# Estimating metrics suitable to an empirical manifold of shapes, using transport against the curse of dimensionality

Guillaume Charpiat

Pulsar Project

INRIA Workshop on Statistical Learning

IHP

05/12/2011

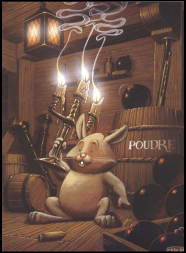
# Map

- **Introduction**
  - Motivation
  - Issues
- **Searching for solutions**
  - Main existing approaches and their limitations
  - Main idea
- **The approach**
  - Shape matching
  - Transport
  - Metric estimation (statistics on deformations)
  - Theory
- **Future work**

# Introduction

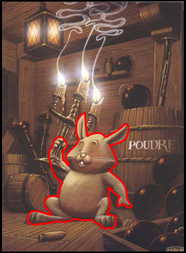
## Image Segmentation

- Find a contour in a given image
- The best curve for a given segmentation criterion
- Criterion based on color homogeneity, texture, edge detectors, etc.




Image

$\longrightarrow$




Segmentation

# Introduction

## Image Segmentation

- Find the best contour for a given criterion

## Variational Method

- Energy $E$ to minimize with respect to a curve $C$
- Compute the derivative of the energy
- Gradient descent: $\partial_t C = -\nabla E(C)$
- Initialization $\longrightarrow$ local minimum
- Other methods: graph cuts (suitable for few energies)

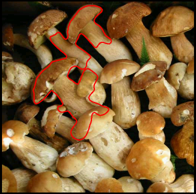
# Introduction

## Image Segmentation

- Find the best contour for a given criterion

## Variational Method

- Minimize criterion by gradient descent with respect to the contour
- Most criteria: no shape information



# Introduction

## Image Segmentation

- Find the best contour for a given criterion

## Variational Method

- Minimize criterion by gradient descent with respect to the contour

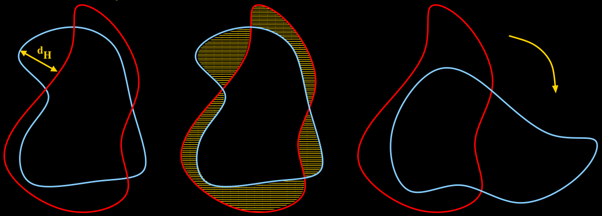
## Shape Statistics

- Sample set of contours from already segmented images
- Shape variability ?
- Shape prior ?

# Introduction

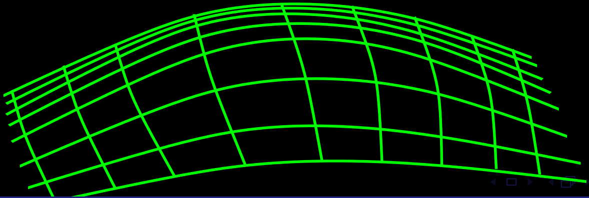
## Motivation

- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)



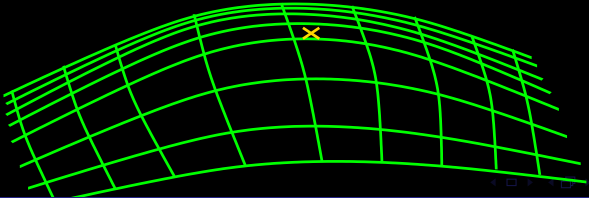
# Introduction
## Motivation

- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)



# Introduction
## Motivation

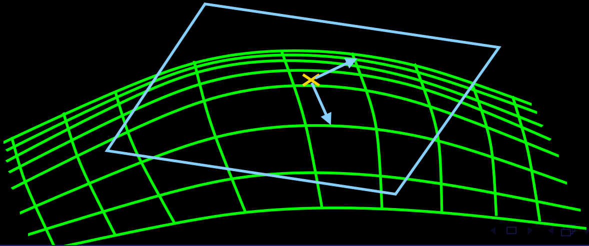
- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)



# Introduction

## Motivation

- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)



# Introduction

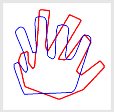
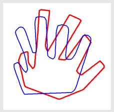
## Motivation

- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)

  - $L^2$ norm of instantaneous deformations
  - $L^2$ + curvature, $H^1$
  - rigid motion more probable $\Longrightarrow$ associated metric

# Introduction
## Motivation

- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)
  - $L^2$ norm of instantaneous deformations
  - $L^2$ + curvature, $H^1$
  - rigid motion more probable $\Longrightarrow$ associated metric

$L^2$ inner product  vs.  rigidifying inner product

# Introduction
## Motivation

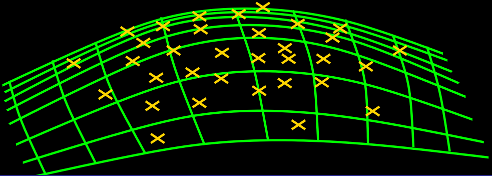
- Shape spaces : which metric ?
  (to define similarity/distance between shapes)
  - Hausdorff distance
  - Symmetric difference area
  - Quotients by transformation groups (rotation, translation, scaling, affine...)
- Shape evolution, morphing : priors on probable deformations ?
  $\Longrightarrow$ Which local metric on deformations ?
  (metric on the manifold of shapes)
  - $L^2$ norm of instantaneous deformations
  - $L^2$ + curvature, $H^1$
  - rigid motion more probable $\Longrightarrow$ associated metric
- $\Longrightarrow$ learn the suitable metric from examples (datasets of shapes)

# Issues

- Sparse sets of highly varying shapes
  - e.g. human silhouettes
  - high intrinsic dimension ($\geqslant 30$)
  - $\Longrightarrow$ no dense training set

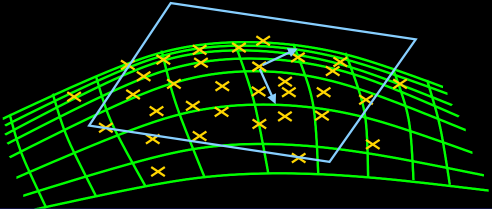
# Issues

- Sparse sets of highly varying shapes
  - e.g. human silhouettes
  - high intrinsic dimension ($\geqslant 30$)
  - $\Longrightarrow$ no dense training set



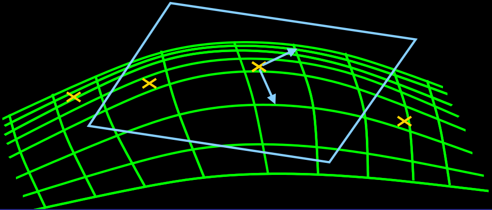
# Issues

- Sparse sets of highly varying shapes
  - e.g. human silhouettes
  - high intrinsic dimension ($\geqslant 30$)
  - $\Longrightarrow$ no dense training set



# Issues

- Sparse sets of highly varying shapes
  - e.g. human silhouettes
  - high intrinsic dimension ($\geqslant 30$)
  - $\Longrightarrow$ no dense training set



# Issues

- Sparse sets of highly varying shapes
  - e.g. human silhouettes
  - high intrinsic dimension ($\geqslant 30$)
  - $\Longrightarrow$ no dense training set

- to compare quantities defined on different shapes :
  need for correspondences
  - match shape with different topologies ?
  - very frequent topological changes

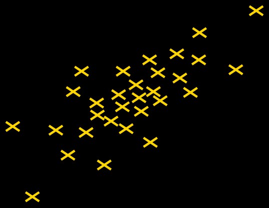
# Searching for solutions

## Main existing approaches and their limitations

Approach 1 : *mean* + *modes* model
Approach 2 : distance-based approaches, such as kernel methods

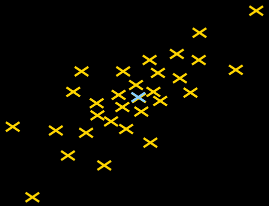
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mean $M$, shapes $S_i$, warpings $W_{M \to S_i}$



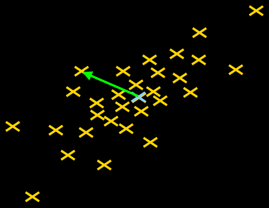
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mean $M$, shapes $S_i$, warpings $W_{M \to S_i}$



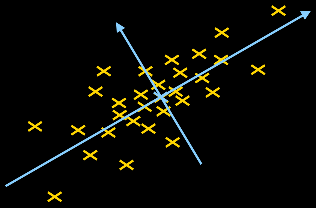
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mean $M$, shapes $S_i$, warpings $W_{M \to S_i}$



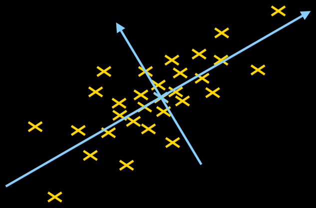
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mean $M$, shapes $S_i$, warpings $W_{M \to S_i}$
- PCA : diagonalize correlation matrix $C$ : $C_{ij} = \left\langle W_{M \to S_i} \middle| W_{M \to S_j} \right\rangle$
  $\implies$ eigenmodes $e_k$ with eigenvalues $\lambda_k$ : best coordinate system



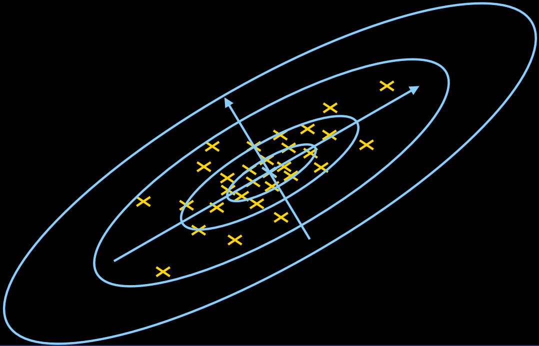
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mean $M$, shapes $S_i$, warpings $W_{M\to S_i}$
- PCA : diagonalize correlation matrix $C$ : $C_{ij} = \left\langle W_{M\to S_i} \middle| W_{M\to S_j} \right\rangle$
  $\Longrightarrow$ eigenmodes $e_k$ with eigenvalues $\lambda_k$ : best coordinate system
- any new deformation $W$ of $M$ :
  $W = \sum_k \alpha_k e_k +$ noise



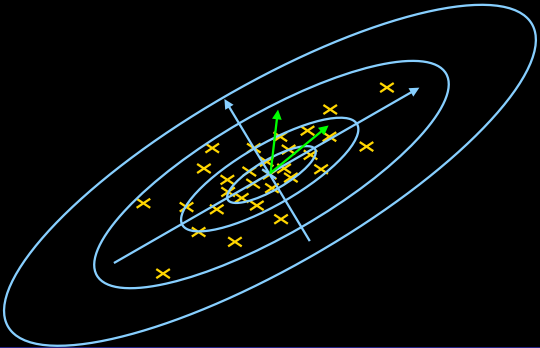
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mahalanobis distance : $d(M+W,(S)) = \sum_k \frac{\alpha_k^2}{\lambda_k^2}$



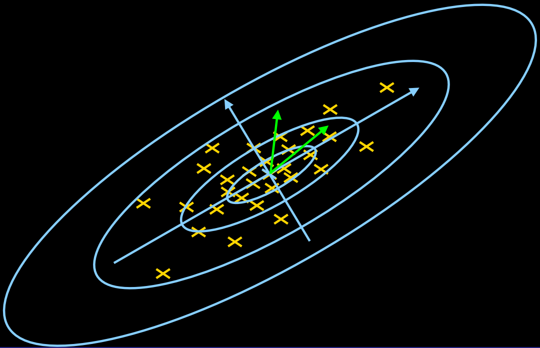
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mahalanobis distance : $d(M+W,(S)) = \sum_k \frac{\alpha_k^2}{\lambda_k^2}$
- associated inner product on deformations, in the tangent space of $M$:
  $\langle W_1 | W_2 \rangle = \sum_k \frac{1}{\lambda_k^2} \alpha_{1,k} \alpha_{2,k}$



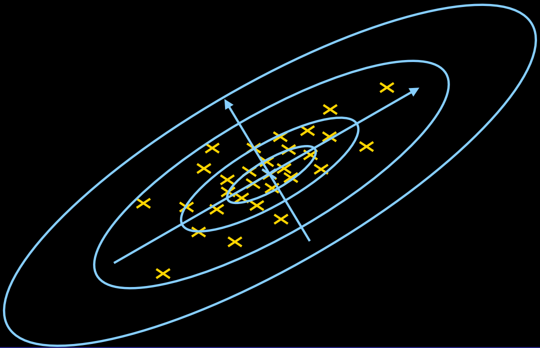
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Mahalanobis distance : $d(M+W,(S)) = \sum_k \frac{\alpha_k^2}{\lambda_k^2}$
- associated inner product on deformations, in the tangent space of $M$:
  $\langle W_1 | W_2 \rangle = \sum_k \frac{1}{\lambda_k^2} \alpha_{1,k} \alpha_{2,k}$
- defines a deformation cost $\|W\|^2 = \langle W | W \rangle$



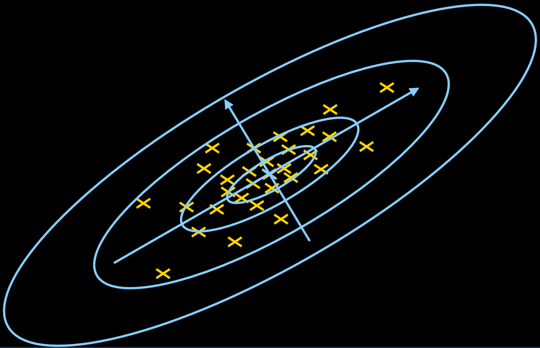
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- probability $p(W) \propto exp(-\sum_k \frac{\alpha_k^2}{2\lambda_k^2})$ : Gaussian distribution



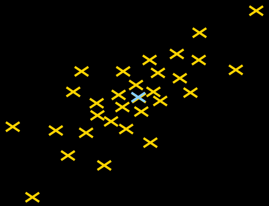
## Mean + modes model : PCA in tangent space (Gaussian distribution)

- probability $p(W) \propto exp(-\sum_k \frac{\alpha_k^2}{2\lambda_k^2})$ : Gaussian distribution
- defines a Gaussian shape prior



## Mean + modes model : PCA in tangent space (Gaussian distribution)

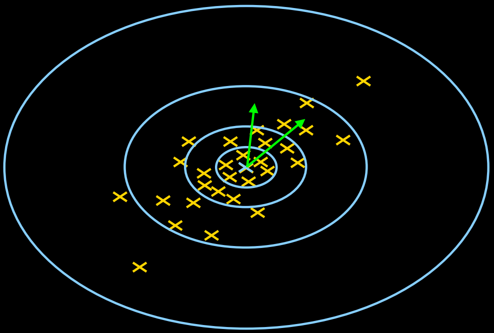
- Empirical distribution : $\mathcal{D}_{emp} = \sum_i \delta_{W_{M \to S_i}}$
  (possibly smoothed by a kernel)



## Mean + modes model : PCA in tangent space (Gaussian distribution)

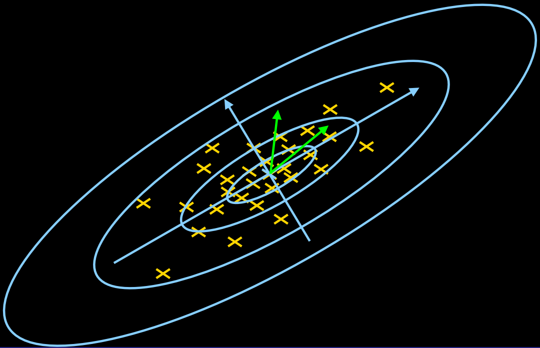
- Empirical distribution : $\mathcal{D}_{emp} = \sum_i \delta_{W_{M \to S_i}}$

  (possibly smoothed by a kernel)
- Any inner product $< | >_P$ in tangent space of the mean
  $\Longrightarrow$ Gaussian distribution $\mathcal{D}_P(W) \propto \exp(-\|W\|_P^2)$



## Mean + modes model : PCA in tangent space (Gaussian distribution)

- Empirical distribution : $\mathcal{D}_{emp} = \sum_i \delta_{W_{M \to S_i}}$
  (possibly smoothed by a kernel)
- Any inner product $< | >_P$ in tangent space of the mean
  $\Longrightarrow$ Gaussian distribution $\mathcal{D}_P(W) \propto \exp(-\|W\|_P^2)$
- Best $P$ for Kullback-Leibler$(\mathcal{D}_P|\mathcal{D}_{emp})$ : PCA!



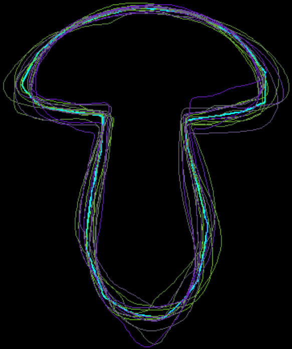
# Approach 1 : mean + modes model

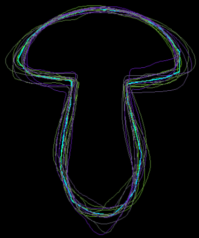
$\hookrightarrow$ example from my PhD thesis


Automatic alignment

$\longrightarrow$

and average shape computation

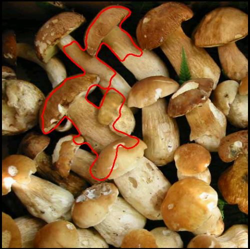




Statistics (PCA) on deformation fields
$\longrightarrow$
between the mean shape and each sample

modes of deformation
= deformation prior
= Gaussian probabilistic model

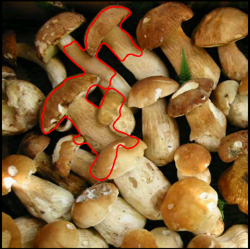
Example of application : image segmentation with shape prior



without shape prior

with shape prior

Example of application : image segmentation with shape prior

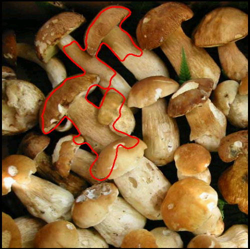


without shape prior

with shape prior

- requires a mean shape (does not always make sense, e.g. person walking)

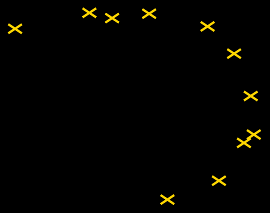
Example of application : image segmentation with shape prior


without shape prior


with shape prior

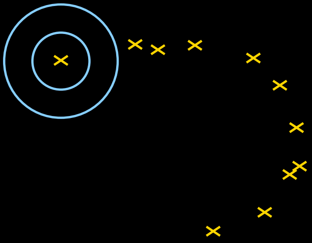
- requires a mean shape (does not always make sense, e.g. person walking)
- requires all deformations between the mean and samples :
  $\Longrightarrow$ relatively similar sample shapes (otherwise, not reliable)

# Approach 2 : distance-based methods (e.g. kernel methods)



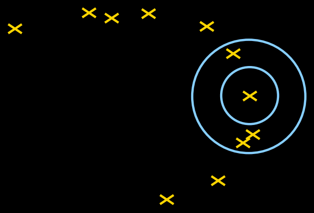
- Kernel : symmetric definite positive function $k(x, y)$
- Expresses the similarity between $x$ and $y$
- Typically, the Gaussian kernel : $k(x, y) = exp(-d(x, y)^2)$
- For each point $x_i$ : $k_i(y) := k(x_i, y)$

# Approach 2 : distance-based methods (e.g. kernel methods)



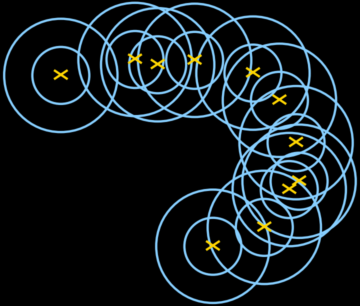
- Kernel : symmetric definite positive function $k(x, y)$
- Expresses the similarity between $x$ and $y$
- Typically, the Gaussian kernel : $k(x, y) = exp(-d(x, y)^2)$
- For each point $x_i$ : $k_i(y) := k(x_i, y)$

## Approach 2 : distance-based methods (e.g. kernel methods)



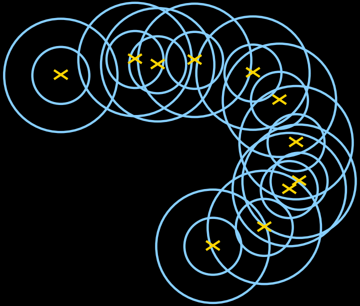
- Kernel : symmetric definite positive function $k(x, y)$
- Expresses the similarity between $x$ and $y$
- Typically, the Gaussian kernel : $k(x, y) = exp(-d(x, y)^2)$
- For each point $x_i$ : $k_i(y) := k(x_i, y)$

## Approach 2 : distance-based methods (e.g. kernel methods)



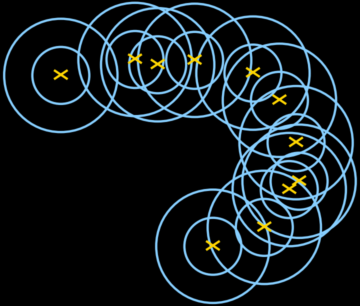
- Kernel : symmetric definite positive function $k(x, y)$
- Expresses the similarity between $x$ and $y$
- Typically, the Gaussian kernel : $k(x, y) = exp(-d(x, y)^2)$
- For each point $x_i$ : $k_i(y) := k(x_i, y)$

# Approach 2 : distance-based methods (e.g. kernel methods)



- choice of a distance, of a kernel ?
- distance between 2 shapes : not much informative (wrt deformations)
- rebuild geometry of space of shapes from distances ?
- distances are not reliable/meaningful for far shapes
- $\Longrightarrow$ needs for a representative neighborhood, i.e. a high dataset density (not affordable)

# Approach 2 : distance-based methods (e.g. kernel methods)



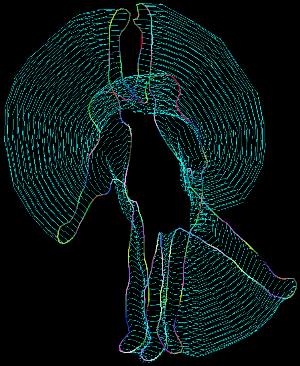
- choice of a distance, of a kernel ?
- distance between 2 shapes : not much informative (wrt deformations)
- rebuild geometry of space of shapes from distances ?
- distances are not reliable/meaningful for far shapes
- $\Longrightarrow$ needs for a representative neighborhood, i.e. a high dataset density
- in a high-dimensional manifold, all distances are similar, and all points are on the boundary of the manifold
- $\Longrightarrow$ cannot work, need for more information than distances

# Main idea

- consider deformations (not just distances)
- should not require high density of training set
- no magic (to handle/interpolate sparse sets) : add a prior

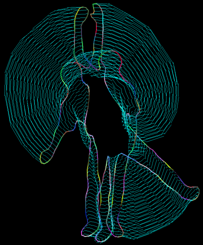
# Main idea

- consider deformations (not just distances)
- should not require high density of training set
- no magic (to handle/interpolate sparse sets) : add a prior
- prior chosen : transported deformations make sense,
  i.e. a deformation observed on one shape can be applied to other shapes



# Main idea

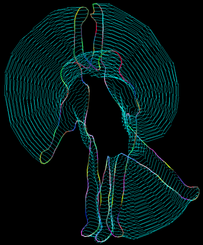
- consider deformations (not just distances)
- should not require high density of training set
- no magic (to handle/interpolate sparse sets) : add a prior
- prior chosen : transported deformations make sense, i.e. a deformation observed on one shape can be applied to other shapes



- transport requires correspondences

# Main idea

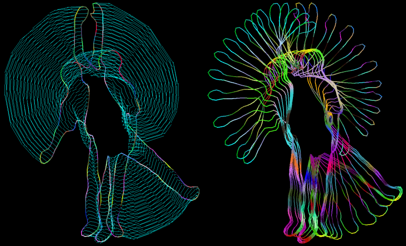
- consider deformations (not just distances)
- should not require high density of training set
- no magic (to handle/interpolate sparse sets) : add a prior
- prior chosen : transported deformations make sense,
  i.e. a deformation observed on one shape can be applied to other shapes



- transport requires correspondences
- but shape matching reliable only for close shapes

## Main idea

- consider deformations (not just distances)
- should not require high density of training set
- no magic (to handle/interpolate sparse sets) : add a prior
- prior chosen : transported deformations make sense,
  i.e. a deformation observed on one shape can be applied to other shapes



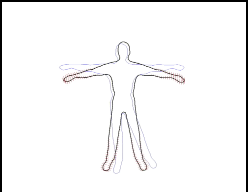
- transport requires correspondences
- but shape matching reliable only for close shapes
- $\Longrightarrow$ compute correspondences between close shapes only, and combine small steps of reliable correspondences to build longer-distance correspondences

# Map

- Close shape matching
- Transport
- Metric estimation (statistics on transported deformations)
- Theoretical justifications

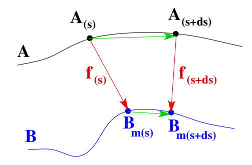
# Shape matching

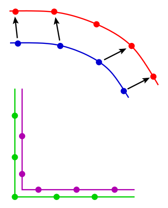
## Simple case : two shapes, A and B, with one connected component



$$\inf_{f:A\to B} \int_A \|f\|^2 + \alpha \|\nabla f\|^2 dA$$

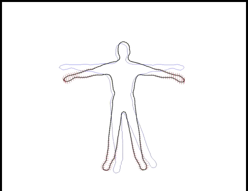
- shape sampling
- dynamic time warping
- theory & experiments :
  higher sampling rate on target



# Shape matching

## Simple case : two shapes, A and B, with one connected component

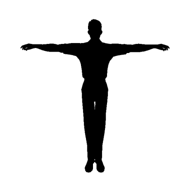


$$\inf_{f:A\to B} \int_A \|f\|^2 + \alpha \|\nabla f\|^2 dA$$

- shape sampling
- dynamic time warping
- theory & experiments :

higher sampling rate on target

## Usual case : random topologies



Usual cases = more complex
(more than 10 connected components in this silhouette)
but
one connected component $\to \bigcup_i$ connected components
= the same

## Further possible improvements

- as such, allows appearing points (mismatches)
- allows disappearing points : matching to $\emptyset$ with a fixed high cost
- pb : better matchings, but energy value loses meaning

### Further possible improvements

- as such, allows appearing points (mismatches)
- allows disappearing points : matching to $\varnothing$ with a fixed high cost
- pb : better matchings, but energy value loses meaning

### Drawbacks

- specific to planar curves
- not symmetric : $m_{A \to B} \neq m^{-1}_{B \to A}$

# Transport

## Local transport

- Set of shapes $(S_i)_{i \in I}$ (e.g. from a video segmentation)

# Transport

## Local transport

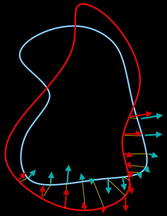
- Set of shapes $(S_i)_{i \in I}$ (e.g. from a video segmentation)
- Two shapes $S_i$ and $S_j$ $\Longrightarrow$ their correspondence field $m_{i \to j}$

# Transport

## Local transport

- Set of shapes $(S_i)_{i\in I}$ (e.g. from a video segmentation)
- Two shapes $S_i$ and $S_j$ $\implies$ their correspondence field $m_{i\to j}$
- Transport (translation, naive) :

$$\begin{aligned} \forall\, h : S_j \to \mathcal{X}, \quad T^L_{j\to i}(h) : \quad S_i \quad &\to \quad \mathcal{X} \\ \left(T^L_{j\to i}(h)\right)(s) \quad &= \quad h\left(m_{i\to j}(s)\right) \end{aligned}$$



# Transport

## Local transport

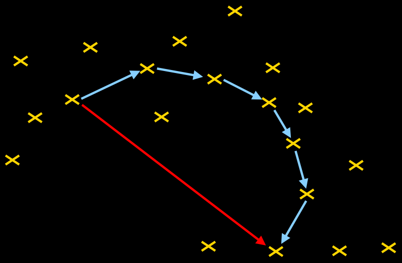
- Set of shapes $(S_i)_{i\in I}$ (e.g. from a video segmentation)
- Two shapes $S_i$ and $S_j$ $\implies$ their correspondence field $m_{i\to j}$
- Transport (translation, naive) :

$$\forall\, h : S_j \to \mathcal{X}, \quad T^L_{j\to i}(h) : \begin{array}{rcl} S_i & \to & \mathcal{X} \\ \left(T^L_{j\to i}(h)\right)(s) & = & h\left(m_{i\to j}(s)\right) \end{array}$$

- Associated cost : $E(m_{i\to j}) \implies$ reliability $w^L_{i\to j} \propto \exp\left(-\alpha E(m_{i\to j})\right)$

## Global transport

- Associated cost : $E(m_{i \to j}) \implies$ reliability $w^L_{i \to j} \propto \exp\left(-\alpha E(m_{i \to j})\right)$
- close shapes : reliable; distant shapes : not reliable
- $\implies$ search for paths of small steps in the training set $(S_i)$
- graph : nodes = shapes, edges = transport, weights = transport cost
- shortest path between pairs of shapes : global transport

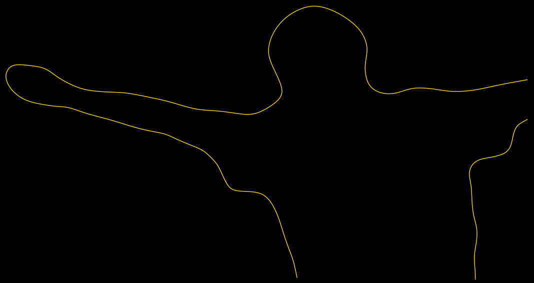


# Global transport

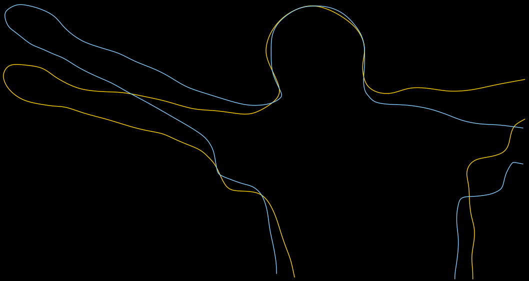
- Associated cost : $E(m_{i\to j}) \implies$ reliability $w^L_{i\to j} \propto \exp\left(-\alpha E(m_{i\to j})\right)$
- close shapes : reliable; distant shapes : not reliable
- $\implies$ search for paths of small steps in the training set $(S_i)$
- graph : nodes = shapes, edges = transport, weights = transport cost
- shortest path between pairs of shapes : global transport
- compose : $T^G_{i\to j} = T^L_{i_n\to j}\ o\ T^L_{i_{n-1}\to i_n}\ o\ \dots\ o\ T^L_{i_1\to i_2}\ o\ T^L_{i\to i_1}$
- reliability : $w^G_{i\to j} = \prod_k w^L_{i_k\to i_{k+1}}$
- use transport to propagate information

Example : colored walker

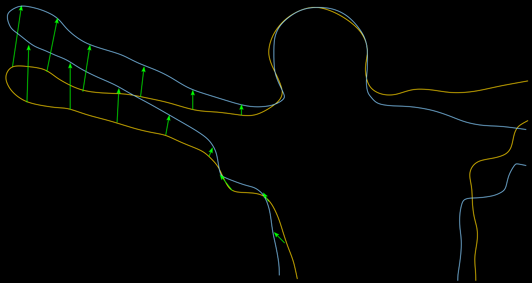
## Global transport



## Global transport

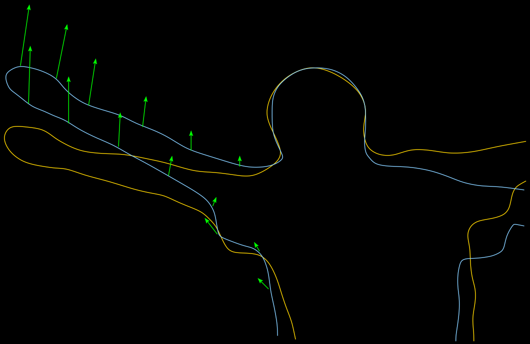


## Global transport



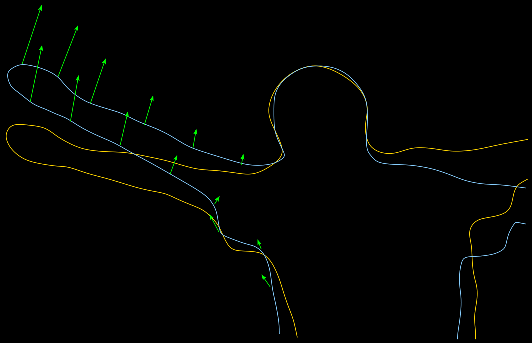
Correspondence field

## Global transport



Transported arm rotation (translation)

## Global transport

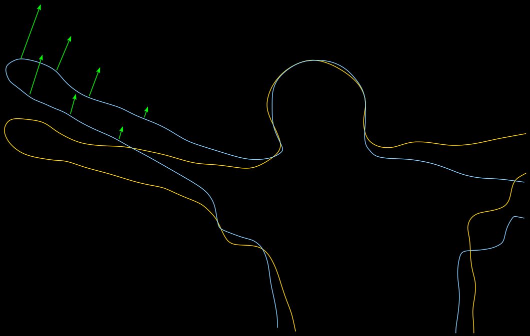


Transported
arm rotation
(better)

## Global transport

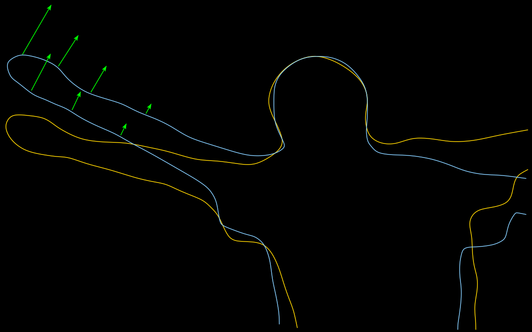
Forearm rotation

## Global transport



Transported
forearm
rotation

## Global transport

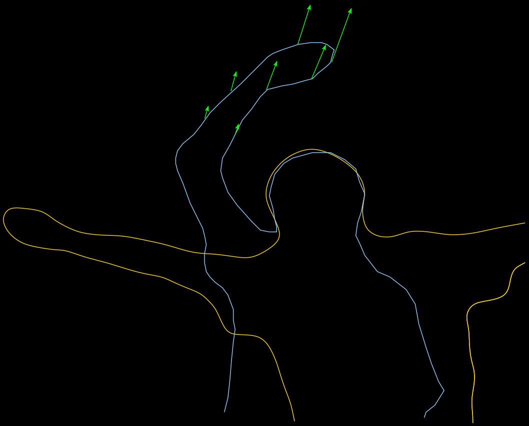


Transported forearm rotation (better)

## Global transport

Transport to another shape

## Global transport

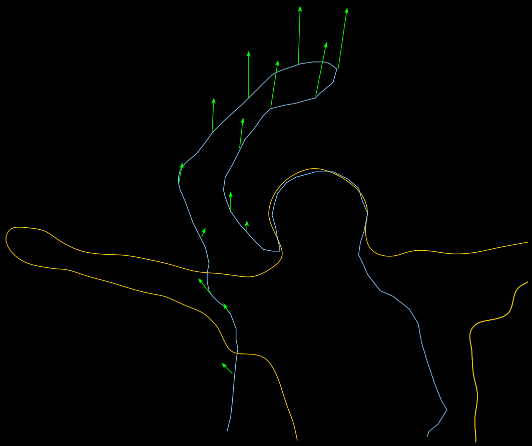


Transported forearm rotation (translation)

## Global transport

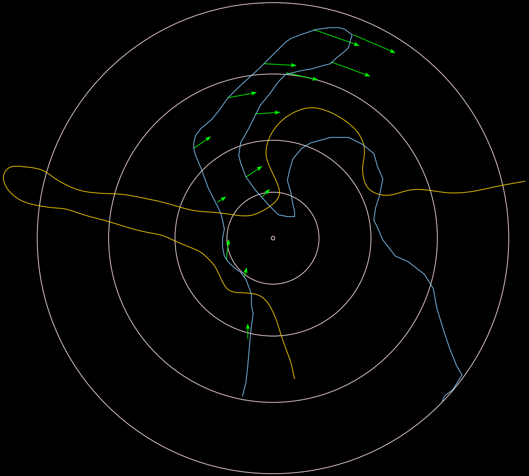
Transported forearm rotation (better)

## Global transport



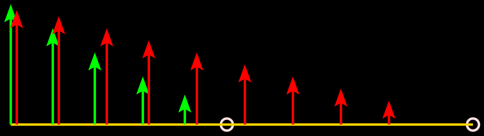
Transported arm rotation (translation)

## Global transport



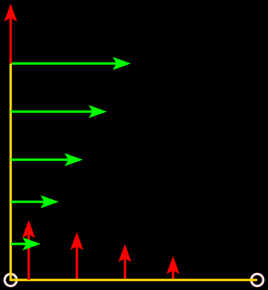
Transported
arm rotation
(better)

## Global transport



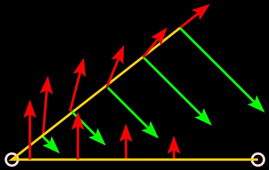
$L^2$
inner-product
$> 0$

## Global transport



$L^2$
inner-product
$= 0$

## Global transport



$L^2$
inner-product
$< 0$

## Remarks about transport

Why ?

- transport : of deformations : needed to increase training set density

Which ?

- "translation" : ok for short pathes
- transport : not obvious (muscles + T-shirt artifacts)
- criterion to assess transport quality / suitability ?
- transport : should be learned (from video sequences ?)
- path could depend on deformation transported

What properties ?

- probability of a deformation transported : can differ
- inner product : no reason to be transport-invariant

## Transport in differential geometry

A connection $\nabla$ is **Riemannian** if the parallel transport it defines preserves the metric $g$. Metric connection :

$$\nabla_X \, g(\cdot,\cdot) = 0 \quad \text{for all vector fields } X \text{ on } \mathcal{M}$$

- not satisfied (probability of a deformation depends on the shape)
  $\Longrightarrow$ **not Riemannian** : transport and metric are independent

## Transport in differential geometry

A connection $\nabla$ is **Riemannian** if the parallel transport it defines preserves the metric $g$. Metric connection :

$$\nabla_X \, g(\cdot,\cdot) = 0 \quad \text{for all vector fields } X \text{ on } \mathcal{M}$$

- not satisfied (probability of a deformation depends on the shape)
  $\Longrightarrow$ **not Riemannian** : transport and metric are independent

## Transport $\Longrightarrow$ connection

Given transport, under few hypotheses (e.g. smoothness), it is possible to recover the associated infinitesimal connection :

$$\nabla_X V = \lim_{h \to 0} \frac{T_\gamma^{h\to 0} V_{\gamma(h)} - V_{\gamma(0)}}{h} = \frac{d}{dt} T_\gamma^{t \to 0} V_{\gamma(t)} \Big|_{t=0} .$$

## Transport in differential geometry

A connection $\nabla$ is **Riemannian** if the parallel transport it defines preserves the metric $g$. Metric connection :

$$\nabla_X\, g(\cdot,\cdot) = 0 \quad \text{for all vector fields } X \text{ on } \mathcal{M}$$

- not satisfied (probability of a deformation depends on the shape)
  $\Longrightarrow$ **not Riemannian** : transport and metric are independent

## Transport $\Longrightarrow$ connection

Given transport, under few hypotheses (e.g. smoothness), it is possible to recover the associated infinitesimal connection :

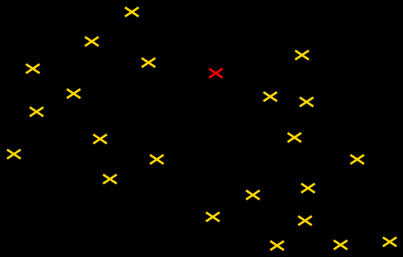
$$\nabla_X V = \lim_{h\to 0} \frac{T_\gamma^{h\to 0} V_{\gamma(h)} - V_{\gamma(0)}}{h} = \frac{d}{dt}\, T_\gamma^{t\to 0} V_{\gamma(t)}\Big|_{t=0}.$$

## Connection $\Longrightarrow$ transport :

Given a covariant derivative $\nabla$, the transport along a curve $\gamma$ is obtained by integrating the condition $\nabla_{\dot\gamma} = 0$.

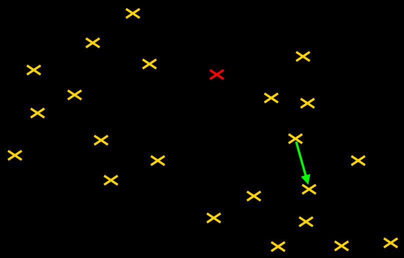
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i \to j}$, transport $T^G_{i \to k}$
- $\implies$ transport deformations to a particular shape $S_k$ :
  $f^{i \to k}_{i \to j} = T^G_{i \to k}(m_{i \to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i \to k}\, w^L_{i \to j}$



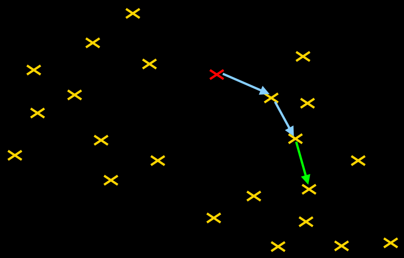
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i\to j}$, transport $T^G_{i\to k}$
- $\implies$ transport deformations to a particular shape $S_k$ :
  $f^{i\to k}_{i\to j} = T^G_{i\to k}(m_{i\to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i\to k}\, w^L_{i\to j}$



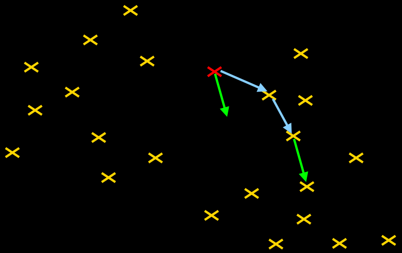
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i \to j}$, transport $T^G_{i \to k}$
- $\implies$ transport deformations to a particular shape $S_k$ :
  $f^{i \to k}_{i \to j} = T^G_{i \to k}(m_{i \to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i \to k}\, w^L_{i \to j}$



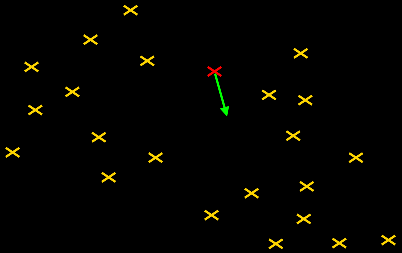
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i\to j}$, transport $T^G_{i\to k}$
- $\implies$ transport deformations to a particular shape $S_k$ :
  $f^{i\to k}_{i\to j} = T^G_{i\to k}(m_{i\to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i\to k}\, w^L_{i\to j}$



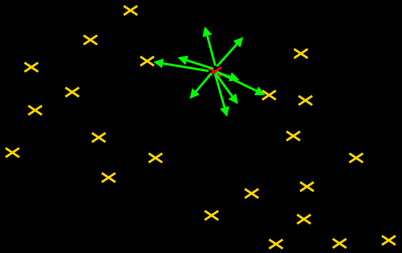
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i\to j}$, transport $T^G_{i\to k}$
- $\implies$ transport deformations to a particular shape $S_k$ :
  $f^{i\to k}_{i\to j} = T^G_{i\to k}(m_{i\to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i\to k}\, w^L_{i\to j}$



# Metric estimation (statistics on deformations)

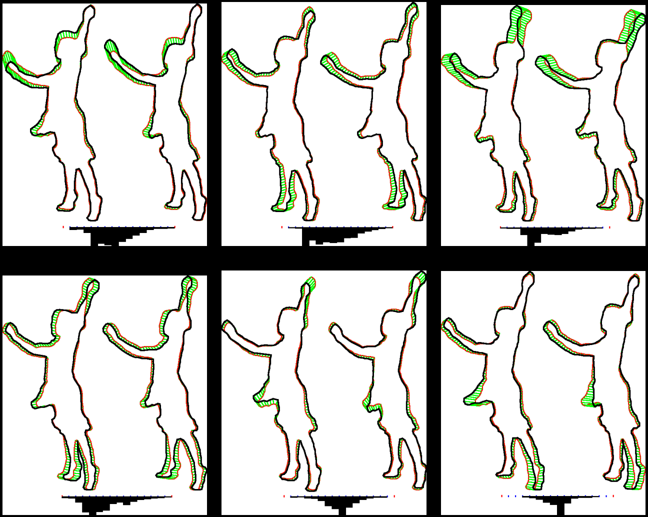
- set of shapes $(S_i)$, local deformations $m_{i\to j}$, transport $T^G_{i\to k}$
- $\Longrightarrow$ transport deformations to a particular shape $S_k$ :
  $f^{i\to k}_{i\to j} = T^G_{i\to k}(m_{i\to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$
  with reliability weights $w^k_{ij} = w^G_{i\to k}\, w^L_{i\to j}$



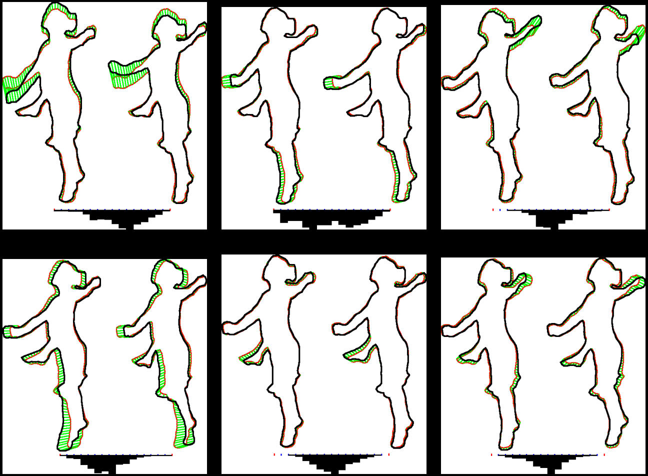
# Metric estimation (statistics on deformations)

- set of shapes $(S_i)$, local deformations $m_{i\to j}$, transport $T^G_{i\to k}$
- $\Longrightarrow$ transport deformations to a particular shape $S_k$ : $f^{i\to k}_{i\to j} = T^G_{i\to k}(m_{i\to j})$ are, $\forall i, j$, deformations defined on the same shape $S_k$ with reliability weights $w^k_{ij} = w^G_{i\to k}\, w^L_{i\to j}$
- statistics, for $k$ fixed : PCA
- PCA with weights, and with $H^1$-norm
- $\Longrightarrow$ eigenmodes $e_n$ (= principal deformations) with eigenvalues $\lambda_n$
- $\Longrightarrow$ defines an inner product $P_k$ = metric in the tangent space of the shape $S_k$
- $P_k$ varies smoothly as a function of $k$

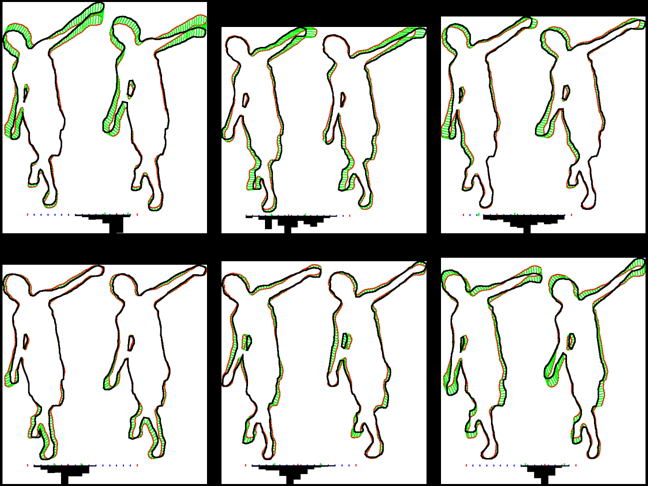
## Example of results : dancing sequence (9s, 24Hz), shape 1



## Example of results : shape 2



## Example of results : shape 3



# Theoretical justifications

## The best metric ?

Searching for principal modes of deformations which vary smoothly (as a function of the shape $S_k$) ?

- vain quest : hairy ball theorem $\Longrightarrow$ no best smooth direction field (or then it has to vanish sometimes)

# Theoretical justifications

## The best metric ?

Searching for principal modes of deformations which vary smoothly (as a function of the shape $S_k$) ?

- vain quest : hairy ball theorem $\Longrightarrow$ no best smooth direction field (or then it has to vanish sometimes)

Best metric for a given distribution (on one shape) ?

- $\Longrightarrow$ PCA gives the best metric for a criterion based on Kullback-Leibler divergence between distributions

# Theoretical justifications

## The best metric ?

Searching for principal modes of deformations which vary smoothly (as a function of the shape $S_k$) ?

- vain quest : hairy ball theorem $\Longrightarrow$ no best smooth direction field (or then it has to vanish sometimes)

Best metric for a given distribution (on one shape) ?

- $\Longrightarrow$ PCA gives the best metric for a criterion based on Kullback-Leibler divergence between distributions

Best metric for a given empirical manifold (all shapes together) ?

- needs a smoothness criterion ( $\Longrightarrow$ transport)

# Theoretical justifications

## The best metric ?

Searching for principal modes of deformations which vary smoothly (as a function of the shape $S_k$) ?

- vain quest : hairy ball theorem $\Longrightarrow$ no best smooth direction field (or then it has to vanish sometimes)

Best metric for a given distribution (on one shape) ?

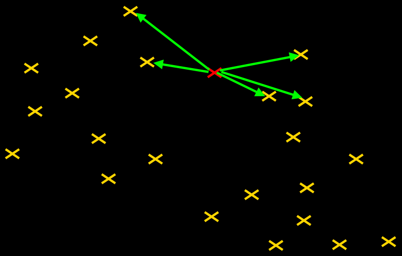
- $\Longrightarrow$ PCA gives the best metric for a criterion based on Kullback-Leibler divergence between distributions

Best metric for a given empirical manifold (all shapes together) ?

- needs a smoothness criterion ( $\Longrightarrow$ transport)
- $\Longrightarrow$ best metric for a criterion involving transport & K-L divergence.
- $\Longrightarrow$ best metric for another criterion involving transport & $L^2$-norm of distributions.

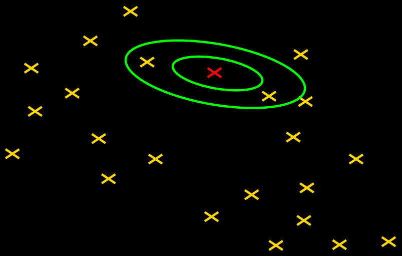
# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i \to j}$, transport $T^G_{i \to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j w^L_{i \to j} \, \delta_{\mathbf{f}_{i \to j}}$



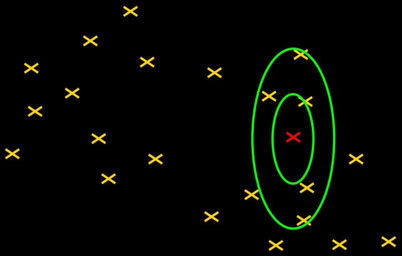
# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i\to j}$, transport $T^G_{i\to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j w^L_{i\to j}\, \delta_{\mathbf{f}_{i\to j}}$



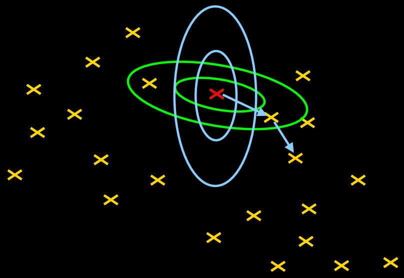
# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i\to j}$, transport $T^G_{i\to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j w^L_{i\to j}\, \delta_{\mathbf{f}_{i\to j}}$



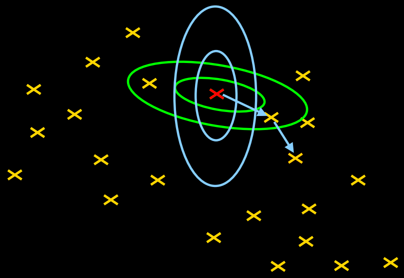
# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i\to j}$, transport $T^G_{i\to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j w^L_{i\to j}\, \delta_{\mathbf{f}_{i\to j}}$



# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i\to j}$, transport $T^G_{i\to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j \; w^L_{i\to j} \; \delta_{\mathbf{f}_{i\to j}}$
- Transported distribution : via $T_{i\to k}(\delta_{\mathbf{f}}) = \delta_{T_{i\to k}(\mathbf{f})}$.
- Criterion : best $(P_k)$ for $\displaystyle\sum_{i,k} \; w^G_{ik} \;\; KL\left(\mathcal{D}_{P_k} \mid T_{i\to k}(\mathcal{D}_{emp_i})\right)$



# Best metric for a given empirical manifold

- set of shapes $(S_i)$, local deformations $\mathbf{f}_{i\to j}$, transport $T^G_{i\to k}$
- Empirical distributions : $\mathcal{D}_{emp_i} = \sum_j w^L_{i\to j}\,\delta_{\mathbf{f}_{i\to j}}$
- Transported distribution : via $T_{i\to k}(\delta_{\mathbf{f}}) = \delta_{T_{i\to k}(\mathbf{f})}$.
- Criterion : best $(P_k)$ for $\displaystyle\sum_{i,k} w^G_{ik}\;KL\left(\mathcal{D}_{P_k} \mid T_{i\to k}(\mathcal{D}_{emp_i})\right)$
- $=$ best $(P_k)$ for $\displaystyle\sum_k KL(\mathcal{D}_{P_k}|\mathcal{D}^T_{emp_k})$

  where $\mathcal{D}^T_{emp_k} = \displaystyle\sum_{i,j} w^k_{i\to j}\;\delta_{\mathbf{f}^k_{i\to j}}$
- Transported deformations to any shape $S_k$ : $f^k_{i\to j} = T^G_{i\to k}(f_{i\to j})$ with reliability weights $w^k_{i\to j} = w^G_{i\to k}\,w^L_{i\to j}$
- $=$ the one obtained by weighted PCA on transported deformations

## Best metric for a given empirical manifold (again!)

- empirical distributions : $\mathcal{D}_{emp_i}$
- kernel-smoothed empirical distributions : $\mathcal{D}^{\mathcal{K}}_{emp_i} = g_i^0 \, d\mu$
- $g_i^0$ : density functions in the tangent space of the shape $S_i$

# Best metric for a given empirical manifold (again!)

- empirical distributions : $\mathcal{D}_{emp_i}$
- kernel-smoothed empirical distributions : $\mathcal{D}^{\mathcal{K}}_{emp_i} = g_i^0 \, d\mu$
- $g_i^0$ : density functions in the tangent space of the shape $S_i$
- search for $g_i$ : close to $g_i$ and smooth from shape to shape
- $E(g) = \sum_i \|g_i - g_i^0\|^2_{L^2(T_i)} + \sum_{ij} w_{ij} \, \| T_{i \to j}(g_i) - g_j \|^2_{L^2(T_j)}$

# Best metric for a given empirical manifold (again!)

- empirical distributions : $\mathcal{D}_{emp_i}$
- kernel-smoothed empirical distributions : $\mathcal{D}^{\mathcal{K}}_{emp_i} = g_i^0 \; d\mu$
- $g_i^0$ : density functions in the tangent space of the shape $S_i$
- search for $g_i$ : close to $g_i$ and smooth from shape to shape
- $E(g) = \sum_i \|g_i - g_i^0\|^2_{L^2(T_i)} + \sum_{ij} w_{ij} \; \| T_{i \to j}(g_i) - g_j \|^2_{L^2(T_j)}$
- minimization $\implies$ $Ag = g^0$ with :

$$\begin{cases} A_{ii} = 1 + \sum_j w_{ij} \; T^*_{i \to j} \; T_{i \to j} + w_{ji} \\ A_{ij} = -w_{ij} \; T^*_{i \to j} - w_{ji} \; T_{j \to i} \qquad \text{for} \quad i \neq j \end{cases}$$

- $A = Id + \varepsilon \Delta$ where $\Delta$ = graph Laplacian (with transports)

## Best metric for a given empirical manifold (again!)

- empirical distributions : $\mathcal{D}_{emp_i}$
- kernel-smoothed empirical distributions : $\mathcal{D}^{\mathcal{K}}_{emp_i} = g_i^0 \, d\mu$
- $g_i^0$ : density functions in the tangent space of the shape $S_i$
- search for $g_i$ : close to $g_i$ and smooth from shape to shape
- $E(g) = \sum_i \|g_i - g_i^0\|^2_{L^2(T_i)} + \sum_{ij} w_{ij} \, \| T_{i\to j}(g_i) - g_j\|^2_{L^2(T_j)}$
- minimization $\implies$ $Ag = g^0$ with :
  $\begin{cases} A_{ii} = 1 + \sum_j w_{ij} \; T^*_{i\to j} \; T_{i\to j} + w_{ji} \\ A_{ij} = -w_{ij} \; T^*_{i\to j} - w_{ji} \; T_{j\to i} \qquad \text{for} \;\; i \neq j \end{cases}$
- $A = Id + \varepsilon\Delta$ where $\Delta$ = graph Laplacian (with transports)
- $g = A^{-1} g^0 \; = \; (Id + \varepsilon\Delta)^{-1} g^0 \; \simeq (Id - \varepsilon\Delta) g^0 \; \simeq \; \mathcal{N}_\varepsilon * g^0.$
- $g = (Id - \varepsilon\Delta) \, g^0$ coincides with the $\mathcal{D}^T_{emp}$
  and the inner products $(P_i)$ which suit $g = (g_i)$ the best (for K-L) are the ones we computed

# Conclusion

- transport is useful to reduce required training set size
- transport is useful to propagate information between shapes
- globally optimal metrics (and low complexity)

# Conclusion

- transport is useful to reduce required training set size
- transport is useful to propagate information between shapes
- globally optimal metrics (and low complexity)

[NORDIA 2009 : *Learning Shape Metrics based on Deformations and Transport*]

# Future works

- learning functions defined on shape spaces / with values in shape spaces
- statistics on image patches through correspondences/transport

## Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations $(f_i)$ on one particular shape

## Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations ($f_i$) on one particular shape

- Empirical distribution of deformations : $\mathcal{D}_{emp} = \sum_i w_i \, \delta_{\mathbf{f}_i}$

# Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations $(f_i)$ on one particular shape

- Empirical distribution of deformations : $\mathcal{D}_{emp} = \sum_i w_i \, \delta_{\mathbf{f}_i}$
- Any inner product (= metric) $P$ is associated to a probability distribution: $\mathcal{D}_P(\mathbf{f}) \propto \exp(-\|\mathbf{f}\|_P^2)$

# Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations $(f_i)$ on one particular shape

- Empirical distribution of deformations : $\mathcal{D}_{emp} = \sum_i w_i \, \delta_{\mathbf{f}_i}$
- Any inner product (= metric) $P$ is associated to a probability distribution: $\mathcal{D}_P(\mathbf{f}) \; \propto \; \exp(-\|\mathbf{f}\|_P^2)$
- Given an inner product $P_0$ $(= H^1)$ of reference, with its orthonormal basis $(e_n)$, supposing that $P$ is continuous wrt. $P_0$:
  $\forall \, \mathbf{f} \in T, \quad \|\mathbf{f}\|_P^2 = \sum_n \alpha_n \left\langle \mathbf{f} \,|\, \mathbf{e}_n \right\rangle_{P_0}^2$

## Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations $(f_i)$ on one particular shape

- Empirical distribution of deformations : $\mathcal{D}_{emp} = \sum_i w_i \, \delta_{\mathbf{f}_i}$
- Any inner product (= metric) $P$ is associated to a probability distribution: $\mathcal{D}_P(\mathbf{f}) \propto \exp(-\|\mathbf{f}\|_P^2)$
- Given an inner product $P_0$ $(= H^1)$ of reference, with its orthonormal basis $(e_n)$, supposing that $P$ is continuous wrt. $P_0$: $\forall\, \mathbf{f} \in T, \quad \|\mathbf{f}\|_P^2 = \sum_n \alpha_n \left\langle \mathbf{f} \,|\, \mathbf{e}_n \right\rangle_{P_0}^2$
- $\implies \mathcal{D}_P$ is Gaussian : $\mathcal{D}_P(\mathbf{f}) := \prod_n \left(\frac{\alpha_n}{\pi}\right)^{\frac{1}{2}} \exp(-\alpha_n \left\langle \mathbf{f} \,|\mathbf{e}_n \right\rangle_{P_0}^2)$

## Link between PCA and Kullback-Leibler divergence

Aim : to find a metric suitable for a given distribution of deformations $(f_i)$ on one particular shape

- Empirical distribution of deformations : $\mathcal{D}_{emp} = \sum_i w_i \, \delta_{\mathbf{f}_i}$
- Any inner product (= metric) $P$ is associated to a probability distribution: $\mathcal{D}_P(\mathbf{f}) \propto \exp(-\|\mathbf{f}\|_P^2)$
- Given an inner product $P_0$ $(= H^1)$ of reference, with its orthonormal basis $(e_n)$, supposing that $P$ is continuous wrt. $P_0$:
  $\forall\, \mathbf{f} \in T, \quad \|\mathbf{f}\|_P^2 = \sum_n \alpha_n \left\langle \mathbf{f} \,|\, \mathbf{e}_n \right\rangle_{P_0}^2$
- $\implies \mathcal{D}_P$ is Gaussian : $\mathcal{D}_P(\mathbf{f}) := \prod_n \left(\frac{\alpha_n}{\pi}\right)^{\frac{1}{2}} \exp(-\alpha_n \left\langle \mathbf{f} \,|\mathbf{e}_n \right\rangle_{P_0}^2)$
- $\implies$ search over inner products = search over Gaussian distributions

## Link between PCA and Kullback-Leibler divergence (bis)

- Gaussian distribution that fits $\mathcal{D}_{emp}$ the best ?

## Link between PCA and Kullback-Leibler divergence (bis)

- Gaussian distribution that fits $\mathcal{D}_{emp}$ the best ?
- search for best Gaussian (= for best $P$) that minimize $KL(\mathcal{D}_P|\mathcal{D}_{emp})$

## Link between PCA and Kullback-Leibler divergence (bis)

- Gaussian distribution that fits $\mathcal{D}_{emp}$ the best ?
- search for best Gaussian (= for best $P$) that minimize $KL(\mathcal{D}_P | \mathcal{D}_{emp})$
- best inner product $P$ is the one given by weighted PCA with norm $P_0$ !

## Link between PCA and Kullback-Leibler divergence (bis)

- Gaussian distribution that fits $\mathcal{D}_{emp}$ the best ?
- search for best Gaussian (= for best $P$) that minimize $KL(\mathcal{D}_P|\mathcal{D}_{emp})$
- best inner product $P$ is the one given by weighted PCA with norm $P_0$ !
- similar result for kernel-smoothed distributions :
  $\mathcal{D}^{\mathcal{K}}_{emp}(\mathbf{f}) = \sum_j w_j \, \mathcal{K}(\mathbf{f}_j - \mathbf{f})$.

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i\to j} \left\| \mathbf{f}^k_{i\to j} - \sum_n \left\langle \mathbf{f}^k_{i\to j} \middle| \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i \to j} \left\| \mathbf{f}^k_{i \to j} - \sum_n \left\langle \mathbf{f}^k_{i \to j} \middle| \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

- $$\sup_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_n \sum_{i,j} w^k_{i \to j} \left\langle \mathbf{f}^k_{i \to j} \middle| \mathbf{e}_n \right\rangle^2_{H^1_\alpha}$$

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i\to j} \left\| \mathbf{f}^k_{i\to j} - \sum_n \left\langle \mathbf{f}^k_{i\to j} \middle| \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

- $$\sup_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_n \sum_{i,j} w^k_{i\to j} \left\langle \mathbf{f}^k_{i\to j} \middle| \mathbf{e}_n \right\rangle^2_{H^1_\alpha}$$

- $$\sup_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_n \mathbf{e}_n \, HFH \, \mathbf{e}_n$$

  where $F = \sum_{i,j} w^k_{i\to j} \mathbf{f}^k_{i\to j} \otimes \mathbf{f}^k_{i\to j}$ = weighted covariance matrix,
  and $H = Id - \alpha\Delta$ = symmetric definite operator s.t.
  $\langle a | b \rangle_{H^1_\alpha} = \langle H a | b \rangle_{L^2}$

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i \to j} \left\| \mathbf{f}^k_{i \to j} - \sum_n \left\langle \mathbf{f}^k_{i \to j} | \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

- $$\sup_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_n \sum_{i,j} w^k_{i \to j} \left\langle \mathbf{f}^k_{i \to j} | \mathbf{e}_n \right\rangle^2_{H^1_\alpha}$$

- $$\sup_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_n \mathbf{e}_n \, HFH \, \mathbf{e}_n$$

  where $F = \sum_{i,j} w^k_{i \to j} \mathbf{f}^k_{i \to j} \otimes \mathbf{f}^k_{i \to j}$ = weighted covariance matrix,
  and $H = Id - \alpha\Delta$ = symmetric definite operator s.t.
  $\langle a | b \rangle_{H^1_\alpha} = \langle H a | b \rangle_{L^2}$

- Change of variables: $\mathbf{d}_n = H^{1/2}\mathbf{e}_n$ : $\sup_{\langle \mathbf{d}_n | \mathbf{d}_{n'} \rangle_{L^2} = \delta_{n=n'}} \sum_n \mathbf{d}_n \, H^{1/2} F H^{1/2} \, \mathbf{d}_n$

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i \to j} \left\| \mathbf{f}^k_{i \to j} - \sum_n \left\langle \mathbf{f}^k_{i \to j} \middle| \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

- classical PCA problem, with correlation matrix :

$$M_{(i,j),(i',j')} = \left\langle \sqrt{w^k_{i \to j}}\, \mathbf{f}^k_{i \to j} \middle| \sqrt{w^k_{i' \to j'}}\, \mathbf{f}^k_{i' \to j'} \right\rangle_{H^1_\alpha}$$

## Weighted PCA with $H^1$ norm

- PCA = find the best axes (to project data on this subspace)
- Minimize projection error :

$$\inf_{\langle \mathbf{e}_n | \mathbf{e}_{n'} \rangle_{H^1_\alpha} = \delta_{n=n'}} \sum_{i,j} w^k_{i\to j} \left\| \mathbf{f}^k_{i\to j} - \sum_n \left\langle \mathbf{f}^k_{i\to j} \middle| \mathbf{e}_n \right\rangle_{H^1_\alpha} \mathbf{e}_n \right\|^2_{H^1_\alpha}$$

- classical PCA problem, with correlation matrix :

$$M_{(i,j),(i',j')} = \left\langle \sqrt{w^k_{i\to j}}\, \mathbf{f}^k_{i\to j} \middle| \sqrt{w^k_{i'\to j'}}\, \mathbf{f}^k_{i'\to j'} \right\rangle_{H^1_\alpha}$$

- eigenvectors :

$$\mathbf{e}_n = \sum_{ij} \gamma_n^{(i,j)} \sqrt{w^k_{i\to j}}\, \mathbf{f}^k_{i\to j}$$