Sparse Approximation of Currents for Statistics on Curves and Surfaces

Stanley Durrleman, Xavier Pennec, Alain Trouvé, Nicholas Ayache

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Sparse Approximation of Currents for Statistics on Curves and Surfaces

Stanley Durrleman — Xavier Pennec — Alain Trouvé — Nicholas Ayache

N° 6571

June 2008

Thème BIO

ISSN 0249-6399 ISRN INRIA/RR--6571--FR+ENG

Sparse Approximation of Currents for Statistics on Curves and Surfaces

Stanley Durrleman* † , Xavier Pennec* , Alain Trouvé† , Nicholas Ayache*

Thème BIO — Systèmes biologiques
Projet Asclepios

Rapport de recherche n° 6571 — June 2008 — 12 pages

Abstract: Computing, processing, visualizing statistics on shapes like curves or surfaces is a real challenge with many applications ranging from medical image analysis to computational geometry. Modelling such geometrical primitives with currents avoids feature-based approach as well as point-correspondence method. This framework has been proved to be powerful to register brain surfaces or to measure geometrical invariants. However, if the state-of-the-art methods perform efficiently pairwise registrations, new numerical schemes are required to process groupwise statistics due to an increasing complexity when the size of the database is growing. Statistics such as mean and principal modes of a set of shapes often have a heavy and highly redundant representation. We propose therefore to find an adapted basis on which mean and principal modes have a sparse decomposition. Besides the computational improvement, this sparse representation offers a way to visualize and interpret statistics on currents. Experiments show the relevance of the approach on 34 sets of 70 sulcal lines and on 50 sets of 10 meshes of deep brain structures.

Key-words: shape statistics , curve , surface , current , matching pursuit , approximation , sparse representation , adapted basis , computational anatomy , brain , sulcal lines , deep brain structures

* Asclepios team project, INRIA Sophia Antipolis, France
† Centre de Mathématiques et Leurs Applications (CMLA), ENS-Cachan, France

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Approximations parcimonieuses de courants pour l'étude statistique d'ensemble de courbes et de surfaces

Résumé : Traiter, manipuler ou représenter des statistiques sur des ensemble de courbes ou de surfaces reste un véritable defi dont les applications potentielles sont nombreuses autant en analyse d'image médicale qu'en géometrie algorithmique. La modélisation des primitives géometriques à l'aide de courants permet d'éviter de baser l'analyse sur des caractéristiques extraites *a priori* ou sur des correspondances de points entre les objets. Ce modèle a montré son efficacité et sa pertinence pour le recalage de surfaces cérébrales ou la mesure d'invariants géometriques. Cependant, si les méthodes les plus récentes permettent de recaler efficacement deux observations l'une sur l'autre, des schema numériques nouveaux sont nécessaires pour manipuler des statistiques déduites d'un ensemble d'observations. La complexité de la représentation des statistiques (moyenne ou modes propres de variations) croît linéairement avec la taille de la base de donnés, rendant ainsi caduque l'utilisation des algorithmes existants. Cependant, ces représentations sont souvent très redondantes à l'échelle d'analyse considéré. Nous proposons donc d'approximer ces statistiques en les décomposant sur des bases de vecteurs adaptés. Au delà de l'amélioration des temps de calcul, ces représentations parcimonieuses permettent de visualiser et donc d'interpréter plus facilement les statistiques sur les courants. Cette approche est testée sur les bases de données anatomiques comportant 34 ensembles de 70 lignes sulcales et 50 ensembles de 10 structures sous-corticales du cerveau.

Mots-clés : statistiques de formes , courbe , surface , courant , poursuite de vecteurs , representation éparse , representation parcimonieuse , base adaptée , anatomie computationnelle , cerveau , lignes sulcales , structures sous-corticales

1 Introduction

There is a substantial need for shape statistics in many communities ranging from medical imaging to computer vision or computational geometry, for example to incorporate shape priors in image segmentation, to analyze geometrical or anatomical differences between groups, to classify new observations according to some characteristics, to shape recognition purpose, etc. The method we propose here precisely enables to learn a statistical shape model (mean and principal modes) from a database of curves or surfaces. Over the last years, many methods have been proposed for encoding statistical priors on curves, quite less were proposed for surfaces. Among them, level-sets for instance [1], use geometrical constraints (length, area, volume, curvature, etc.) as priors. However, it is difficult to figure out how to automatically learn such priors from typical datasets. Modelling based on Point Distribution Model [2] assumes exact or at least fuzzy correspondences between point sets. Medial axis representations [3] requires to completely specify the topology of shapes.

To define a mean shape, a "covariance" matrix and the principal modes via principal component analysis (PCA), one usually requires some metric properties on the shape space itself. Currents were precisely introduced as an interesting alternative to do that for shape registrations [4, 5, 6]. This modelling enables to define an inner product and hence an easy to compute distance between shapes *without assuming any point correspondences* between discrete structures and without *selecting arbitrary features*. Since the space of currents is linear, one could *theoretically* compute directly standard statistics on shapes such as mean or PCA. However, there is a drawback to this mathematically well grounded method: the polynomial computational complexity in the number of points in shapes (which was not critical for pairwise registration) becomes a clear bottleneck for groupwise statistics. We propose in this paper an efficient computational framework that overcomes these limitations by providing a sparse representation of currents at any desired accuracy. Our algorithm builds on ideas from the approximation theory previously developed to decompose images in wavelet bases [7, 8]. To the very best of our knowledge, this is the first time that these signal processing techniques are used in geometric shape analysis. Experimental results clearly demonstrate the interest of our method: the deformation of a mean obtained from 3 shape instances such as in Fig.3, each with 15 000 points, which was previously taking 10 hours, is now taking about 5 minutes (using the same code as in [4, 6]). For a set of 50 instances, representing the mean requires 1.2 Kb in our framework, versus 8 Mb originally. Deforming the former still requires 5 minutes while it is not feasible to deform the later without high performance computing. This offers a way to the automated learning of shape priors from typical databases, which can be used, for instance, to adapt the generic model to each particular observation for image segmentation. This also gives a way to visualize and interpret mean and principal modes of a set of anatomical structures.

In the first section, we present how shapes like curves and surfaces may be modeled geometrically as currents. We then introduce a sparse deconvolution scheme to approximate currents. In the last section we show on real anatomical data how this tool may compress shapes and may help to visualize and interpret statistics on shapes.

2 Non-parametric Representation of Shapes as Currents

In the framework of currents, a curve or a surface is characterized by the way it integrates vector fields. Their construction is based therefore on a space of square integrable vector fields W. The space of currents W^* is the set of continuous mapping from W to $\mathbb{R}$ (see [5, 9, 10] for more details). A piecewise smooth surface S in $\mathbb{R}^3$ like a set of smooth surfaces or meshes defines a current thanks to:

$$S : \omega \longrightarrow \int_S \langle \omega(x), u_x \times v_x \rangle_{\mathbb{R}^3} \, d\sigma(x) \tag{1}$$

where (u_x, v_x) is an orthogonal basis of the tangent plane of the surface at the point x and $n_x = u_x \times v_x$ the normal of the surface defined almost everywhere. This measures the flux of the vector field through the surface. Similarly, a piecewise smooth curve L in $\mathbb{R}^3$ such as a set of smooth curves or polygonal lines is a current thanks to:

$$\forall \omega \in W, L(\omega) = \int_L \langle \omega(x), \tau(x) \rangle_{\mathbb{R}^3} \, dx \tag{2}$$

where $\tau(x)$ is the tangent vector at point x defined almost everywhere. This is the integral of the vector field along the path defined by L. To ensure that the mappings 1 and 2 are continuous (and hence $S, L \in W^*$) we need to assume that any $\omega \in W$ is bounded and that for a fixed constant C:$\|\omega\|_\infty \leq C \|\omega\|_W$. This technical condition guarantees in particular that small errors measured in W are numerically small. As a consequence, W is reproducing kernel Hilbert space (r.k.h.s.) with kernel K[10]: for every $\omega \in W$ and $\alpha \in \mathbb{R}^3$:

$$\langle \omega(x), \alpha \rangle_{\mathbb{R}^3} = \langle \omega, K(., x)\alpha \rangle_W \tag{3}$$

We choose here W as the r.k.h.s. whose kernel is Gaussian: $K(x, y) = \exp(-\|x - y\|^2 / \lambda_W^2)\mathrm{Id}$.

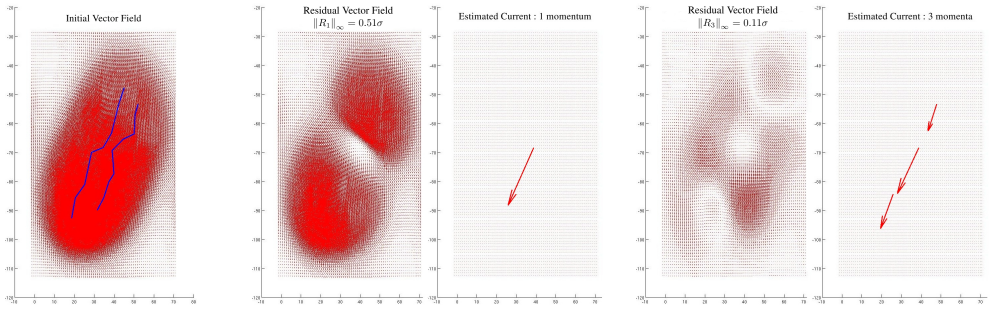
In this setting, the infinite dimensional space W^* is the dense span of Dirac currents: δ_x^α that may be seen as a tangent (or normal) α entirely concentrated at point x:

$$\delta_x^\alpha(\omega) = \langle \omega(x), \alpha \rangle_{\mathbb{R}^3} \tag{4}$$

for any $\omega \in W$. A curve (resp. a surface) is decomposed into its infinite set of tangents (resp. normals). Combining Eq.4 and 3 leads to: $\delta_x^\alpha(\omega) = \langle \omega, K(x, .)\alpha \rangle_W$ for any ω. This shows that we can associate to every basis element δ_x^α a **dual representation** in terms of vector field: $K(x, .)\alpha$ which is simply the Gaussian convolution of the vector α. Thanks to the theory [5, 9], we can extend this result: there is a one-to-one mapping $\mathcal{L}_W : W \longrightarrow W^*$ that maps any vector field ω to a current T such that $T(\omega) = \left\langle \mathcal{L}_W^{-1}(T), \omega \right\rangle_W$. With this notation, $\mathcal{L}_W(K(., x)\alpha) = \delta_x^\alpha$ showing that K is the Green function of the operator $\mathcal{L}_W$. δ_x^α is called the momentum associated to the vector field $K(., x)\alpha$. $\mathcal{L}_W$ enables also to provide W^* with an inner product:

$$\langle T, T' \rangle_{W^*} = \left\langle \mathcal{L}_W^{-1}(T), \mathcal{L}_W^{-1}(T') \right\rangle_W \tag{5}$$

On basis elements, this gives: $\left\langle \delta_x^\alpha, \delta_y^\beta \right\rangle_{W^*} = \alpha^t K(x, y)\beta$.



a-Initial Configuration b-iteration 1 c-iteration 3

Figure 1: A sparse deconvolution scheme: Gaussian convolution of the mean of the two blue lines with $\lambda_W = 15$ (a). First (b) and third (c) iterations: estimated momenta on the right, residual vector field on the left. σ is the standard deviation $\|L - L'\|_\infty / \sqrt{2}$. The momenta converge to the true solution while the residue tends to zero.

Suppose that we have a set of N discretized shapes (polygonal lines or meshes). These shapes are close in W^* to the finite set of their tangent (resp. normal) at the center of the segments (resp. the center of mass of the mesh cell). Their mean or principal modes are linear combinations of the input shapes and therefore of the form: $\tilde{T} = \sum_{i=1}^{N} \lambda_i T_i = \sum_{k=1}^{N_T} \delta_{x_k}^{\alpha_k}$ where N_T is the total number of tangents or normals in the dataset and may be therefore very large. This representation, if exact, is far from being optimal: it may be highly redundant at the scale λ_W. The Gaussian convolution of the N_T momenta: $\gamma = \mathcal{L}_W^{-1}(\tilde{T})$ precisely integrates this redundancy (See Fig.1-a). The two representation (T and $\mathcal{L}_W^{-1}(T)$) are equivalent: theoretically the deconvolution of γ would retrieve exactly T. But the deconvolution is an ill-posed problem and requires specific numerical schemes as the matching pursuit algorithm we propose here. This method estimates iteratively adapted Dirac currents $\delta_{x_{k'}}^{\alpha_{k'}}$ such that the series $\sum_{k'} \delta_{x_{k'}}^{\alpha_{k'}}$ approximates the true solution $T = \mathcal{L}_W(\gamma)$ but with faster decreasing terms than the initial decomposition. The first terms of this series provide therefore an approximation of the true solution with an increasing precision (see Fig.1-b,c). We can then expect that the deconvolution of γ will return a sparser representation of $\tilde{T}$.

3 A Sparse Deconvolution Scheme

Matching Pursuit Algorithms were proposed in [11, 8] to find adapted wavelets bases for image decomposition. We adapt here the idea of how to find a basis adapted to a particular signal to our framework based on currents. Given a vector field $\gamma \in W$, we want to find an approximation of T that solves $\mathcal{L}_W^{-1}(T) = \gamma$. The goal is to estimate a set of N points (x_i) and N vectors (α_i) such that the current $\Pi(T) = \sum_{i=1}^{N} \delta_{x_i}^{\alpha_i}$ is the closest possible to T. If one knows the optimal points, $\Pi(T)$ is the orthogonal projection of T onto Span $\left(\delta_{x_i}^{\epsilon_k}; k = 1, 2, 3, i = 1 \dots N\right)$ where $(\epsilon_k)_{k=1,2,3}$ is the canonical basis of $\mathbb{R}^3$. (In the following the index k always takes values $k = 1, 2, 3$). The

orthogonality conditions are $\left\langle T, \delta_{x_i}^{\epsilon_k} \right\rangle_{W^*} = \left\langle \Pi(T), \delta_{x_i}^{\epsilon_k} \right\rangle_{W^*}$, which gives (applying $\mathcal{L}_W^{-1}$ and Eq.5) the linear set of $3N$ equations:

$$\sum_{p=1}^{N} \left(K(x_i, x_p)\alpha_p\right)_k = \gamma(x_i)_k \tag{6}$$

The search for the optimal points (x_i) has been proved to be NP-hard in general [7]. The orthogonal matching pursuit algorithm is a suboptimal greedy approach to this problem: the first point x_1 is the one for which the projection of T on $\delta_x^{\varepsilon_k}$ is maximal. Since $\langle \mathcal{L}_W(\gamma), \delta_x^{\varepsilon_k} \rangle_{W^*} = \langle \gamma, K(.,x)\varepsilon_k \rangle_W = \gamma(x)_k$, x_1 is a point where $\|\gamma(x)\|$ is maximal. Solving Eq.6 gives $\alpha_1^1 = K(x_1,x_1)^{-1}\gamma(x_1)$. We then remove from γ its orthogonal projection on $\mathcal{L}_W^{-1}(\delta_{x_1}^{\varepsilon_k})$ and we iterate the procedure on the residue: $\gamma_1 = \gamma - K(.,x_1)\alpha_1^1$. Finally, the algorithm can be then written as:

- Input: a vector field γ, a threshold $\eta > 0$
- Initialization: $\gamma_0 = \gamma$, $N = 0$
- While $\|\gamma_N\|_\infty \geq \eta$ do:
  1. $x_{N+1} = \text{argmax}_{x \in \mathbb{R}^3} \|\gamma_N(x)\|_{\mathbb{R}^3}$
  2. Find $(\alpha_i^{N+1})_{1 \leq i \leq N+1}$ by solving $\sum_{p=1}^{N+1} \left(K(x_i,x_p)\alpha_p^{N+1}\right)_k = \gamma(x_i)_k$
  3. $\gamma_{N+1} = \gamma - \sum_{i=1}^{N+1} K(.,x_i)\alpha_i^{N+1}$
  4. $N = N+1$

After N steps, the algorithm gives an approximation of $T = \mathcal{L}_W(\gamma)$ with N momenta: $\Pi_N(T) = \sum_{i=1}^{N} \delta_{x_i}^{(\alpha_i^N)}$. We prove in appendix A that $\Pi_N(T)$ converges to T as N tends to infinity (e.g. $\|\Pi_N(T) - T\|_{W^*} \longrightarrow 0$). The auxiliary variable $\gamma_N = \gamma - \mathcal{L}_W^{-1}(\Pi_N)$ stores the residual vector field that remains to be explained. We prove also in appendix A that the L^∞-norm of this residue tends towards zeros as N tends towards infinity (e.g. $\|\gamma_N\|_\infty \longrightarrow 0$). This means, in particular, that the norm of the residue is below any positive threshold in finite time, thus proving that the algorithm finishes.

From a computational point of view, we suppose that the input vector field γ is sampled on a linearly spaced grid Λ (supposed to be large enough to assume periodic boundary conditions). It can be stored therefore as an image of $3D$ vectors. At step 1, we restrict ourselves to find the maximum of $\|\gamma(x)\|_{\mathbb{R}^3}$ on the grid's points, such that the estimated position x_i always belong to the grid. At step 3, one computes the vector field: $\mathcal{L}_W^{-1}(\Pi_N)(x) = \sum_{i=1}^{N} K(x,x_i)\alpha_i^N$ at grid's point. Since K is translation-invariant, this Gaussian convolution can be efficiently computed by FFT's. γ_N is also stored as an image of vectors. The output $\Pi_N(T)$ is stored as a list of (position,vector).

One wants to use this algorithm to find a sparse representation of an input current $T = \sum_{s=1}^{N_T} \delta_{y_s}^{\beta_s}$ where N_T is a priori very large. For this purpose, we computes $\gamma(x) = \mathcal{L}_W^{-1}(T)(x) = \sum_{s=1}^{n} K(x,x_s)\beta_s$. Since γ is band-limited, we sample it on a linearly spaced grid Λ with a step Δ such that Δ/λ_W is small (typically less than $1/5$). We fix the grid such that the data are further than λ_W of the

grid's border. This sampling leads to an image of vectors used as input of the algorithm. The threshold η has to be specified for every applications. If T is a linear combination of N_{obs} currents $T_1, \ldots, T_N$ (such as the mean current, a principal mode, the difference between two currents, etc.), we can choose η as a fixed ratio of the standard deviation of the set of currents: $\eta = \tau\sigma$ where $\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} \left\|T_i - \bar{T}\right\|^2$ and $\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T_i$. This means that the algorithm finishes when the approximation's error is smaller than $\tau\%$ of the variance.

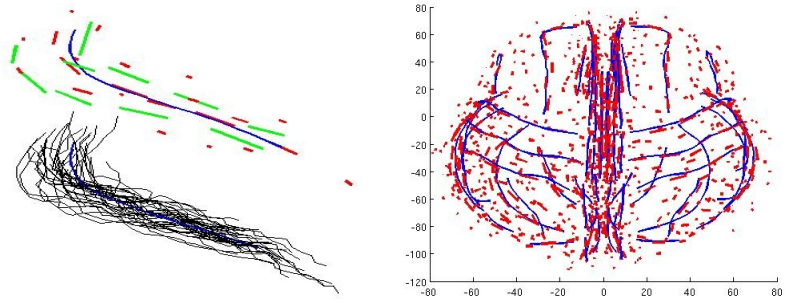
Finally, for a given current T we have 3 distinct representations: the initial one with N_T momenta (the total number of segments or mesh's cells in the set of currents), the projection of T onto the grid with *a priori* N_{grid} momenta (which depends on both the spreading of the input data and the step Δ) and our sparse representation with N_{mom} momenta. Depending on the number of input points, their spreading and redundancy at the scale λ_W, these 3 representations can vary dramatically in size.

4 Numerical Experiments

Curves in 3D We use a set of 70 sulci delineated in $N_{obs} = 34$ subjects. The sulci are the fissures on the brain surface and they are often used to measure anatomical differences between subjects [12]. For each sulcal line, we approximate the mean current $\bar{L} = \frac{1}{N_{obs}}\sum_{i=1}^{N_{obs}} L_i$ for $\lambda_W = 12$mm and $\tau = 5\%$. Results are shown in Fig.2-a for the Sylvian Fissure of the right hemisphere and 2-b for all 70 sulci. The initial number of momenta for the mean fissure was $N_T = 899$ (i.e. the number of segments of all lines) whereas the final approximation needs only $N_{mom} = 54$ momenta. In this case, the compression ratio is of 94%. Considering all sulci, the compression rate is on average: $94.8\% \pm 0.02$. Each lines' grid have a step $\Delta = 2$mm and typically $N_{grid} = 10^5$ points. Our mean is visually in good agreement with other mean curves computing from B-spline representation [13].

We then compute the eigenmodes of the lines sets by PCA. We find the eigenvectors (V_k) of the $N_{obs} \times N_{obs}$ matrix $\Sigma = \left(\left\langle L_i - \bar{L}, L_j - \bar{L}\right\rangle_{W^*}\right)_{i,j=1\ldots N_{obs}}$. The k^{th} eigenmode is given by the linear combination of input currents: $m_k = \sum_i (V_k)_i (L_i - \bar{L})$. We approximate the first eigenmode of the Sylvian Fissure of the right hemisphere (Fig.2-a): this mode captures the spreading of the lines set.

Surfaces We compute the mean current of 10 meshes of subcortical brain structures (Caudate, Putamen, Globus Pallidus, Amygdala and Hippocampus for each hemisphere) for $N_{obs} = 50$ subjects [14]. For surfaces, we represent the estimated momenta (normal of an infinitesimal mesh cell) as equilateral triangles whose normals correspond to the momenta. Results are shown in Fig.3. The compression ratio between N_T and N_{mom} for the 10 structures is on average of $99.96\% \pm 10^{-4}$. Each structures' grid has a step $\Delta = 1$mm and for one structure we have the following typical values: $N_T = 50 \times 3000 = 1.5e^5$, $N_{grid} = 3e^5$ and $N_{mom} = 100$. Fig.3-c shows that the quality of approximation remains good until very high compression ratio. Similarly, we compute mean from meshes of 7 controls. The difference between both means is a current that we approximate: the arrows of Fig.4-b are the 10 first estimated momenta of this difference, suggesting that the autistic mean is more curved at the Hippocampus' extremity and thicker in the middle. Such visual results need to be confirmed by rigorous statistical results.



a- Sylvian Fissure b- All 70 sulci (top view of the brain)

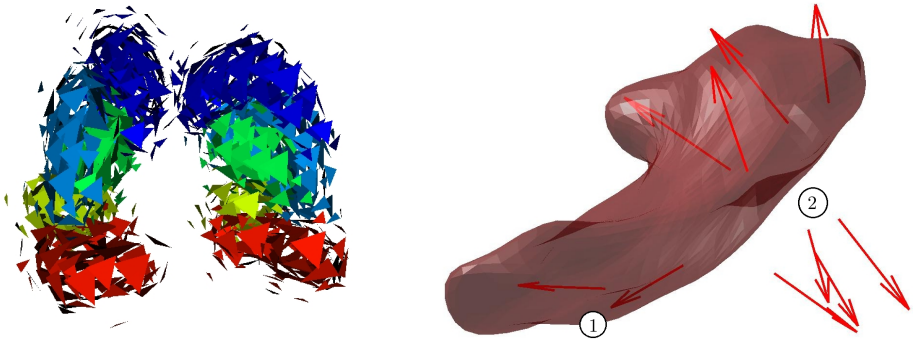
Figure 2: Statistics for 70 sulci in 34 subjects ($\lambda_W = 12$mm, $\tau = 5\%$). Left: all set of lines (black), their mean (red) and first eigenmode at σ (green) showing the spreading of the lines set. Right: Mean currents (red) compared to the mean lines (blue) computed from B-spline parametrization of curves [13]. Results are in good agreement.

a- structures of 2 autistics b- Mean of 50 autistics c-approx. vs. compression

Figure 3: Approximation of the mean current for 10 structures segmented in 50 subjects (b) ($\lambda_W = 5$mm, $\tau = 5\%$ with data of diameter 60mm). Good approximation's quality can be achieved until very high compression ratio (c). Red points correspond to the approximation shown in b, for which the error equals $\tau = 5\%$ of the variance.

5 Discussion and Conclusion

The method presented here enables to approximate, at any desired accuracy, the mean and principal modes of a set of curves or surfaces. On the one hand, this approximation gives a way to visualize and hence to interpret such statistics on shapes as emphasized by our results on real anatomical datasets. On the other hand, the very high compression ratio we achieve on real data offers a way to include such statistics on registration scheme for example. A registration algorithm can indeed deforms an input shape with a spatial and temporal complexity of $f(N_T)$ [4] (f is between linear



a- Mean of 7 controls b- Difference between both means on Hippocampus

Figure 4: a-Mean of 7 controls.b-difference between mean of autistics and mean of controls (arrows) superimposed with the Hippocampus of a control. Mean from autistics is more curved at hippocampus' extremity (area 1) and thicker in area 2.

and quadratic function) or solve evolution differential equations on a grid [15] with a complexity $f(N_{grid})$. In the first case our representation with N_{mom} momenta improves greatly both spatial and temporal complexity of the method while guaranteeing a fixed approximation error. This makes now possible to fit statistical surfaces to image data for segmentation purpose (deformable models). This would also be useful to adapt the template estimation framework set up in [16] in case of images and small deformations to estimate shape prototypes from a large set of geometrical primitives. Our method could have also a significant impact in several other fields including surface rendering and animation in computer graphics, surface reconstruction and re-meshing in computational geometry, surface registration and fusion for object modeling in 3D computer vision, statistics on the cortex surfaces over multiple subjects for detecting activations from fMRI in neuroscience.

For pure visualization purposes, however, the method presented here need further improvements for a better rendering of the approximated currents. For this purpose, we investigate how to bridge the gap with standard computer graphics methods like splats for instance.

Acknowledgments We warmly thank Prof. Thompson (UCLA) for providing the sulcal lines database through the Asclepios-LONI associated team `BrainAtlas`, as well as Prof. Gerig (University of Utah) for giving access to the database of deep brain structures. The work was partly supported by the European IP projet Health-e-Child (IST-2004-027749).

A Proof of convergence of the orthogonal matching pursuit algorithm

Given a vector field $\gamma = \mathcal{L}_W^{-1}(T)$, the algorithm, as explained in Section 3, estimates at each step n a current Π_n that is supposed to approximate the true solution T. To prove this fact, we show that the infinite norm of the residual vector field $\left\|\mathcal{L}_W^{-1}(\Pi_n) - \gamma\right\|_\infty$ as well as the W^*-norm of the residual current $\|T - \Pi_n\|_{W^*}$ ($= \left\|\gamma - \mathcal{L}_W^{-1}(\Pi_n)\right\|_W$) tends to zero as n tends towards infinity.

We denote by $E_n \doteq \text{Span}(\delta_{x_i}^\alpha; 1 \leq i \leq n, \alpha \in \mathbb{R}^d)$ for the iteratively estimated points (x_i). The estimated current Π_n is defined as the W-orthogonal projection onto E_n. We denote R_n the residue so that $\mathcal{L}_W(\gamma) = \Pi_n + R_n$. The corresponding residual error in W is denoted by $\gamma_n = \mathcal{L}_W^{-1}(R_n)$.

A.1 Convergence for L^∞ norm

We assume that $\langle K(x,x)u, u\rangle \geq c|u|^2$ for any $x, u \in \mathbb{R}^d$. This is the case if K is translation invariant, since then $K(x,x) = K(0)$. At each iteration, due to the numerical implementation, we choose $x_{n+1} = x_*$ so that

$$|\gamma_n(x_*)| \geq \max_{x \in \mathbb{R}^3} |\gamma_n(x)| \,/2 \tag{7}$$

instead of the point that reaches exactly the maximum.

For any $\alpha \in \mathbb{R}^d$, since $\Pi_n + \delta_{x_*}^\alpha \in E_{n+1}$

$$\|R_{n+1}\|_{W^*}^2 \leq \left\|\Pi_n + \delta_{x_*}^\alpha - T\right\|_{W^*}^2 \tag{8}$$

Minimizing the right-hand side of Eq.8 with respect to α leads to $\alpha_* = K(x_*, x_*)^{-1}\gamma_n(x_*)$ for which $\left\|\Pi_n + \delta_{x_*}^{\alpha_*} - T\right\|_{W^*}^2 = \left\|R_n + \delta_{x_*}^{\alpha_*}\right\|_{W^*}^2 = \|R_n\|_{W^*}^2 - \langle K(x_*,x_*)^{-1}\gamma_n(x_*), \gamma_n(x_*)\rangle \leq \|R_n\|_{W^*}^2 - c|\gamma_n(x_*)|^2$. Thus we get from (7) and (8) that $\|R_{n+1}\|_{W^*}^2 \leq \|R_n\|_{W^*}^2 - c\|\gamma_n\|_\infty^2/2$. Therefore the series $\|R_n\|_{W^*}$ is monotonically decreasing and hence converges. Moreover, $\frac{c}{2}\sum_{k=1}^\infty \|\gamma_k\|_\infty^2 \leq \|T\|_{W^*}^2$, from which we deduce that $\|\gamma_n\|_\infty \to 0$.

A.2 Convergence for W^*-norm

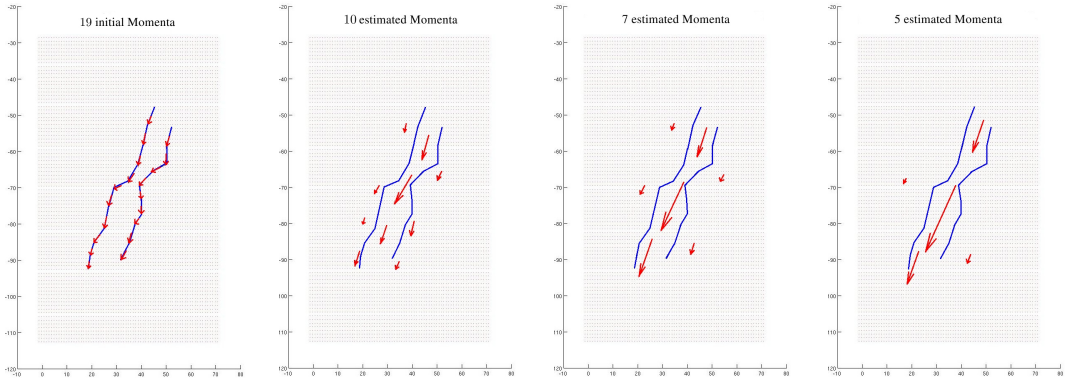
Introducing F_n such that $E_{n+1} = E_n \overset{\perp}{\oplus} F_n$ and p_n the W-orthogonal projection on F_n we have for $n \geq m$: $\|R_n - R_m\|_{W^*}^2 = \|\Pi_n - \Pi_m\|_{W^*}^2 = \sum_{k=m}^{n-1}\|p_k\|_{W^*}^2 \leq \|T\|_{W^*}^2$. Writing $R_m = \sum_{i=1}^m \delta_{x_i}^{\alpha_i}$, we get $\|R_n\|_{W^*}^2 \leq \langle R_n, R_m\rangle_{W^*} + \langle R_n, R_n - R_m\rangle_{W^*} \leq \sum_{i=1}^m |\alpha_i|\|\gamma_n\|_\infty + \|T\|_{W^*}^2 \sum_{k=m}^\infty \|p_k\|_{W^*}^2$. Thus, $\lim \|R_n\|_{W^*}^2 \leq \|T\|_{W^*}^2 \sum_{k=m}^\infty \|p_k\|_{W^*}^2$. Since $\sum_{k=0}^\infty \|p_k\|_{W^*}^2 < \infty$ we get the result for $m \to \infty$.

B A Toy Example in 2D

We apply the orthogonal matching pursuit to represent the mean current $\bar{L} = (L + L')/2$ of two polygonal lines L and L'. Fig.5-a shows the Dirac approximation of the mean current $\bar{L} =$

$\sum_{i=1}^{N} \delta_{c_i}^{\tau_i/2} + \sum_{j=1}^{N'} \delta_{c'_j}^{\tau'_j/2}$ projected on the closest grid's points. Fig.5-b,c show different approximations for different kernel's size λ_W with a fixed ratio $\tau = 5\%$ defined as the ratio between the norm of the residual vector field and the standard deviation $\sigma = \|L - L'\|_\infty / \sqrt{2}$. The greater λ_W, the more redundant the initial representation, the sparser our approximation for a fixed accuracy.

Fig.1-a shows the dual representation of the mean of two curves (Fig.5-a) which is the Gaussian convolution of Fig.5-a with $\lambda_W = 15$: $\mathcal{L}_W^{-1}(\bar{L})(x) = \frac{1}{2}\sum_{i=1}^{N} K(x, c_i)\tau_i + \frac{1}{2}\sum_{j=1}^{N} K(x, c'_j)\tau'_j$. First iterations of the algorithm for $\lambda_W = 15$ are shown Fig.1-b,c.



a-Initial mean current b-$\lambda_W = 10$ c-$\lambda_W = 15$ d-$\lambda_W = 20$

Figure 5: Initial mean of two lines ($\lambda_W \to 0$) and their approximation for increasing kernel's size λ_W and fixed rate $\tau = 5\%$. Original two lines are superimposed in blue. The greater λ_W, the more redundant the initial representation, the sparser our approximation for a fixed accuracy.

References

[1] Baillard, C., Hellier, P., Barillot, C.: Segmentation of brain 3D MR images using level sets and dense registration. Medical Image Analysis **5**(3) (2001) 185–194

[2] Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. Computer Vision and Image Understanding **89**(2-3) (2003) 114–141

[3] Gorczowski, K., Styner, M., Jeong, J.Y., Marron, J.S., Piven, J., Hazlett, H.C., Pizer, S.M., Gerig, G.: Statistical shape analysis of multi-object complexes. In: Computer Vision and Pattern Recognition CVPR, IEEE (2007) 1–8

[4] Vaillant, M., Glaunès, J.: Surface matching via currents. In: Proceedings of Information Processing in Medical Imaging. Volume 3565 of LNCS., Springer (2005) 381–392

[5] Glaunès, J.: Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique. PhD thesis, Université Paris 13, http://cis.jhu.edu/ joan/TheseGlaunes.pdf (2005)

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

[7] Davis, G., Mallat, S., Avellaneda, M.: Greedy adaptive approximations. Journal of Constructive Approximation **13**(1) (1997) 57–98

[8] Pati, Y., R.Rezaifar, Krishnaprasad, P.: Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Conf. Record of the 27th Asilomar Conference on Signals, Systems and Computers. Volume 1. (November 1993) 40–44

[9] Schwartz, L.: Sous espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). J. Analyse Math. **13** (1964) 115–256

[10] Aronszajn, N.: Theory of reproducing kernels. Transactions of the American Mathematical Society (68) (1950) 337–404

[11] Mallat, S., Zhang, Z.: Matching pursuits with time-frequency dictionaries. IEEE Transactions on Signal Processing **41**(12) (1993) 3397–3415

[12] Thompson, P., Schwartz, C., Lin, R., Khan, A., Toga, A.: 3D statistical analysis of sulcal variability in the human brain. Journal of Neuroscience **16**(13) (1996) 4261–4274

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

[14] Hazlett, H., Poe, M., Gerig, G., Smith, R., Provenzale, J., Ross, A., Gilmore, J., Piven, J.: Magnetic resonance imaging and head circumference study of brain size in autism. The Archives of General Psychiatry **62** (2005) 1366–1376

[15] Aubert, G., Kornprobst, P.: Mathematical problems in image processing - Partial differential equations and the calculus of variations. Volume 147 of Applied Mathematical Sciences. Springer (2001)

[16] Allassonnière, S., Amit, Y., Trouvé, A.: Towards a coherent statistical framework for dense deformable template estimation. Journal of the Royal Statistical Society Series B **69**(1) (2007) 3–29

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399