Statistical Computing on Non-Linear Spaces for Computational Anatomy

Xavier Pennec and Pierre Fillard

ABSTRACT
Computational anatomy is an emerging discipline that aims at analyzing and modeling the individual anatomy of organs and their biological variability across a population. However, understanding and modeling the shape of organs is made difficult by the absence of physical models for comparing different subjects, the complexity of shapes, and the high number of degrees of freedom implied. Moreover, the geometric nature of the anatomical features usually extracted raises the need for statistics on objects like curves, surfaces and deformations that do not belong to standard Euclidean spaces. We explain in this chapter how the Riemannian structure can provide a powerful framework to build generic statistical computing tools. We show that few computational tools derive for each Riemannian metric can be used in practice as the basic atoms to build more complex generic algorithms such as interpolation, filtering and anisotropic diffusion on fields of geometric features. This computational framework is illustrated with the analysis of the shape of the scoliotic spine and the modeling of the brain variability from sulcal lines where the results suggest new anatomical findings.

1 Computational Anatomy: Aims and Methods

Anatomy is the science that studies the structure and the relationship in space of different organs and tissues in living systems. Before the renaissance, anatomical descriptions were mainly based on animal models and the physiology was more philosophical than scientific. Modern anatomy really began with the authorized dissection of human cadavers, giving birth to the "De humani corporis fabrica" published by in 1543 by Vesale (1514-1564), and was strongly pushed by the progresses in surgery, as exemplified by the "Universal anatomy of the human body" (1561-62) of the great surgeon Ambroise Paré (1509-1590). During the following centuries, many progresses were done in anatomy thanks to new observation tools like microscopy and histology, going down to the level of cells in the 19th and 20th centuries. However, in-vivo and in-situ imaging is radically renewing the field since the 1980ies. An ever growing number of imaging modalities allows observing both the anatomy and the function at many spatial scales (from cells to the whole body) and at multiple time scales: milliseconds (e.g. beating heart), years (growth or aging), or even ages (evolution of species). Moreover, the non-invasive aspect allows repeating the observations on multiple subjects. This has a strong impact on the goals of the anatomy which are changing from the description of a *representative individual* to the description of the structure and organization of organs at the *population level.* The huge amount of information generated also raises the need for computerized methods to extract and structure information. This led in the last 10 to 20 years to the gradual evolution of *descriptive atlases* into interactive and *generative models*, allowing the simulation of new observations. Typical examples are given for the brain by the MNI 305 [25] and ICBM 152 [46] templates that are the basis of the Brain Web MRI simulation engine [20]. In the orthopedic domain, one may cite the "bone morphing" method [34, 63] that allows to simulate the shape of bones.

The combination of these new observation means and of the computerized methods is at the heart of computational anatomy, an emerging discipline at the interface of geometry, statistics and image analysis which aims at developing algorithms to model and analyze the biological shape of tissues and organs. The goal is not only to estimate representative organ anatomies across diseases, populations, species or ages but also to model the organ development across time (growth or aging) and to establish their variability. Another goal is to correlate this variability information with other functional, genetic or structural information (e.g. fiber bundles extracted from diffusion tensor images). From an applicative point of view, a first objective is to understand and to model how life is functioning at the population level, for instance by classifying pathologies from structural deviations (taxonomy) and by integrating individual measures at the population level to relate anatomy and function. For instance, the goal of spatial normalization of subjects in neuroscience is to map all the anatomies into a common reference system. A second application objective is to provide better quantitative and objective measures to detect, understand and correct dysfunctions at the individual level in order to help therapy planning (before), control (during) and follow-up (after).

The method is generally to map some generic (atlas-based) knowledge to patients-specific data through atlas-patient registration. In the case of observations of the same subject, many geometrical and physically based registration methods were proposed to faithfully model and recover the deformations. However, in the case of different subjects, the absence of physical models relating the anatomies leads to a reliance on statistics to learn the geometrical relationship from many observations. This is usually done by identifying anatomically representative geometric features (points, tensors, curves, surfaces, volume transformations), and then modeling their statistical distribution across the population, for instance via a mean shape and covariance structure analysis after a group-wise matching. In the case of the brain, one can rely on a hierarchy of structural models ranging from anatomical or functional landmarks like the AC and PC points [66, 15], curves like crest lines [65] or sulcal lines [44, 41, 29], surfaces like the cortical surface or sulcal ribbons [68, 1, 72], images seen as 3D functions, which lead to voxel-based morphometry (VBM) [6], diffusion imaging or rigid, multi-affine or diffeomorphic transformations [70, 49, 2], leading to Tensor-based morphometry (TBM). However, one crucial point is that these features usually belong to curved manifolds rather than to Euclidean spaces, which precludes the use of classical linear statistics. For instance, the average of points on a sphere is located inside the sphere and not on its surface.

To address this problem, one has to rely on statistical tools that work directly on manifolds in an intrinsic way. We summarize in Section 2 the mathematical bases that are needed to properly work on finite dimensional manifolds. Then, we show in Section 3 that a consistent set of statistical tools, including mean and covariance matrix analysis, can be developed based on the choice of a Riemannian metric. This algorithmic framework to compute on manifolds is then extended to process fields of geometric features (manifold-valued image). In particular, we show that one can perform interpolation, filtering, isotropic and anisotropic regularization and restoration of missing data (extrapolation or in-painting) on manifold valued images by using generalized weighted means and partial derivative equations (PDEs). Finally, the methodology is exemplified in Section 4 with two example applications: the statistical analysis of the anatomic variability of the spine in scoliotic patients, where a set of rigid body transformations is used to model the articulations between the vertebrae; and the modeling of the variability of the brain from a data-set of precisely delineated sulcal lines, where covariance matrices (symmetric positive definite matrices, so-called tensors) are used to describe the individual and joint anatomical variability (Green function) of points in the brain.

2 Mathematical bases of computing on manifolds

Computing on simple manifolds like the 3D sphere or a flat torus (for instance an image with opposite boundary points identified) might seems easy as we can see the geometrical properties (e.g. invariance by rotation or translation) and imagine tricks to alleviate the different problems. However, when it comes to slightly more complex manifolds like tensors, rigid body or affine transformations, without even thinking to infinite dimensional manifolds like spaces of surfaces or diffeomorphisms, computational tricks are much more difficult to find and have to be determined on a case by case basis. The goal of this section is to exemplify with the development of basic but generic statistical tools that the work specific to each manifold can be limited the determination of a few computational tools derive from a chosen Riemannian metric. These tools will then constitute the basic atoms to build more complex generic algorithms in Section 3.

2.1 The basic structure: the Riemannian metric

In the geometric framework, one has to separate the topological and differential properties of the manifold from the metric ones. The first ones determine the local structure of a manifold $\mathcal{M}$ by specifying neighboring points and tangent vectors, which allows to differentiate smooth functions on the manifold. The topology also impacts the global structure as it determines if there exists a connected path between two points. However, we need an additional structure to quantify how far away two connected points are: a distance. By restricting to distances which are compatible with the differential structure, we enter into the realm of Riemannian geometry. A *Riemannian metric* is defined by a continuous collection of scalar products $\langle \, . \mid . \, \rangle_p$ (or equivalently norms $\|.\|_p$) on each tangent space $T_p\mathcal{M}$ at point p of the manifold. Thus, if we consider a curve on the manifold, we can compute at each point its instantaneous speed vector (this operation only involves the differential structure) and its norm to obtain the instantaneous speed (the Riemannian metric is needed for this operation). To compute the length of the curve, this value is integrated as usual along the curve. The distance between two points of a connected Riemannian manifold is the minimum length among the curves joining these points. The curves realizing this minimum are called geodesics. The calculus of variations shows that geodesics are the solutions of a system of second order differential equations depending on the Riemannian metric. In the following, we assume that the manifold is geodesically complete, i.e. that all geodesics can be indefinitely extended. This means that the manifold has neither boundary nor any singular point that we can reach in a finite time. As an important consequence, the Hopf-Rinow-De Rham theorem states that there always exists at least one minimizing geodesic between any two points of the manifold (i.e. whose length is the distance between the two points).

2.2 Exponential chart

Let p be a point of the manifold that we consider as a local reference and $\vec{v}$ a vector of the tangent space $T_p\mathcal{M}$ at that point. From the theory of second order differential equations, we know that there exists one and only one geodesic $\gamma_{(p,\vec{v})}(t)$ starting from that point with this tangent vector. This allows to wrap the tangent space onto the manifold, or equivalently to develop the manifold in the tangent space along the geodesics (think of rolling a sphere along its tangent plane at a given point), by mapping to each vector $\vec{v} \in T_p\mathcal{M}$ the point q of the manifold that is reached after a unit time by the geodesic $\gamma_{(p,\vec{v})}(t)$. This mapping $\mathrm{Exp}_p(\vec{v}) = \gamma_{(p,\vec{v})}(1)$ is called the *exponential map* at point p. Straight lines going through 0 in the tangent space are transformed into geodesics going through point p on the manifold and distances along these lines are conserved (Fig. 1).

The exponential map is defined in the whole tangent space $T_p\mathcal{M}$ (since the manifold is geodesically complete) but it is generally one-to-one only

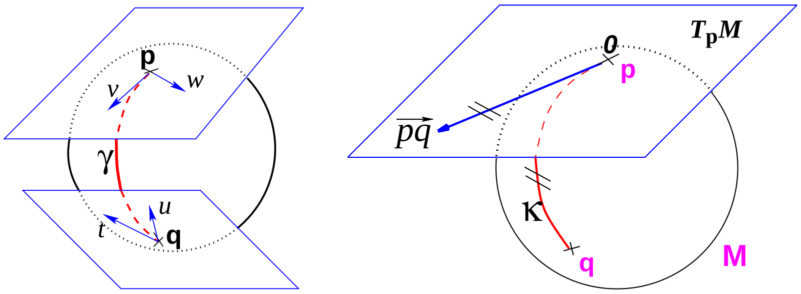

FIGURE 1. **Left:** The tangent planes at points p and q of the sphere $\mathcal{S}_2$ are different: the vectors v and w of $T_p\mathcal{M}$ cannot be compared to the vectors t and u of $T_q\mathcal{M}$. Thus, it is natural to define the scalar product on each tangent plane. **Right:** The geodesics starting at x are straight lines in the exponential map and the distance along them is conserved.

locally around 0 in the tangent space (i.e. around p in the manifold). In the sequel, we denote by $\overrightarrow{pq} = \text{Log}_p(q)$ the inverse of the exponential map: this is the smallest vector (in norm) such that $q = \text{Exp}_p(\overrightarrow{pq})$. If we look for the maximal definition domain, we find out that it is a star-shaped domain delimited by a continuous curve C_p called the *tangential cut-locus*. The image of C_p by the exponential map is the cut locus $\mathcal{C}_p$ of point p. This is (the closure of) the set of points where several minimizing geodesics starting from p meet. On the sphere $\mathcal{S}_2(1)$ for instance, the cut locus of a point p is its antipodal point and the tangential cut locus is the circle of radius π.

The exponential and log maps within this domain realizes a chart (a local parameterization of the manifold) called *the exponential chart at point* p. It covers all the manifold except the cut locus of the reference point p, which has a null measure. In this chart, geodesics starting from p are straight lines, and the distance from the reference point are conserved. This chart is somehow the "most linear" chart of the manifold with respect to the reference point p. The set of all the exponential charts at each point of the manifold realize an atlas which allows working very easily on the manifold, as we will see in the following.

2.3 Practical implementation

The exponential and logarithmic maps (from now on Exp and Log maps) are obviously different for each manifold and for each metric. Thus they have to be determined and implemented on a case by case basis. Example for rotations, rigid body transformations can be found for the left invariant metric in [60], and examples for tensors in [58, 4]. Exponential charts constitute very powerful atomic functions in terms of implementation on which we will be able to express practically all the geometric operations: the implementation of Log_p and Exp_q is the basis of programming on Riemannian manifolds, as we will see in the following.

In a Euclidean space, the exponential charts are nothing but one orthonormal coordinates system translated at each point: we have in this case $\overrightarrow{pq} = \text{Log}_p(q) = q - p$ and $\text{Exp}_p(\vec{v}) = p + \vec{v}$. This example is more than a simple coincidence. In fact, most of the usual operations using additions and subtractions may be reinterpreted in a Riemannian framework using the notion of *bi-point*, an antecedent of vector introduced during the 19th Century. Indeed, vectors are defined as equivalent classes of bi-points in a Euclidean space. This is possible because we have a canonical way (the translation) to compare what happens at two different points. In a Riemannian manifold, we can still compare things locally (by parallel transportation), but not any more globally. This means that each "vector" has to remember at which point of the manifold it is attached, which comes back to a bi-point.

A second way to see the vector $\overrightarrow{pq}$ is as a vector of the tangent space at point p. Such a vector may be identified to a point on the manifold using the exponential map $q = \text{Exp}_p(\overrightarrow{pq})$. Conversely, the logarithmic map may be used to map almost any bi-point (p, q) into a vector $\overrightarrow{pq} = \text{Log}_p(q)$ of $T_p\mathcal{M}$. This reinterpretation of addition and subtraction using logarithmic and exponential maps is very powerful to generalize algorithms working on vector spaces to algorithms on Riemannian manifolds, as illustrated in Table 1.1 and the in following sections.

	Euclidean space	Riemannian manifold
Subtraction	$\overrightarrow{pq} = q - p$	$\overrightarrow{pq} = \text{Log}_p(q)$
Addition	$p = q + \vec{v}$	$q = \text{Exp}_p(\vec{v})$
Distance	$\text{dist}(p, q) = \lVert q - p \rVert$	$\text{dist}(p, q) = \lVert \overrightarrow{pq} \rVert_p$
Mean value (implicit)	$\sum_i (p_i - \bar{p}) = 0$	$\sum_i \overrightarrow{\bar{p}p_i} = 0$
Gradient descent	$p_{t+\varepsilon} = p_t - \varepsilon \overrightarrow{\nabla C(p_t)}$	$p_{t+\varepsilon} = \text{Exp}_{p_t}(-\varepsilon \overrightarrow{\nabla C(p_t)})$
Geodesic interpolation	$p(t) = p_0 + t\, \overrightarrow{p_0 p_1}$	$p(t) = \text{Exp}_{p_0}(t\, \overrightarrow{p_0 p_1})$

TABLE 1.1. Re-interpretation of standard operations in a Riemannian manifold.

2.4 *Example of Metrics on Covariance matrices*

Let us take an example with positive definite symmetric matrices, called tensors in medical image analysis. They are used for instance to encode the covariance matrix of the Brownian motion (diffusion) of water in Diffusion Tensor Imaging (DTI) [8, 40] or to encode the joint variability at different places (Green function) in shape analysis (see [30, 29, 31] and Section 4). They are also widely used in image analysis to guide the segmentation, grouping and motion analysis [47, 73, 16, 74].

The main problem is that the tensor space is a manifold that is not a vector space with the usual additive structure. As the positive definiteness constraint delimits a convex half-cone in the vector space of symmetric matrices, convex operations (like the mean) are stable in this space but problems arise with more complex operations. For instance, there is inevitably a point in the image where the time step is not small enough when smoothing fields of tensors with gradient descents, and this results into negative eigenvalues.

To answer that problem, we proposed in [58] to endow the space of tensors with a Riemannian metric invariant by any change of the underlying space coordinates, i.e. invariant under the action of affine transformations of covariance matrices. A few mathematical developments showed that the Exp and Log and distance maps were given with quite simple formulas involving the matrix logarithm exp and log:

$$\begin{aligned}\mathrm{Exp}_\Sigma(W) &= \Sigma^{1/2} \exp\left(\Sigma^{-1/2} W \Sigma^{-1/2}\right) \Sigma^{1/2}\\ \mathrm{Log}_\Sigma(\Lambda) &= \Sigma^{1/2} \log\left(\Sigma^{-1/2} \Lambda \Sigma^{-1/2}\right) \Sigma^{1/2}\\ \mathrm{dist}^2(\Sigma, \Lambda) &= \mathrm{Tr}\left(\log(\Sigma^{-1/2} \Lambda \Sigma^{-1/2})^2\right)\end{aligned}$$

This metric leads to a very regular Hadamard manifold structure, a kind of hyperbolic space without cut-locus, which simplifies the computations. Tensors with null and infinite eigenvalues are both at an infinite distance of any positive definite symmetric matrix: the cone of positive definite symmetric matrices is changed into a space of "constant" (homogeneous) non-scalar curvature without boundaries. Moreover, there is one and only one geodesic joining any two tensors, the mean of a set of tensors is uniquely defined, and we can even define globally consistent orthonormal coordinate systems of tangent spaces. Thus, the structure we obtain is very close to a vector space, except that the space is curved.

This affine-invariant Riemannian metric derives from affine invariant connections on homogeneous spaces [54]. It has been introduced in statistics to model the geometry of the multivariate normal family (the Fisher information metric)[18, 64, 19] and in simultaneously by many teams in medical image analysis to deal with DTI [33, 52, 9, 42, 58]. In [58], we showed that this metric could be used not only to compute distances between tensors, but also as the basis of a complete computational framework on manifold-valued images as will be detailed in Section 3.

By trying to put a Lie group structure on the space of tensors, Vincent Arsigny observed later that the matrix exponential was a diffeomorphism from the space of symmetric matrices to the tensor space. Thus, one can seamlessly transport all the operations defined in the vector space of symmetric matrices to the tensor space [5, 4]. This gives a commutative Lie group structure to the tensors, and the Euclidean metric on symmetric matrices is transformed into a bi-invariant Riemannian metric on the tensor manifold. As geodesics are straight lines in the space of symmetric matrices, the expression of the Exp, Log and distance maps for the Log-Euclidean metric is easily determined:

$$\begin{aligned}\mathrm{Exp}_{\Sigma}(W) &= \exp(\log(\Sigma) + \partial_W \log(\Sigma))\\ \mathrm{Log}_{\Sigma}(\Lambda) &= D\exp(\log(\Sigma))\,(\log(\Lambda) - \log(\Sigma))\\ \mathrm{dist}^2_{LE}(\Sigma_1, \Sigma_2) &= \mathrm{Tr}\left((\log(\Sigma_1) - \log(\Sigma_2))^2\right)\end{aligned}$$

These formulas look more complex than for the affine invariant metric because they involve the differential of the matrix exponential and logarithm in order to transport tangent vectors from one space to another [59]. However, they are in fact nothing but the transport of the addition and subtraction through the exponential of symmetric matrices. In practice, the log-Euclidean framework consist in taking the logarithm of the tensor data, computing like usual in the Euclidean space of symmetric matrices, and coming back at the end to the tensor space using the exponential [5, 3].

From a theoretical point of view, geodesics through the identity are the same for both metrics, but this is not true any more in general at other points of the tensor manifold [4]. A careful comparison of both metrics in practical applications [3, 5] showed that there was very few differences on the results (of the order of 1%) on real DTI images, but that the log-Euclidean computations where 4 to 10 times faster. For other types of applications, like adaptive re-meshing [53], the anisotropy of the tensors can be much larger, which may lead to larger differences. In any case, initializing the iterative optimizations of affine-invariant algorithms with the log-Euclidean result drastically speeds-up the convergence. Important application example of this tensor computing framework were provided in [27, 28] with a statistically grounded estimation and regularization of DTI images. The white matter tractography that was allowed by these methods in clinical DTI images with very poor signal to noise ratios could lead to new clinical indications of DTI, for instance in the spinal chord [23].

3 Statistical Computing on Manifolds

The previous section showed how to derive the atomic Exp and Log maps from a Riemannian metric. We now summarize in this section how one generalizes on this basis many important statistical notions, like the mean, covariance and Principal Component Analysis (PCA), as well as many image processing algorithms like interpolation, diffusion and restoration of missing data (extrapolation). For details about the theory of statistics on Riemannian manifolds in itself, we refer the reader to [56, 57] and reference therein. Manifold-valued image processing is detailed in [58] with the example of tensors.

3.1 First statistical moment: the mean

The Riemannian metric induces an infinitesimal volume element on each tangent space, and thus a reference measure $d\mathcal{M}(p)$ on the manifold that can be used to measure random events on the manifold and to define the probability density function (the function ρ such that $dP(p) = \rho(p) d\mathcal{M}(p)$, if it exists). It is worth noticing that the induced measure $d\mathcal{M}$ represents the notion of *uniformity* according to the chosen Riemannian metric. This automatic derivation of the uniform measure from the metric gives a rather elegant solution to the Bertrand paradox for geometric probabilities [62, 38]. With the probability measure of a random element, we can integrate functions from the manifold to any vector space, thus defining the expected value of this function. However, we generally cannot integrate manifold-valued functions. Thus, one cannot define the mean or expected "value" of a random manifold element that way.

One solution is to rely on a distance-based variational formulation: the Fréchet (resp. Karcher) expected features minimize globally (resp. locally) the variance:

$$\sigma^2(q) = \int \operatorname{dist}(p,q)^2 \, dP(p) = \frac{1}{n} \sum_{i=1}^{n} \operatorname{dist}(p_i, q)^2,$$

written respectively in the continuous and discrete forms. One can generalize the variance to a dispersion at order α by changing the L_2 with an α-norm: $\sigma_\alpha(p) = (\int \operatorname{dist}(p,q)^\alpha dP(p))^{1/\alpha}$. The minimizers are called the central Karcher values at order α. For instance, the median is obtained for $\alpha = 1$ and the modes for $\alpha = 0$, exactly like in the vector case. It is worth noticing that the median and the modes are not unique in general in the vector space, and that even the mean may not exists (e.g. for heavy tailed distribution). In Riemannian manifolds, the existence and uniqueness of all central Karcher values is generally not ensured as they are obtained through a minimization procedure. However, for a finite number of discrete samples at a finite distance of each other, which is the practical case in statistics, a mean value always exists and it is unique as soon as the distribution is sufficiently peaked [37, 39].

Local minima may be characterized as particular critical points of the cost function: at Karcher mean points, the gradient of the variance should be null. However, the distance is continuous but not differentiable at cut locus points where several minimizing geodesic meets. For instance, the distance from a point of the sphere to its antipodal point is maximal, but decrease continuously everywhere around it. One can show [56, 57] that the variance it differentiable at all points where the cut locus has a null measure and has gradient: $\overrightarrow{\nabla \sigma^2}(q) = -2 \int \overrightarrow{qp} \, dP(p) = \frac{-2}{n} \sum_{i=1}^{n} \overrightarrow{qp_i}$ respectively in the continuous (probabilistic) and discrete (statistical) formulations. In practice, this gradient is well defined for all distributions that have a pdf since the cut locus has a null measure. For discrete samples, the gradient exists if there is no sample lying exactly on the cut-locus of the current test point. Thus, we end up with the implicit characterization of Karcher mean points as exponential barycenters which was presented in Table 1.1.

To practically compute the mean value, we proposed in [60] for rigid body transformations and in [56, 57] for the general Riemannian case to use a Gauss-Newton gradient descent algorithm. It essentially alternates the computation of the barycenter in the exponential chart centered at the current estimation of the mean value, and a re-centering step of the chart at the point of the manifold that corresponds to the computed barycenter (geodesic marching step). This gives the Newton iteration: $\bar{p}^{t+1} = \mathrm{Exp}_{\bar{p}^t}\left(\frac{1}{n}\sum_{i=1}^{n} \overrightarrow{\bar{p}^t p_i}\right)$. One can actually show that its convergence is locally quadratic towards non degenerated critical points [55, 43, 22].

3.2 Covariance matrix and Principal Geodesic Analysis

Once the mean point is determined, using the exponential chart at the mean point is particularly interesting as the random feature is represented by a random vector with null mean in a star-shaped domain. With this representation, there is no difficulty to define the covariance matrix:

$$\Sigma = \int \overrightarrow{\bar{p}q}.\overrightarrow{\bar{p}q}^{\mathrm{T}} \, dP(q) = \frac{1}{n}\sum_{i=1}^{n} \overrightarrow{\bar{p}q_i}.\overrightarrow{\bar{p}q_i}^{\mathrm{T}}$$

and potentially higher order moments. This covariance matrix can then be used to defined the Mahalanobis distance between a random and a deterministic feature: $\mu_{(\bar{p},\Sigma)}(q) = \overrightarrow{\bar{p}q}^{\mathrm{T}}\Sigma^{(-1)}\overrightarrow{\bar{p}q}$. Interestingly, the expected Mahalanobis distance of a random element is independent of the distribution and is equal to the dimension of the manifold, as in the vector case. This statistical distance can be used as a basis to generalize some statistical tests such as the mahalanobis D^2 test [57].

To analyze the results of a set of measurements in a Euclidean space, one often performs a principal component analysis (PCA). A generalization to Riemannian manifolds called Principal Geodesic Analysis (PGA) was proposed in [32] to analyze shapes based on the medial axis representations (M-reps). The basic idea is to find a low dimensional sub-manifold generated by some geodesic subspaces that best explain the measurements (i.e. such that the squared Riemannian distance from the measurements to that sub-manifold is minimized). Another point of view is to assume that the measurements are generated by a low dimensional Gaussian model. Estimating the model parameters amounts to a covariance analysis in order to find the k-dimensional subspace that best explains the variance. In a Euclidean space, these two definitions correspond thanks to Pythagoras's theorem. However, in the Riemannian setting, geodesic subspaces are generally not orthogonal due to the curvature. Thus, the two notions differ:

while the Riemannian covariance analysis can easily be performed in the tangent space of the mean, finding Riemannian sub-manifolds turns out to become an almost intractable problem. As a matter of fact, the solution retained by [32] was finally to rely on the covariance analysis.

When the distribution is unimodal and sufficiently peaked, we believe that covariance analysis is anyway much better suited. However, for many problems, the goal is rather to find a sub-manifold on which measurements are more or less uniformly distributed. This is the case for instance for features sampled on a surface or points sampled along a trajectory (time sequences). While the one dimensional case can be tackled by regression [21], the problem for higher dimensional sub-manifolds remains quite open. Some solutions may come from manifold embedding techniques as exemplified for instance in [17].

3.3 Interpolation and filtering as weighted means

One of the important operations in geometric data processing is to interpolate values between known measurements. The standard way to interpolate on a regular lattice is to make a linear combination of samples f_k at integer (lattice) coordinates $k \in \mathbb{Z}^d$: $f(x) = \sum_k w(x-k)\, f_k$. A typical example is the sinus cardinal interpolation. With the nearest-neighbor, linear (or tri-linear in 3D), and higher order spline interpolations, the kernel is piecewise polynomial, and has a compact support [67, 48]. With normalized weights, this interpolation can be seen as a weighted mean. Thus, it can be generalized in the manifold framework as an optimization problem: the interpolated value $p(x)$ on our feature manifold is the point that minimizes $C(p(x)) = \sum_{i=1}^n w_i(x)\, \mathrm{dist}^2(p_i, p(x))$. This can easily be solved using the iterative Gauss-Newton scheme proposed for the Karcher mean. The linear interpolation is interesting and can be written explicitly since it is a simple geodesic walking scheme: $p(t) = \mathrm{Exp}_{p_0}(t\, \overrightarrow{p_0 p_1}) = \mathrm{Exp}_{p_1}((1-t)\, \overrightarrow{p_1 p_0})$.

Many other operators can be rephrased as weighted means. For instance approximations and convolutions like Gaussian filtering can be viewed as the average of the neighboring values weighted by a (Gaussian) function of their spatial distance. For instance, $\hat{F}(x) = \int_{\mathbb{R}^n} K(u)\ F(x+u)\ du$ is the minimizer of $C(\hat{F}) = \int_{\mathbb{R}^n} K(u)\ \ \mathrm{dist}^2(F(x+u), \hat{F}(x))\ du$. In a Riemannian manifold, this minimization problem is still valid, but instead of a closed-form solution, we have once again a Gauss-Newton iterative gradient descent algorithm to reach the filtered value:

$$\hat{p}^{t+1}(x) = \int_{\mathbb{R}^n} K(u)\, \mathrm{Log}_{\hat{p}^t(x)}(p(x+u))\, du.$$

We can also use anisotropic and non-stationary kernels $K(x,u)$. For instance, it can be modulated by the norm of the derivative of the field in the direction u. We should notice that for a manifold-value field $p(x)$, the directional derivatives $\partial_u p(x)$ is a tangent vector of $T_{p(x)}\mathcal{M}$ which can be practically approximated using finite "differences" in the exponential chart: $\partial_u p(x) \simeq \mathrm{Log}_{p(x)}(p(x+u)) + O(\|u\|^2)$. However, to measure the norm of this vector, we have to use the Riemannian metric at that point: $\|\partial_u p\|_p$.

3.4 Harmonic diffusion and anisotropic regularization

An alternative to kernel filtering is to consider a regularization criterion that penalizes the spatial variations of the field. A measure of variation is the spatial gradient (the linear form that maps to any spatial direction u the directional derivative $\partial_u p(x)$), which can be robustly computed as the matrix that best approximates the directional derivatives in the neighborhood (e.g. 6, 18 or 26 connectivity in 3D). The simplest criterion based on the gradient is the Harmonic energy

$$Reg(p) = \frac{1}{2}\int_\Omega \|\nabla p(x)\|^2_{p(x)}\ dx = \frac{1}{2}\sum_{i=1}^{d}\int_\Omega \|\partial_{x_i} p(x)\|^2_{p(x)}\ dx.$$

The Euler-Lagrange equation with Neumann boundary conditions is as usual $\nabla Reg(p)(x) = -\Delta p(x)$. However, the Laplace-Beltrami operator on the manifold $\Delta p(x)$ is the sum of the usual flat Euclidean second order directional derivatives $\partial^2_{x_i} p(x)$ in a locally orthogonal system and an additional term due to the curvature of the manifold that distorts the orthonormality of this coordinate system. To practically compute this operator, we proposed in [58] an efficient and general scheme based on the observation that the Christoffel symbols and their derivatives along the geodesics vanish at the origin of the exponential chart. This means that the correction for the curvature is in fact already included: by computing the standard Laplacian *in that specific map*, one gets the directional Laplace-Beltrami operator for free: $\Delta_u p = \mathrm{Log}_{p(x)}(p(x+u)) + \mathrm{Log}_{p(x)}(p(x-u)) + O(\|u\|^4)$. Averaging over all the directions in a neighborhood finally gives a robust and efficient estimation.

A very simple scheme to perform Harmonic diffusion is to use a first order geodesic gradient descent. At each iteration and at each point x, one walks a little bit along the geodesic which start at the current point with the opposite of the gradient of the regularization criterion:

$$p^{t+1}(x) = \mathrm{Exp}_{p^t(x)}\left(-\varepsilon \Delta p^t(x)\right) \quad \text{with} \quad \Delta p(x) \propto \sum_{u\in\mathcal{V}} \frac{1}{\|u\|^2}\mathrm{Log}_{p(x)}(p(x+u))$$

In order to filter within homogeneous regions but not across their boundaries, an idea is to penalize the smoothing in the directions where the derivatives are important [61, 35]. This can be realized directly in the discrete implementation of the Laplacian by weighting the directional Laplacian by a decreasing function of the norm $\|\partial_u p\|_p$ of the gradient in that direction.

For instance, we used $\Delta_u p = \sum_u c(\|\partial_u p\|_p)\, \Delta_u p$ with $c(x) = \exp\left(-x^2/\kappa^2\right)$ in [58]. As the convergence of this scheme is not guaranteed (anisotropic regularization "forces" may not derive from a well-posed energy), the problem may be reformulated as the optimization of a ϕ-function of the Riemannian norm of the spatial gradient (a kind of robust M-estimator): $Reg_\phi(p) = \frac{1}{2}\int_\Omega \phi\left(\|\nabla p(x)\|_{p(x)}\right) dx$. By choosing an adequate ϕ-function, one can give to the regularization an isotropic or anisotropic behavior [7]. The main difference with a classical Euclidean calculation is that we have to take the curvature into account by using the Laplace-Beltrami operator, and by measuring the length of directional derivatives using the Riemannian metric at the right point [26]. Using $\Psi(x) = \phi'(x)/x$, we get:

$$\nabla Reg_\phi(p) = -\Psi(\|\nabla p\|_p)\Delta p - \sum_{i=1}^{d} \partial_{x_i} \Psi(\|\nabla p\|_p) \partial_{x_i} p.$$

3.5 Diffusion-based interpolation and extrapolation

The pure diffusion reduces the noise in the data but also the amount of information. Moreover, the total diffusion time that controls the amount of smoothing is difficult to estimate. At an infinite diffusion time, the field will be completely homogeneous. Thus, it is more interesting to consider the data as noisy observations and the regularization as a prior on the spatial regularity of the field. Usually, one assumes a Gaussian noise independent at each position, which leads to a least-squares criterion through a maximum likelihood approach. For a dense data field $q(x)$, the similarity criterion that is added to the regularization criterion is simply $Sim(p) = \int_\Omega \operatorname{dist}^2\left(p(x)\, ,\, q(x)\right)\, dx$. The only difference here is that it uses the Riemannian distance. It simply adds a linear (geodesic) spring $\nabla_p \operatorname{dist}^2(p,q) = -2\, \overrightarrow{pq}$ to the global gradient to prevent the regularization from pulling to far away from the original data.

For sparse measures, using directly the maximum likelihood on the observed data leads to deal with Dirac (mass) distributions in the derivatives, which is a problem for the numerical implementation. One solution is to consider the Dirac distribution as the limit of the Gaussian function G_σ when σ goes to zero, which leads to the regularized derivative [58]: $\nabla Sim(x) = -2\, \sum_{i=1}^{n} G_\sigma(x - x_i)\, \overrightarrow{p(x) p_i}$.

4 Modeling the Anatomy

4.1 A statistical shape model of the scoliotic spine

Now that we have the methodology to work with geometric features, let us see how it can be used to model the anatomy. A first interesting example was proposed by Jonathan Boisvert [13, 12] with a 3D articulated model of the spine. The model gathers the relative configurations of the vertebrae along the spinal chord (the parameters are the rigid transforms that superpose neighboring vertebrae) rather than the position and orientation of each vertebra in a global reference frame. As small local motions at one point of the spine may have a large impact of the position at another point, this local representation is better capturing information that may get unnoticed in a global reference frame. However, this requires making statistics on geometric objects (rigid body transformation parameters) rather than on just points.

The statistical model of the spine was established in a population of 307 untreated scoliotic patients. Each vertebra was reconstructed in 3D from anatomical landmarks in bi-planar radiographies. Posture during data acquisition was normalized but individual factors such as the age, sex or type of scoliotic curve were not taken into account. Thus, the statistics capture the anatomical variability inherent to the pathology but also the growth stage. The Fréchet mean and the generalized covariance of the articulated model was then computed. As there are 102 degrees of freedom (5 lumbar and 12 thoracic vertebrae), the analysis of the covariance matrix could hardly be performed by hand. Thus, the most meaningful modes of variation were extracted using a PCA on the tangent plane.

A visual inspection reveals that the first modes had clinical meanings and were explaining curve patterns that are routinely used in different clinical classifications of scoliosis (see [14, 11] for details). For instance, the first mode appears to be associated with the patient growth with a mild thoracic curve (King's type II or III depending on the amplitude of the mode) and the second could be described as a double thoraco-lumbar curve (King's type I), see Fig. 2. A more quantitative analysis showed that there is a statistically significant link between the 4 principal modes and King's classes, although each class is generally linked to a combination of modes rather than only one mode [11].

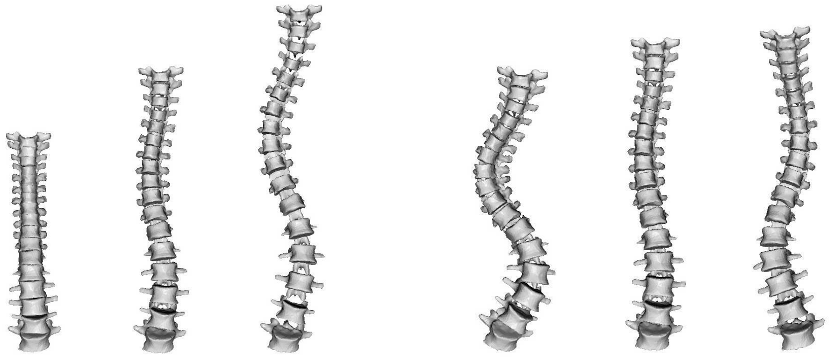

FIGURE 2. First (left) and second (right) modes of variation of the statistical spine model depicted at -3, 0 (mean) and 3 times its standard deviation. Images courtesy of Jonathan Boisvert, Polytechnique School of Montreal, Canada.

4.2 Learning Brain Variability from Sulcal Lines

A second interesting shape analysis application is the statistical modeling of the brain variability in a given population of 3D images. In such a process, a first step is to measure the variability of each anatomical position independently by identifying for instance corresponding points among each individual anatomy (structural homologies). This allows us to encode the brain variability by covariance matrices that we call *variability tensors*. The reason why we should not simplify these tensors into simpler scalar values is that there are evidences that structural variations are larger along certain preferred directions [69].

As exemplify in introduction, a hierarchy of anatomical features may be used to abstract the brain anatomy. We chose sulcal lines as they are low dimensional structures easily identified by neuroscientists. Moreover, a certain number of sulcal landmarks consistently appear in all normal individuals and allow a consistent subdivision of the cortex into major lobes and gyri [45]. In the framework of the associated team program Brain-Atlas between Asclepios at INRIA and LONI at UCLA, we use a data-set of sulcal lines manually delineated in 98 subjects by expert neuroanatomists according to a precise protocol[1]. We used the 72 sulcal curves that consistently appear in all normal subjects (abusively called *sulci* in the sequel).

To find the corresponding points between the 98 instances of each of the 72 sulci, we proposed in [30, 29] an original methodology which alternatively compute the matches that minimize the distance between the mean curve and the instances, and re-estimates the mean curve from the updated matches. As a result, we obtain the mean sulcal curves, along with a variability tensor which encodes the covariance matrix of the anatomical positions in each subject corresponding to each mean point. To optimally adjust the number of tensor needed to represent the variability information along each sulcus, we proposed a *tensor picking* method. The principle is to approximate our tensor measurements using a linear interpolation in-between N tensors picked along the line. The optimal subset of tensors is determined by optimizing the distance between interpolated and measured tensors along the line. The number of tensors picked along each line is adjusted so that the interpolation error does not exceed a prescribed value. In this process, we used the Riemannian metrics presented in Section 2.4 for their very good theoretical properties, and the algorithmic framework developed in Section 3.3. Tensor picking is illustrated in Fig. 3. We were able to show that selecting only 366 variability tensors was sufficient to encode the variability of the 72 sulci without a significant loss of accuracy.

The result is a sparse field of tensors, which can naturally be extrapolated to the whole space using the framework described in Section 3.5 (Fig. 4). This dense map of tensors was shown to be in good agreement with previous published results: the highly specialized and lateralized areas such as the planum parietale and the temporo-parietal areas consistently show the

---

[1] `http://www.loni.ucla.edu/~khayashi/Public/medial_surface/`

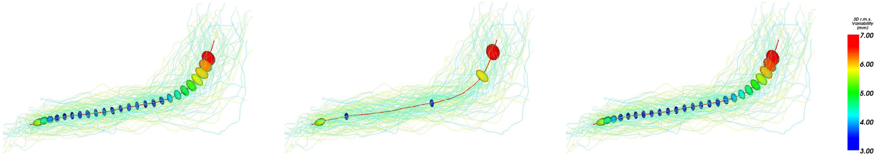

FIGURE 3. **Measuring variability tensors along the Sylvian Fissure.** Left: Covariance matrices (ellipsoids at one standard deviation) are overlaid at regularly sampled spatial positions along the mean sulci. Middle: Tensors selected by our tensor picking operation. Right: Tensors reconstructed by linear interpolation in-between them. Notice that only 5 tensors in that case nicely represent the variability of the entire sulcus.

highest amount of variability. The lowest amount of variability is consistently found in phylogenetically older areas (e.g. orbitofrontal cortex) and primary cortices that myelinate earliest during development (e.g., primary somatosensory and auditory cortex). However, our variability map gives more than just the amount of variability: we can extract from the tensors the spatial directions where the variability is the greatest at every single anatomical position. We refer the reader to [30, 29] for a more detailed explanation of the method and for the neuroscience interpretation of these results.

Modeling independently the variability at each point may not be sufficient as we may overlook potential statistical relationships between different brain regions. Indeed, long range relationships may arise from common genetic and trophic influences across brain regions (e.g., brain regions that develop together). In our framework, such relationships can be revealed by an analysis of the correlation between spatially close anatomical positions along the lines (neighboring points), but also distant points (e.g., a point and its symmetric counterpart in the opposite hemisphere).

In [31], we studied the correlation between two points $\bar{x}$ and $\bar{y}$ of the mean sulcal lines by canonical correlation analysis. This analysis is based on the total covariance matrix (TCM) of the corresponding points x_i and

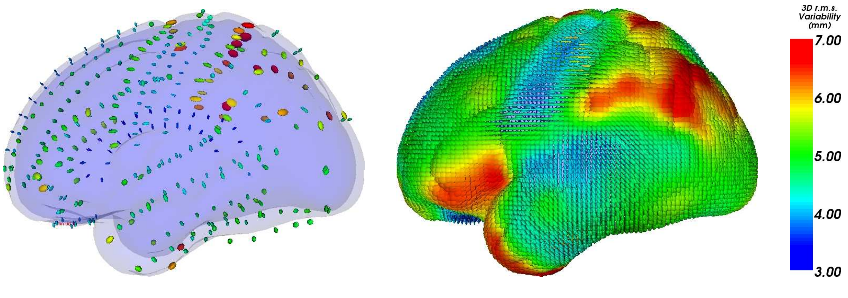

FIGURE 4. Variability tensor extrapolation. Left: The 366 tensors retained for our model. Right: Result of the extrapolation. Each point of this average brain shape contains a variability tensor.

y_i in each subject anatomy:

$$\mathrm{TCM}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}\begin{pmatrix} x_i-\bar{x} \\ y_i-\bar{y} \end{pmatrix}\begin{pmatrix} x_i-\bar{x} \\ y_i-\bar{y} \end{pmatrix}^{\top} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^{t} & \Sigma_{yy} \end{pmatrix}.$$

The TCM being a 6-dimensional tensor, our Riemannian processing framework naturally allows us to extrapolate this tensor from the sulcal lines to the whole cortex surface. To statistically assess the correlation of the variability at any two anatomical positions, we used the Bartlett-Lawley test which tests the rank of the correlation matrix $\Gamma = \Sigma_{xx}^{-1/2}\Sigma_{xy}\Sigma_{yy}^{-1/2}$. The singular values of this matrix are the correlation coefficients between the spatial directions given by corresponding eigenvector. A rank of at least 1 means that the variability at $\bar{x}$ and at $\bar{y}$ are significantly correlated in at least one direction. To account for multiple comparisons, we used the very conservative Bonferroni correction. In addition to the expected local correlation, results indicates that there was generally a long-range correlations with the symmetric point in the other hemisphere, but also unexpected long-range correlations with other parts of the brain as shown for instance in Fig. 5 with the superior temporal sulcus.

5 Challenges

We have shown in this chapter that the choice of a Riemannian metric and the implementation of a few tools derived from it, namely the Exp and Log maps, provide the bases for building a consistent algorithmic framework to compute on manifolds. In particular, we showed that one can compute consistent statistics, perform interpolation, filtering, isotropic and anisotropic regularization and restoration of missing data.

We also showed that powerful computational models of the anatomy could be built thanks to this Riemannian computing framework. For instance, Section 4.1 demonstrates that using a proper non-linear model of

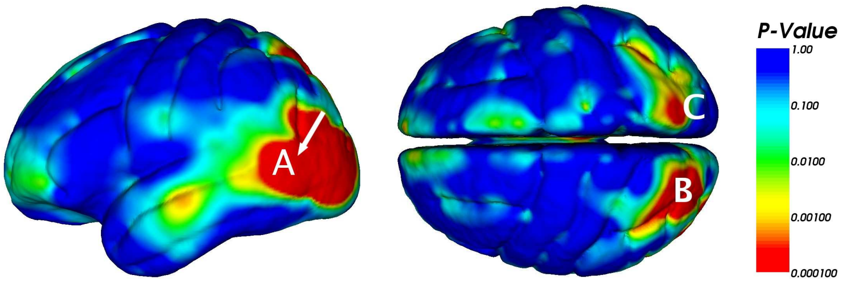

FIGURE 5. **Map of anatomical correlations.** The tip of the superior temporal sulcus (marked A) was picked as a reference point. The map indicates regions which are spotted as correlated with this reference position (hot colors mean correlation). The most correlated points include the parietal sulci (marked B and C), a very interesting neuroscience finding.

the spine allows to find a good separation of different physiological phenomena such as pathological deformations and normal growth. In this example, using the generic tools of the Riemannian computing framework particularly simplifies both the inception of the experimental setup, its implementation and the exploitation of the results. Section 4.2 also proposes a brain variability model that is able to recover the estimation of the sulcal variability with a very low number of parameters. By pushing further this statistical investigation, we showed that the same Riemannian framework could be used to measure and model the anatomical correlations between any two positions. These correlations are very interesting from a neuroscientific point of view since they can reveal factors of dependence between brain regions (like regions that develop or fail to develop together).

However, there are many challenges left open both from the theoretical and application point of views. For instance, it would be necessary to extend the computing framework presented here to infinite dimensional manifolds in order to deal properly with curves, surfaces and diffeomorphisms. For the case of diffeomorphism, we already known how to provide Riemannian metrics for which the geodesics can be computed by optimization [10, 49, 50, 36]. Through the so called EPDiff equation (Euler-Poincarré equation for diffeomorphisms), this optimization framework has been recently rephrased in an exponential/logarithm framework similar to the one developed here [51]. Thus, the basic algorithmic tools are the same, except that optimizing each time to compute the exponential and the logarithm has a deep impact on the computational times. Moreover, the infinite number of dimensions forbids the use of many tools like the probability density functions! For instance, the computation of simple statistics like the mean and the principal component analysis of diffeomorphism raises practical representation problems [71, 24].

From a computational anatomy standpoint, the huge number of degrees of freedom involved in the estimation of the anatomical variability will require to aggregate information coming from many different sources in order to improve the statistical power. As there is no gold standard, we should also be careful that many biases may be hidden in the results. Thus, methods to compare and fuse statistical information coming from many different anatomical features will need to be developed in order to confirm anatomical findings. For the brain variability, one could for instance add to the sulci other cortical landmarks like sulcal ribbons and gyri, the surface of internal structures like the ventricles, the hippocampus or the corpus callosum, or fiber pathways mapped from DTI. These sources of information are individually providing a partial and biased view of the whole variability. Thus, we expect to observe a good agreement in some areas, and complementary measures in other areas. This will most probably lead in a near future to new anatomical findings and more robust medical image analysis applications.

6 REFERENCES

[1] A. Andrade, F. Kherif, J.-F. Mangin, K. Worsley, A.-L. Paradis, O. Simon, S. Dehaene, and J.-B. Poline. Detection of fMRI activation using cortical surface mapping. *Human Brain Mapping*, 12:79–93, 2001.

[2] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-Euclidean framework for statistics on diffeomorphisms. In *Proc. of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I*, number 4190 in LNCS, pages 924–931, 2-4 October 2006.

[3] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple calculus on tensors in the log-Euclidean framework. In J. Duncan and G. Gerig, editors, *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part I*, volume 3749 of *LNCS*, pages 115–122, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag. PMID: 16685836.

[4] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2006.

[5] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, August 2006. PMID: 16788917.

[6] J. Ashburner and K. J. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 2000.

[7] G. Aubert and P. Kornprobst. *Mathematical problems in image processing - Partial differential equations and the calculus of variations*, volume 147 of *Applied Mathematical Sciences*. Springer, 2001.

[8] P. Basser, J. Mattiello, and D. L. Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.

[9] P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53:221–225, 2005.

[10] M. Beg, M. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. Journal of Computer Vision*, 61(2):139–157, 2005.

[11] J. Boisvert, F. Cheriet, X. Pennec, N. Ayache, and H. Labelle. A novel framework for the 3D analysis of spine deformation modes. In *Research into Spinal Deformities*, volume 123 of *Studies in Health Technology and Informatics*, pages 176–182, 2006. PMID: 17108423.

[12] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache. Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Transactions on Medical Imaging*, 2008. In press.

[13] J. Boisvert, X. Pennec, N. Ayache, H. Labelle, and F. Cheriet. 3D anatomic variability assesment of the scoliotic spine using statistics on Lie groups. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, pages 750–753, Crystal Gateway Marriott, Arlington, Virginia, USA, April 2006. IEEE.

[14] J. Boisvert, X. Pennec, H. Labelle, F. Cheriet, and N. Ayache. Principal spine shape deformation modes using Riemannian geometry and articulated models. In *Proc of the IV Conference on Articulated Motion and Deformable Objects, Andratx, Mallorca, Spain, 11-14 July*, volume 4069 of *LNCS*, pages 346–355. Springer, 2006. AMDO best paper award 2006.

[15] F. Bookstein. *The Measurement of Biological Shape and Shape Change*, volume 24 of *Lecture Notes in Biomathematics*. Springer-Verlag, 1978.

[16] T. Brox, J. Weickert, B. Burgeth, and P. Mrázek. Nonlinear structure tensors. *Image and Vision Computing*, 24(1):41–55, 2006.

[17] A. Brun. *Manifolds in Image Science and Visualization*. PhD thesis, Linköping University, 2007. Linköping Studies in Science and Technology Dissertions No 1157.

[18] J. Burbea and C. Rao. Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. *Journal of Multivariate Analysis*, 12:575–596, 1982.

[19] M. Calvo and J. Oller. An explicit solution of information geodesic equations for the multivariate normal model. *Statistics and Decisions*, 9:119–138, 1991.

[20] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, June 1998.

[21] B. Davis, P. Fletcher, E. Bullitt, and S. Joshi. Population shape regression from random design data. In *Proc. of ICCV'07*, 2007.

[22] J.-P. Dedieu, G. Malajovich, and P. Priouret. Newton method on Riemannian manifolds: Covariant alpha-theory. *IMA Journal of Numerical Analysis*, 23:395–419, 2003.

[23] D. Ducreux, P. Fillard, D. Facon, A. Ozanne, J.-F. Lepeintre, J. Renoux, M. Tadié, and P. Lasjaunias. Diffusion tensor magnetic resonance imaging and fiber tracking in spinal cord lesions: Current and future indications. *Neuroimaging Clinics of North America*, 17(1):137–147, February 2007.

[24] S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Measuring brain variability via sulcal lines registration: a diffeomorphic approach. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 4791 of *LNCS*, pages 675–682, Brisbane, Australia, October 2007. Springer.

[25] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.

[26] P. Fillard, V. Arsigny, N. Ayache, and X. Pennec. A Riemannian framework for the processing of tensor-valued images. In O. F. Olsen, L. Florak, and A. Kuijper, editors, *Deep Structure, Singularities, and Computer Vision (DSSCV)*, number 3753 in LNCS, pages 112–123. Springer Verlag, June 2005.

[27] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, pages 786–789, Crystal Gateway Marriott, Arlington, Virginia, USA, April 2006.

[28] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. *IEEE Transactions on Medical Imaging*, 26(11):1472–1482, Nov. 2007.

[29] P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage*, 34(2):639–650, January 2007.

[30] P. Fillard, V. Arsigny, X. Pennec, P. M. Thompson, and N. Ayache. Extrapolation of sparse tensor fields: Application to the modeling of brain variability. In G. Christensen and M. Sonka, editors, *Proc. of Information Processing in Medical Imaging 2005 (IPMI'05)*, volume 3565 of *LNCS*, pages 27–38, Glenwood springs, Colorado, USA, July 2005. Springer.

[31] P. Fillard, X. Pennec, P. Thompson, and N. Ayache. Evaluating brain anatomical correlations via canonical correlation analysis of sulcal

lines. In *Proc. of MICCAI'07 Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation*, Brisbane, Australia, 2007.

[32] P. Fletcher, S. Joshi, C. Lu, and S. Pizer. Gaussian distributions on Lie groups and their application to statistical shape analysis. In C. Taylor and A. Noble, editors, *Proc. of Information Processing in Medical Imaging (IPMI'2003)*, volume 2732 of *LNCS*, pages 450–462. Springer, 2003.

[33] P. T. Fletcher and S. C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis, ECCV 2004 Workshops CVAMIA and MMBIA, Prague, Czech Republic, May 15, 2004*, volume 3117 of *LNCS*, pages 87–98. Springer, 2004.

[34] M. Fleute and S. Lavalle. Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. In Springer, editor, *Proc. of Medical Image Computing and Computer-Assisted Interventation (MICCAI'98)*, volume 1496 of *LNCS*, pages 879–887, 1998.

[35] G. Gerig, R. Kikinis, O. Kübler, and F. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–232, June 1992.

[36] S. C. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Processing*, 9(8):1357–1370, 2000.

[37] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications in Pure and Applied Mathematics*, 30:509–541, 1977.

[38] M. Kendall and P. Moran. *Geometrical probability.* Number 10 in Griffin's statistical monographs and courses. Charles Griffin & Co. Ltd., 1963.

[39] W. Kendall. Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence. *Proc. London Math. Soc.*, 61(2):371–406, 1990.

[40] D. Le Bihan, J.-F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. *Journal Magnetic Resonance Imaging*, 13(4):534–546, 2001.

[41] G. Le Goualher, E. Procyk, D. Collins, R. Venugopal, C. Barillot, and A. Evans. Automated extraction and variability analysis of sulcal neuroanatomy. *IEEE Transactions on Medical Imaging*, 18(3):206–217, 1999.

[42] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision*, 25(3):423–444, Oct. 2006.

[43] R. Mahony and R. Manton. The geometry of the newton method on non-compact lie groups. *Journal of Global Optimization*, 23:309–327, 2002.

[44] J.-F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, D. L. Collins, A. C. Evans, and J. Régis. Object-based morphometry of the cerebral cortex. *IEEE Transactions on Medical Imaging*, 23(8):968–982, Aug. 2004.

[45] J.-F. Mangin, D. Rivière, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, P. Scifo, T. Ochiai, F. Brunelle, and J. Régis. A framework to study the cortical folding patterns. *NeuroImage*, 23(Supplement 1):S129–S138, 2004.

[46] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci*, 356:1293–1322, 2001.

[47] G. Medioni, M.-S. Lee, and C.-K. Tang. *A Computational Framework for Segmentation and Grouping*. Elsevier, 2000.

[48] E. Meijering. A chronology of interpolation: From ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, 90(3):319–342, March 2002.

[49] M. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Re-view of Biomedical Engineering*, pages 375–405, 2003.

[50] M. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1/2):61–84, 2001.

[51] M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 2006.

[52] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal of Matrix Analysis and Applications*, 26(3):735–747, 2005.

[53] B. Mohammadi, H. Borouchaki, and P. George. Delaunay mesh generation governed by metric specifications. Part II: applications. *Finite Elements in Analysis and Design*, pages 85–109, 1997.

[54] K. Nomizu. Invariant affine connections on homogeneous spaces. *American J. of Math.*, 76:33–65, 1954.

[55] B. Owren and B. Welfert. The newton iteration on Lie groups. *BIT Numerical Mathematics*, 40(1):121–145, 2000.

[56] X. Pennec. Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements. In A. Cetin, L. Akarun, A. Ertuzun, M. Gurcan, and Y. Yardimci, editors, *Proc. of Nonlinear Signal and Image Processing (NSIP'99)*, volume 1, pages 194–198, June 20-23, Antalya, Turkey, 1999. IEEE-EURASIP.

[57] X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, July 2006. A preliminary appeared as INRIA RR-5093, January 2004.

[58] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.

[59] X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. Riemannian elasticity: A statistical regularization framework for non-linear registration. In J. Duncan and G. Gerig, editors, *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part II*, volume 3750 of *LNCS*, pages 943–950, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag. PMID: 16686051.

[60] X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, December 1997.

[61] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 12(7):629–639, 1990.

[62] H. Poincaré. *Calcul des probabilités.* 2nd edition, Paris, 1912.

[63] K. Rajamani, S. Joshi, and M. Styner. Bone model morphing for enhanced surgical visualization. In IEEE, editor, *Proc of IEEE Symp. on Biomedical Imaging: Nano to Macro (ISBI) 2004*, volume 2, pages 1255–1258, Apr. 2004.

[64] L. Skovgaard. A Riemannian geometry of the multivariate normal model. *Scand. J. Statistics*, 11:211–223, 1984.

[65] G. Subsol, J.-P. Thirion, and N. Ayache. A scheme for automatically building 3D morphometric anatomical atlases: application to a skull atlas. *Medical Image Analysis*, 2(1):37–60, 1998.

[66] J. Talairach and P. Tournoux. *Co-Planar Stereotaxic Atlas of the Human Brain: 3-dimensional Proportional System : an Approach to Cerebral Imaging.* Thieme Medical Publishers, New York, 1988.

[67] P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758, July 2000.

[68] P. Thompson, D. MacDonald, M. Mega, C. Holmes, A. Evans, and A. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21(4):567–581, 1977.

[69] P. Thompson, M. Mega, R. Woods, C. Zoumalan, C. Lindshield, R. Blanton, J. Moussai, C. Holmes, J. Cummings, and A. Toga. Cortical change in alzheimer's disease detected with a disease-specific population-based brain atlas. *Cerebral Cortex*, 11(1):1–16, January 2001.

[70] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.

[71] M. Vaillant, M. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23(Supp. 1):S161–S169, 2004.

[72] M. Vaillant, A. Qiu, J. Glauns, and M. Miller. Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus. *NeuroImage*, 34(3):1149–1159, 2007.

[73] J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix-valued images. In M. Nashed and O. Scherzer, editors, *Inverse Problems, Image Analysis, and Medical Imaging.*, volume 313 of *Contemporary Mathematics*, pages 251–268, Providence, 2002. AMS.

[74] J. Weickert and H. Hagen, editors. *Visualization and Processing of Tensor Fields*. Mathematics and Visualization. Springer, 2006.