A Novel Gauss–Markov Random Field Approach for Regularization of Diffusion Tensor Maps

CHAPTER · JANUARY 2003

DOI: 10.1007/978-3-540-45210-2_46

4 AUTHORS, INCLUDING:

Marcos Martin-Fernandez
Universidad de Valladolid
137 PUBLICATIONS **568** CITATIONS
SEE PROFILE

Carl-Fredrik Westin
Harvard University
363 PUBLICATIONS **10,080** CITATIONS
SEE PROFILE

Carlos Alberola-López
Universidad de Valladolid
87 PUBLICATIONS **618** CITATIONS
SEE PROFILE

A Novel Gauss-Markov Random Field Approach for Regularization of Diffusion Tensor Maps

Marcos Martín-Fernández[1], Raul San José-Estépar[1,2], Carl-Fredrik Westin[2], and Carlos Alberola-López[1⋆]

[1] E.T.S. Ingenieros de Telecomunicación,
Universidad de Valladolid,
Cra. Cementerio s/n, 47011 Valladolid, SPAIN
{marcma, caralb}@tel.uva.es
rjosest@lpi.tel.uva.es
[2] Surgical Planning Laboratory,
Brighman and Women's Hospital and
Harvard Medical School,
75 Francis St., Boston, MA 02115, USA
{rjosest,westin}@bwh.harvard.edu

Abstract. In this paper we propose a novel Gaussian MRF approach for regularization of tensor fields for fiber tract enhancement. The model follows the Bayesian paradigm: prior and transition. Both models are given by Gaussian distributions. The prior and the posterior distributions are Gauss-MRFs. The prior MRF promotes local spatial interactions. The posterior MRF promotes that local spatial interactions which are compatible with the observed data. All the parameters of the model are estimated directly from the data. The regularized solution is given by means of the Simulated Annealing algorithm. Two measures of regularization are proposed for quantifying the results. A complete volume DR-MRI data have been processed with the current approach. Some results are presented by using some visually meaningful tensor representations and quantitatively assessed by the proposed measures of regularization.

1 Introduction

Diffusion Tensor (DT) Magnetic Resonance Imaging (MRI) is a volumetric imaging modality in which the quantity assigned to each voxel of the volume scanned is not a scalar, but a tensor that describes local water diffusion. Tensors have direct geometric interpretations, and this serves as a basis to characterize local structure in different tissues. The procedure by which tensors are obtained can be consulted elsewhere [9];. The result of such a process is, ideally speaking, a 3×3 symmetric positive-semidefinite (PSD) matrix [4].

Tensors support information of the underlying anisotropy within the data. As a matter of fact, several measures of such anisotropy have been proposed out of tensors to make things easier to interpret; see, for instance, [1,9]. However,

⋆ To whom correspondence should be addressed.

R. Moreno-Díaz and F. Pichler (Eds.): EUROCAST 2003, LNCS 2809, pp. 506–517, 2003.

these measures rely of the ideal behavior of the tensors, which may be in some cases far from reality due to the presence of noise in the imaging process itself. As was pointed out in [7], periodic beats of the cerebro-spinal fluid and partial volume effects may add a non-negligible amount of noise to the data, and the result is that the hypothesis of PSD may not be valid. Authors are aware of this fact, so some regularization procedures have been proposed in the past [5,6,7,9].

In this paper we focus on regularization of DT maps using Markov Random Fields (MRFs); other regularization philosophies exist (see, for instance, [8] and [10] and references therein) although they will not be discussed in the paper. About MRFs we are aware of the existence of other Markovian approaches to this problem [6,7], in which the method presented is called by the authors the *Spaghetti model*. These papers propose an interesting optimization model for data regularization. However, some issues could be a matter of discussion. The authors build their MRF on the basis of a neighborhood system that may change through the optimization process, a fact that is not theoretically correct, though acceptable practical results may be obtained. In addition, the transition model used by the authors does not seem to have a clear probabilistic interpretation, but, in our opinion, only a functional interpretation.

The work we are about to present is an extended and detailed version of the short paper [3]. The model is built upon a Bayesian philosophy in which the two terms, namely, the prior and the likelihood function, have a clear physical meaning. Closed-form expressions have been obtained for the posterior, so the resulting model has a solid probabilistic foundation and, in our opinion, mathematical elegance. The maximum *a posteriori* (MAP) estimator will be found by means of the Simulated Annealing algorithm [2].

The paper is structured as follows. In section 2 we provide some background information on MRFs. This section provides most of the terminology that will be used in the paper. In section 3 we present the Bayesian model upon which the regularization process will be carried out. Then, section 4 proposes two quantitative measures of roughness that will be used to evaluate the regularization achieved by the proposed method. In section 5 several results are presented and, finally, in section 6 we present some conclusions and several issues that are still unaddressed.

2 Background on MRFs

A MRF models a multidimensional random variable with local (generally spatial) interactions. The MRF is defined on a finite grid or lattice $\mathbf{S}$ of sites (pixels or voxels). For a field to be a MRF, two conditions have to be satisfied:

– Positivity property: all the configurations are always possible, i.e., the probability density function must satisfy:

$$p(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \Re^{|\mathbf{S}|}, \tag{1}$$

where $|\cdot|$ stands for the cardinality of its set argument.

– Markov property: the probability density function of one site conditioned to all the others sites must be equal to the probability density function of this site conditioned to its neighbors only, i.e.:

$$p(x_s/x_u, u \in \mathbf{S}, u \neq s) = p(x_s/x_u, u \in \delta(s)), \tag{2}$$

where $\delta(s)$ is a predefined fixed neighborhood for site s. Neighborhood systems must satisfy:
- $t \in \delta(s) \Longleftrightarrow s \in \delta(t)$, and
- $s \notin \delta(s)$.

The conditional probability function $p(x_s/x_u, u \in \delta(s))$ is called the local characteristic of the field.

A multivariate Gaussian distribution defined over a lattice $\mathbf{S}$ with local spatial interactions is a MRF for the defined neighborhood system. This MRF so defined is called Gauss-MRF or autonormal distribution.

Typically the cardinality of $\mathbf{S}$ is too great so sampling the field directly is not feasible. The same problem occurs when the MAP estimation is sought; a *brute-force* search for the field mode is also unfeasible. Two algorithms, the convergence properties of which have been theoretically demonstrated, are used to solve each of the two problem, namely, the Gibbs Sampler algorithm, which provides field realizations, and the Simulated Annealing algorithm which provides field maximizers [2].

3 The Model

The model presented in this paper operates on each slice of the volume data; it will be applied independently to each component of the tensor field. We have only considered the 2D case, so the third row and the third column of each tensor have been discarded. However, as will be soon clear, the model here proposed carries over trivially to the 3D case.

A Bayesian framework consists of two ingredients: the prior and the transition (or the likelihood) models; these are connected in the posterior, the third probability model involved, by means of the Bayes theorem. The solution is some sort of parameter on the posterior (the mean, the median or the mode). In our case, we have resorted to a MAP solution, i.e., to finding the mode of the posterior.

Thus we have divided the exposition of the model in three subsections, as follows.

3.1 The Prior

Let $\mathbf{x} \in \Re^{|\mathbf{S}|}$ denote a (non-observable) realization of the field consisting of one of the tensor components for every pixel; if the slice consists of M rows and N columns, vector $\mathbf{x}$ is a column vector with $M \times N$ components. The prior captures the knowledge about the field in the absence of any data. For smoothness to be

guaranteed, a reasonable assumption is to consider a Gauss-MRF or autonormal model:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{|\mathbf{S}|}\det(\boldsymbol{\chi})}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\kappa})^T\boldsymbol{\chi}^{-1}(\mathbf{x}-\boldsymbol{\kappa})\right], \tag{3}$$

with $\boldsymbol{\kappa}$ and $\boldsymbol{\chi}$ the mean vector and covariance matrix respectively. The structure of latter must obey the MRF property. As previously stated, sampling this distribution directly is not a sensible choice. So, an alternative is to find the local characteristics which, in this case, turns out to be a one dimensional Gaussian distribution at site $s \in \mathbf{S}$ given by:

$$p(x_s/x_u, u \in \delta(s)) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{(x_s-\eta_s)^2}{2\sigma_s^2}\right], \tag{4}$$

for a given homogeneous neighborhood system $\delta(s)$ for each site $s \in \mathbf{S}$. The local characteristic parameters are the local mean η_s and the local variance σ_s^2, which are functions of the neighboring sites $\delta(s)$ exclusively. This dependence will not be explicitly written to ease the notation. This prior parameters are estimated from those neighboring sites by means of the maximum likelihood (ML) method, which in the Gaussian case gives the common estimators:

$$\eta_s = \frac{1}{|\delta(s)|}\sum_{u\in\delta(s)} x_u \qquad \text{and} \qquad \sigma_s^2 = \frac{1}{|\delta(s)|}\sum_{u\in\delta(s)}(x_u-\eta_s)^2. \tag{5}$$

3.2 The Transition Model

In the transition model (or the noise model) the observed field $\mathbf{y} \in \Re^{|\mathbf{S}|}$ is given by:

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{6}$$

where $\mathbf{n} \in \Re^{|\mathbf{S}|}$ is the noise field. The observed field is assumed to be a Gaussian field whose mean is equal to the non-observable field, thus, the probability density function of the observed field $\mathbf{y}$ conditioned to $\mathbf{x}$ is:

$$p(\mathbf{y}/\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{|\mathbf{S}|}\det(\boldsymbol{\Upsilon})}} \exp\left[-\frac{1}{2}(\mathbf{y}-\mathbf{x})^T\boldsymbol{\Upsilon}^{-1}(\mathbf{y}-\mathbf{x})\right], \tag{7}$$

with $\boldsymbol{\Upsilon}$ is the noise covariance matrix. This matrix is commonly accepted to be proportional to the identity matrix, so the noise field is conditionally white and stationary. Therefore, we can write for each pixel:

$$p(y_s/x_s, x_u, u \in \delta(s)) = p(y_s/x_s) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{(y_s-x_s)^2}{2\sigma_n^2}\right]. \tag{8}$$

The noise variance has to be estimated from the observed noisy data $\mathbf{y}$. Two possible solutions can be considered:

– The minimum of the local ML estimated variances:

$$\sigma_1^2 = \min_{s\in\mathbf{S}} \left[\frac{1}{|\delta(s)|} \sum_{u\in\delta(s)} \left(y_u - \frac{1}{|\delta(s)|} \sum_{t\in\delta(s)} y_t \right)^2 \right]. \tag{9}$$

– The mean of the local ML estimated variances:

$$\sigma_2^2 = \frac{1}{|\mathbf{S}|} \sum_{s\in\mathbf{S}} \left[\frac{1}{|\delta(s)|} \sum_{u\in\delta(s)} \left(y_u - \frac{1}{|\delta(s)|} \sum_{t\in\delta(s)} y_t \right)^2 \right]. \tag{10}$$

We have observed that the minimum σ_1^2 tends to underestimate the value, while the mean σ_1^2 overestimates it; a value in between have been considered a good choice.

3.3 The Posterior Model

Bayes theorem allows us to write:

$$p(\mathbf{x}/\mathbf{y}) = \frac{p(\mathbf{y}/\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}, \tag{11}$$

where we know that:

$$p(\mathbf{y}) = \int_{\Re^{|\mathbf{S}|}} p(\mathbf{y}/\mathbf{x})p(\mathbf{x})d\mathbf{x}. \tag{12}$$

The distribution $p(\mathbf{x}/\mathbf{y})$ is known to be a Gauss-MRF with joint density function:

$$p(\mathbf{x}/\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^{|\mathbf{S}|}\det(\boldsymbol{\Lambda}_\mathbf{y})}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\xi}_\mathbf{y})^T \boldsymbol{\Lambda}_\mathbf{y}^{-1}(\mathbf{x}-\boldsymbol{\xi}_\mathbf{y}) \right], \tag{13}$$

with $\boldsymbol{\xi}_\mathbf{y}$ and $\boldsymbol{\Lambda}_\mathbf{y}$ the mean and the covariance matrix of the posterior, respectively. The dependence on the observed field is explicitly noted.

The MAP gives an estimation of the non-observable regularized field $\mathbf{x}$ given the noisy observation $\mathbf{y}$. As mentioned in the section 2, the Simulated Annealing algorithm [2] gives a practical procedure to find the MAP; to that end, the only piece of information that we need is the posterior local characteristic. This can be easily obtained by:

$$p(x_s/y_s, x_u, u\in\delta(s)) = \frac{p(y_s/x_s, x_u, u\in\delta(s))p(x_s/x_u, u\in\delta(s))}{p(y_s/x_u, u\in\delta(s))}, \tag{14}$$

where we know that:

$$p(y_s/x_u, u\in\delta(s)) = \int_{\Re} p(y_s/x_s, x_u, u\in\delta(s))p(x_s/x_u, u\in\delta(s))dx_s. \tag{15}$$

Since both the prior and the likelihood are Gaussian, so is the posterior, thus the posterior local characteristic at site $s \in \mathbf{S}$ is:

$$p(x_s/y_s, x_u, u \in \delta(s)) = \frac{1}{\sqrt{2\pi}\rho_s} \exp\left[-\frac{(x_s - \mu_s)^2}{2\rho_s^2}\right], \tag{16}$$

where μ_s and ρ_s^2 are the posterior local mean and variance, which are given, after some algebra, by:

$$\mu_s = \frac{\sigma_s^2 y_s + \sigma_n^2 \eta_s}{\sigma_s^2 + \sigma_n^2} \qquad \text{and} \qquad \rho_s^2 = \frac{\sigma_s^2 \sigma_n^2}{\sigma_s^2 + \sigma_n^2}. \tag{17}$$

Finally, to implement the Simulated Annealing algorithm a partially parallel visiting scheme and a logarithmic cooling schedule for the temperature T are adopted, thus fixing $\rho_s^2(T) = T\rho_s^2$ (the implementation details of this algorithm can be consulted elsewhere [2]). The Simulated Annealing algorithm is applied in parallel to the three distinguishable components of the symmetric matrix tensor. After the algorithm visits each site simultaneously for the three components, the PSD condition is tested. If this test is passed the algorithm proceeds with the next site, otherwise, the sample is discarded and generated again. In practice, this event rarely occurs.

4 Regularization Measures

Two measures of regularization based on local roughness have been proposed to quantify the results achieved:

- The Frobenius norm of tensor differences, defined by:

$$R_f = \sum_{s \in \mathbf{S}} \sum_{u \in \delta'(s)} ||\mathbf{A}_s - \mathbf{A}_u||, \tag{18}$$

  where $\mathbf{A}_s$ is the tensor matrix for site s and for the neighborhood system $\delta'(s)$ which is different from the model neighborhood system $\delta(s)$ so that regularization results are measured with neighbors not involved in the regularization process.
- The entropy of discretized tensor linear component, defined by:

$$R_e = -\sum_{m=1}^{M} \sum_{n=1}^{N} p_{mn} \log_2 p_{mn}, \tag{19}$$

  where the joint probabilities p_{mn} are to be estimated from data. They are defined by:

$$p_{mn} = p(\Gamma = \Gamma_m, \Theta = \Theta_n), \tag{20}$$

  where Γ_m and Θ_n are the M and N discrete possible values for the random variables Γ and Θ. This variables are defined as:

$$\Gamma = \begin{cases} \frac{c_s}{c_u} & c_s \leq c_u \\ \\ \frac{c_u}{c_s} & c_s > c_u \end{cases} \qquad \text{and} \qquad \Theta = \angle \mathbf{v}_s - \angle \mathbf{v}_u, \tag{21}$$

where the $u \in \delta'(s)$ are any pair of neighboring sites. $\mathbf{v}_s$ is the main eigenvector of the tensor matrix $\mathbf{A}_s$ for site s. c_s represents the linear component of the tensor defined as [9]:

$$c_s = \frac{\lambda_{\max_s}}{\lambda_{\max_s} - \lambda_{\min_s}}, \tag{22}$$

with $\lambda_{\max_s}$ and $\lambda_{\min_s}$ the maximum and minimum eigenvalues of the tensor matrix $\mathbf{A}_s$. The Γ random variable ranges in the interval $(0, 1)$ while the Θ random variable in the interval $(0, \pi)$.

5 Some Experimental Results

We have processed a DT-MRI volumetric data of size $185 \times 70 \times 20 \times 2 \times 2$, i.e., twenty sections of size 185×70. The model neighborhood system $\delta(s)$ is fixed to the 12 closer sites surrounding s and the neighborhood system $\delta'(s)$ is set to the difference set between the 20 closer sites and the 12 closer sites surrounding s for the quantitative measures of roughness. For the noise variance we have chosen three values in the interval (σ_1^2, σ_2^2) by fixing:

$$\sigma_n^2 = k(\sigma_2^2 - \sigma_1^2) + \sigma_1^2, \tag{23}$$

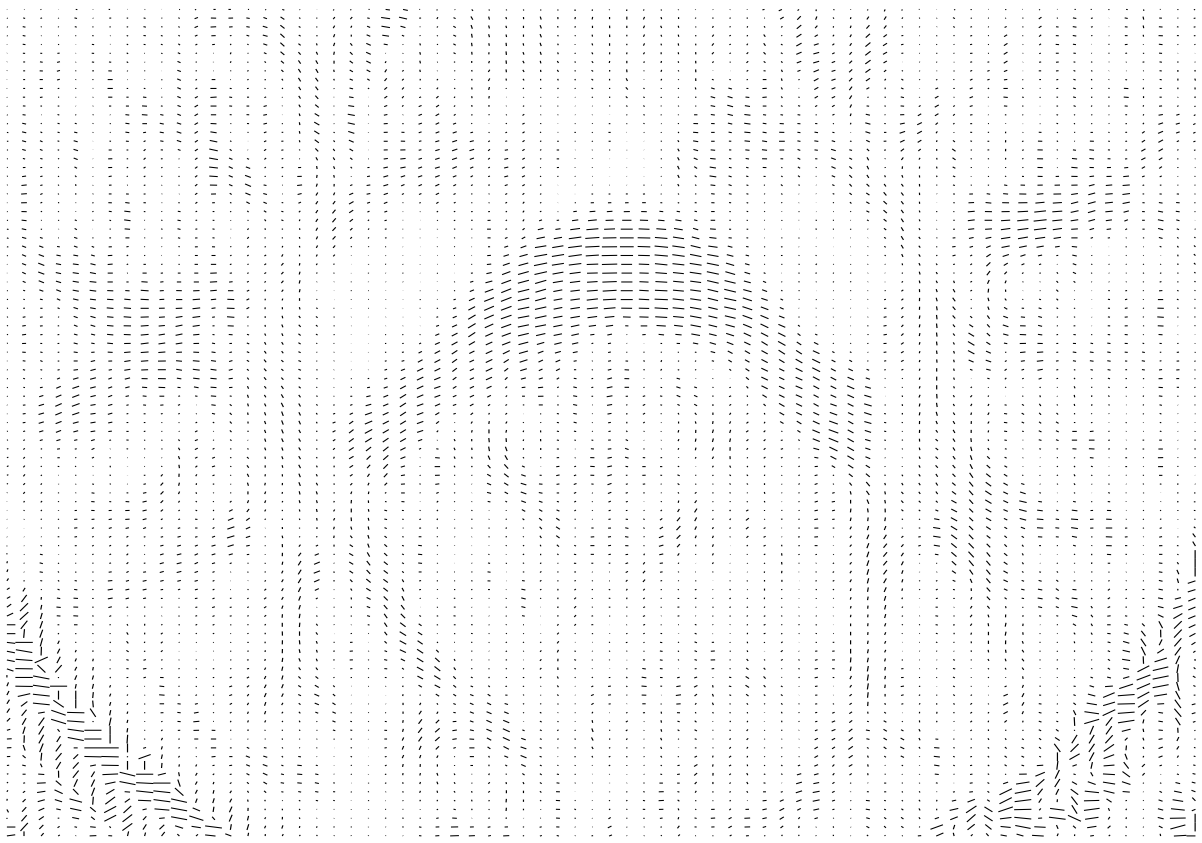
with $k \in (0, 1)$, yielding:

1. Low variance and thus low regularization, with $k = 0.25$.
2. Moderate variance and regularization, with $k = 0.5$.
3. High variance and regularization, with $k = 0.75$.

Figure 1(a) shows a detail of the original tensor map for slice number 20. Every vector is shown as a segment at site s, the angle and the length of which are given by the steering of the main eigenvector $\mathbf{v}_s$ and linear component c_s, respectively. In figure 1(b) a detail of the regularized tensor map is shown for the moderate regularization. It can be seen that the background noise is removed and the regions are more homogeneous without blurring the boundaries.

In figure 2(a) the Frobenius norm of the original tensor matrices for slice 20 is shown as an image in which the intensity of pixel s is given by $||\mathbf{A}_s||$. In figure 2(b) the same image representation is given for the moderate regularization. Comparing both images, the regularization achieved is perhaps now clearer than in figures 1(a) and 1(b). This leads us to propose this kind of image representation as a good procedure for visually assessing regularization results.

In figure 2(c) another representation is also proposed for the original tensor data for slice 20. In this case the image is a color coding of the image in figure 1(a) for which the RGB coordinated for site s are given by $R_s = \gamma c_s R(\angle \mathbf{v}_s)$, $G_s = \gamma c_s G(\angle \mathbf{v}_s)$ and $B_s = \gamma c_s B(\angle \mathbf{v}_s)$, with γ a gain factor (currently set to 2) and periodic functions $R(\cdot)$, $G(\cdot)$ and $B(\cdot)$ with period π so that the mapping between colors and angles is one-to-one. With this coding, the darker the pixel, the lower the linear component. In figure 2(d) the same color-coded



(a)

(b)

Fig. 1. (a) Original tensor map for slice number 20 (detail). (b) Regularized tensor map for slice number 20 (detail).

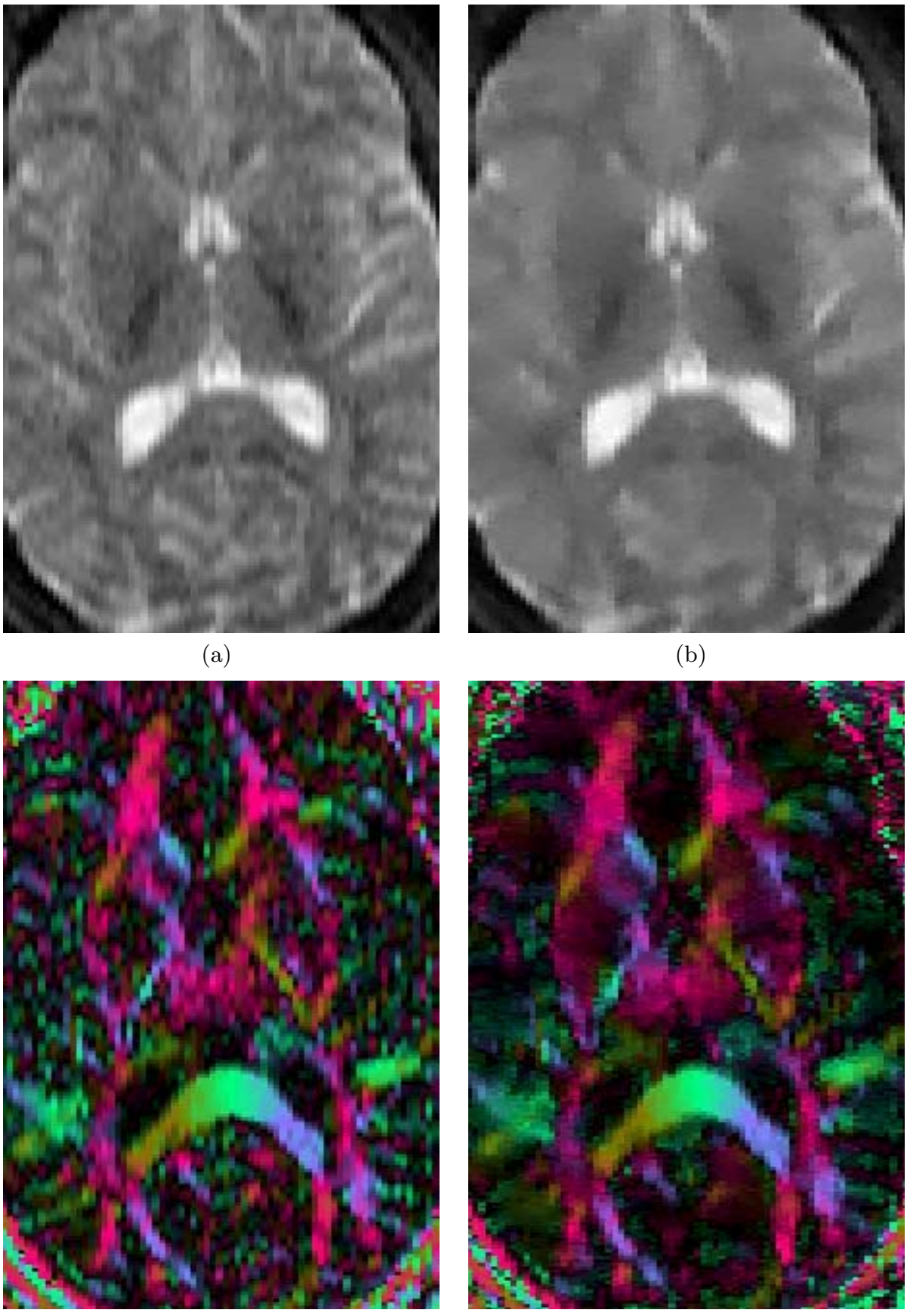
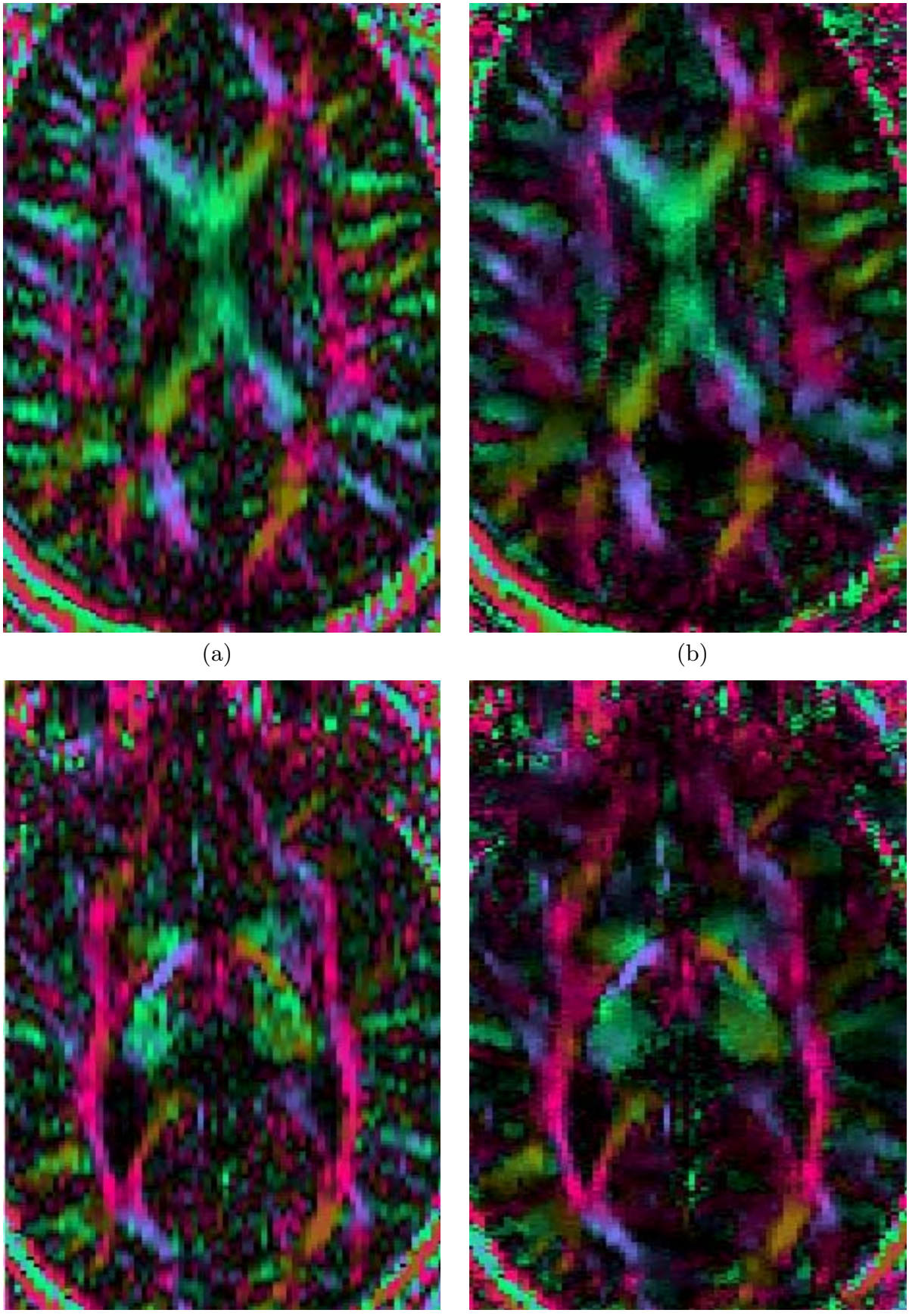


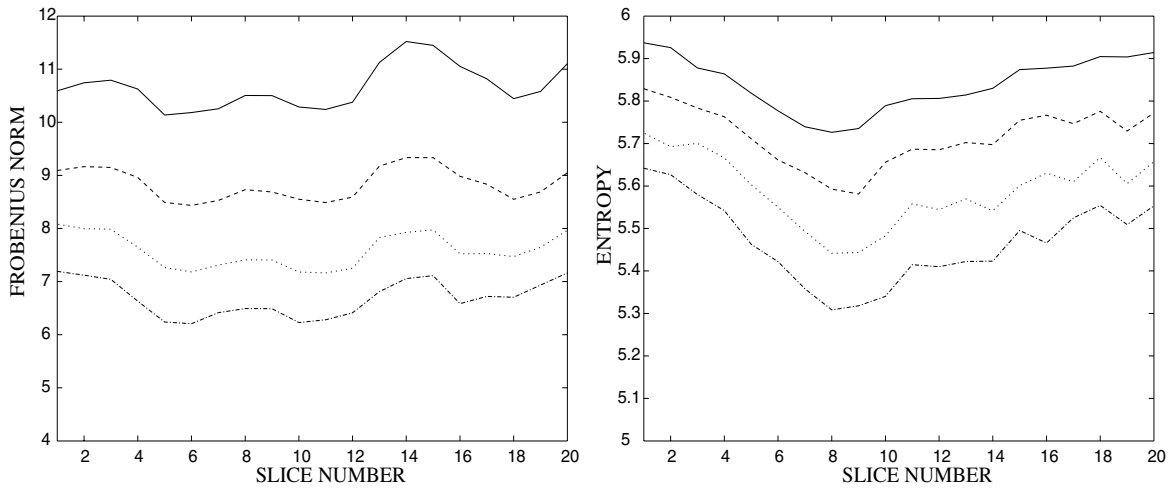
Fig. 2. For slice number 20: (a) original Frobenius norm image, (b) regularized Frobenius norm image, (c) original color-coded image, and (d) regularized color-coded image.



(a) (b)

(c) (d)

Fig. 3. For slice number 7: (a) original color-coded image, and (d) regularized color-coded image. For slice number 12: (c) original color-coded image, and (d) regularized color-coded image.



(a) (b)

Fig. 4. Quantitative measures of regularization for the twenty slices with several noise variances: (a) Frobenius norm of tensor differences and (b) entropy.

image representation is given for the moderate regularization applied on slice 20. In this case the same comments as before apply.

The images shown in figures 3(a)-(b) and in figures 3(c)-(d) are analogous to the ones shown in figures 2(c)-(d) but in this case they represent the original and the moderate regularized color-coded images for slices 7 and 12, respectively.

We have calculated the two quantitative regularization measures proposed in section 4 for the 20 slices and for the three degrees of regularization. The results achieved are shown in figures 4(a)-(b). Figure 4(a) plots the Frobenius norm of tensor differences, the R_f value, as a function of the number of slice and parameterized by the degree of regularization. The lower curve is for high, the second for moderate and the third for low regularization respectively. The upper curve is for the original data. In figure 4(b) the entropy, the R_e value, is also plotted versus the slice number and parameterized by the degree of regularization with the same sorting scheme. The results obtained show clearly in a quantitative manner the degree of regularization which can be achieved with the presented approach.

6 Conclusions

In this paper we have described a novel probabilistic Bayesian model for the regularization of DT maps. The proposed model uses Gauss-MRFs for modeling the spatial relation of each component of the tensors along the field. All the model parameters are estimated directly from the data. A volumetric data set of twenty slices has been regularized by using the proposed method giving good results, the quality of which have been both visually assessed and quantified by the proposed measures of roughness. Also, color-coded images have been introduced to visualize the linear component of the tensors. Finally, the Frobenius norm

images are an alternative procedure to visually check the level of regularization that is achieved.

Some issues are still unaddressed:

- Exploiting the existing correlation between tensor elements.
- Studying the sensitivity with respect to the size of the neighborhood, both for the model and for the measures of regularization.
- Developing a procedure to guarantee the PSD condition without discarding samples in the optimization process.

Acknowledgments. The authors acknowledge the Comisión Interministerial de Ciencia y Tecnología for research grant TIC2001-3808-C02, the Junta de Castilla y León for research grant VA91/01, NIH grant P41-RR13218 and CIMIT.

References

1. P. Basser, C. Pierpaoli, Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI, *Journal of Magnetic Resonance*, Ser. B, Vol. 111, No. 3, June 1996, pp. 209–219.
2. S. Geman, D. Geman, Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, Nov. 1984, pp. 721–741.
3. M. Martín-Fernández, R. San José Estépar, C.F. Westin, C. Alberola-López, A Novel Gauss-Markov Random Field Approach for Regularization of Diffusion Tensor Maps, *Proc. of the NeuroImaging Workshop, Eurocast 2003*, Las Palmas de Gran Canaria, Feb. 2003, pp. 29–32.
4. G.H. Golub, C.F. Van Loan, Matrix Computations, *John Hopkins University Press*, Baltimore, Maryland, USA, 1996.
5. G.J.M. Parker, J.A. Schnabel, M.R. Symms, D.J. Werring, G.J. Barker, Nonlinear Smoothing for Reduction of Systematic and Random Errors in Diffusion Tensor Imaging, *Journal of Magnetic Resonance Imaging*, Vol. 11, No. 6, 2000, pp. 702–710.
6. C. Poupon, J.F. Mangin, V. Frouin, J. Regis, F. Poupon, M. Pachot-Clouard, D. Le Bihan, I. Bloch, Regularization of MR Diffusion Tensor Maps for Tracking Brain White Matter Bundles, in *Lecture Notes in Computer Science*, W. M. Wells, A. Colchester, S. Delp, Eds., Vol. 1946, Oct. 1998, pp. 489–498.
7. C. Poupon, C.A. Clark, V. Frouin, J. Regis, D. Le Bihan, I. Bloch, J.F. Mangin, Regularization Diffusion-Based Direction Maps for the Tracking of Brain White Matter Fascicles, *NeuroImage*, Vol. 12, 2000, pp. 184–195.
8. D. Tschumperlé, R. Deriche, DT-MRI Images: Estimation, Regularization and Application, *Proc. of the NeuroImaging Workshop, Eurocast 2003*, Las Palmas de Gran Canaria, Feb. 2003, pp. 46–47.
9. C.F. Westin, S.E. Maier, H. Mamata, A. Nabavi, F.A. Jolesz, R. Kikinis, Processing and Visualization for Diffusion Tensor MRI, *Medical Image Analysis*, Vol. 6, No. 2, June 2002. pp. 93–108.
10. C.F. Westin, H. Knuttson, Tensor Field Regularization using Normalized Convolution, *Proc. of the NeuroImaging Workshop, Eurocast 2003*, Las Palmas de Gran Canaria, Feb. 2003, pp. 67–70.