the T2 volume of the T_{MM} and and T_{T2} templates did appear slightly sharper than T_{DT}. This is in line with what might be expected: the contrast between signal from grey matter and from fluid in the sulci and ventricles in the DTI volume may be insufficient to overcome regularisation. A clear example of this in the lemur is shown in the bottom left of Fig. 1. Here, T_{DT} has been unable to correctly align one of the sulci, whereas both T_{MM} and T_{T2} have had no such difficulty.

The DTI volumes are of excellent quality. High FA within white matter indicates that not only are the anisotropic regions of the individual subjects brought to the correct positions, but that they arrive there with a consistent orientation. Perhaps somewhat surprisingly, the T_{T2} DTI template is not clearly worse than the other methods within these high FA regions. We suggest two possible reasons
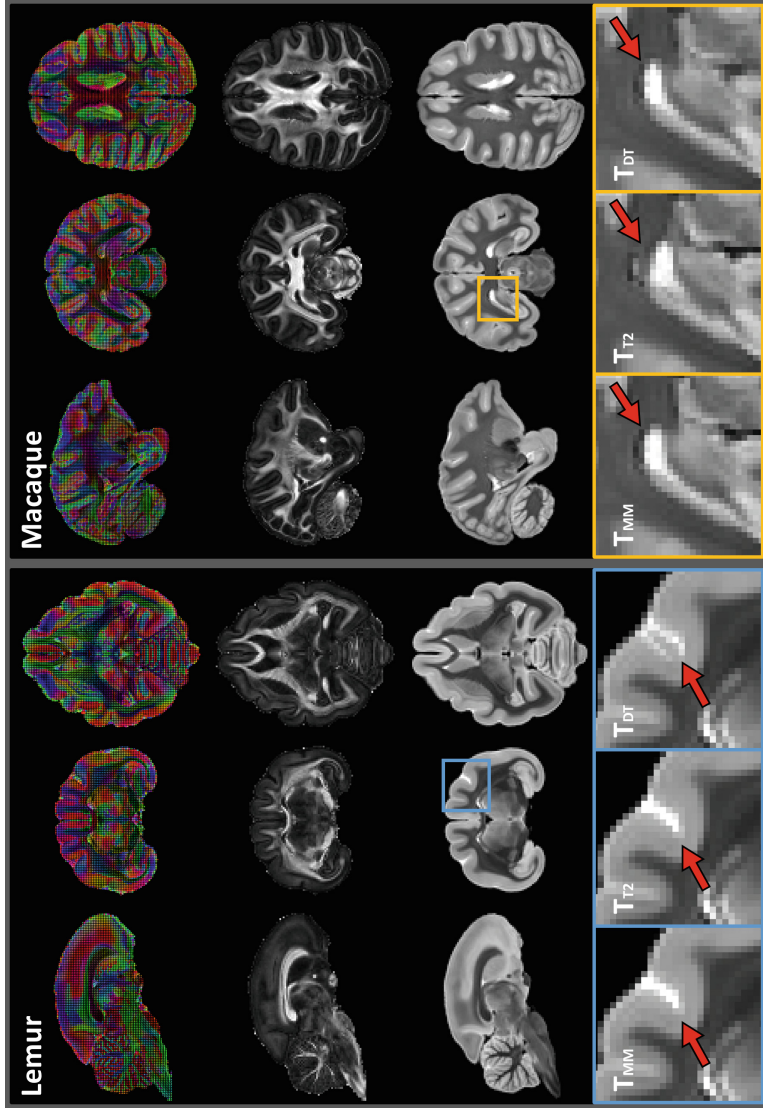
Fig. 1. DTI (1st row), FA (2nd row) and T2 (3rd row) images from the T_{MM} template for lemur (left) and macaque (right). In the lemur we highlight a region where T_{DT} was unable to correctly align the sulci due to insufficient contrast between grey matter and the exterior fluid. In the macaque we highlight a region where T_{T2} was unable to correctly align the ventricles due to differences in T2 contrast of the ventricular fluid between subjects. In both cases the T_{MM} template has aligned the region correctly.

for this. Firstly, the regularisation method is primarily focussed on maintaining plausible deformations. This means that the interior of the white matter is transformed in a sensible manner as it follows the tissue boundaries during registration of the T2 images. Secondly, the areas of anisotropy are always quite near tissue boundaries due to the small size of the lemur and macaque brains. Therefore, the possible deformations which correctly match tissue types are constrained in terms of allowable rotational effects. However, within some comparatively isotropic regions we do observed larger differences. In the macaque, the fluid in the ventricle of one subject had a significantly different (darker) T2 contrast than the others. The bottom right of Fig. 1 shows how this has led to ghosting around the ventricle in T_{T2}. In contrast, information in the DTI modality has allowed both T_{MM} and T_{DT} to correctly align the structure.

From this we can see a clear benefit in terms of anatomical correctness that the multimodal registration approach provides over its unimodal counterparts.

4.2   Quantitative Evaluation

Quantification of how well the subject data aligns to the templates is shown for the lemur in Table 1 and for the macaque in Table 2. Both species show similar trends between the methods indicating that these results generalise well.

Unsurprisingly, T_{T2} has the highest average MI, followed by T_{MM} and finally T_{DT}. This is in line with the visual observations of the T2 image sharpness in the respective templates. However, as shown by the ventricular ghosting above, the higher MI value does not necessarily mean better global anatomical correctness.

Whole brain OVL is consistently highest for T_{MM} indicating that both high and low FA regions are being well aligned by this method. The lower performance of unimodal methods can be attributed on the one hand to T_{T2} having no knowledge of rotations introduced by the warps, and on the other hand to poorer overlap of tissue types in T_{DT}.

When the OVL calculation is restricted to high-FA regions, either though FA-weighting or explicit FA masking, the results for all methods is higher than for the brain as a whole. As might be expected, T_{DT} performance is very similar to T_{MM} in these regions, and yet the multimodal approach still has an advantage. Interestingly, the performance deficit between T_{T2} and the other methods is greater for the macaque brain compared to the lemur. This may be due to an increase in the amount of uncontrolled rotation induced by the T_{T2} warps during alignment of the comparatively more complex cortical folding patterns in the macaque.

4.3   Overall Evaluation

Using multimodal data to drive the registration successfully overcame shortcomings in both the scalar and tensor only methods, whilst preserving the best aspects of each. DTI similarity measures improved across all areas of the template brains, and we therefore believe that a multimodal approach to registration can only be an advantage in the analysis of diffusion MRI data. Whilst scalar

Table 1. Average MI and OVL measures in the lemur

	MM template	T2 template	DTI template
T2 Mutual Information	1.357	1.407	1.268
OVL Whole Brain	0.859	0.854	0.850
OVL FA Weighted	0.968	0.966	0.967
OVL FA Masked	0.946	0.940	0.944

Table 2. Average MI and OVL measures in the macaque

	MM template	T2 template	DTI template
T2 Mutual Information	1.151	1.210	1.072
OVL Whole Brain	0.784	0.769	0.774
OVL FA Weighted	0.964	0.949	0.960
OVL FA Masked	0.891	0.870	0.880

registration appears to have a slight advantage over multimodal in terms of MI scores, the lower OVL scores and some clear anatomical inconsistencies in the T2 driven template suggest that this might be attributable to over-fitting. As the goal of registration in neuroimaging is anatomical consistency rather than outright image similarity, this regularisation effect of the tensor data in the multimodal method is in fact a desirable quality.

5   Conclusion and Future Work

We have shown that multimodal registration is a powerful tool for template creation, capable of leveraging complimentary imaging contrasts to find a common space which is truly representative of the group data. Our combination of scalar, tensor and regularisation cost functions allows us to optimise multiple aspects of the deformations simultaneously such that the most anatomically plausible mapping to this common space can be found. The multimodal-driven lemur and macaque templates we created show improved consistency with individual subject scans compared to both unimodal scalar and tensor driven templates.

The data used to create these templates is quite unique, with the lemur template in particular being the first of its kind for this species. We hope that by having a multimodal template available, any future analysis done on this species will be able to take advantage of increased anatomical consistency when analysing and reporting results of both individual and group studies.

Future work will focus on applying these same techniques to large human cohorts such as the Human Connectome Project and UK Biobank where we hope the benefits of multimodal templating will be even more apparent.

Ethical Approval. All brains were obtained from animals that were euthanised for causes unrelated to research and treated in accordance with appropriate guidelines.

References

1. Irfanoglu, M.O., et al.: DR-TAMAS: diffeomorphic registration for tensor accurate alignment of anatomical structures. Neuroimage **132**, 439–454 (2016)
2. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. **12**(1), 26–41 (2008)
3. Chen, P.: Hessian matrix vs. Gauss-Newton Hessian matrix. SIAM **49**(4), 1417–1435 (2011)
4. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vis. **61**(2), 139–157 (2005)
5. Chun, I.Y., Fessler, J.A.: Convolutional dictionary learning: acceleration and convergence. IEEE Trans. Image Process. **27**(4), 1697–1712 (2018)
6. Peeters, T.H.J.M., Rodrigues, P.R., Vilanova, A., ter Haar Romeny, B.M.: Analysis of distance/similarity measures for diffusion tensor imaging. In: Laidlaw, D., Weickert, J. (eds.) Visualization and Processing of Tensor Fields. MATHVISUAL, pp. 113–136. Springer, Heidelberg (2009)
7. Yeo, B., et al.: DT-REFinD: diffusion tensor registration with exact finite-strain differential. IEEE Trans. Med. Imaging **28**(12), 1914–1928 (2009)
8. Ashburner, J., Andersson, J.L., Friston, K.J.: Image registration using a symmetric prior-in three dimensions. Hum. Brain Mapp. **9**(4), 212–25 (2000)
9. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.: FSL. Neuroimage **62**(2), 782–790 (2012)
10. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging **20**(1), 45–57 (2001)
11. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. Med. Image Anal. **5**(2), 143–156 (2001)
12. Karacali, B., Davatzikos, C.: Estimating topology preserving and smooth displacement fields. IEEE Trans. Med. Imaging **23**(7), 868–880 (2004)
13. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Magn. Reson. Med. **44**(1), 41–50 (2000)