CLINICAL DT-MRI ESTIMATION, SMOOTHING AND FIBER TRACKING WITH LOG-EUCLIDEAN METRICS.

Pierre Fillard, Vincent Arsigny, Xavier Pennec and Nicholas Ayache

INRIA Sophia Antipolis - EPIDAURE / ASCLEPIOS Project,
06902 Sophia Antipolis Cedex, France.

ABSTRACT

Diffusion tensor MRI is an imaging modality that is gaining importance in clinical applications. However, in a clinical environment, data have to be acquired rapidly, often at the detriment of the image quality. We propose a new variational framework that specifically targets low quality DT-MRI. The Rician nature of the noise on the images leads us to a maximum likelihood strategy to estimate the tensor field. To further reduce the noise, we optimally exploit the spatial correlation by adding to the estimation an anisotropic regularization term. This criterion is easily optimized thanks to the use of recently introduced Log-Euclidean metrics. Results on real clinical data show promising improvements of fiber tracking in the brain and the spinal cord.

1. INTRODUCTION

Diffusion tensor Imaging (DTI) [1] is a unique tool to assess in vivo oriented structures within tissues via the directional measure of water diffusion. However, using such an imaging modality in a clinical environment is difficult: data must be acquired in a short amount of time due to pathologies that often prevent the patient to stay in a static position for too long. This results in acquisitions with a limited number of encoding gradients and low signal-to-noise ratios (SNR). The estimation of the diffusion tensor field from diffusion weighted images (DWI) is noise-sensitive and thus clinical DTI is often not suitable for fiber tracking. For these reasons, there has been a growing interest in the regularization of tensor images.

The Stejskal-Tanner diffusion equation [1] relates the diffusion tensor D to each noise-free DWI with $S_i = S_0 \exp(-b\; g_i^T D g_i)$, where S_i is the original DWI corresponding to the encoding gradient g_i, S_0 an image with a null gradient, and b the diffusion factor. To get a linear system, one usually takes the logarithm [2] of the DWI. Solving the system in a least square (LS) sense leads to the minimization of a quadratic criterion with algebraic methods. Doing this implicitly assumes a Gaussian noise on the images logarithm, which is justified for high SNR. Also for high SNR, the MRI noise is well approximated by a Gaussian on images directly. However, when working with clinical DTI, SNR is very low. In that case, the real nature of the noise in the images is Rician. Wang et al. [3] proposed an estimation criterion on the complex DWI signal that is adapted to a Rician noise. However, one generally cannot access the full complex signal but only its magnitude. To overcome this limitation, we propose a maximum likelihood (ML) strategy which exploits the a priori knowledge on the probability density function (pdf) of the Rician noise. We will show that considering such a noise leads to an unbiased estimator that corrects for the *shrinking effect*: tensors tend to be smaller when estimated on the signal directly when the noise is Rician. However, solving this non-linear criterion requires an adapted framework to work with tensors.

For the ultimate application targeted in this paper, i.e. fiber tracking, the tensor field needs to be regularized without blurring the transitions between distinct tracts, which delimit anatomical and functional brain regions. Most of the regularization methods proposed so far rely on a feature of tensors that belong to a vector space: [4] uses the spectral decomposition of tensors to independently regularize their eigenvectors and eigenvalues, while [3] smooths the Cholesky factors of tensors. In this paper, we propose to use Log-Euclidean (LE) metrics for tensors [5], that have excellent theoretical and practical properties for tensor processing.

Affine-invariant Riemannian metrics [6, 7, 8] overcome the limitations of the Euclidean calculus on tensors, like the appearance of negative eigenvalues and the swelling effect as described in [5]. With these metrics, tensor space is turned into a regular manifold where matrices with null and negative eigenvalues are at an infinite distance of any tensor. However, computations with these metrics are time-consuming since they extensively use the matrix exponential, logarithm, square root and inverse. A novel family of metrics, called Log-Euclidean, detailed in [5], combines most of the properties of the affine-invariant family with a computational cost close to the Euclidean case, and thus are more suitable for tensor processing. To a tensor D is associated a unique *logarithm* L which is *symmetric*. It verifies $D = \exp(L)$ where $\exp$ is the matrix exponential. Conversely, a symmetric matrix is associated to a tensor thanks to the exponential map. L is obtained by taking the scalar logarithm of D eigenvalues. Since their is a one-to-one mapping between the tensor space and the space of symmetric matrices, one can give the tensor space a *vec-*

tor space structure by transporting the addition and the scalar multiplication onto the space of symmetric matrices with the exponential. Practically, the use of LE metrics consists in taking the matrix logarithm of tensors, running computations on these vectors, and mapping the result back to the tensor space with the matrix exponential.

The rest of the paper is organized as follows. In Sec. 2, we detail the variational method of the joint estimation and smoothing of DTI. In Sec. 3, we present very promising results on a brain dataset, and the first successful reconstruction of a spinal cord tract.

2. JOINT ESTIMATION AND SMOOTHING OF DTI

The joint estimation and regularization of DTI can be tackled by a variational formulation :

$$E(L) = \frac{1}{2}\text{Sim}(L) + \frac{\lambda}{2}\text{Reg}(L), \quad (1)$$

with Sim(.) being the data attachment term (estimation from the DWIs), Reg(.) being the regularization term, and λ a normalization factor. To use the LE framework, we work on the tensor logarithm $L = \log(D)$. Next, we first present the estimation term, then the regularization term.

2.1. Two Criteria for Tensor Estimation

2.1.1. A LS Criterion on the DWI signal

Generally, the noise is supposed to be Gaussian on the DWI signal. The maximum likelihood (ML) estimator is the LS:

$$\text{Sim}_{sig}(L) = \sum_{i=0}^{N} \left(\hat{S}_i - S_0 \exp(-b g_i^T \exp(L) g_i)\right)^2. \quad (2)$$

For more clarity, we denote by $S_i(L) = S_0 \exp(-b g_i^T \exp(L)$ $g_i)$ the noise-free DWI signal with parameter L and by $\hat{S}_i$ the measured one. Traditionally, one uses $S_0 = \hat{S}_0$. Experiments to re-estimate S_0 in our ML framework show no significant difference. Energy 2 is minimized with a simple first order gradient descent. Then, one needs to differentiate it. Relying on the property that $\text{Trace}(A\partial_B \exp) = \text{Trace}(B\partial_A \exp)$, we obtain: $\nabla\text{Sim}_{sig}(L) = 2b\sum_{i=0}^{N}(\hat{S}_i - S_i)\partial S/\partial L$, with $\partial S/\partial L = S_i \partial_{g_i g_i^T} \exp(L)$. A practical implementation of the directional derivatives $\partial_G \exp(L)$ is given in [9].

2.1.2. A Maximum Likelihood Estimator

If we assume a Rician noise of variance σ^2 on the data, the pdf of the measured signal $\hat{S}$ knowing the expected signal S is [10]:

$$p(\hat{S}/S) = \frac{\hat{S}}{\sigma^2} \exp\left(-\frac{\hat{S}^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{S\hat{S}}{\sigma^2}\right), \quad (3)$$

where I_0 is the modified 0th order Bessel function of the first kind. Such a noise induces a bias on the DWI signal. Indeed, one can show that in this case the DWI signal is shifted by approximatively $\sigma^2/(2S)$ [10]. Thus, the DWI signal is greater than it should be, and tensors when estimated without correcting for this bias will be smaller than they actually are (a higher signal means a lower diffusion). To correct for this shrinking effect, we propose the ML estimator for the pdf 3:

$$\text{Sim}_{ML}(L) = -\sum_{i=0}^{N} \log\left(p(\hat{S}_i/S_i)\right). \quad (4)$$

The differentiation of equation 4 in the LE framework gives: $\nabla\text{Sim}_{ML}(L) = -1/\sigma^2 \sum_{0=1}^{N}(S_i - \alpha\hat{S}_i)\partial S/\partial L$, with $\alpha = I_0'/I_0(\hat{S}_i S_i/\sigma^2)$. The formulation is similar to the gradient of Eq. 2, except that a correcting factor depending on the signal and the noise variance appears. A simple estimator of the noise variance is based on the following. MRI includes an empty region outside the patient. Considering the fact that the square magnitude of such region is null, taking its mean gives us an estimation of $2\sigma^2$.

To reduce the influence of potential outliers, one could use M-Estimators: they consist in replacing the residuals $(\hat{S}_i - S_i)$ of Eq. 2 and $(S_i - \alpha\hat{S}_i)$ of Eq. 4 by a function of these residuals, which zeroes out the influence of one measure if it is aberrant. In practice, we did not observe outliers justifying the use of M-Estimators. We now investigate the regularization term of the global criterion.

2.2. An Anisotropic Regularization Term

The anisotropic regularization of the tensor field can be handled through the minimization of a ϕ-functional: $\text{Reg}(L) = \int_\Omega \phi\left(\|\nabla L\|\right)$. The ϕ-function will give an anisotropic behavior to the regularization, i.e. it will preserve the edges of the tensor field while smoothing homogeneous regions. Similarly

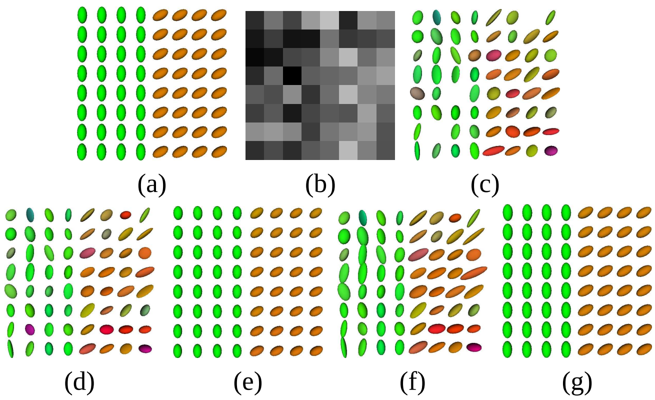

Fig. 1. **Estimations of a synthetic dataset. (a):** the original field. **(b):** a DWI ($g = (-0.22, 0.85, 0.49)^T$). **(c):** the classic estimation (non-displayed tensors are not positive). **(d):** the SigGD, **(e):** the SigGD+Reg, **(f):** the ML, **(g):** the ML+Reg.

to criteria 2 and 4, one has to differentiate the regularization energy to minimize it by gradient descent:

$$\nabla \text{Reg}(L) = -2\psi\left(\|\nabla L\|\right)\Delta L - 2\nabla^T\left(\psi\left(\|\nabla L\|\right)\right)\nabla L, \quad (5)$$

with $\psi(s) = \phi\prime(s)/s$. Directional derivatives, gradient and Laplacian are estimated with a finite differences scheme like with a scalar image (see [9] for details). For the experiments, we used $\psi(s) = (1 + s^2/\kappa^2)^{-1/2}$ as in [4]. κ is a normalization factor for the gradient. Finally, by combining the gradient of Eq. 2 or 4 with gradient 5, one obtains the evolution equation of the joint estimation and smoothing of DTI: $L_{t+1} = L_t - dt/2(\nabla\text{Sim}(L_t) + \lambda\nabla\text{Reg}(L_t))$. Of course, one has to exponentiate the solution to obtain a tensor.

Implementation: We used a quasi Newton (BFGS) optimization strategy for the minimization [11]. The integration step was adapted at each iteration with the Wolfe linear research [11]. The initialization was done by the classical estimation where non-positive tensors were replaced by the mean of positive neighbors. A threshold on the norm of the full criterion gradient was set to determine the convergence.

3. RESULTS ON SYNTHETIC AND CLINICAL DATA

3.1. Synthetic Data

We synthetically generated a 16x16x16 tensor field containing two homogeneous regions with anisotropic tensors (Fig. 1 a), like in [3]. The DWI are artificially produced using the Stejskal-Tanner equation and 25 diffusion gradients (b_0=10.0). A Rician noise of variance $\sigma_n = 1.5$ is finally added to each DWI including the b_0 (Fig. 1 b). We compared 5 estimations: a classic estimation with an algebraic resolution (Classic), a gradient descent on the DWI signal with a Gaussian noise (Eq. 2) (SigGD) and with our ML framework on the Rician noise. (Eq. 4) (ML). These two similarity criteria are also combined with an anisotropic regularization term ($\kappa = 0.1$ and $\lambda = 1.0$). To quantitatively evaluate the benefits, we computed the number of non-positive tensors (NPT), the mean, minimum and maximum error between each estimation and the original data with the LE metric. The shrinking effect is highlighted with the ratio of the mean tensor volume after estimation and the mean tensor volume of the original data. Results are summarized below.

	NPT	Mean	Min	Max	Vol. ratio
Classic	224	∞	0.131	∞	0.82
SigGD	0	0.520	0.136	1.193	0.70
SigGD+Reg	0	0.223	0.112	0.442	0.73
ML	0	0.481	0.116	1.113	0.96
ML+Reg	0	0.056	0.030	0.09	0.99

Since non positive tensors appear with a classic estimation (fig. 1 c), the LE metric gives an infinite error. On the contrary, our ML estimator (fig. 1 f) insures a positive definite result and corrects for the shrinking effect visible in Fig. 1 d and e. Finally, the regularization term smooths the field while preserving the boundary between the 2 distinct regions (fig. 1 g). Note that no tensor swelling effect occurs.

We have verified with these experiments that using our ML estimator corrects for the shrinking effect caused by the Rician noise, which makes our estimator suitable for clinical datasets with low SNR.

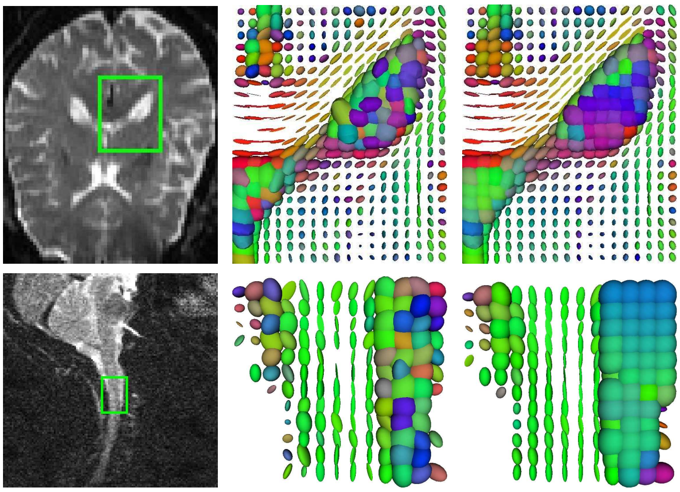

Fig. 2. Tensor field estimation of a brain (top row) and spinal cord (bottom row) datasets. Left: A slice of the b_0 image. **Middle:** The classic estimation of the green square region. Missing tensors in the splenium region are non-positive. **Right:** The ML + Reg estimation of the same region.

3.2. Clinical Data

We tested our ML method on 2 clinical datasets of medium and low quality. First, we used a brain dataset (Fig. 2 top) acquired with 7 encoding gradients (Basser sequence [1], b-value of 1000 s.mm^{-2}) on a 1.5T scanner. Second, we used an experimental acquisition of the spinal cord on a 1.5T scanner (Fig. 2 bottom) obtained with 25 encoding gradients and the same b-value as previously (acquisition is coronal). This new type of acquisition is currently actively investigated in clinical research.The parameters used for the processing are: $\lambda = 1.0, \kappa = 0.1$ for the brain dataset and $\lambda = 2.0, \kappa = 0.1$ for the spinal cord.

Figures 2 middle and right show a closeup of the splenium region and around the middle of the spinal cord. We clearly see that missing tensors of a classic estimation (Fig. 2 middle) are replaced by correct ellipsoids using our ML framework (fig. 2 right). The regularization perfectly smooths homogeneous regions like the ventricle or within the spinal cord without blurring the transitions with nearby structures like the splenium tract in the brain dataset. Moreover, one notices a slightly higher anisotropy when using the ML estimator but this effect should be further investigated. Finally, we observed an increase of about 30% of the tensor volume in the spinal cord dataset, and about 10% in the brain dataset, compared to the classic estimation.

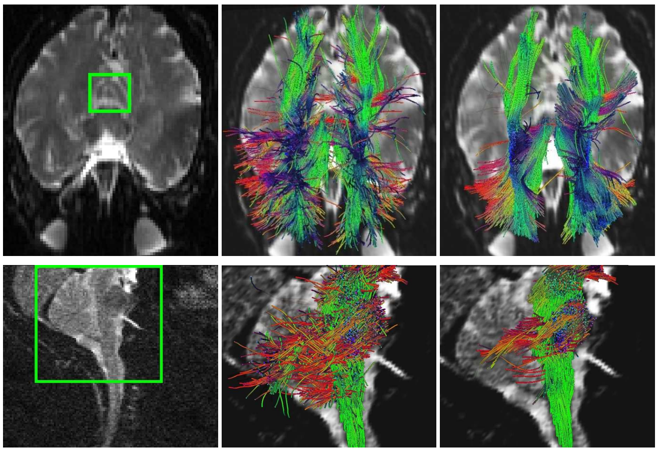

Fig. 3. **Improvement of fiber reconstruction. Top:** The corticospinal tract reconstructed after a classic estimation (middle) and our ML framework (right). **Bottom:** The spinal cord reconstruction after the same estimations. Left images show the region of interests where the tracking is initiated.

3.3. Improvement of Tractography

At the very end of the DTI processing pipeline resides the tractography. Among the numerous available methods for fiber tracking, we choose a relatively fast and easy to implement method [12] to exemplify how the tracking is improved by our joint estimation and regularization. Prior to the tracking, tensor fields are resampled so that voxels are isotropic (interpolation method is described in [5]). We tracked the fibers from the previous estimations. Results of tracking in the brain and the spinal cord are shown in Fig. 3. With our estimation, the tracking is qualitatively much smoother in both cases and shows less dispersion. The overall number of fibers reconstructed is also larger. The smoothness of the tensor field indeed leads to more regular and longer fibers: tracts that were stopped due to the noise are now fully reconstructed. However, a more quantitative analysis of the influence of the method on the tracking result would be necessary.

4. CONCLUSION

This paper presents a new methodology to process noisy DT-MRI typical of clinical applications. The estimation, which assumes for the first time a Rician noise, is achieved with a ML strategy. Solving this non-linear criterion requires an adapted tool to process tensors, and Log-Euclidean metrics are a perfect candidate. The ML estimator has the advantage to correct for the bias in the DWI images which causes tensors to shrink if estimated from the images directly. We have shown that adding an anisotropic regularization term to the estimation smooths homogeneous regions while preserving boundaries with fiber tracts. Finally, the promising improvement of the fiber reconstruction on 2 clinical datasets shows that even clinical DTI can be used for tractography.

In the future, the questions of validation and reproducibility need to be answered. We could think of repeating scans of the same patient in various orientations and scanners to estimate the reliability of our methodology. One also could think of using histological data and phantoms as in [13]. Finally, the impact on the tracking must be quantified. We are currently working on developing a dispersion measure of the fibers for that purpose.

5. REFERENCES

[1] P. Basser, J. Mattiello, and D. Le Bihan, “MR diffusion tensor spectroscopy and imaging,” *Biophysical Journal*, vol. 66, pp. 259–267, 1994.

[2] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz, and R. Kikinis, “Processing and visualization for diffusion tensor MRI,” *MedIA*, vol. 6(2), pp. 93–108, 2002.

[3] Z. Wang, B.C. Vemuri, Y. Chen, and T.H. Mareci, “A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI,” *IEEE TMI*, vol. 23, no. 8, August 2004.

[4] C. Chefd’hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, “Constrained flows of matrix-valued functions: Application to diffusion tensor regularization,” in *ECCV’02*, LNCS 2350, pp. 251–265.

[5] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, “Fast and simple calculus on tensors in the Log-Euclidean framework,” in *MICCAI’05*, LNCS.

[6] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, “Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor MRI,” Research Report 5242, INRIA, 2004.

[7] P. Fletcher and S. Joshi, “Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors.,” in *CVAMIA and MMBIA*, 2004, LNCS 3117, pp. 87–98.

[8] X. Pennec, P. Fillard, and N. Ayache, “A Riemannian framework for tensor computing,” *IJCV*, vol. 66, no. 1, Jan. 2006, Also as INRIA Research Report (RR) 5255.

[9] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, “Joint estimation and smoothing of clinical DT-MRI with a log-euclidean metric,” RR 5607, INRIA, 2005.

[10] J. Sijbers, A.J. den Dekker, P. Scheunders, and D. Van Dyck, “Maximum likelihood estimation of Rician distribution parameters,” *TMI*, vol. 17, no. 3, June 1998.

[11] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and minimization Algorithms*, Springer-Verlag, 1993.

[12] P. Fillard and G. Gerig, “Analysis tool for diffusion tensor MRI,” in *MICCAI’03 (2)*, LNCS 2878, pp. 967–968.

[13] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. Le Bihan, and J.-F. Mangin, “Fiber tracking in q-ball fields using regularized particle trajectories,” in *IPMI’05*, July, vol. 1 of *LNCS 3565*, pp. 52–63.