Statistics of Shape: Eigen Shapes "PCA and PGA"

Sarang Joshi

Departments of Radiation Oncology, Biomedical Engineering and Computer Science

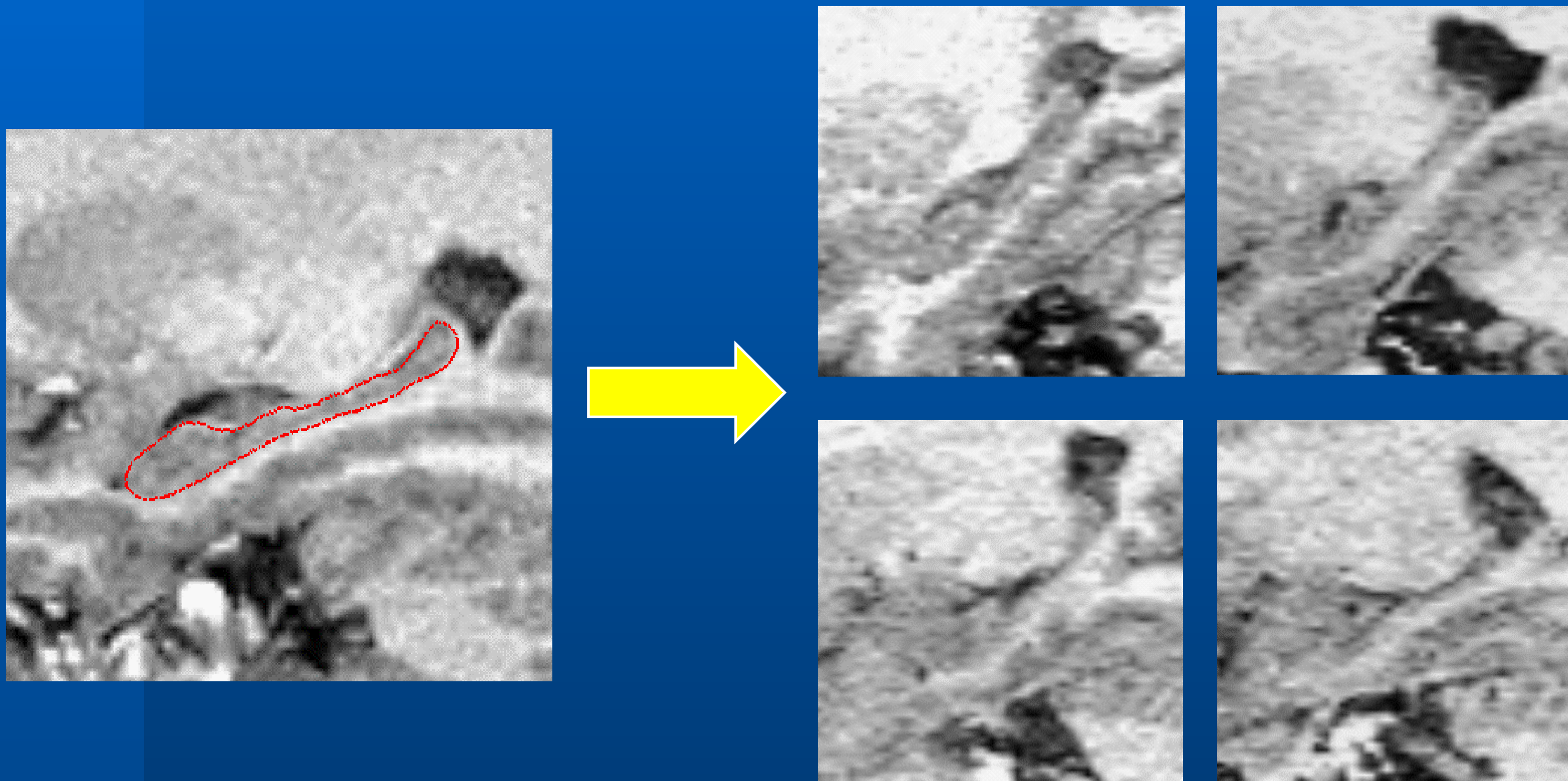
University of North Carolina at Chapel Hill

Study of the Shape of the Hippocampus

- Two approaches to study the shape variation of the hippocampus in populations:
  - Statistics of deformation fields using “Principal Components Analysis”
  - Statistics of medial descriptions using Lei Groups:- “Principal Geodesic Analysis”

Statistics of Deformation Fields

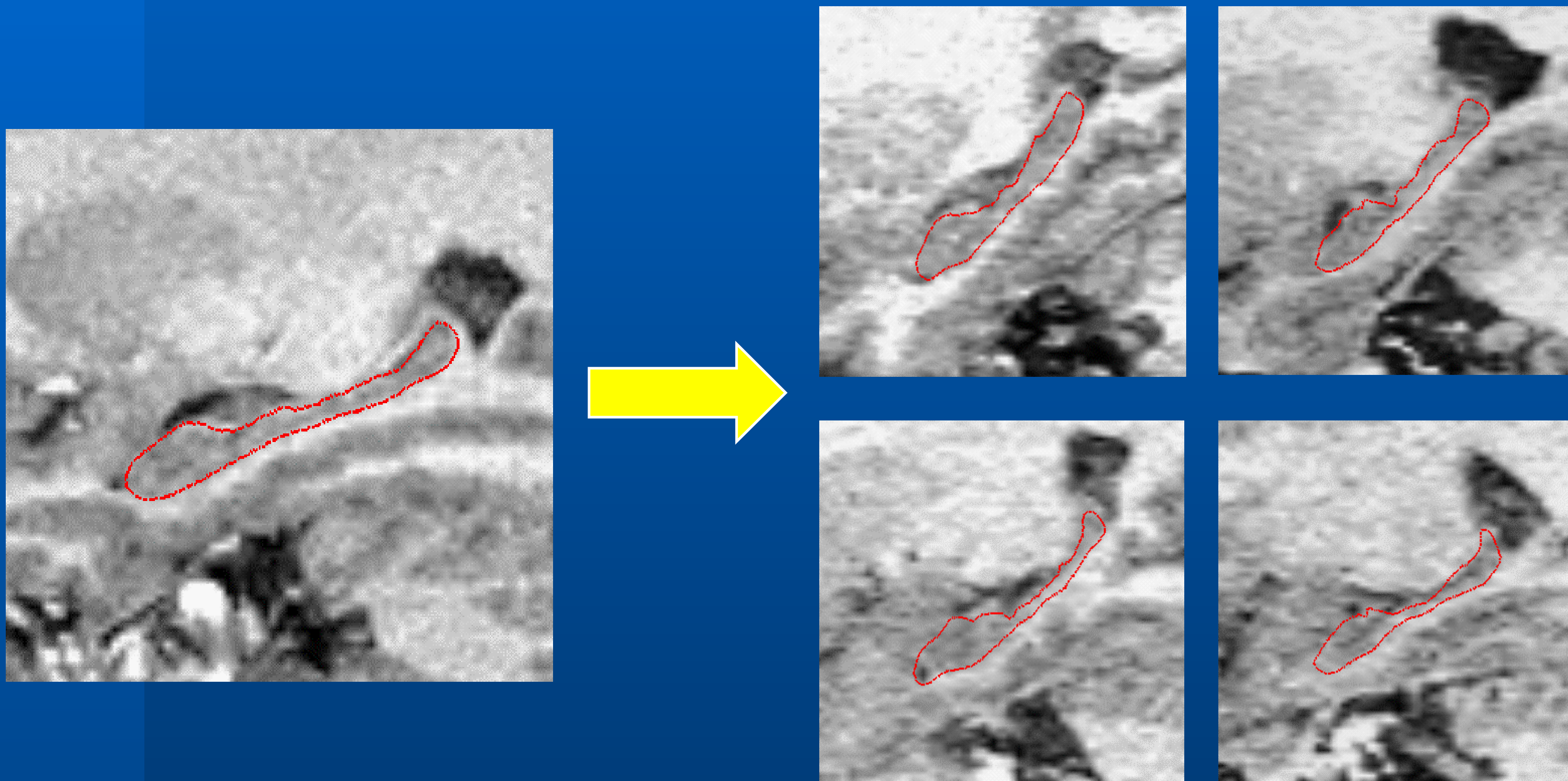
Hippocampal Mapping



Atlas

Patients

Hippocampal Mapping

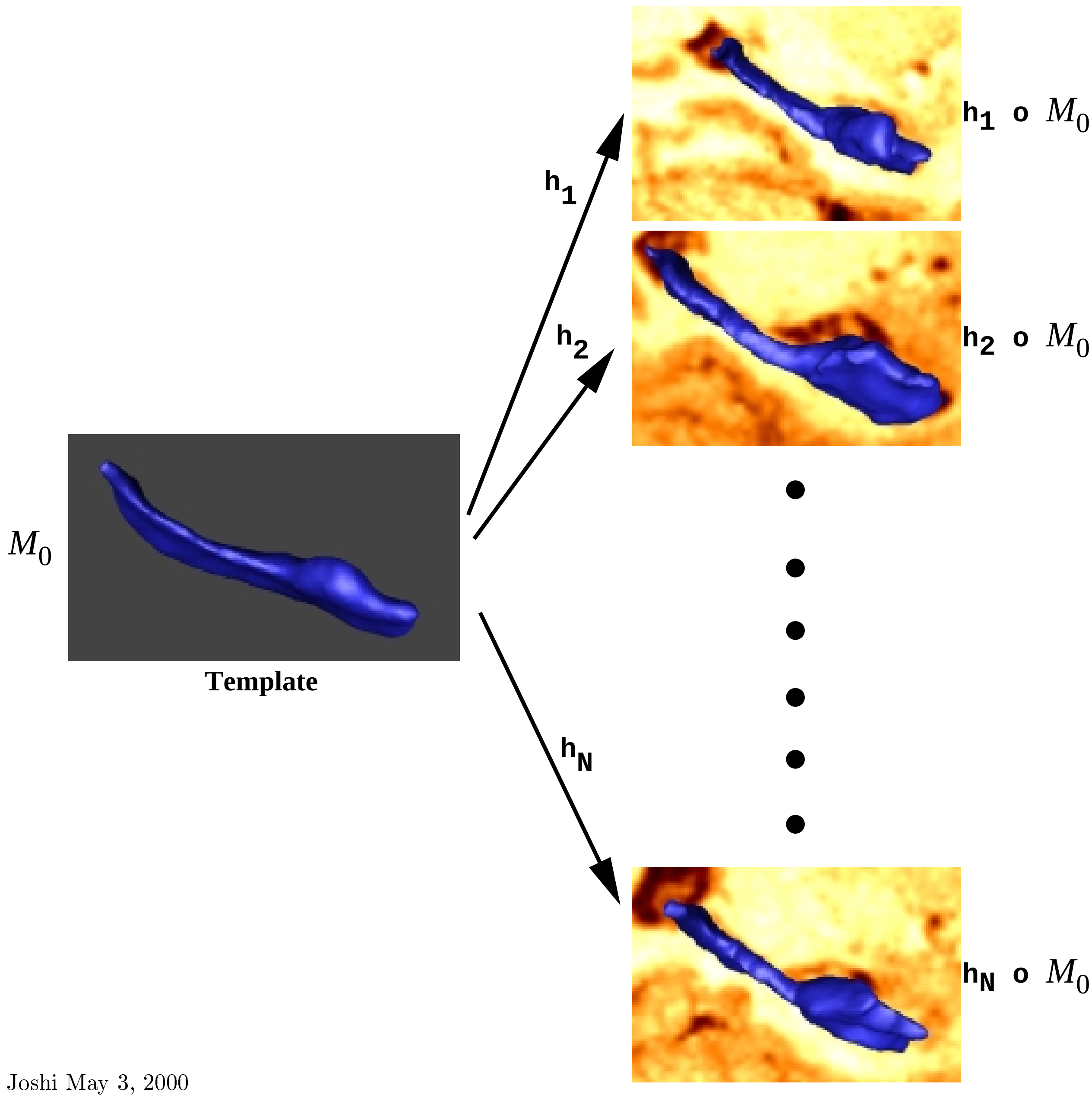


Atlas Subjects

Shape of 2-D Sub-Manifolds of the Brain: Hippocampus.

The provisory template hippocampal surface $\mathcal{M}_0$ is carried onto the family of targets:

$$\mathcal{M}_0 \underset{h_1^{-1}}{\overset{h_1}{\rightleftarrows}} \mathcal{M}^1, \mathcal{M}_0 \underset{h_2^{-1}}{\overset{h_2}{\rightleftarrows}} \mathcal{M}^2, \cdots, \mathcal{M}_0 \underset{h_N^{-1}}{\overset{h_N}{\rightleftarrows}} \mathcal{M}^N \; .$$

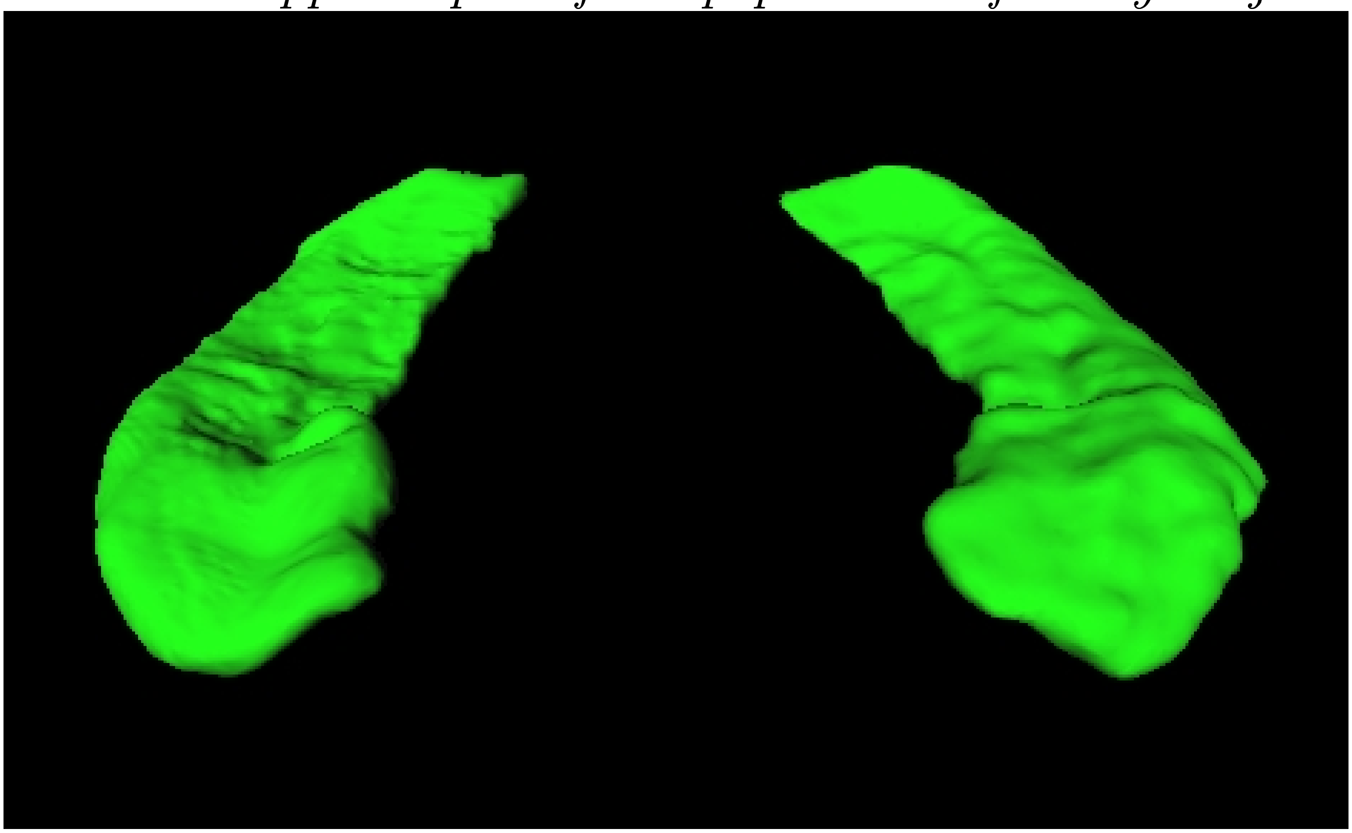


Shape of 2-D Sub-Manifolds of the Brain: Hippocampus.

- The mean transformation and the template representing the entire population:

$$\bar{h} = \frac{1}{N}\sum_{i=1}^{N} h_i \quad , \quad \mathcal{M}_{temp} = \bar{h} \circ \mathcal{M}_0 \ .$$

The mean hippocampus of the population of thirty subjects.



Shape of 2-D Sub-Manifolds of the Brain: Hippocampus.

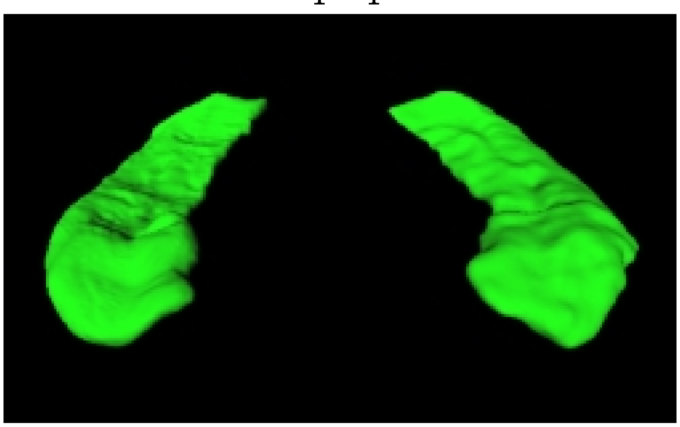
- Mean hippocampus representing the control population:

$$\bar{h}_{control} = \frac{1}{N_{control}} \sum_{i=1}^{N_{control}} h_i^{control} \ , \ \mathcal{M}_{control} = \bar{h}_{control} \circ \mathcal{M}_0 \ .$$

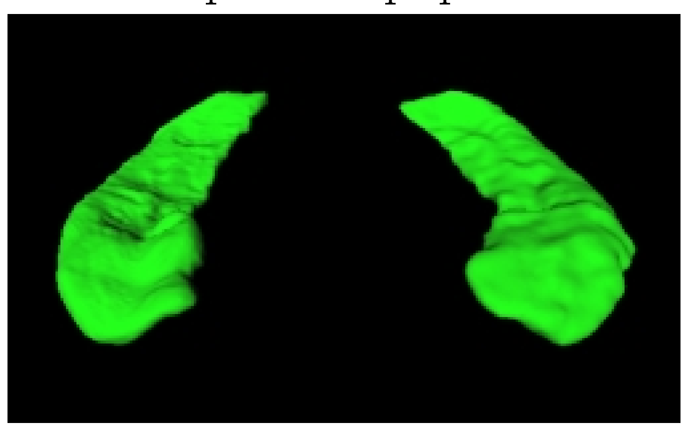
- Mean hippocampus representing the Schizophrenic population:

$$\bar{h}_{schiz} = \frac{1}{N_{schiz}} \sum_{i=1}^{N_{schiz}} h_i^{schiz} \ , \ \mathcal{M}_{schiz} = \bar{h}_{schiz} \circ \mathcal{M}_0 \ .$$

Control population



Schizophrenic population



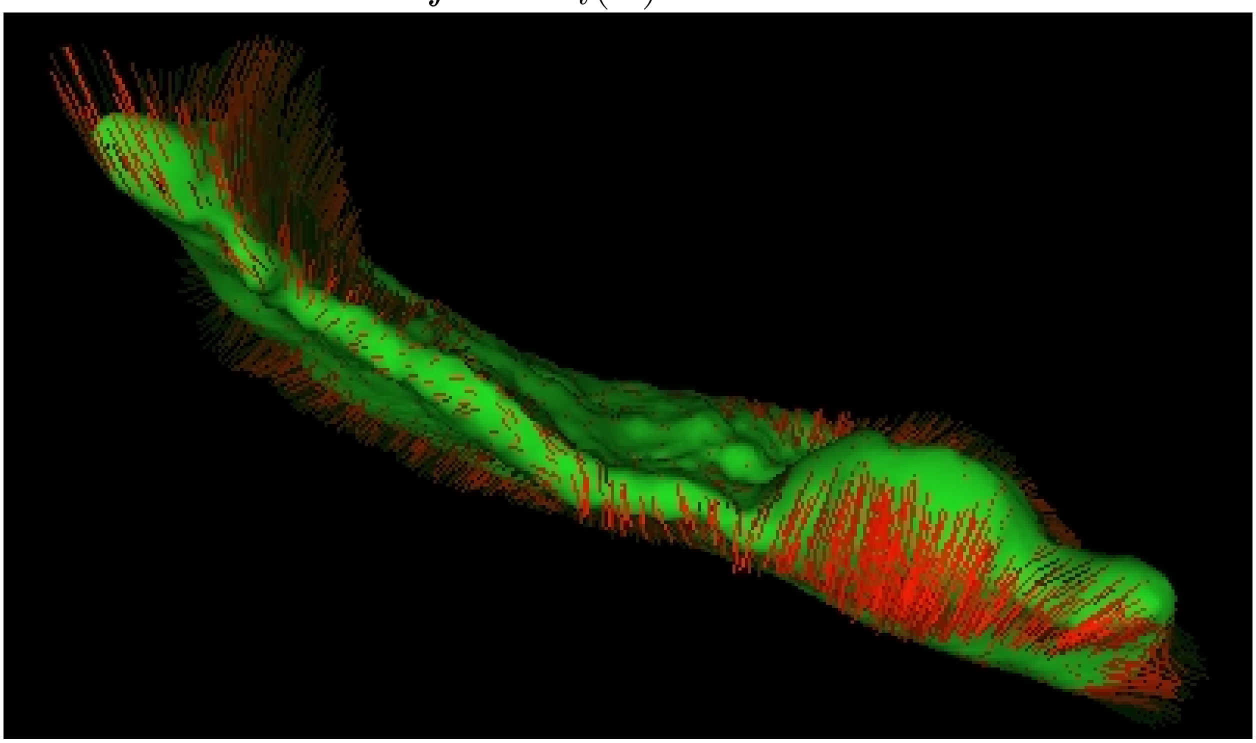
Gaussian Random Vector Fields on 2-D Sub-Manifolds.

- Hippocampi $\mathcal{M}^i, i = 1, \cdots, N$ deformation of the mean $\mathcal{M}_{temp}$:

$$\mathcal{M}^i : \{y | y = x + u_i(x)\ , x \in \mathcal{M}_{temp}\}$$

$$u_i(x) = h_i(x) - x, x \in \mathcal{M}_{temp}\ .$$

Vector field $u_i(x)$ shown in red.



- Construct Gaussian random vector fields over sub-manifolds.

Gaussian Random Vector Fields on 2-D Sub-Manifolds.

• Let $\mathcal{H}(\mathcal{M})$ be the Hilbert space of square integrable vector fields on $\mathcal{M}$. Inner product on the Hilbert space $\mathcal{H}(\mathcal{M})$:

$$\langle f, g \rangle = \sum_{i=1}^{3} \int_{\mathcal{M}} f^i(x) g^i(x) d\nu(x)$$

where $d\nu$ is a measure on the oriented manifold $\mathcal{M}$.

Definition 1 *The random field* $\{U(x), x \in \mathcal{M}\}$ *is a* **Gaussian random field** *on a manifold* $\mathcal{M}$ *with mean* $\mu_u \in \mathcal{H}(\mathcal{M})$ *and covariance operator* $K_u(x, y)$ *if* $\forall f \in \mathcal{H}(\mathcal{M})$, $\langle f, \cdot \rangle$ *is normally distributed with mean* $m_f = \langle \mu_u, f \rangle$ *and variance* $\sigma_f^2 = \langle K_u f, f \rangle$

• Gaussian field is completely specified by it's mean μ_u and the covariance operator $K_u(x, y)$.

• Construct Gaussian random fields as a quadratic mean limit using a complete $\mathbb{R}^3$-valued orthonormal basis

$$\{\phi_k, k = 1, 2, \cdots\} \ , \ \langle \phi_i, \phi_j \rangle = 0 \ , i \neq j$$

Gaussian Random Vector Fields on 2-D Sub-Manifolds.

Theorem 1 *Let $\{U(x), x \in \mathcal{M}\}$ be a Gaussian random vector field with mean $m_U \in \mathcal{H}$ and covariance K_U of finite trace. There exists a sequence of finite dimensional Gaussian random vector fields $\{U_n(x)\}$ such that*

$$U(x) \stackrel{\text{q.m.}}{=} \lim_{n\to\infty} U_n(x)$$

where

$$U_n(x) = \sum_{k=1}^{n} Z_k(\omega)\phi_k(x) \ ,$$

$\{Z_k(\omega), k = 1, \cdots\}$ are independent Gaussian random variables with fixed means $E\{Z_k\} = \mu_k$ and covariances $E\{|Z_i|^2\} - E\{Z_i\}^2 = \sigma_i^2 = \lambda_i, \sum_i \lambda_i < \infty$ and (hk, λ_k) are the eigen functions and the eigen values of the covariance operator K_U:

$$\lambda_i\phi_i(x) = \int_{\mathcal{M}} K_U(x,y)\phi_i(y)d\nu(y) \ ,$$

where $d\nu$ is the measure on the manifold $\mathcal{M}$.

If $d\nu$, the surface measure on $\bar{\mathcal{M}}_{temp}$ is atomic around the points x_k then $\{\phi_i\}$ satisfy the system of linear equations

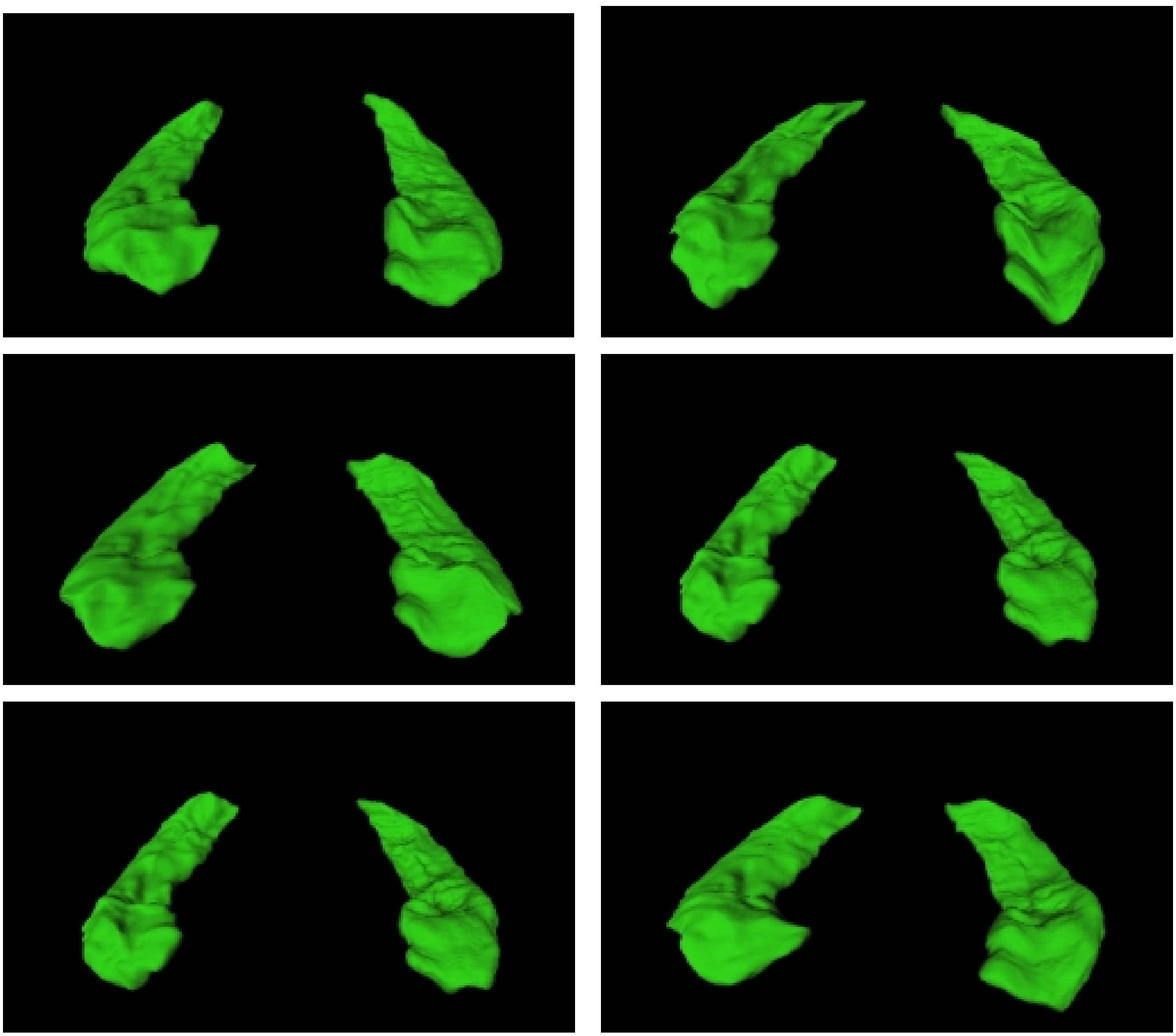
$$\lambda_i\phi_i(x_k) = \sum_{j=1}^{M} \hat{K}_U(x_k, y_j)\phi_i(y_j)\nu(y_j) \quad , i = 1, \cdots, N \ ,$$

where $\nu(y_j)$ is the surface measure around point y_j.

Eigen Shapes of the Hippocampus.

- Eigen shapes $\mathcal{E}^i, i = 1, \cdots, N$ defined as:

$$\mathcal{E}^i = \{x + (\lambda_i)\phi_i(x) : x \in \bar{\mathcal{M}}_{temp}\} \ .$$



- Eigen shapes completely characterize the variation of the sub-manifold in the population.

Statistical Significance of Shape Difference Between Populations.

• Assume that $\{u_j^{schiz}, u_j^{control}\}, j = 1, \cdots, 15$ are realizations from a Gaussian process with mean $\bar{u}_{schiz}$ and $\bar{u}_{control}$ and common covariance K_U.

Statistical hypothesis test on shape difference:

$$\begin{aligned} H_0 : \bar{u}_{norm} &= \bar{u}_{schiz} \\ H_1 : \bar{u}_{norm} &\neq \bar{u}_{schiz} \end{aligned}$$

•Expand the deformation fields in the eigen functions ϕ_i:

$$\begin{aligned} u_N^{schiz(j)}(x) &= \sum_{i=1}^{N} Z_i^{schiz(j)} \phi_i(x) \\ u_N^{control(j)}(x) &= \sum_{i=1}^{N} Z_i^{control(j)} \phi_i(x) \end{aligned}$$

• $\{Z_j^{schiz}, Z_j^{control}, j = 1, \cdots, 15\}$ Gaussian random vectors with means $\bar{Z}_{schiz}$ and $\bar{Z}_{control}$ and covariance Σ.

Hotelling's T^2 test:

$$T_N^2 = \frac{M}{2} (\hat{\bar{Z}}_{norm} - \hat{\bar{Z}}_{schiz})^T \hat{\Sigma}^{-1} (\hat{\bar{Z}}_{norm} - \hat{\bar{Z}}_{schiz}) \ .$$

N	T_N^2	p-value : $P_N(H_0)$
3	9.8042	0.0471
4	14.3086	0.0300
5	14.4012	0.0612
6	19.6038	0.0401

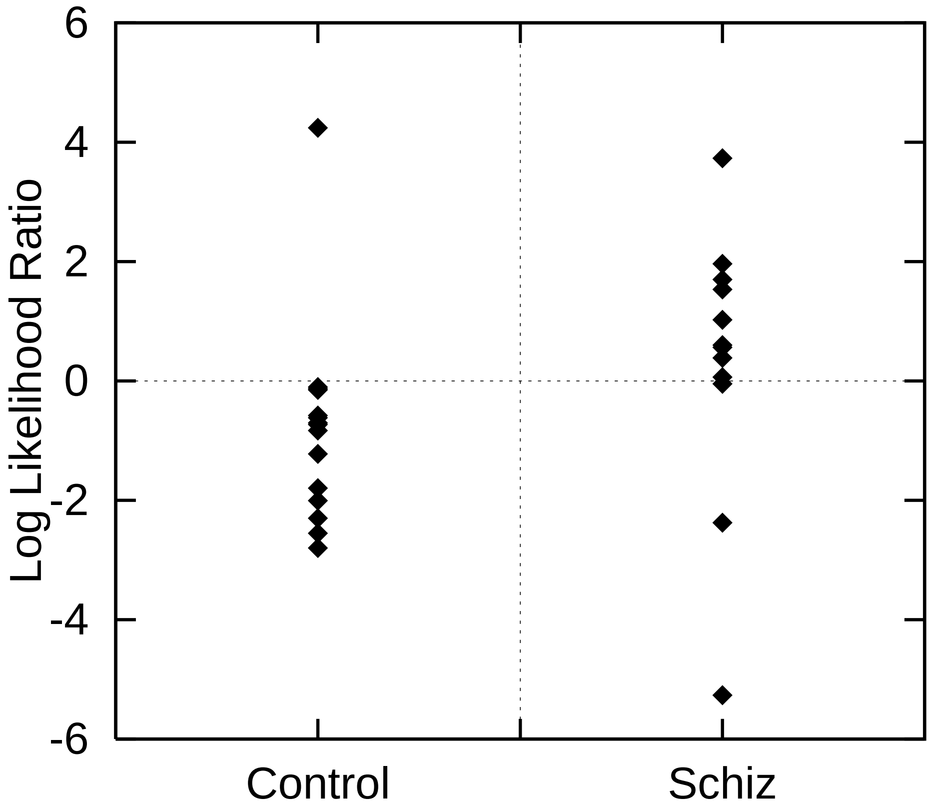
N: number of eigen functions.

Bayesian Classification on Hippocampus Shape Between Population.

• Bayesian log-likelihood ratio test: H_0: normal hippocampus, H_1: schizophrenic hippocampus.

$$\Lambda_N = -(Z - \hat{\bar{Z}}_{schiz})^{\dagger}\hat{\Sigma}^{-1}(Z - \hat{\bar{Z}}_{schiz}) + (Z - \hat{\bar{Z}}_{norm})^{\dagger}\hat{\Sigma}^{-1}(Z - \hat{\bar{Z}}_{norm}) \underset{H_1}{\overset{H_0}{\lessgtr}} 0$$

• Use Jack Knife for estimating probability of classification:

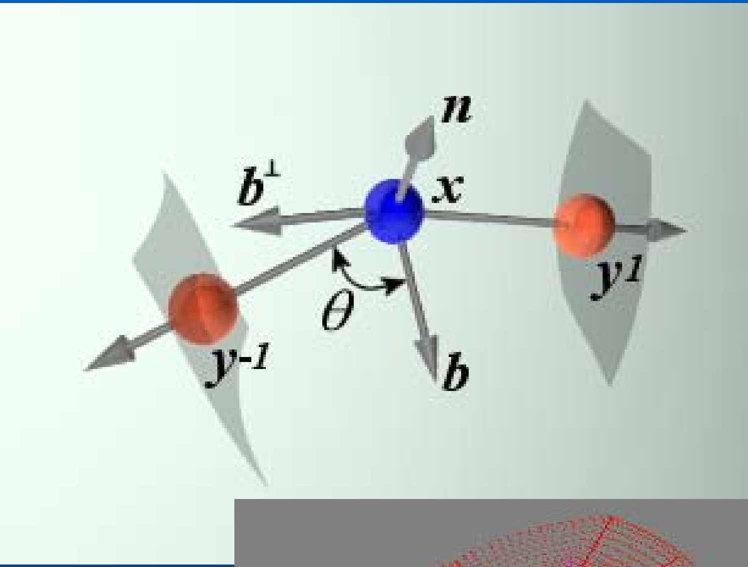


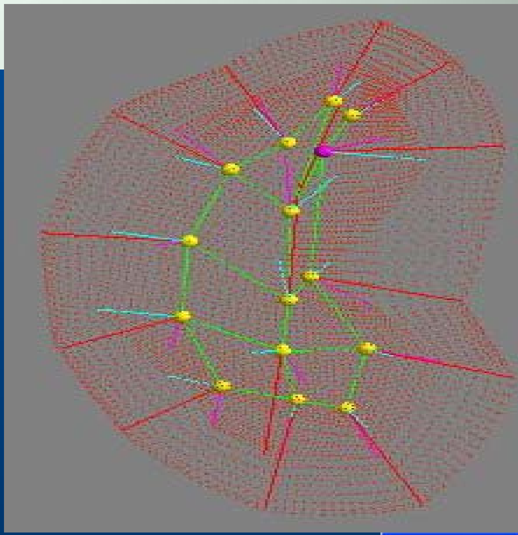
Statistics of Medial descriptions

- Each figure a quad mesh of medial atoms:

$$\{m^0_{i,j} : i = 1 \cdots N, j = 1 \cdots M\}$$

$$m^0_{i,j} = (x_{i,j}, r, F, \theta)$$



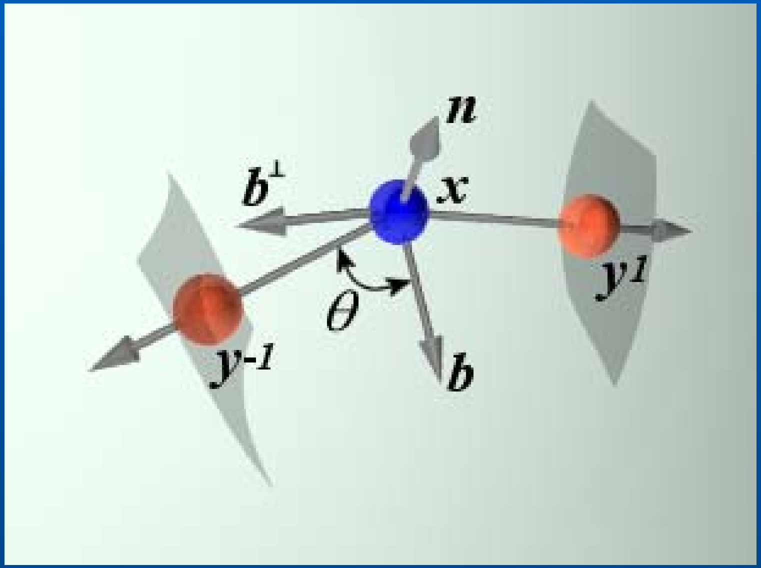


- Medial atom parameters include angles and rotations.
- Medial atoms do not form a Hilbert Space
  - Cannot use “Eigen Shape” for statistical characterization!!

MIDAG@UNC

Statistics of Medial descriptions

- Set of all Medial Atoms forms Lie-Group



$$m = (x_{i,j}, r, F, \theta)$$

$$m \in \mathbb{R}^3 \times \mathbb{R}^+ \times SO(3) \times SO(2)$$

$\mathbb{R}^3$:Position x

$\mathbb{R}^+$:Radius r

$SO(3)$:Frame

$SO(2)$:Object angle

Lie Groups

- A Lie group is a group G which is also a differential manifold where the group operations are differential maps.

$$\mu : (x, y) \mapsto xy : G \times G \mapsto G$$

$$\iota : x \mapsto x^{-1} : G \mapsto G$$

•Both composition and the inverse are differential maps

$$\mathrm{R}^3 : \mu(x, y) = x + y, x^{-1} = -x$$

$$\mathrm{R}^+ : Multiplica\,tive\ reals \quad \mu(x, y) = xy, x^{-1} = \frac{1}{x}$$

$$SO(3) : 3\times 3\ Orthogonal\ \ Matrix\ \ Group$$

$$SO(2) : 2\times 2 Orthogonal\ \ Matrix\ \ Group$$

Lie Group Means

- Algebraic mean not defined on Lie Groups
- Use geometric definition:
  - Remanian Distance well defined on a Manifold.
- Given N medial atoms $\{m_i : i = 1 \cdots N\}$
  the mean $\overline{m}$ is defined as the group element that minimizes the average squared distance to the data.

$$\overline{m} = \arg\min \frac{1}{N} \sum_{i=1}^{N} | d(m, m_i) |^2$$

- No closed form solution need to use Lie-Group optimization techniques.

Geodesic Curves

- Medial manifold is curved and hence no straight lines.
- Distance minimizing **Geodesic** curves are analogous to straight lines in Euclidian Space.
- Geodesics in Lie Groups are given by the exponent map:

$$g(t) = \exp(tA)$$

- Geodesics are one parameter sub-groups analogous to 1-dimensional subspaces in R^N

Principal Geodesics

- Since the set of all medial atoms is a curved manifold linear PCA is not defined as well.

- Principal Geodesics are defined as the geodesics that minimize residual distance.
  - No closed form solution: Needs non linear optimization.