A multi-resolution framework for diffusion tensor images

Florack, L.M.J.; Astola, L.J.

Published: 01/01/2008

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):
A Multi-Resolution Framework for Diffusion Tensor Images

Luc Florack
Department of Mathematics and Computer Science & Department of Biomedical Engineering
Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

Laura Astola
Department of Mathematics and Computer Science
Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

{L.M.J.Florack,L.J.Astola}@tue.nl

Abstract

A new scale space paradigm is proposed for multi-resolution analysis of diffusion tensor images (DTI). An a priori consistency requirement is stipulated, which precludes a linear model. A nonlinear adaptation is proposed to remedy the problem. Subsequently it is shown how differentiation can be operationalized.

Considerations in this paper are relevant for DTI analysis in a differential geometric framework, in which the DTI image imposes a Riemannian structure. As such it adds further support in favor of the “geometric rationale”, opening the door for a multi-resolution geometric approach towards fibre tracking, connectivity analysis, and so forth.

Keywords. Scale space, diffusion tensor imaging, Riemannian geometry, log-Euclidean framework, differentiation.

1. Introduction

Recent literature advocates differential geometry\(^1\) as a powerful mathematical framework for analyzing diffusion tensor images (DTI) [2, 8, 23, 34, 36]. However, operationalization of differential concepts is hampered by the ill-posed nature of classical differential operators. This has led to the introduction of regularization techniques [30, 34, 40], not only in DTI, but in image analysis in general.

An intrinsic coupling of regularization and differentiation is achieved by treating images as tempered distributions [37]. In the framework of distribution theory the act of differentiation (including zeroth order) requires a class of smooth test functions (linear filters), by virtue of which it becomes well-posed. In scale space theory the class of test functions is axiomatically restricted so as to arrive at an operational definition of an image derivative [7, 9, 10, 15, 17, 18, 19, 20, 21, 22, 24, 31, 38, 44, 45]. The paradigmatic (zeroth order) filter is the normalized Gaussian of a priori arbitrary scale. This naturally produces a scale space representation, i.e. a continuous family of images intended to capture raw image data at arbitrary levels of resolution (within physically reasonable limits). This holds a fortiori for any image derivative.

Linear scale space theory provides a natural framework for image differentiation in the absence of constraints that are incompatible with linearity. An instance of incompatibility is encountered in image processing by nonlinear diffusion [16, 35, 42, 43], in which the desire to “preserve edges” precludes linearity. Another instance in which linearity conflicts with fundamental demands—and the subject of this paper—is encountered in the context of DTI.

A DTI sample is a symmetric positive definite matrix (more precisely, a contravariant 2-tensor [26, 39]). In the differential geometric rationale its inverse defines the components of the Riemannian metric tensor. The first order structure of the DTI image then induces a unique, so-called affine connection, which is the geometric construct for “fitting together” neighbouring tangent spaces, allowing e.g. fibre tracking via parallel transport (geodesics). In turn, the first order structure of this connection determines the Riemann curvature tensor\(^2\), which can be seen as a measure of geodesic deviation (inclination of neighbouring fibres to

---

\(^1\) Geometric considerations have been applied to domain and codomain of a DTI image. Our focus is the former but on the fly we obtain a natural connection to the latter.

\(^2\) In 3D the Riemann tensor is equivalent to the so-called Ricci tensor.
We shall require $F$ in analogy with the relative acceleration experienced by freely falling particles in an inhomogeneous gravitational field [26].

In view of the above, we propose a multi-resolution framework for DTI images, so as to admit well-posed differentiation at any level of resolution. The space of symmetric positive definite matrices is closed under inversion. A DTI image and its pointwise inverse are dual metrics, and their matrix representations belong to the same class of positive definite symmetric matrix fields. A scale space generator for this class must therefore be such that it manifestly preserves inverse relationships at all scales. In other words, the acts of blurring and inversion of matrix fields should commute. One easily verifies that this a priori commutativity requirement is inconsistent with linear blurring.

After establishing some notation in the next section we address the problem of how to set up a consistent scale space paradigm for DTI in Section 3. In Section 4 we operationalize differential structure, and we conclude in Section 5.

2. Notation

We denote a DTI-image by $f : \mathbb{R}^n \to S_n^+$, where $S_n^+ \subset S_n \subset M_n$ denotes the set of $R$-valued symmetric positive definite $n \times n$ matrices, $S_n$ the set of $R$-valued symmetric $n \times n$ matrices, and $M_n$ the set of all $n \times n$ matrices. Its pointwise inverse is $f^{-1} : \mathbb{R}^n \to S_n^+$, so that $(f^{-1} f)(x) = (f f^{-1})(x) = I$, the identity matrix, at each point $x \in \mathbb{R}^n$.

$C^\omega(\mathbb{R}^n, M_n)$ denotes the class of analytical functions $f : \mathbb{R}^n \to M_n$. Self-explanatory definitions hold for $C^\omega(\mathbb{R}^n, S_n^+) \subset C^\omega(\mathbb{R}^n, S_n) \subset C^\omega(\mathbb{R}^n, M_n)$.

The isotropic Gaussian scale space kernel in $n$ dimensions is given by

$$
\phi_\sigma(x) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left(-\frac{1}{2} \frac{\|x\|^2}{\sigma^2}\right) .
$$

The scale space representation of $f \in C^\omega(\mathbb{R}^n, S_n^+)$ is obtained by the blurring operator (detailed in the next section)

$$
\mathcal{F} : C^\omega(\mathbb{R}^n, S_n^+) \times \mathbb{R}^+ \to C^\omega(\mathbb{R}^n, S_n^+) : (f, \sigma) \mapsto \mathcal{F}(f, \sigma) .
$$

We shall require $\mathcal{F}(f, 0) = f$ for all $f \in C^\omega(\mathbb{R}^n, S_n^+)$, and sometimes use the shorthand notation $f_\sigma \equiv \mathcal{F}(f, \sigma)$.  

3. The Basic Paradigm

3.1. Formulation of Consistency

Our goal is to formulate a consistent scale space representation for symmetric positive definite matrix fields $f \in C^\omega(\mathbb{R}^n, S_n^+)$. Since inversion is automorphic on $C^\omega(\mathbb{R}^n, S_n^+)$, there is no mathematical justification for treating $f$ and $f^{-1}$ differently. Thus consistency means that if $f_\sigma = \mathcal{F}(f, \sigma)$, then for each $\sigma \in \mathbb{R}^+$ we must have $f_\sigma^{-1} = \mathcal{F}(f^{-1}, \sigma)$, i.e. blurring and inversion must commute. Note that commutativity would not hold if we would define $\mathcal{F}(f, \sigma) = f * \phi_\sigma$, since $f^{-1} * \phi_\sigma \neq (f * \phi_\sigma)^{-1}$ for $\sigma > 0$. Consistency requires the scale space generator $\mathcal{F}$ to be nonlinear. In the spirit of previous work [4, 11]—applicable to scalar images—we will look for a pseudo-linear scale space representation with the desired commutativity property. We start with an intermezzo on relevant results in the context of matrix-valued functions.

3.2. Matrix Exponential and Logarithm

The exponential map $\exp : M_n \to GL_n$ maps an arbitrary $n \times n$ matrix to a nonsingular matrix, i.e. an element of the general linear group [12, 13, 27]. For later convenience we define $M_n^+ = \exp(M_n) \subset GL_n$. For our purpose it suffices to consider diagonalizable matrices. In fact, we need only consider elements of $S_n \subset M_n$, which are diagonalizable with real eigenvalues, in which case the range of the exponential map equals $\exp(S_n) = S_n^+$. So we will employ the prototype

$$
\exp : S_n \to S_n^+ : A \mapsto \exp A .
$$

An operational representation of a general analytical matrix function is given by Sylvester’s formula:

$$
F(A) \equiv \sum_{i=1}^m F(\lambda_i) A_i ,
$$

in which the $\lambda_i$, $i = 1, \ldots, m$, are all distinct eigenvalues of $A$. In Eq. (4) the left hand side—with intentional abuse of notation—is defined by virtue of the analytical scalar function $F \in C^\omega(\mathbb{R}, \mathbb{R})$ on the right hand side, i.e. $F \equiv \exp$, and the so-called Frobenius covariants are given by

$$
A_i = \prod_{j=1, j \neq i}^m \frac{1}{\lambda_i - \lambda_j} (A - \lambda_j I) .
$$

In practice $n = 3$. For the sake of illustration we also consider $n = 2$.

Generically one expects $m = n$ a.e. within the DTI image domain.
The logarithmic map, restricted to $S^+_n$ for our purposes, has prototype
\[ \ln : S^+_n \rightarrow S_n : B \mapsto \ln B, \]
and is the unique inverse of the exponential map restricted to $S_n$: \( \ln(S^+_n) = S_n \). In this case we may apply Eqs. (4–5) with $F \equiv \ln$.

### 3.3. Multi-Scale Representation

We shall be primarily interested in a second order multi-scale representation. This suffices for nearly all applications that exploit differential geometry, such as the computation of Christoffel symbols [23, 34, 36] and the Riemann curvature tensor [2] compatible with the metric. Generalization to higher orders is straightforward.

Let us start by considering the basic zeroth order scale space paradigm, accounting for commutativity of blurring and inversion. Subsequently we turn to first and second order differential structure of the established representation.

\[
\begin{align*}
g_\sigma &= \exp \left( \ln g * \phi_\sigma \right) \overset{\text{inv}}{\longrightarrow} h_\sigma = \exp \left( \ln h * \phi_\sigma \right) \\
\ln g * \phi_\sigma &\quad \overset{\exp}{\longleftarrow} \ln h * \phi_\sigma \\
\ln g &\quad \overset{\exp}{\longleftarrow} \ln h \\
g &\quad \overset{\text{inv}}{\longleftarrow} h
\end{align*}
\]

Figure 1. Commuting diagram for blurring and inversion.

Fig. 1 shows how a consistent multi-scale representation can be obtained. Indeed, if $f \in C^\omega([R, M_n])$, then $f_\sigma = F(f, \sigma)$ constructed according to
\[ F(f, \sigma) = \exp \left( \ln f * \phi_\sigma \right), \tag{7} \]
satisfies the desired commutativity property,
\[ F(f^m, \sigma) = F(f, \sigma)^m. \tag{8} \]
This follows immediately by inspection of Fig. 1 and Eq. (7), using the identities
\[ \exp(-A) = (\exp A)^m \quad \text{and} \quad \ln B^m = -\ln B, \tag{9} \]
for $A \in S_n$, $B \in S^+_n$. The zeroth order scheme, Eq. (7), is in line with the so-called log-Euclidean framework proposed by Arsigny et al., Pennec et al., and Fillard et al. [1, 8, 34], but is here based on a different motivation, viz. the fundamental commutativity axiom. In this sense the zeroth order result comes out with no surprise. However, the differential geometric rationale also compels us to consider derivatives of this zeroth order multi-scale representation. In the next section we show that this is not quite trivial, but that closed-form expressions can be obtained in this respect as well.

### 4. Differentiation

A complication arises when differentiating the zeroth order representation, Eq. (7), due to the fact that matrices do not in general commute. In particular this complicates the chain rule. To appreciate this, we recall a few nontrivial results on derivatives of a matrix exponential function. (These results are formulated for arbitrary, sufficiently regular $n \times n$ matrix fields [3, 5, 13, 28, 29, 32].)

**Definition 1** Let $X \in C^\omega([R, M_n])$. The parametric derivative of the matrix exponential function $\exp X \in C^\omega([R, M_n^+])$ is defined as
\[ \frac{d}{dt} \exp \left( X(t) + hX(t) \right) = \frac{\exp \left( X(t) + hX(t) \right) - \exp \left( X(t) \right)}{h}. \]

**Theorem 1** The parametric derivative of a one-parameter matrix-valued function $X \in C^\omega([R, M_n])$ can be operationalized as follows:
\[ \frac{d}{dt} \exp \left( X(t) \right) = \int_0^1 \exp \left( (1-\alpha)X(t) \right) \frac{dX(t)}{dt} \exp \left( \alpha X(t) \right) d\alpha. \]

In order to prove this the following lemma turns out useful, the proof of which follows immediately by substitution.

**Lemma 1** The unique solution of the inhomogeneous matrix-ODE
\[ \frac{dY}{d\lambda} = XY + F \quad \text{with initial condition } Y(0) = Y_0, \]
with $X \in M_n$ constant, and $F \in C^\omega([R, M_n])$, is given by
\[ Y(\lambda) = \exp (\lambda X) \left( Y_0 + \int_0^\lambda \exp (-\mu X) F(\mu) d\mu \right). \]

**Proof of Theorem 1.** Consider the following matrix-ODE, with $X, V \in M_n$ constant, and $h \in R$:
\[ \frac{dY}{d\lambda} = (X + hV) Y \quad \text{with initial condition } Y(0) = I. \]
The solution—obtained by identifying the term $hVY$ with the inhomogeneous term $F$ in Lemma 1, with $Y_0 = I$—can be represented in the form of a Volterra integral equation:

$$Y(\lambda) = \exp(\lambda X) + h \int_0^\lambda \exp((\lambda - \alpha)X) V Y(\alpha) \, d\alpha.$$ 

This gives rise to an $O(h)$ approximation

$$Y(\lambda) = \exp(\lambda X) + h \int_0^\lambda \exp((\lambda - \alpha)X) V \exp(\alpha X) \, d\alpha + O(h^2).$$

On the other hand we also have the exact solution,

$$Y(\lambda) = \exp(\lambda (X + hV)).$$

Identifying the last two expressions, putting $\lambda = 1$, yields

$$\exp(X + hV) - \exp(X) = h \int_0^1 \exp((1 - \alpha)X) V \exp(\alpha X) \, d\alpha + O(h^2),$$

from which the proof readily follows. \qed

We define the gradient of a multivariate matrix-valued function $X \in C^\infty(\mathbb{R}^k, M_{n \times n})$, $\nabla \exp(X)$, in a similar fashion. We may generalize Theorem 1 accordingly.

**Theorem 2** Let $X \in C^\infty(\mathbb{R}^k, M_{n \times n})$. The gradient of the matrix exponential function $\exp X \in C^\infty(\mathbb{R}^k, M_{n \times n})$ can be represented component-wise as follows:

$$\partial_\nu \exp(X(x)) = \int_0^1 \exp((1 - \alpha)X(x)) \partial_\nu X(x) \exp(\alpha X(x)) \, d\alpha.$$

**Proof of Theorem 2.** Freeze all arguments except $x^\nu$, and apply Theorem 1. \qed

By repetitive application of Theorem 2, using the product rule, we obtain higher order derivatives. (The product rule readily generalizes to matrix-valued functions, as long as one respects the ordering of non-commuting matrices.) We refer to Appendix A for the second order case, presenting the operational formula for the Hessian, $\nabla^2 \exp(X)$. It is clear that complexity increases rapidly with order.

In the context of DTI we set $k = n$, and restrict ourselves to $S_n \subset M_{n \times n}$, $S_{n-1} \subset M_{n \times n}$. In the following, logarithm takes precedence over convolution product.

**Result 1** Setting $X = \ln f \ast \phi_\sigma$, we have obtained formulas for the gradient—Theorem 2—and Hessian—Theorem 3. Appendix A—of the multi-scale representation of a matrix field $f \in C^\infty(\mathbb{R}^n, S^n)$. Moreover, if $X = \ln f \ast \phi_\sigma$, then $\partial_\mu X = \ln f \ast \partial_\mu \phi_\sigma$, $\partial_\mu \partial_\nu X = \ln f \ast \partial_\mu \partial_\nu \phi_\sigma$. These are the familiar expressions for derivatives in linear scale space theory, applied to the logarithm of the input image, and can be computed in the same way (by direct convolution, via DFT, or recursively [6, 14, 17, 41]). However, to obtain the desired multi-scale derivatives requires us to compute matrix logarithms and exponentials, and to compute the aforementioned integral representations (Theorems 2–3) over non-commuting matrix products, e.g. using a standard numerical Riemann sum approximation for the parameter integrals [3, 5]. Although technically straightforward, this is a numerical complication, but it is the price we must pay for consistency. Up to second order, the relevant case in the differential geometric rationale, this fortunately turns out relatively unproblematic. Figs. 2–3 illustrate the zeroth order scale space representation of a synthetic tensor field $f \in C^\infty(\mathbb{R}^2, S^2)$. (Higher order results are difficult to visualize, and best studied in the context of a particular application, such as tractography or connectivity analysis.) Asymptotics are as expected. Given suitable boundary conditions, the blurred DTI matrix field tends to an isotropic, homogeneous field given by the pointwise identity matrix as scale tends to infinity. The smooth connection between Euclidean geometry at large scales and the Riemannian geometry induced by the raw DTI image at data scale opens the door to a coarse-to-fine approach in differential geometric analysis.

5. Conclusion and Summary

We have proposed an operational scale space paradigm for symmetric positive definite matrix fields, such as encountered in DTI. The automorphic nature of matrix inversion within this class led us to impose a basic consistency demand, viz. that blurring and inversion should commute. This demand is incompatible with linear scale space theory, but is manifest in the proposed nonlinear modification. Perhaps not surprisingly, the commutativity axiom reproduces the zeroth order representation introduced byArsigny et al., Fillard et al., and Pennec et al. As such it adds further support to their so-called log-Euclidean framework [1, 8, 34].

A complication is the nontrivial nature of the chain rule for nested matrix-valued functions. We have derived novel, explicit integral expressions for first and second order derivatives, and provided an operational scheme to compute these at any physically meaningful scale. Extension to higher orders is straightforward, but produces cumbersome expressions, with a concomitant computational price due to the fact that the complexity of the resulting integral expressions grows rapidly with order. We have argued, however, that second order structure may well be sufficient in a differen-
tial geometric framework, since it allows one to obtain the most important geometric objects, notably the affine connection or Christoffel symbols, and the Riemann curvature tensor. Numerical implementation is straightforward.

Interestingly, the proposed multi-resolution scheme naturally joins the differential geometric paradigms on the DTI domain (Riemannian geometry) and codomain (log-Euclidean framework). An intriguing, hitherto unexplored possibility enabled by this scheme is a coarse-to-fine (or Euclidean-to-Riemannian) approach towards differential geometric DTI analysis (tractography, connectivity analysis, and so forth). In this respect one should always keep in mind that resolution limitations in diffusion imaging preclude a microscopic analysis of biological tissue at cellular level, so that data scale is a priori as arbitrary as any other empirical scale. One is thus confined to a mesoscopic analysis, in which tissue organization induces the “right scales” for analysis rather than technical acquisition limitations (pixels). In a multi-scale representation all scales are a priori equivalent, thus providing a natural basis for probing the “deep structure” of DTI data [21]. This is a subject for much future work.

The proposed multi-resolution framework is also relevant for high angular resolution diffusion imaging (HARDI), since in this context, too, the differential geometric rationale is applicable, cf. the active contour approach based on Finsler geometry by Melonakos et al. [25], and the intimate analytical connection between HARDI and DTI pointed out by Özarslan and Mareci [33].

Acknowledgements.

The Netherlands Organisation for Scientific Research (NWO) is gratefully acknowledged for financial support.

References


Figure 2. Two-dimensional synthetic image showing a positive symmetric tensor field unperturbed (left) and perturbed by noise (right). The second is subject to the scale space extension illustrated in Fig. 3.
A. Hessian of the Matrix Exponential Function

Theorem 3 Let $X \in C^\omega(\mathbb{R}^k, M_n)$. The Hessian of the matrix exponential function $\exp X \in C^\omega(\mathbb{R}^k, M_n^+)$ can be represented component-wise as follows:

$$\partial_\mu \partial_\nu \exp (X(x)) = \int_0^1 \exp ((1-\sigma)X(x)) \partial_\mu \partial_\nu X(x) \exp (\sigma X(x)) \, d\sigma$$

$$+ \int_0^1 \int_0^\sigma \exp ((1-\sigma)X(x)) \partial_\nu X(x) \exp (\sigma - \rho)X(x)) \partial_\mu X(x) \exp (\rho X(x)) \, d\rho \, d\sigma$$

$$+ \int_0^1 \int_0^\sigma \exp ((1-\sigma)X(x)) \partial_\mu X(x) \exp ((\sigma - \rho)X(x)) \partial_\nu X(x) \exp (\rho X(x)) \, d\rho \, d\sigma.$$ 

Note that if $[\nabla X, X] = 0$ (always the case if $n = 1$, hardly ever otherwise), this expression reduces to $\partial_\mu \partial_\nu \exp (X) = (\partial_\mu \partial_\nu X + \partial_\nu X \partial_\mu X) \exp (X)$, as it should.

Proof of Theorem 3. Apply the Theorem 2 twice, and perform a change of variables so as to obtain the integral expression stated in the theorem.

2

Figure 3. Illustration of Eq. (7) for twelve levels of exponentially increasing scales. The two-dimensional synthetic data are of course not very realistic, and are only provided for the sake of illustration. Qualitative behaviour and asymptotics will be similar in practice.