Shape Analysis and Computational Anatomy: A Geometrical Perspective on the Statistical Analysis of Population of Manifolds

Alain Trouvé

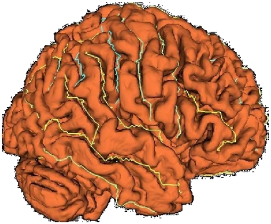
CMLA, Ecole Normale Supérieure de Cachan

Grenoble, Statlearn'11
2011-03-17

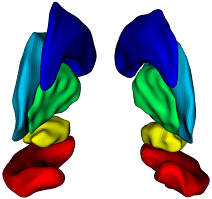
Outline

Anatomical shapes

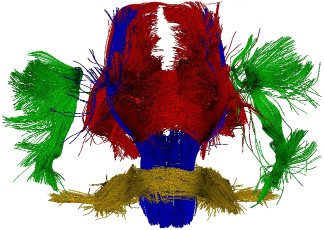
Few anatomical structures segmented in MRI



Sulcal Lines



Internal Structures

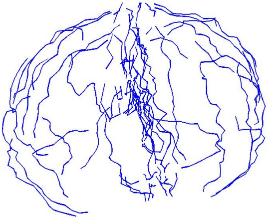


Fiber Bundles

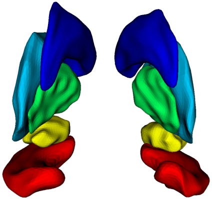
Various shape spaces: points, surfaces, pieces of submanifolds, grey-level images, tensor fields, etc.

Anatomical shapes

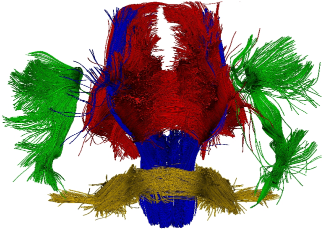
Few anatomical structures segmented in MRI



Sulcal Lines



Internal Structures



Fiber Bundles

Various shape spaces: points, surfaces, pieces of submanifolds, grey-level images, tensor fields, etc.

Trivial metric don't work

- Differentiable structure should be compatible with "smooth" transformation of a shape (e.g. geometrical transformation)
- Not true for the L^2 metric on image and "smooth" transformations

$$\tau \to \tau\dot{f}(x) \doteq f(x+\tau)$$

- $\tau \to t \cdot f$ is not smooth (but $\tau \to \tau$ is !).

A global geometrical setting through Riemannian submersion from a group of transformation onto a homogeneous space

- Consider a transitive group action i.e.

$$\begin{array}{ccc} G \times M & \to & M \\ (g, m) & \to & g.m \end{array}$$

  where $M = G\dot{m}_0$ with $m_0 \in M$ ("template").
- If G is equipped with a G_0 (isotropy group) equivariant metric then

$$d_M(m, m') = \inf\{\, d_G(g, g') \mid g.m_0 = m,\ g'.m_0 = m'\}$$

  is a distance on M (coming from the projected riemannian distance on M)

$$G/G_0 \simeq M$$

- Simple framework: Right invariant metric on G (standard construction on finite dimensional Lie group).

(Video) Example on the sphere (landmarks matching) (J. Glaunes)

Thm (T.)

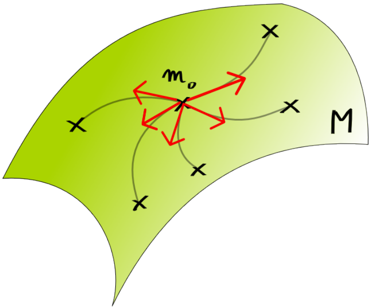
For any $V \hookrightarrow C^1_0(\mathbb{R}^d, \mathbb{R}^d)$ there exists $G_V \subset \mathrm{Diff}^1(\mathbb{R}^d)$ with a right invariant distance d_G for which

- G_V is complete
- There exists a minimizing geodesic between any two elements in G_V

$$d_G(\mathrm{Id}, \phi)^2 = \inf\{\int_0^1 |v_t|^2 dt \mid \dot{\phi} = v \circ \phi, \phi_1 = \phi\}$$

Outline

Common representation framework



- Exponential mapping

$$\begin{array}{cccc} \exp_{m_0}: & T_{m_0}M & \to & M \\ & \delta m_0 & \to & \exp_{m_0}(\delta m_0) \end{array}$$

  such that $t \to \exp_{m_0}(t\delta m_0)$ is the solution of the geodesic equation

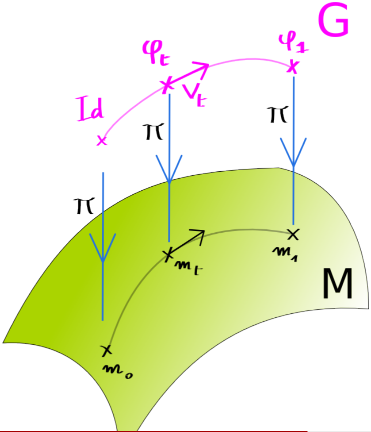
$$\nabla_{\dot m}\dot m = 0$$

  starting from m_0 with initial velocity δm_0.
- Local diffeomorphism (when $\dim M < \infty$)

Feasibility of the normal coordinate computation

- Optimal control problem, $\pi(\phi) = \phi.m_0$:

$$\left|\begin{array}{l} \min \frac{1}{2}\int |v_t|_V^2 dt + g(m_1) \quad g(m) = r(m, m_{\mathrm{obs}}) \\ \dot{m} = d\pi(m).v = v.m \end{array}\right.$$



- Associated hamiltonian

$$H(m, p, v) = (p|v.m) - \frac{1}{2}|v|_V^2$$

- Metric operator:

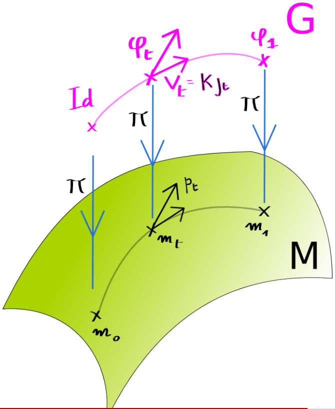
$$\frac{1}{2}|v|_V^2 = (Lv|v) \text{ with } L : V \to V^*$$

Reduction via momentum map

$$\frac{\partial H}{\partial v} = 0 \Rightarrow Lv = j(m, p)$$

same dimension than m (dim < ∞ for landmarks)

momentum map



- $v = Kj(m, p)$ (Horizontal lift)
- Maximum Pontryagin Principle

$$H_r(m, p) = \max_v H(m, p, v)$$

- Hamiltonian evolution:

$$\left| \begin{array}{lcl} \dot{m} & = & \frac{\partial H_r}{\partial p} \\ \dot{p} & = & -\frac{\partial H_r}{\partial q} \end{array} \right.$$

Reduction and statistical power

- A geodesic on M comes from a geodesic on G but its initial velocities v_0 or its momentum Lv_0 belongs to a subspace

$$V_0^* = j(m_0, T_{m_0}^* M) \subset V^*$$

So, geodesic optimization $\Rightarrow$ dimensionality reduction $\Rightarrow$ better statistical power.

- Curve example: $\gamma_0 : [0,1] \to \mathbb{R}^d$ a continuous curve. Then p_0 is a vectorial measure $\mathcal{M}_f([0,1],\mathbb{R}^d) = C([0,1],\mathbb{R}^d)^*$ and

$$v_0(x) = \int_0^1 K(x,\gamma_0(s))dp_0(s) \quad K(x,y) \in \mathcal{M}_d(\mathbb{R}) \text{ kernel}$$

- Moreover, if the geodesic comes from a smooth inexact matching problem e.g.

$$g(\gamma_1) = \int |\gamma_{\text{obs}} - \gamma_1(s)|^2 ds$$

then $p_1 + \frac{\partial g}{\partial \gamma}(\gamma_1) = 0$ and $p_1 \in C([0,1],\mathbb{R}^d)$. *Same is true for* p_0

- For images, if the template is smooth and the data attachment term is smooth e.g. $g(I_1) = \int |I_{obs} - I_1|^2(x)dx$ then

$$v_0(x) = \int K(x,y)p_0(y)\nabla I_0(y)dy$$

$Lv_0 = p_0 \nabla I_0$ distribution of vector fields normal to the level set of I_0.

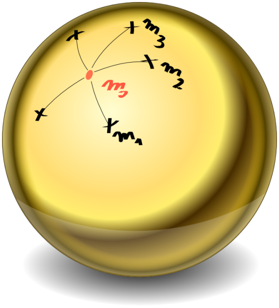
Outline

Karcher means

For a data set $\{m_i\}$ $1 \leq i \leq n$ the *Karcher mean* is the point m_0 minimizing

$$V(m_0) \doteq \sum_{i=1}^{n} d_M(m_0, m_1)^2$$

- Existence and uniqueness for finite dimensional M and $\{m_i\}$ sufficiently closed or under negative curvature condition. Situation unclear for $\dim M = \infty$.



- Usually observations are y_i belongs to an observation space $\mathcal{Y}$ different from M. Need the introduction of a data attachment term.

$$\underset{m_0, m_1, \cdots, m_n}{\text{minimize}} \sum_{i=1}^{n} d_M(m_0, m_1)^2 + \lambda \sum_{i=1}^{n} r(m_i, y_i)$$

or equivalently (lift on the group G)

$$\underset{m_0, \phi_1, \cdots, \phi_n}{\text{minimize}} \sum_{i=1} d_G(\text{Id}, \phi_i)^2 + \lambda \sum_{i=1}^{n} r(\phi_i.m_i, y_i)$$

Hypertemplate setting

$$\underset{m_0,\phi_1,\cdots,\phi_n}{\text{minimize}} \sum_{i=1} d_G(\mathrm{Id},\phi_i)^2 + \lambda \sum_{i=1}^{n} r(\phi_i.m_i, y_i)$$

- Variational problem: if r is smooth enough we have existence of a solution $\hat{\phi}_i, \cdots, \hat{\phi}_n$ for m_0 fixed: n pairwise matching problems.
- If m_0 is let free, existence issues if $\dim(M) = +\infty$.
- Introduce an hypertemplate m_h and look for $m_0 = \psi.m_h$ solution of

$$\underset{\psi,\phi_1,\cdots,\phi_n}{\text{minimize}} \; d_G(\mathrm{Id},\psi)^2 + \sum_{i=1} d_G(\mathrm{Id},\phi_i)^2 + \lambda \sum_{i=1}^{n} r(\phi_i \circ \psi.m_h, y_i)$$

$\hat{m}_0 = \hat{\psi}.m_h$, $\hat{m}_1 = \hat{\phi}_i.\hat{m}_0$ Actually used to build atlases in medical imaging

Atlas learning through hypertemplate

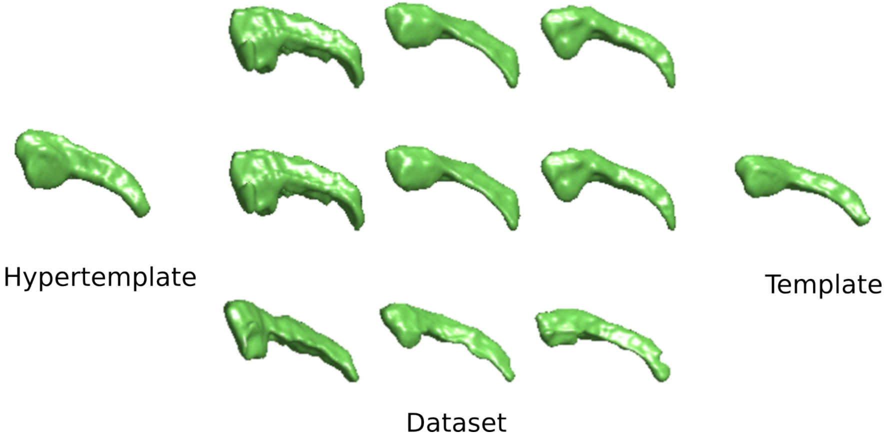


Figure: 3D hippocampuses data.- Ma, Miller, T., Younes Neuroimage'08

Outline

Building a differentiable and riemannian setting on shape spaces

Normal coordinates for statistical analysis

Means and Atlases

Currents and Manifold Representation

Statistics and statistical models

Challenges

Why using currents ?

- Challenging problem for submanifold data :
- if X is a submanifold, it does not depend as a manifold on any particular parametrization (up to smooth chart changes)
- Noisy observation of manifolds may not be a smooth manifold or a manifold at all !

What is a current ?

- Goes back to De Rham. Introduced in this setting by Glaunes and Vaillant.
- Currents integrate differential forms

$$\omega \in \underbrace{\Omega_0^p(\mathbb{R}^d)}_{\text{cont. } p\text{-form}} \to C_X(\omega) = \int_X \omega \in \mathbb{R}$$

On a chart $\gamma : U \to X \subset \mathbb{R}^d$

$$\int_{\gamma(U)} \omega \doteq \int_U \omega_{\gamma(s)}\left(\frac{\partial \gamma}{\partial s_1} \wedge \cdots \wedge \frac{\partial \gamma}{\partial s_p}\right) ds$$

but the expression in independent of a smooth positively oriented reparametrization of the coordinate space $\psi : U \to U$

$$\frac{\partial \gamma \circ \psi}{\partial s_1} \wedge \cdots \wedge \frac{\partial \gamma \circ \psi}{\partial s_p} = \mathrm{Jac}(\psi)\left(\frac{\partial \gamma}{\partial s_1} \wedge \cdots \wedge \frac{\partial \gamma}{\partial s_p}\right) \circ \psi$$

- X can be seen as an element of $(\Omega_0^p(\mathbb{R}^d))^*$. Depends on the orientation of X (X is an orientable manifold).

RKHS norms on currents

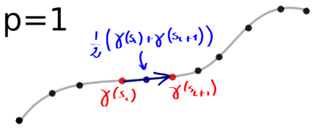
- Idea: if W is a RKHS on the space on p-forms and $W \hookrightarrow \Omega_0^p(\mathbb{R}^d)$ then $\Omega_0^p(\mathbb{R}^d)^* \hookrightarrow W^*$ (W dense subset of $\Omega_0^p(\mathbb{R}^d)$). We get an hilbertian structure on currents (dual norm).
- A kernel for p-forms is given as $K : \mathbb{R}^d \times \mathbb{R}^d \to (\Lambda^p\mathbb{R}^d \otimes \Lambda^p\mathbb{R}^d)^*$ and if X, Y are two orientable sub-manifolds

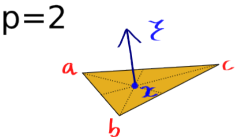
$$\langle C_X, X_Y \rangle_{W^*} \doteq \int_{X\times Y} K$$

Punctual currents and approximations

- $\xi \in \Lambda^p \mathbb{R}^d$, $x \in \mathbb{R}^d$ define a punctual current $\xi \otimes \delta_x \in W^*$ such that

$$(\xi \otimes \delta_x | \omega) = \omega_x(\xi)$$





- Discretization : line $\gamma : [0,1] \to \mathbb{R}^d$

$$C_\gamma \simeq \sum_{i=1}^{n} (\gamma(s_{i+1} - s_i)) \otimes \delta_{(\gamma(s_i)+\gamma(s_{i+1}))/2}$$

- Triangulated surface :

$$T(a,b,c) \to \xi \otimes \delta_x$$

with $x = (a+b+c)/3$ and $\xi = \vec{ab} \wedge \vec{ac}$.

- Make sense for arbitrary dimensions p and d.

Matching pursuit algorithm

Start from an initial manifold and $C_0 = C_X$ and iterate :

$$\left| \begin{array}{l} (x_{n+1}, \xi_{n+1}) = argmax_{x,\xi} \langle \xi \otimes \delta_x, \underbrace{C_X - C_n}_{\text{residual}} \rangle_{W^*} \\ C_{n+1} = C_n + \xi_{n+1} \otimes \delta_{x_{n+1}} \end{array} \right.$$

- Very convenient to get compressed representation of manifolds
- Complexity control via sparse non parametric approximations

-Durrleman, Pennec, T., Ayache MICCAI'08.

Currents and means (Atlas construction for manifolds)

1. Define a *parametrization invariant* data attachment term: for X and Y two orientable submanifold (same dimension)

$$r(X, Y) = h(|C_X - C_Y|_{W^*})\,.$$

2. Easy to consider noisy observation as currents $\mathcal{Y} = W^*$ and to defined generalized Karcher means :

$$\underset{C_0,\phi_1,\cdots,\phi_n}{\text{minimize}} \sum_{i=1} d_G(\mathrm{Id}, \phi_i)^2 + \lambda \sum_{i=1}^{n} |\phi_i.C_0 - y_i|^2_{W^*}$$

   where $(\phi, C) \to C.\phi$ is the push forward action.

3. For $\phi_1, \cdots, \phi_n$ fixed feasible computation of C_0 even with a large number of observations y_i (control of the number of points via matching pursuit).

Atlas construction for fiber tracts

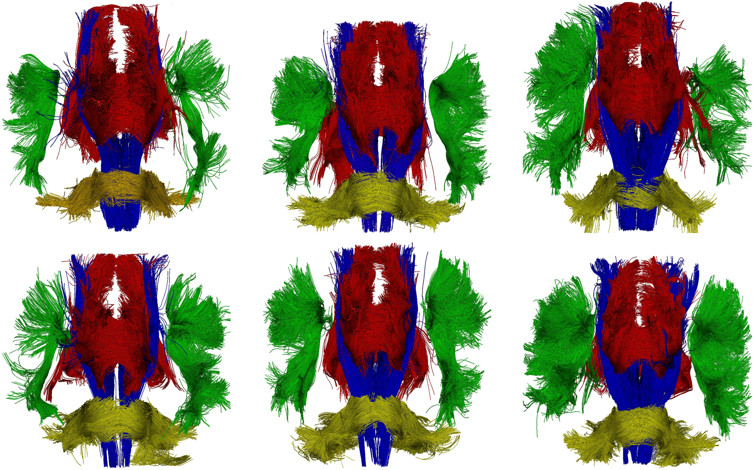
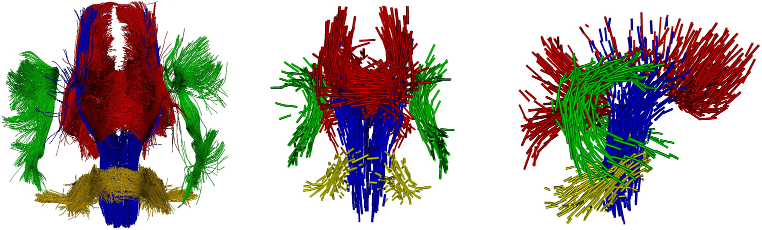


Figure: 5 fiber tracts segmented in 6 subjects

Atlas construction for fiber tracts



(a) One subject

(b) template (occipital view)

(c) template (lateral view)

Figure: Computed Template - Durrleman, Fillard, Pennec, T., Ayache Neuroimage'11

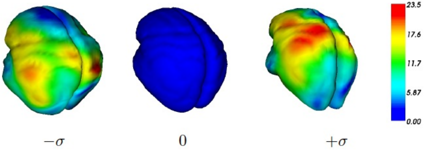
Outline

Deformation analysis

- $\hat{\phi}_1, \cdots, \hat{\phi}_n$ gives initial momentum $\hat{p}_1, \cdots, \hat{p}_n$ the statistical analysis in the (finite dimensional) tangent space.
- PCA analysis (using the induced metric)

The analysis can be lifted on the space of diffeomorphism by horizontal lift giving generative model via shooting.

Sulcal lines: pairwise registration first mode



Deformation + textures (for manifolds)

- Deformations
$$\hat{\phi}_1, \cdots, \hat{\phi}_n \longrightarrow \hat{p}_1, \cdots, \hat{p}_n \,.$$
- Texture : Residues
$$\hat{R}_i = y_i - \hat{\phi}.C_0 \in W^*$$
- Analysis of the joint model

Fiber tracts: Deformation pairwise registration first mode (corticobulbar tract)

def. mode at $-\sigma$ $\phi^{-m_v} \star \bar{B}$ template $\bar{B}$ def. mode at $+\sigma$ $\phi^{+m_v} \star \bar{B}$

Deformation + textures (for manifolds)

- Deformations

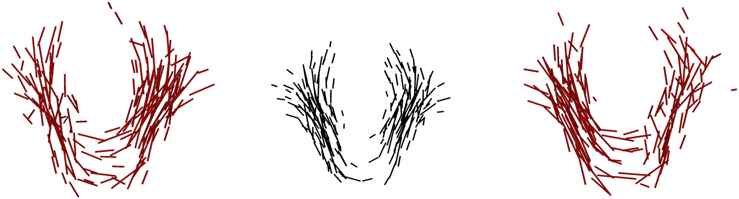
$$\hat{\phi}_1, \cdots, \hat{\phi}_n \longrightarrow \hat{p}_1, \cdots, \hat{p}_n .$$

- Texture : Residues

$$\hat{R}_i = y_i - \hat{\phi}.C_0 \in W^*$$

- Analysis of the joint model

Fiber tracts: First Mode of Residues (corticobulbar tract)



texture mode at $-\sigma$ $\bar{B} - m_\varepsilon$ | template $\bar{B}$ | texture mode at $+\sigma$ $\bar{B} + m_\varepsilon$

Discrimination study workflow

- Learning of a common template + analysis in normal coordinates
- Application there of stander classification/discrimination methods
- Analysis of $\log(\mathrm{Jac}(\Phi))$ on the template (atrophy patterns).

Population analysis via bayesian mixed-effects hierarchical model

$$y_i = \phi_i.(m_0 + T_i) + \sigma n_i$$

- Fixed effects (population effects): template m_0,*law* of the deformations ϕ, *law* of the texture model T, noise level σ: $\theta = (m_0, p_\phi, p_T, \sigma)$.
- Random effects (individual effects): individual deformation ϕ_i, texture T_i.
- Many hidden-variables: ϕ_i's,T_i's, n_i's.
- Hyperparameters: priors on the fixed effect distribution.

$$\hat{\theta} = \underset{\theta}{\text{argmax}}\, P(\theta | y_1, \cdots, y_n)$$

Deterministic algorithm as approximations of stochastic ones

$$y_i = \phi_i.(m_0 + T_i) + \sigma n_i$$

- Usual deterministic method for atlas learning appear to be straightforward approximations of EM-type algorithms (where the E-step is replaced by the mode approximation of the posterior distribution on the hidden variables)- Allassonnière, Amit, T. JRSS'07
- We could use this statistical setting to propose better algorithms: MCMC methods for the posterior distribution. -Kuhn and Lavielle, Compt. Stat. and Datata Analysis'05, Allassonniere, Kuhn, T. Bernoulli'10

Current state-of-the-art: Finite dimensional setting pour the fixed effects, no texture, SAEM-MCMC algorithms, linear deformation model, images.

Templates estimation through SAEM-MCMC algorithm

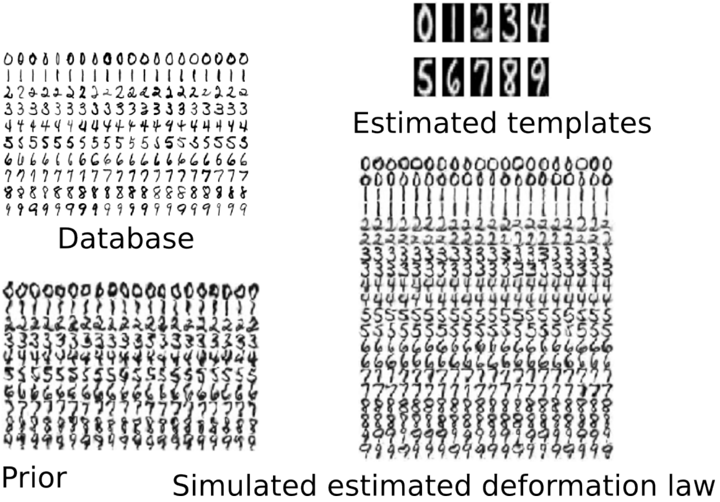


Figure: US-Postal database, from Allassonnière et al., Bernoulli'10. Single model, SAEM-MCMC algorithm

Templates estimation through SAEM-MCMC algorithm

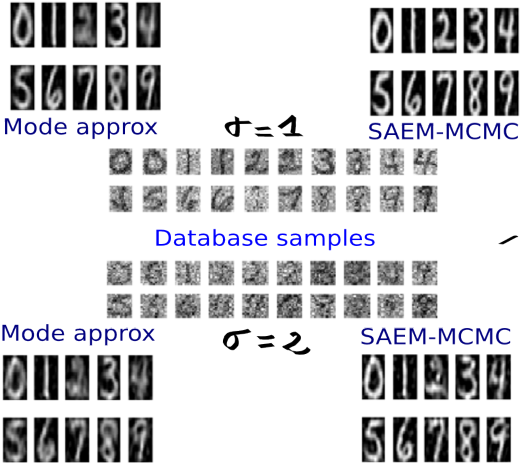


Figure: US-Postal database, from Allassonnière et al., Bernoulli'10. Robustness to noise: Mode approximation Versus SAEM-MCMC algorithm

Mixture models

Statistical models extended to mixture models allowing multi-template estimation

Figure: US Postal database - multi-template learning via mixture model. From Allassonniere, Kuhn ESAIM P&S'10

Outline

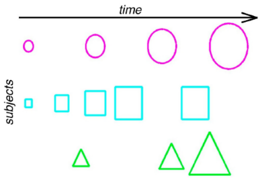
- Proper statistical modeling and model estimation point of view appears to be a promising avenue in this high dimensional setting compared to more purely model free deterministic point of view
- In particular, integration of the posterior distribution of deformations instead of mode approximation is necessary for consistent model estimation and to work in noisy situations.

- However, still a high computational burden for MCMC sampling since the hidden variable are living in a high dimensional space. This implies the need of better adapted sampling scheme using the implicit low dimensionality of the high posterior log-likelihood curved submanifold (*Riemannian Manifold Hamiltonian Monte Carlo*, Girolami, Calderhead, Chin JRSS'10).
- Extension to the non linear situation using momentum representation and geodesic shooting is possible (work in progress)
- Manifolds: extension to the manifold setting is not done yet. Nor the estimation in this framework of a texture part (fiber tracks).
- Current implementation and theoretical work for the statistical modeling is restricted to the finite dimensional setting. Needs to understand the limit to the infinite dimensional setting.

Next frontier: Modeling shape evolution and growth

An emerging question

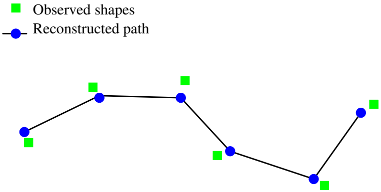
- An **emerging** question of interest is now to study the time dependent data of shapes (images, landmarks, surfaces or tensors).
- Main target application: Growth studies, longitudinal studies.



with specific needs and challenges

- Flexibility: more or less non parametric models
- Versatility (various data and contexts)
- Robustness (noise, time sampling)
- Interpretability (ideally generative stochastic models)

Piecewise Geodesics Models



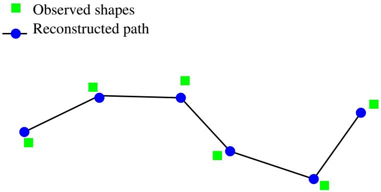
Miller's Growth Model (TS-LDDMM)

$$\hat{\mathbf{x}}_t = \phi_t \cdot \mathbf{x}_0, \; (\partial_t \phi = v_t \circ \phi)$$

$$\inf_v \int_0^1 |v_t|^2 dt + \sum_{k=1}^{n} g_k(\phi_t \cdot \mathbf{x}_0, \mathbf{x}_{t_k}^{\text{obs}})$$

where $\mathbf{x}_{t_k}^{\text{obs}}$ are observed shapes (one subject).

Piecewise Geodesics Models



Miller's Growth Model (TS-LDDMM)

$$\hat{\mathbf{x}}_t = \phi_t \cdot \mathbf{x}_0, \ (\partial_t \phi = v_t \circ \phi)$$

$$\inf_v \int_0^1 |v_t|^2 dt + \sum_{k=1}^{n} g_k(\phi_t \cdot \mathbf{x}_0, \mathbf{x}_{t_k}^{\text{obs}})$$

where $\mathbf{x}_{t_k}^{\text{obs}}$ are observed shapes (one subject).

Piecewise Geodesics Models

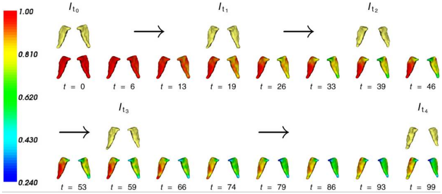


Figure: Longitudinal growth model results for Huntington's Disease examining the caudate nucleus, From A. Khan and M. F. Beg, ISBI 2008.

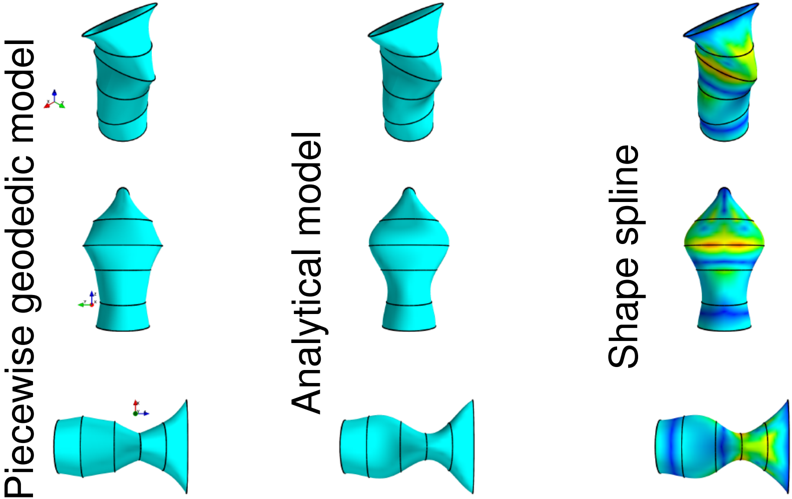
Miller's Growth Model (TS-LDDMM)

$$\hat{\mathbf{x}}_t = \phi_t \cdot \mathbf{x}_0, \; (\partial_t \phi = v_t \circ \phi)$$

$$\inf_v \int_0^1 |v_t|^2 dt + \sum_{k=1}^n g_k(\phi_t \cdot \mathbf{x}_0, \mathbf{x}_{t_k}^{\text{obs}})$$

where $\mathbf{x}_{t_k}^{\text{obs}}$ are observed shapes (one subject).

2nd order model: Shape Spline V Piecewise Geodesic



A few references

Shapes and Diffeomorphisms

- L. Younes, "*Shapes and Diffeomorphisms*", Springer, May 2010.
- A. Trouvé and L. Younes, "*Shape Spaces*", Handbook of Mathematical Methods in Image Processing., Springer 2010.

Computational Anatomy, Statistic modeling, Currents and Growth

- M.I. Miller, "*Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms*", NeuroImage, ps19-s23, 2004.
- S. Allassonnière, Y. Amit, and A. T. *Towards a coherent statistical framework for dense deformable template estimation.* JRSS, Series B, 69(1) :3-29, 2007.
- S. Allassonnière, E. Khun, A. T. *Bayesian deformable models building via stochastic approximation algorithm : A convergence study.* Bernoulli, 2010
- J. Glaunes, *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*, PhD Thesis, Univ Paris 13, 2005.
- S. Durrleman, *Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution*, PhD Thesis, Univ. Nice, 2010.
- A. T. and F.-X. Vialard *"Shape Splines and Stochastic Shape Evolution: A Second Order Point of View*", QAM 2010.