Diffeomorphic Evolution Equations

Laurent Younes

Johns Hopkins University

IPAM, July 2008

Partially supported by NSF,NIH,ONR

Collaborators

M. Miller, A. Trouvé, D. Holm, T. Ratnanather, J. Zweck, S. Zhang, A. Qiu, M-F Beg, F. Arrate, . . .

Part 1: Diffeomorphic Metrics

Notation

- d is the dimension of the underlying space (1,2 or 3).
- We fix an open subset Ω in $\mathbb{R}^d$.
- A *diffeomorphism* of Ω is an invertible transformation from Ω to itself, which is continuously differentiable (C^1) with a differentiable inverse.
- Diffeomorphisms form a group for composition, denoted $\mathrm{Diff}(\Omega)$.

Actions

- Diffeomorphisms being transformations of Ω, they alter quantities that are included in, or defined on Ω.
- Examples: collections of points, functions, densities, measures, tensor fields, etc.
- Generic notation: $m =$ a deformable object, belonging to some space M.
- Diffeomorphisms act on M if, for any diffeomorphism φ and any object m, a *deformed object* $\varphi \cdot m$ can be defined. The requirements are $\mathrm{id} \cdot m = m$ and $\varphi \cdot (\psi \cdot m) = (\varphi \circ \psi) \cdot m$.

Examples of Actions

- For configurations of points, $m = (m_1, \ldots, m_N)$, we have $\varphi \cdot m = (\varphi(x_1), \ldots, \varphi(x_N))$.
- For images, $m : \Omega \to \mathbb{R}^k$, we have $\varphi \cdot m = m \circ \varphi^{-1}$.
- For densities $m : \Omega \to \mathbb{R}^+$, we have $\varphi \cdot m = m \circ \varphi^{-1} | \det D(\varphi^{-1})|$, where D is the differential.

Vector Fields and ODE's

- Given a time-dependent vector field, i.e., a function $v : (t, y) \mapsto v(t, y)$ defined on $[0, 1] \to \Omega$, define the ordinary differential equation (ODE) on Ω

$$\partial_t y = v(t, y).$$

- The flow associated to this equation is the function $(t, x) \mapsto \varphi(t, x)$ such that, for a given x, $t \mapsto \varphi(t, x)$ is the solution of the ODE with initial condition $y(0) = x$.
- It is uniquely defined under standard smoothness assumptions on v, and for fixed t, $x \mapsto \varphi(t, x)$ is a diffeomorphism of Ω (under suitable boundary conditions).

The Basic Hilbert Space

- Consider a Hilbert space space V of vector fields on Ω with dot product, $(v, w) \mapsto \left\langle v\,,\, w\right\rangle_V$ and norm $\|v\|_V^2 = \left\langle v\,,\, v\right\rangle_V$.
- Assume that the dot product is associated to an operator, L, with

$$\|v\|_V^2 = \int_\Omega Lv^T v dx.$$

(to be understood in a generalized sense: $Lv \in V^*$ can be singular).

- L is the duality operator between V and V^*, with inverse $K = L^{-1}$.

- For all our models, K is a smooth kernel operator: there exists a function $(x, y) \mapsto K(x, y)$ such that

$$(Kv)(x) = \int_\Omega K(x, y) v(y).$$

(We make the common abuse of notation of using the same letter for the operator and the kernel.)

- In full generality, $K(., .)$ is matrix valued, although all our applications use a scalar kernel.
- Typical example:

$$K(x, y) = \exp(-|x - y|^2 / 2\sigma^2).$$

Riemannian Metric on Diffeomorphisms

- Use $\langle . , . \rangle_V$ for the metric at $\varphi = \mathrm{id}$ on $\mathrm{Diff}(\Omega)$, and use the associated right invariant metric

$$\|w\|_\varphi = \|w \circ \varphi^{-1}\|_V.$$

- With this metric, the kinetic energy of a path in diffeomorphisms $t \mapsto \varphi(t, .), t \in [0, 1]$ is

$$\int_0^1 \|v(t, .)\|_V^2 dt$$

with $\partial_t \varphi(t, .) = v(t, \varphi(t, .))$.

Part 2: Diffeomorphic Gradient Evolution

Gradient descent on $\mathrm{Diff}(\Omega)$

- Assuming it exists, the gradient of a function E on a manifold M, denoted $\nabla E(m)$, is defined by

$$\partial_\varepsilon E(m(\varepsilon))_{|_{\varepsilon=0}} = \left\langle \nabla E(m(0)) \, , \, \partial_\varepsilon m_{|_{\varepsilon=0}} \right\rangle_{m(0)}$$

  (It is a map from M to TM.)

- Gradient descent is

$$\partial_t m = -\nabla E(m).$$

Case of diffeomorphisms

- When $M = \mathrm{Diff}(\Omega)$, the gradient is defined by

$$\partial_\varepsilon E(\varphi + \varepsilon\delta\varphi)_{|\varepsilon=0} = \left\langle \nabla E(\varphi) \circ \varphi^{-1}, \delta\varphi \circ \varphi^{-1} \right\rangle_V.$$

- Letting $v = \delta\varphi \circ \varphi^{-1}$,

$$\partial_\varepsilon E((\mathrm{id} + \varepsilon v) \circ \varphi)_{|\varepsilon=0} = \left\langle \nabla E(\varphi) \circ \varphi^{-1}, v \right\rangle_V.$$

- The left-hand side is the *shape derivative* of E at φ.
- Letting $S(\varphi) = \nabla E(\varphi) \circ \varphi^{-1}$, gradient descent becomes

$$\partial_t \varphi = -S(\varphi) \circ \varphi.$$

- This generates a flow.

Example: Greedy Image Matching (Christensen et al.)

- Fix two images m and m' and minimize

$$E(\varphi) = \int_\Omega (m \circ \varphi^{-1} - m')^2 dy.$$

- Then (using the kernel on V)

$$S(\varphi)(x) = -2 \int_\Omega K(x, y)(m \circ \varphi^{-1}(y) - m'(u))\nabla(m \circ \varphi^{-1})(u).$$

- When $K^{-1} = L = (\alpha \mathrm{Id} - \Delta)^k$ (where Δ is the Laplacian), the associated gradient descent evolution is the equation proposed by Christensen et al (1993). The gradient descent formulation is due to Trouvé (1998).

Adding a regularization term

- Define
$$E(\varphi) = \int_\Omega (m \circ \varphi^{-1} - m')^2 dy + \lambda \int_\Omega |D\varphi|^2 dx$$
(with the matrix norm $|A|^2 = \mathrm{trace}(AA^T)$ in the penalty term).
- The resulting function $S(\varphi)$ decomposes in $S(\varphi) = S_1(\varphi) + \lambda S_2(\varphi)$, with S_1 as before.
- For S_2, we have
$$S_2(\varphi) = 2 \int_\Omega (\det D\varphi \circ \varphi^{-1})^{-1} D\varphi \circ \varphi^{-1} (D\varphi \circ \varphi^{-1})^T \nabla_1 K dx.$$

Diffeomorphic Segmentation

- Let m be a template (e.g. a binary volume).
- Define
$$E(\varphi) = -\int_\Omega m \circ \varphi^{-1}(x) f(y) dy$$
where f is some function, associated to an observed image, quantifying the relative likelihood for a point with coordinate y to belong in the interior of the shape rather than in the exterior.
- Then
$$S(\varphi)(.) = \int_{\varphi(\partial m)} K(., y) N^\varphi d\sigma^\varphi(y).$$

Finitely Generated Gradient (F. Arrate)

- Define N control points $x_1, \ldots, x_N$ relative to the template geometry.
- Restrict the "gradient" to take the form $\tilde{\nabla} E(\varphi) = \tilde{S}(\varphi) \circ \varphi$ with

$$\tilde{S}(\varphi)(.) = \sum_{k=1}^{N} K_0(., \varphi(x_k))\alpha_k.$$

Finitely Generated Gradient (F. Arrate)

- Define N control points $x_1, \ldots, x_N$ relative to the template geometry.
- Restrict the "gradient" to take the form $\tilde{\nabla} E(\varphi) = \tilde{S}(\varphi) \circ \varphi$ with

$$\tilde{S}(\varphi)(.) = \sum_{k=1}^{N} K_0(., \varphi(x_k))\alpha_k.$$

- Take $\tilde{S}(\varphi)$ as the orthogonal projection of $S(\varphi)$ on

$$W = \left\{ \sum_{k=1}^{N} K_0(., \varphi(x_k))\alpha_k : \ \alpha_1, \ldots, \alpha_N \in \mathbb{R}^d \right\}$$

for the dot product $\langle . \, , \, . \rangle_V$.

- Let K and K_0 be Gaussian kernels (with $\Omega = \mathbb{R}^d$).
- With $g_{\sigma^2}(u) = (2\pi\sigma^2)^{-d/2} \exp(-|u|^2/2\sigma^2)$, take

$$K(x, y) = g_{\sigma^2}(x - y) \text{ and } K_0(x, y) = g_{\sigma_0^2}(x - y)$$

with $\sigma_0 > \sigma$.

- Then solve

$$\int_\Omega g_{\sigma_0^2-\sigma^2}(y - \varphi(x_k)) e_j^T S(\varphi)(y) dy = \sum_{l=1}^N g_{2\sigma_0^2-\sigma^2}(\varphi(x_k) - \varphi(x_l)) e_j^T \alpha_l.$$

- Use

$$\frac{d\varphi_t}{dt}(y) = -\sum_{k=1}^N g_{\sigma_0^2}(\varphi_t(y) - \varphi_t(x_k))\alpha_k(t).$$

Example

Segmentation in the hippocampal region

Example

Segmentation in the hippocampal region

Diffeomorphic MCF (Zheng et al.)

- Let $E(m) = \text{area}(m)$.
- Use an area-normalized metric:

$$\|\xi\|_m^2 = \inf\left\{\left\langle \rho_m v\,,\, \rho_m v\right\rangle_V, v\cdot m = \xi\right\}.$$

with

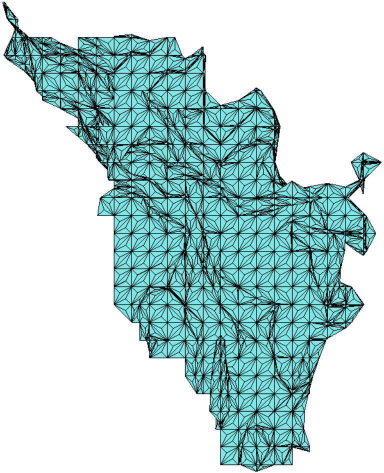
$$\rho_m(y) = \sqrt{\int_m K(y,x)d\sigma(x)}.$$

- The flow for this metric is (for a point $p \in M$)

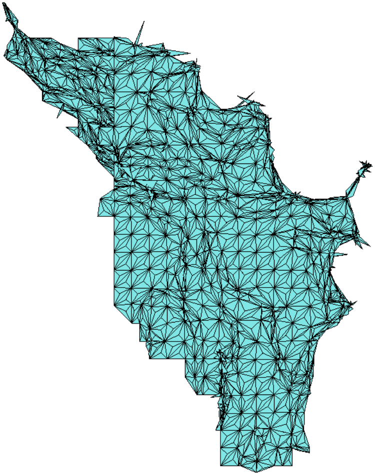
$$\frac{dp}{dt} = \rho_m(p)^{-1}\int_m K(p,y)\rho_m(y)^{-1}H(y)N(y)d\sigma(y).$$

where H is the mean curvature and N the normal to the surface.

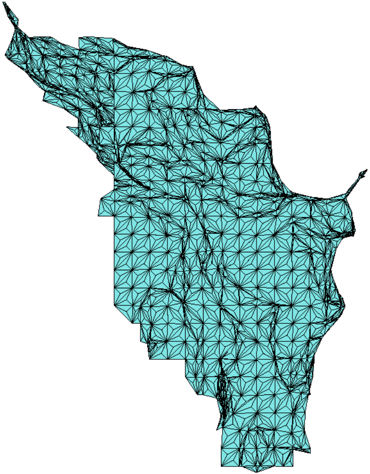
Experiments



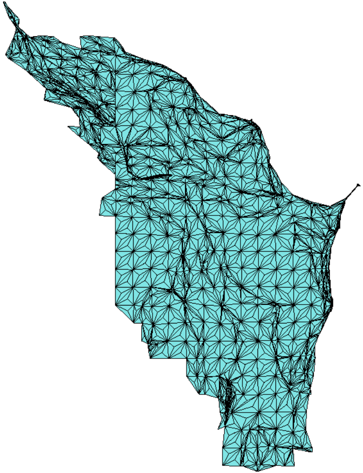
Original cortical surface



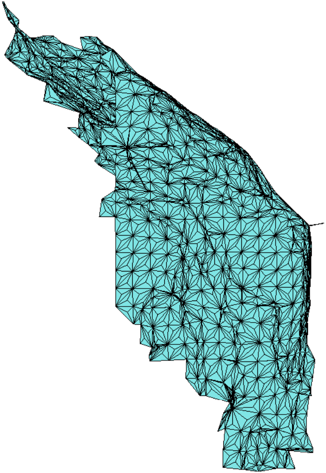
MCF at $t = 1$.



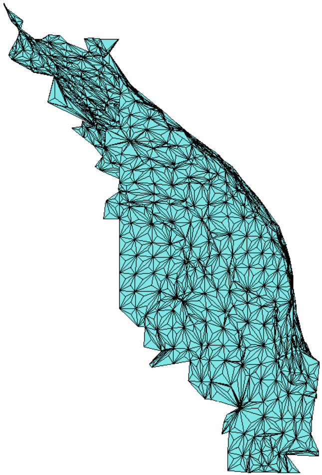
DMCF at $t = 1$.



DMCF at $t = 2$.



DMCF at $t = 5$.



DMCF at $t = 10$.

Part 3: Geodesics

EPDiff

- Geodesic equations for our metric on groups of diffeomorphisms have a long history, starting with Arnold 1966, re-analyzed by Marsden, Ratiu, Holm (Euler-Poincaré Reduction).
- They are evolution equations that describe time-dependent diffeomorphisms $t \mapsto \varphi(t, .)$ that have no acceleration in the group metric.
- This is a necessary (and locally sufficient) condition for being energy minimizing.

EPDiff

- Geodesic equations for our metric on groups of diffeomorphisms have a long history, starting with Arnold 1966, re-analyzed by Marsden, Ratiu, Holm (Euler-Poincaré Reduction).
- They are evolution equations that describe time-dependent diffeomorphisms $t \mapsto \varphi(t, .)$ that have no acceleration in the group metric.
- This is a necessary (and locally sufficient) condition for being energy minimizing.
- The 'classical' (non-singular) form of this "EPDiff" equation is

$$\partial_t Lv + D(Lv)v + Lv\nabla \cdot v + (Dv)^T Lv = 0$$

with $\partial_t \varphi(t, .) = v(t, .) \circ \varphi(t, .)$.

- Another, more general, characterization is: for all smooth vector field w,

$$\partial_t \int_\Omega Lv^T w dx = -\int_\Omega Lv^T (Dv\, w - Dw\, v) dx.$$

- Another, more general, characterization is: for all smooth vector field w,

$$\partial_t \int_\Omega Lv^T w dx = -\int_\Omega Lv^T (Dv\, w - Dw\, v) dx.$$

- Letting $a_t = Lv_t$, EPDiff can also be written as

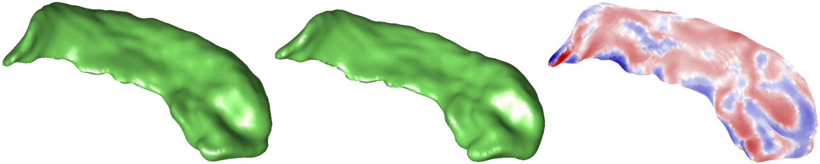
$$\begin{cases} v_t = Ka_t \\ \partial_t \varphi_t = v_t \circ \varphi_t \\ \partial_t a_t = -Da_t v_t - a_t \nabla \cdot v_t - (Dv_t)^T a_t \end{cases}$$

Momentum Representation

- EPDiff provides an important representation of diffeomorphisms via the *Riemannian exponential map*.
- This map takes in input an initial velocity v_0 (or the corresponding momentum Lv_0) and solves the EPDiff equation until time $t = 1$.
- The output being the final diffeomorphism, φ_1, this yields a transformation

$$P : Lv_0 \mapsto \varphi_1 = P(Lv_0) \in G$$

Example



Visualizing the momentum on a binary volume. Left: template image; center: target image; right: momentum superimposed on the template image. The momentum can then be used as a deformation signature for shape studies.

EPDiff and LDDMM

- EPDiff coincides with the Euler-Lagrange equation for the *Large Deformation Diffeomorphic Metric Matching* algorithms (LDDMM).
- LDDMM minimizes energy of the form

$$E(\varphi) = d_V(\mathrm{id}, \varphi)^2 + E_{\mathrm{data}}(\varphi \cdot m_0)$$

- So, solving the LDDMM problem automatically provides the momentum representation of the optimal diffeomorphism φ.

Geodesics in M

- The EPDiff equation "projects" on shapes and images when combined with the notion of horizontality, which are preferred directions of deformation relative to the object.
- A geodesic starting from $m_0 \in M$ must take the form $m_t = \varphi_t \cdot m_0$ with
  - (i) $t \mapsto \varphi_t$ is a geodesic on G_V (it satisfies EPDiff).
  - (ii) The initial velocity at time $t = 0$ is horizontal at m_0.
- Property (ii) propagates over time: if it is true, then v_t is horizontal at m_t for all t.
- Because horizontal vector fields have specific forms for specific types of objects, this simplifies in specific object-dependent forms of EPDiff.

Case of point sets

- Horizontal vector fields at $m = (x_1, \ldots, x_N)$ must take the form
$$v(.) = \sum_{k=1}^{N} K(., x_k)\alpha_k.$$
- EPdiff becomes
$$\begin{cases} v_t(.) = \sum_{l=1}^{N} K(., x_l)\alpha_l \\ \partial_t x_k - v_t(x_k) = 0 \\ \partial_t \alpha_k = -\sum_{l=1}^{N} \nabla_1 K(x_k, x_l)\alpha_k^T \alpha_l \end{cases}$$
Here, $\nabla_1 K$ represent the gradient of $(x, y) \mapsto K(x, y)$ with respect to the first coordinate, x.
- The momentum representation for landmarks is captured by the family of d dimensional vectors, a_0. The initial conditions (m_0, a_0) uniquely specify the values of m_t and a_t at all times.

Case of Images

- When objects are images, horizontal vector fields are characterized by the property that Lv is orthogonal to the level lines of the image.
- Important example: $Lv_0(x) = a_0(x)\nabla m_0(x)$.
- In this case, geodesics satisfy

$$\begin{cases} v_t = -\int_\Omega K(., y) a_t(y) \nabla m_t(y) dy \\ \partial_t m_t + \nabla m_t^T v_t = 0 \\ \partial_t a_t + \nabla \cdot (a_t v_t) = 0 \end{cases}$$

Case of densities

- Horizontal momenta at m are also orthogonal to the level sets of m.
- With $Lv_0 = a_0 \nabla m_0$, the evolution equations are

$$\begin{cases} v_t = \int_\Omega K(.,y) \nabla a_t(y) m_t(y) dy \\ \partial_t m_t + \nabla \cdot (m_t v_t) = 0 \\ \partial_t a_t + \nabla a_t^T v_t = 0 \end{cases}$$

- The roles of m and a are reversed compared to the image case.

A common form for these equations

- All previous geodesic equations have the form

$$\begin{cases} v_t = F(m_t, a_t) \\ \partial_t m_t = H(m_t, v_t) \\ \partial_t a_t = G(m_t, a_t, v_t) \end{cases}$$

  where the functions F, G, H are linear in a and v, but not necessarily in m.
- The function $t \mapsto a_t$ can be interpreted as a representation for horizontal momenta, via the relation $Lv = LF(m, a)$.

Part 4: Parallel translation

Parallel Translation

- The momentum representation (exponential map) is an important tool that can provide a normalized representation of a family of shapes relative to a fixed template.
- Parallel translation can be used to transport this representation from one template to another.
- Applications: asymmetry, longitudinal studies (follow-up vs. baseline), etc...

- Parallel translation on a manifold is an operation that takes a tangent vector at some point in the manifold and translate it along a given curve (for example: a geodesic between two 'templates').
- Assume that the geodesic equation is

$$\begin{cases} v_t = F(m_t, a_t) \\ \partial_t m_t = H(m_t, v_t) \\ \partial_t a_t = G(m_t, a_t, v_t) \end{cases}$$

- Then parallel translation is

$$\begin{cases} H(m_t, F(m_t, \partial_t b_t)) = \frac{1}{2}\Big\{ H_m(m_t, v_t)\eta_t - H_m(m_t, w_t)\xi_t \\ \qquad + H(m_t, F_m(m_t, a_t)\eta_t - F_m(m_t, b_t)\xi_t) \\ \qquad + H(m_t, F(m_t, G(m_t, b_t, v_t) + G(m_t, a_t, w_t)))\Big\} \\ \xi_t = H(m_t, v_t) = \partial_t m_t \\ \eta_t = H(m_t, w_t) \\ v_t = F(m_t, a_t) \\ w_t = F(m_t, b_t) \end{cases}$$

- F_m and H_m the differentials for F and H with respect to the first variable, m.
- The first equation defines db/dt implicitly, involving the inversion of the linear operator $\beta \mapsto H(m, F(m, \beta))$.

Case of diffeomorphisms

$$\begin{aligned}\partial_t b_t &= \frac{1}{2} L(Dv_t.w_t - Dw_t.v_t) \\ &-\frac{1}{2}(Db_t v_t + b_t \nabla \cdot v_t + Dv_t^T b_t + Da_t w_t + a_t \nabla \cdot w_t + Da_t^T w_t)\end{aligned}$$

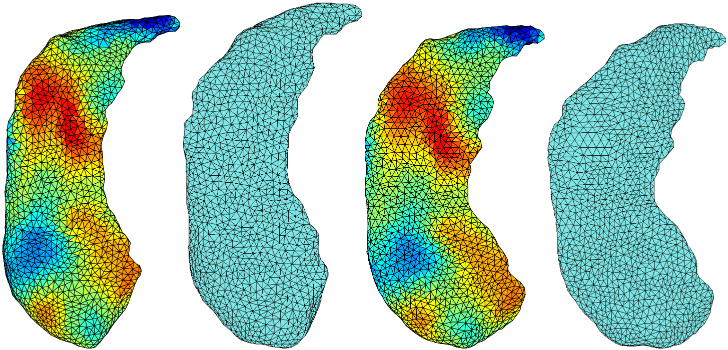
with $a_t = Lv_t$ and $b_t = Lw_t$.

Point sets

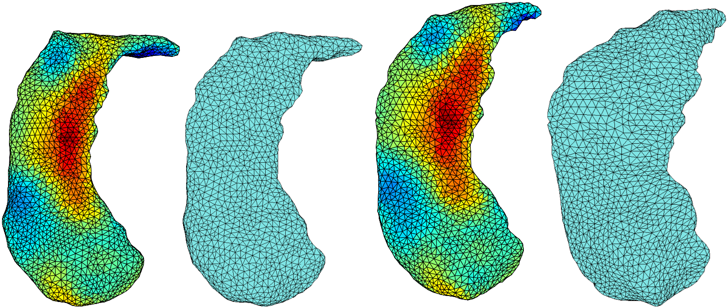
- $m_t = (x_1(t), \dots, x_N(t))$ and $a_t = (\alpha_1(t), \dots, \alpha_N(t))$
-

$$\begin{cases} \displaystyle\sum_{l=1}^N K(x_k, x_l)\partial_t \beta_l = \frac{1}{2}\Big(Dv_t(x_k)\eta_k - Dw_t(x_k)\xi_k \\ \displaystyle\quad + \sum_{l=1}^N (\nabla_2 K(x_k, x_l)^T \eta_l \alpha_l + \nabla_2 K(x_k, x_l)^T \xi_l^T \beta_l) \\ \displaystyle\quad - \sum_{l=1}^N K(x_k, x_l)(Dv_t(x_l)^T \beta_l + Dw_t^T(x_l)\alpha_l)\Big) \\ \displaystyle w_t = \sum_{l=1}^N K(., x_l(t))\beta_l(t) \\ \eta_k = w_t(x_k) \end{cases}$$

Example



Example



Case of images

- Let m_t be a geodesic for the image metric with $v_t = K(a_t \nabla m_t)$. Let $\xi_t = -\nabla m_t^T v_t$, and assume that $w_0 = -K(b_0 \nabla m_0)$ is given with $\eta_0 = -\nabla m_0^T w_0$.
- The parallel translation of η_0 along m_t for the image metric is

$$\begin{cases} \nabla m_t^T K(\partial_t b_t \nabla m_t) = \frac{1}{2}\Big(\nabla \eta_t^T v_t - \nabla \xi_t^T w_t + \nabla m_t^T K(a_t \nabla \eta_t - b_t \nabla \xi_t) \\ \quad -\nabla m_t^T K(\nabla \cdot (b_t v_t + a_t w_t) \nabla m_t)\Big) \\ \eta_t = -\nabla m_t^T K(b_t \nabla m_t) \\ w_t = K(b_t \nabla m_t) \end{cases}$$

- To explicitly compute db_t/dt, one needs to invert the operator $\zeta \mapsto \nabla m_t^T K(\zeta \nabla m_t)$.

Case of densities

$$
\begin{cases}
\nabla \cdot (m_t K(m_t \nabla \partial_t b_t)) = -\frac{1}{2}\Big(\nabla \cdot (\eta_t v_t - \xi_t w_t) - \\
\quad \nabla \cdot (m_t K(\eta_t \nabla a_t - \xi_t \nabla b_t)) + \nabla \cdot (m_t K(m_t \nabla \cdot (\nabla a_t^T w_t + \nabla b_t^T v_t)))\Big) \\
\eta_t = \nabla \cdot (m_t K(m_t \nabla b_t)) \\
w_t = -K(m_t \nabla b_t)
\end{cases}
$$

General characterization of Par. Trans.

- The relation between parallel translation on diffeomorphisms and on objects is
  (i) $(\partial_t w_t - KA(v_t, w_t)) \cdot m_t = 0$,
  (ii) w_t is horizontal at m_t.

  with

$$
\begin{aligned}
A(v_t, w_t) &= \frac{1}{2} L(Dv_t.w_t - Dw_t.v_t) \\
&\quad - \frac{1}{2}(Db_t v_t + b_t \nabla \cdot v_t + Dv_t^T b_t \\
&\quad + Da_t w_t + a_t \nabla \cdot w_t + Da_t^T w_t)
\end{aligned}
$$

  with $a_t = Lv_t$ and $b_t = Lw_t$.

Part 5: The Jacobi Equation

Jacobi Equation

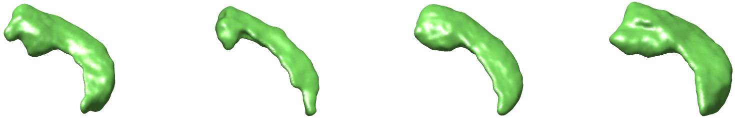
- They are first order variations of geodesics with respect to changes in their initial conditions.
- Important when solving variational problems that involve geodesics: LDDMM, J.Ma's template estimation algorithm,...
- Also provides an alternate implementation of parallel transport.

- The equation is

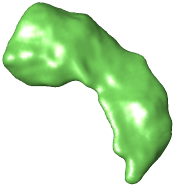
$$\begin{cases} \delta v_t = F_m(m_t, a_t)\delta m_t + F(m_t, \delta a_t) \\ \partial_t \delta m_t = H_m(m_t, v_t)\delta m_t + H(m_t, \delta v_t) \\ \partial_t \delta a_t = G_m(\delta m_t, a_t, v_t)\delta m_t + G(m_t, \delta a_t, v_t) + G(m_t, a_t, \delta v_t) \end{cases}$$

- The Jacobi field is, by definition, $J(t) = t\delta m_t$.

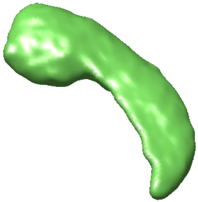
Bayesian Template Estimation: hippocampi (110 images) (Jun Ma)



4 targets out of 110



Hypertemplate



Estimated template

Conclusion

- We have described, in this paper, a whole range of evolution equations (gradient descent, geodesics, parallel transport, Jacobi fields) that are related to important aspects of computational anatomy.
- These equations have all taken part in medical imaging applications, for smoothing, segmentation, registration, longitudinal analysis etc.
- This provides a complementary angle on computational anatomy, which more often focuses on variational formulations, like LDDMM.
- This also brings new open problems: extension to new shape modalities, their numerical implementation, new applications for medical data.