

Locally Orderless Registration for Diffusion Weighted Images

Henrik G. Jensen, Francois Lauze, Mads Nielsen, and Sune Darkner

Department of Computer Science, University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen, Denmark
{henne,madsn,francois,darkner}@di.ku.dk

Abstract. Registration of Diffusion Weighted Images (DWI) is challenging as the data, in contrast to scalar-valued images, is a composition of both directional and intensity information. The DWI signal is known to be influenced by noise and a wide range of artifacts. Therefore, it is attractive to use similarity measures with invariance properties, such as Mutual Information. However, density estimation from DWI is complicated by directional information. We address this problem by extending Locally Orderless Registration (LOR), a density estimation framework for image similarity, to include directional information. We construct a spatio-directional scale-space formulation of marginal and joint density distributions between two DWI, that takes the projective nature of the directional information into account. This accounts for orientation and magnitude and enables us to use a wide range of similarity measures from the LOR framework. Using Mutual Information, we examine the properties of the scale-space induced by the choice of kernels and illustrate the approach by affine registration.

1 Introduction

The registration of Diffusion Weighted Images (DWI) is interesting as it contains information about the fibrous micro-architecture otherwise invisible to structural MRI. Registration of these structures enables us to compare connectivity within and across subjects. However, registration is challenging due to the inherent geometry of DWI; notably high-angular resolution diffusion imaging (HARDI) which models more complex displacement profiles. We extend the Locally Orderless Registration (LOR) [2] density estimation framework for image similarity from scalar-valued images to DWI. The LOR is a scale-space framework for image density estimation that allows us to employ a wide range of similarity measures for registration, including MI. By introducing a spatio-directional kernel, thus including the space of gradient directions, we model the relationship between direction and measurements as histograms. The histograms are mapped to probability density estimates by normalization and marginalization over the deformed space.

Our contribution is a full LOR scale-space formulation for DWI, offering explicitly control of orientation, image, intensity and integration scale. We examine

the effects of the scale-space and illustrate the application of the density estimate by affine image registration of DWI data using Mutual Information.

2 Previous Work and Background

Locally Orderless Images (LOI) [6] is a scale-space representation of intensity distributions in images modeling three inherent scales: the image scale (i.e. image smoothing), the integration scale (local histogram), and the intensity scale (soft bin width). The first mention of LOI in the context of image registration was in by Hermosillo et al. [4] where a variational approach to image registration was presented. The LOR framework [2], an extension of [1], generalized a range of similarity measures as linear and non-linear functions of density estimates for scalar-valued images. One such non-linear similarity measure is Mutual Information (MI) [13]. MI is one of the most frequently used similarity measures in image registration and was introduced as a multi-modal similarity measure. MI is frequently used in MRI due to its invariance properties with respect to intensity values and is associated with scalar-valued images. It is used in the context of DWI for distortion-correction [9] on e.g. b_0 or individual DWI directions. Van Hecke et al. [12] used MI for non-rigid registration of DWI. Under an assumption of alignment, each gradient direction was evaluated separately as well as in a pooled fashion to form a joint density distribution. Interpolation of directional information in DWI was introduced by Tao and Miller [10] for affine registration using SSD and extended by Duarte-Carvajalino et al. [3] to non-rigid B-spline registration. The angular interpolation was extended with a Watson distribution by Rathi et al. [8]. Raffelt et al. [7] used SSD after spherical deconvolution for fiber modeling (FODs), while others, like Yap et al. [14], compared the coefficients of the spherical harmonics.

Image registration is the process of spatially aligning (two) images (I and J) under some transformation Φ given some regularity condition $\mathcal{S}(\Phi)$ and similarity $\mathcal{F}(I \circ \Phi, J)$ such that $\mathcal{M}(I, J, \Phi)$ is minimized

$$\mathcal{M}(I, J, \Phi) = \mathcal{F}(I \circ \Phi, J) + \mathcal{S}(\Phi) \quad (1)$$

In this paper we address the estimation of \mathcal{F} of single shell DWI as an extension of LOR with application to MI. DWI MR attenuation signals at location \mathbf{x} , for a gradient direction \mathbf{v} , are modeled by $S(\mathbf{x}, \mathbf{v}) = S_0(\mathbf{x})e^{-bI(\mathbf{x}, \mathbf{v})}$ [10] and apparent diffusion coefficients volumes are given by $I(\mathbf{x}, \mathbf{v}) = -\frac{1}{b} \log \frac{S(\mathbf{x}, \mathbf{v})}{S_0(\mathbf{x})}$. Gradient directions \mathbf{v} are taken on the unit sphere \mathbb{S}^2 although diffusion are orientation-free and have $I(\mathbf{x}, \mathbf{v}) \approx I(\mathbf{x}, -\mathbf{v})$, i.e., with antipodal symmetry. Such a symmetric function $I(\mathbf{x}, -)$ on the sphere can be represented by a function on the projective space \mathbb{P}^2 of directions of \mathbb{R}^3 , $\mathbb{P}^2 \simeq \mathbb{S}^2 / \{\pm 1\}$.

We start by defining the type of transformation considered for the LOR density estimates for single shell DWI presented in this paper. For any transformation ϕ of a point \mathbf{x} , we consider only diffeomorphic mappings $\phi(\mathbf{x}): \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Under this assumption, ϕ is invertible and its differential, or Jacobian $d_{\mathbf{x}}\phi$ at

\mathbf{x} , gives naturally rise to a projective transformation on \mathbb{P}^2 : $t\mathbf{v} \mapsto t d_{\mathbf{x}}\phi(\mathbf{v})$, $t \in \mathbb{R} \setminus \{0\}$. We drop \mathbf{x} and simply write $d\phi$. Its representation over \mathbb{S}^2 is $\mathbf{v} \in \mathbb{S}^2 \mapsto \pm \frac{d\phi(\mathbf{v})}{|d\phi(\mathbf{v})|}$. This term appears within a spherical kernel Γ_{κ} with antipodal symmetry, making the sign irrelevant. Setting $\psi(\mathbf{v}) = \frac{d\phi(\mathbf{v})}{|d\phi(\mathbf{v})|}$, it corresponds to the transformation proposed by [10]. We therefore extend our transformation to $\tilde{\Phi} : (\mathbf{x}, \mathbf{v}) \mapsto (\phi(\mathbf{x}), \psi(\mathbf{v}))$. This type of transformation is also argued in [7], although neither [10] nor [7] did consider its projective nature. We proceed to describe the LOR.

The LOR framework defines the similarity over three scales: The image scale σ , the intensity scale β , and the integration scale α . In registration, for a transformation ϕ , we get

$$h_{\beta\alpha\sigma}(i, j | \phi, \mathbf{x}) = \int_{\Omega} P_{\beta}(I_{\sigma}(\phi(\mathbf{x})) - i) P_{\beta}(J_{\sigma}(\mathbf{x}) - j) W_{\alpha}(\boldsymbol{\tau} - \mathbf{x}) d\boldsymbol{\tau} \quad (2)$$

$$p_{\beta\alpha\sigma}(i, j | \phi, \mathbf{x}) \simeq \frac{h_{\beta\alpha\sigma}(i, j | \phi, \mathbf{x})}{\int_{\Lambda^2} h_{\beta\alpha\sigma}(k, l | \phi, \mathbf{x}) dk dl} \quad (3)$$

where $i, j \in [a_1, a_2]$ are values in the image intensity range, $I_{\sigma}(\phi(\mathbf{x})) = (I * K_{\sigma})(\phi(\mathbf{x}))$ and $J_{\sigma}(\mathbf{x}) = (J * K_{\sigma})(\mathbf{x})$ are images convolved with the kernel K_{σ} with standard deviation σ , P_{β} is a Parzen-window of scale β , and W_{α} is a Gaussian integration window of scale α . The marginals are trivial and obtained by integration over the appropriate variable. The LOR-approach to similarity lets us use a set of generalized similarity measures, the linear and non-linear

$$\mathcal{F}_{lin} = \int_{\Lambda^2} f(i, j) p(i, j) di dj \quad \mathcal{F}_{non-lin} = \int_{\Lambda^2} f(p(i, j)) di dj \quad (4)$$

where the linear measure $f(i, j)$ includes e.g. sum of squared differences and Huber, and the non-linear $f(p(i, j))$ includes e.g. MI, NMI, see [2] for details.

3 Locally Orderless DWI

To extend the density estimates of LOR to include directional information, we introducing a kernel on the sphere to account for directional smoothing. With that in mind, we extend spatial smoothing to be spatio-directional, where the directional smoothing preserves this symmetry, and thus the projective structure, via a symmetric kernel $\Gamma_{\kappa}(\boldsymbol{\nu}, \mathbf{v})$ on \mathbb{S}^2 . We define the smoothed signal $I_{\sigma, \kappa}$ at scales (σ, κ) by

$$I_{\sigma\kappa}(\mathbf{x}, \mathbf{v}) = \int_{\mathbb{S}^2} \left(\int_{\Omega} I(\boldsymbol{\tau}, \boldsymbol{\nu}) K_{\sigma}(\boldsymbol{\tau} - \mathbf{x}) d\boldsymbol{\tau} \right) \Gamma_{\kappa}(\boldsymbol{\nu}, \mathbf{v}) d\boldsymbol{\nu} = (I * (K_{\sigma} \otimes \Gamma_{\kappa}))(\mathbf{x}, \mathbf{v}) \quad (5)$$

where $K_{\sigma}(\mathbf{x})$ is a Gaussian kernel with σ standard deviation. We use a symmetric Watson distribution [5] as $\Gamma_{\kappa}(\boldsymbol{\nu}, \mathbf{v})$ for directional smoothing on \mathbb{S}^2 , given by

$$\Gamma_{\kappa}(\boldsymbol{\nu}, \mathbf{v}) = C e^{\kappa \langle (\boldsymbol{\nu}, \mathbf{v}) \rangle^2}, \quad C = M\left(\frac{1}{2}, \frac{1}{d}, \kappa\right) = \sum_{i=0 \dots \infty} \frac{\kappa^i}{\left(\frac{3}{2}\right)^i i!} \quad (6)$$

where M is the Kummer function for a d -dimensional unit vector $\boldsymbol{\nu}$, (in this case $d = 3$), $\pm \boldsymbol{v}$ the center of the distribution, and κ the concentration parameter, which is roughly inverse proportional to the variance on the sphere. As one alternative, a symmetrized von Mises-Fisher [5] distribution could be considered.

In order to use the similarity measures provided by the LOR framework, we extend the LOR formulation from scalar-valued images to DWI. The joint histogram, that is, the contribution to $h(i, j): \Lambda^2 \rightarrow \mathbb{R}_+$ of the joint histogram and normalization can be written as

$$h_{\beta\alpha\sigma\kappa}(i, j|\boldsymbol{x}) = \int_{\Omega \times S^2} P_{\beta}(I_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v}) - i)P_{\beta}(J_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v}) - j)W_{\alpha}(\boldsymbol{\tau} - \boldsymbol{x})d\boldsymbol{x} \times d\boldsymbol{v} \tag{7}$$

$$p_{\beta\alpha\sigma\kappa}(i, j|\boldsymbol{x}) = \frac{h_{\beta\alpha\sigma\kappa}(i, j|\boldsymbol{x})}{\int_{\Lambda^2} h_{\beta\alpha\sigma\kappa}(k, l|\boldsymbol{x})dk dl} \tag{8}$$

where $I_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v})$ and $J_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v})$ are defined as in Equation (5), P is a Gaussian Parzen-window with standard deviation β , and W a Gaussian window of integration around \boldsymbol{x} with standard deviation α . The marginals are trivial and obtained by integration over the appropriate variable. The joint and marginal probability densities allow us to apply the generalized similarity measures in Equation (4). In this paper, we use the non-linear MI.

4 Image Registration

We write the joint histogram and density for similarity in image registration as

$$h_{\beta\alpha\sigma\kappa}(i, j|\tilde{\Phi}, \boldsymbol{x}) = \tag{9}$$

$$\int_{\Omega \times S^2} P_{\beta}(I_{\sigma\kappa}(\phi(\boldsymbol{x}), \psi(\boldsymbol{v})) - i)P_{\beta}(J_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v}) - j)W_{\alpha}(\boldsymbol{\tau} - \boldsymbol{x})d\boldsymbol{\tau} \times d\boldsymbol{v}$$

$$p_{\beta\alpha\sigma\kappa}(i, j|\tilde{\Phi}, \boldsymbol{x}) = \frac{h_{\beta\alpha\sigma\kappa}(i, j|\tilde{\Phi}, \boldsymbol{x})}{\int_{\Lambda^2} h_{\beta\alpha\sigma\kappa}(i, j|\tilde{\Phi}, \boldsymbol{x})dl dk} \tag{10}$$

Most similarity measures are global measures, including MI. To make the density estimate global, we let $\alpha \rightarrow \infty$ such that W becomes constant. The first-order structure of the similarity (1) is derived following the approach of [2], denoting differentials as $dg = Dg(\boldsymbol{x})d\boldsymbol{x}$, where D is the partial derivative operator and $d\boldsymbol{x}$ a vector of differentials. We seek $d\mathcal{M}$, the derivative of (1), ignoring the regularization term and omitting irrelevant parameters in the notation. The derivative of MI with respect to $h(i, j)$ is found in [2]. Thus, we seek $dh(i, j)$

$$dh(i, j) = \int_{\Omega \times S^2} dP_{\beta}(I_{\sigma\kappa}(\phi(\boldsymbol{x}), \psi(\boldsymbol{v})) - i)P_{\beta}(J_{\sigma\kappa}(\boldsymbol{x}, \boldsymbol{v}) - j)W_{\alpha}(\boldsymbol{\tau} - \boldsymbol{x})d\boldsymbol{\tau} \times d\boldsymbol{v} \tag{11}$$

with $dP_{\beta}(I_{\sigma\kappa}(\phi(\boldsymbol{x}), \psi(\boldsymbol{v})) - i) = DP_{\beta}(I_{\sigma\kappa}(\phi(\boldsymbol{x}), \psi(\boldsymbol{v})) - i)dI_{\sigma\kappa}(\phi(\boldsymbol{x}), \psi(\boldsymbol{v}))$

$DP_\beta(I_{\sigma\kappa}(\phi(\mathbf{x}), \psi(\mathbf{v})) - i)$ can be found in [1, 2]. Inserting (5), we get

$$dI_{\sigma\kappa}(\phi(\mathbf{x}), \psi(\mathbf{v})) = d \int_{S^2} \left(\int_{\Omega} I(\boldsymbol{\tau}, \boldsymbol{\nu}) K_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x})) d\boldsymbol{\tau} \right) \Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) d\boldsymbol{\nu}. \quad (12)$$

and using the Leibniz integration rule and the product rule, we get

$$dI_{\sigma\kappa}(\phi(\mathbf{x}), \psi(\mathbf{v})) = \int_{S^2} \left(d \int_{\Omega} I(\boldsymbol{\tau}, \boldsymbol{\nu}) K_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x})) d\boldsymbol{\tau} \right) \Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) d\boldsymbol{\nu} + \int_{S^2} \left(\int_{\Omega} I(\boldsymbol{\tau}, \boldsymbol{\nu}) K_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x})) \boldsymbol{\tau} \right) d\Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) d\boldsymbol{\nu} \quad (13)$$

We consider each of the terms on the sum separately. Using Leibniz integration rule on the first term of the sum, we get

$$\int_{S^2} \left(\int_{\Omega} I(\boldsymbol{\tau}, \boldsymbol{\nu}) dK_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x})) \boldsymbol{\tau} \right) \Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) d\boldsymbol{\nu} \quad (14)$$

where $dK_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x})) = DK_\sigma(\boldsymbol{\tau} - \phi(\mathbf{x}))d\phi(\mathbf{x})$ which is trivial in the context of registration. From the second term we get $d\Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) = D\Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v}))d\psi(\mathbf{v})$ and specifying $\Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v}))$ as a Watson distribution gives

$$D\Gamma_\kappa(\boldsymbol{\nu}, \psi(\mathbf{v})) = C e^{\kappa(\langle \boldsymbol{\nu}, \psi(\mathbf{v}) \rangle)^2} 2\kappa \langle \boldsymbol{\nu}, \psi(\mathbf{v}) \rangle d\psi(\mathbf{v}) \quad (15)$$

which leaves $d\phi(\mathbf{x})$ and $d\psi(\mathbf{v})$. The first term $d\phi$ is the Jacobian of ϕ and classical in registration literature. The first-order information on spherical reorientation $d\psi(\mathbf{v})$ is more complicated, as with our definition of $\psi(\mathbf{v})$ as $\frac{d\phi(\mathbf{v})}{|d\phi(\mathbf{v})|}$, this leads to second-order information of ϕ , which is complex but trivial.

5 Experiments and Results

A series of experiments was conducted to illustrate the scales introduced (spatial, intensity, and directional) with respect to MI. We computed the MI between two subjects and plotted the MI as a function of global rotation and translation (Figure 1) as well as local rotation of three random patches of $10 \times 10 \times 10$ voxels (Figure 2). In addition, we performed a few affine registrations of DWI data using the proposed extension of LOR to DWI and MI. We used data from the Human Connectome Project (HCP) database, release Q3, structurally aligned to the MNI-152 template [11], with 90 gradient directions and a b-value of 3000.

The locally orderless structure introduces four explicit scales on DWI: Image K_σ , intensity P_β , and integration W_α , as well as the extension to orientation Γ_κ . To examine the effect of the scales (ignoring integration scale W_α , i.e. $\alpha \rightarrow \infty$), we use Mutual Information, which in itself is a complicated measure. Mutual Information of two observations A, B can be interpreted as the capability of A to encode B . We use this notion of MI to examine the properties of the proposed extension of density estimation to DWI. As a first observation from Figure 1

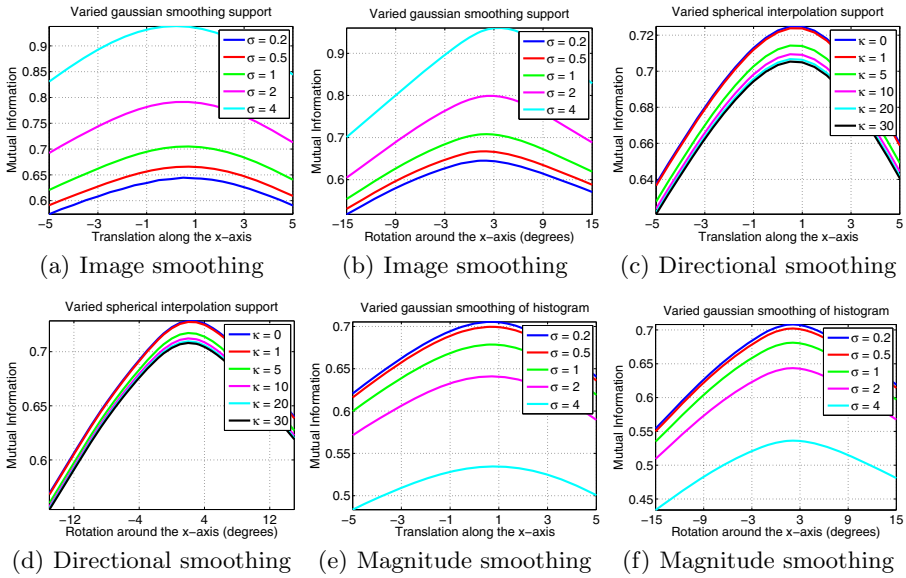


Fig. 1. MI as a function of translation and rotation at different scales. As shown, smoothing in \mathbb{R}^3 (a & b) moves the optima, while the change in the angular or diffusion scale (c & d) preserves the MI, despite increased the angular information. This is a good indication of a substantial information in the directions. Smoothing of diffusion magnitudes (e & f) has a similar effect to that observed for scalar-valued images.

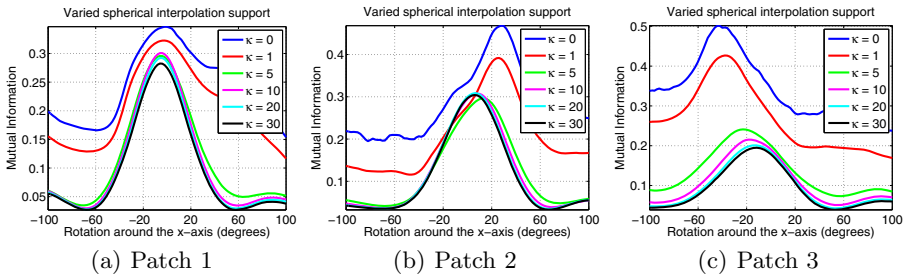


Fig. 2. We computed the MI between two DWI volumes within three random patches of $10 \times 10 \times 10$ voxels as a function of rotation and directional smoothing. This clearly illustrates the change in optima as a function of the directional scale. As the images are reasonably well-aligned, this is a strong indication that directional information is required for proper local alignment

it is clear that the DWI optima does not correspond to the structural optima of the registration (to MNI) provided by the HCP. This is illustrated by the fact that the maxima of Figure 1 are not at 0.

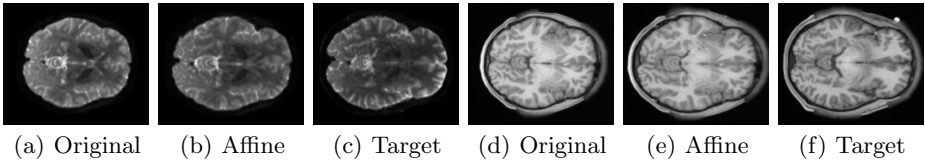


Fig. 3. Two DWI images registered using affine transformation and MI for DWI. (left) $b=0$ gradient images. (Right) T1-weighted images.

The Image Scale influences the MI significantly. Smoothing in the individual directions increases the MI (Figures 1(a) and 1(b)). This increase is not surprising as this smoothing of the intensities will transform the distribution of observed intensities towards the mean of the image. **The Intensity Scale** (i.e. Parzen-window) behaves as reported in [2] where the optima displaces with increased kernel size (Figures 1(e) and 1(f)). Increasing the size of the Parzen-window corresponds to reducing the number of bins. **The Orientation Scale** has an effect similar to image smoothing (Figures 1(c) and 1(d)). We observe that smoothing results in increased Mutual Information as the diffusion measurements of all 90 directions converge towards the rotation-invariant *mean diffusivity* for $\kappa \rightarrow 0$. Note that the corresponding curves of MI using small kernels, i.e. higher angular resolution, only results in a small decrease in the MI. Figures 1(c) and 1(d) shows preservation of the slope of MI towards the optima is observed, revealing a well-defined optima. Locally, Figure 2, we observe a dramatic shift in optima from mean diffusivity $\kappa = 0$ to high directional resolution $\kappa = 30$. As illustrated, the local optima shifts 30-40 degrees as a function of scale, which justify the need for our proper scale-space formulation for similarity of DWI. **To illustrate** the LOR for DWI with MI, we performed a few affine registrations using MI, $\kappa = 30$, a cubic B-spline Parzen-window with 200 bins, and B-Spline image interpolation. A registration can be seen in Figure 3.

6 Discussion and Conclusion

The LOR for DWI includes directional information and so first-order information of the deformation is required. We therefore restrict the deformation model to diffeomorphisms to ensure well-defined derivatives. For gradient-based optimization, this implies that the second-order information of the deformation is required, which severely complicates any implementation. We have chosen the Watson distribution for its simplicity compared to e.g. a symmetrized von Mises-Fisher kernel or symmetrized geodesic distances.

We have presented an extension of the Locally Orderless Registration for DWI by introducing a scale-space which accounts for the projective nature of DWI in a theoretically sound manner. Our experiments show that directional resolution is important in order to obtain proper local alignment in registration. Our formulation allows us to directly control the scales of the information from which we

estimate the similarity. By extending the LOR framework, we can easily apply a wide range of similarity measures. We provided the first-order information of the densities, briefly reviewed the effects of the scales, and illustrated the approach by affine registration of DWI using Mutual Information.

References

1. Darkner, S., Sporring, J.: Generalized partial volume: An inferior density estimator to parzen windows for normalized mutual information. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 436–447. Springer, Heidelberg (2011)
2. Darkner, S., Sporring, J.: Locally Orderless Registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(6), 1437–1450 (2013)
3. Duarte-Carvajalino, J.M., Sapiro, G., Harel, N., Lenglet, C.: A framework for linear and non-linear registration of diffusion-weighted mris using angular interpolation. *Frontiers in Neuroscience* 7 (2013)
4. Hermosillo, G., Chef d’Hotel, C., Faugeras, O.: Variational methods for multimodal image matching. *International Journal of Computer Vision* 50(3), 329–343 (2002)
5. Jupp, P.E., Mardia, K.: A unified view of the theory of directional statistics, 1975–1988. *International Statistical Review*, 261–294 (1989)
6. Koenderink, J., Van Doorn, A.: The structure of locally orderless images. *International Journal of Computer Vision* 31(2), 159–168 (1999)
7. Raffelt, D., Tournier, J., Fripp, J., Crozier, S., Connelly, A., Salvado, O., et al.: Symmetric diffeomorphic registration of fibre orientation distributions. *NeuroImage* 56(3), 1171–1180 (2011)
8. Rathi, Y., Michailovich, O., Shenton, M.E., Bouix, S.: Directional functions for orientation distribution estimation. *Medical Image Analysis* 13(3), 432–444 (2009)
9. Rohde, G., Barnett, A., Basser, P., Marenco, S., Pierpaoli, C.: Comprehensive approach for correction of motion and distortion in diffusion-weighted mri. *Magnetic Resonance in Medicine* 51(1), 103–114 (2004)
10. Tao, X., Miller, J.V.: A method for registering diffusion weighted magnetic resonance images. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 594–602. Springer, Heidelberg (2006)
11. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K.: The wu-minn human connectome project: an overview. *Neuroimage* 80, 62–79 (2013)
12. Van Hecke, W., Leemans, A., D’Agostino, E., De Backer, S., Vandervliet, E., Parizel, P.M., Sijbers, J.: Nonrigid coregistration of diffusion tensor images using a viscous fluid model and mutual information. *IEEE Transactions on Medical Imaging* 26(11), 1598–1612 (2007)
13. Wells, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* 1(1), 35–51 (1996)
14. Yap, P.T., Chen, Y., An, H., Yang, Y., Gilmore, J.H., Lin, W., Shen, D.: Sphere: Spherical harmonic elastic registration of hardi data. *NeuroImage* 55(2), 545–556 (2011)