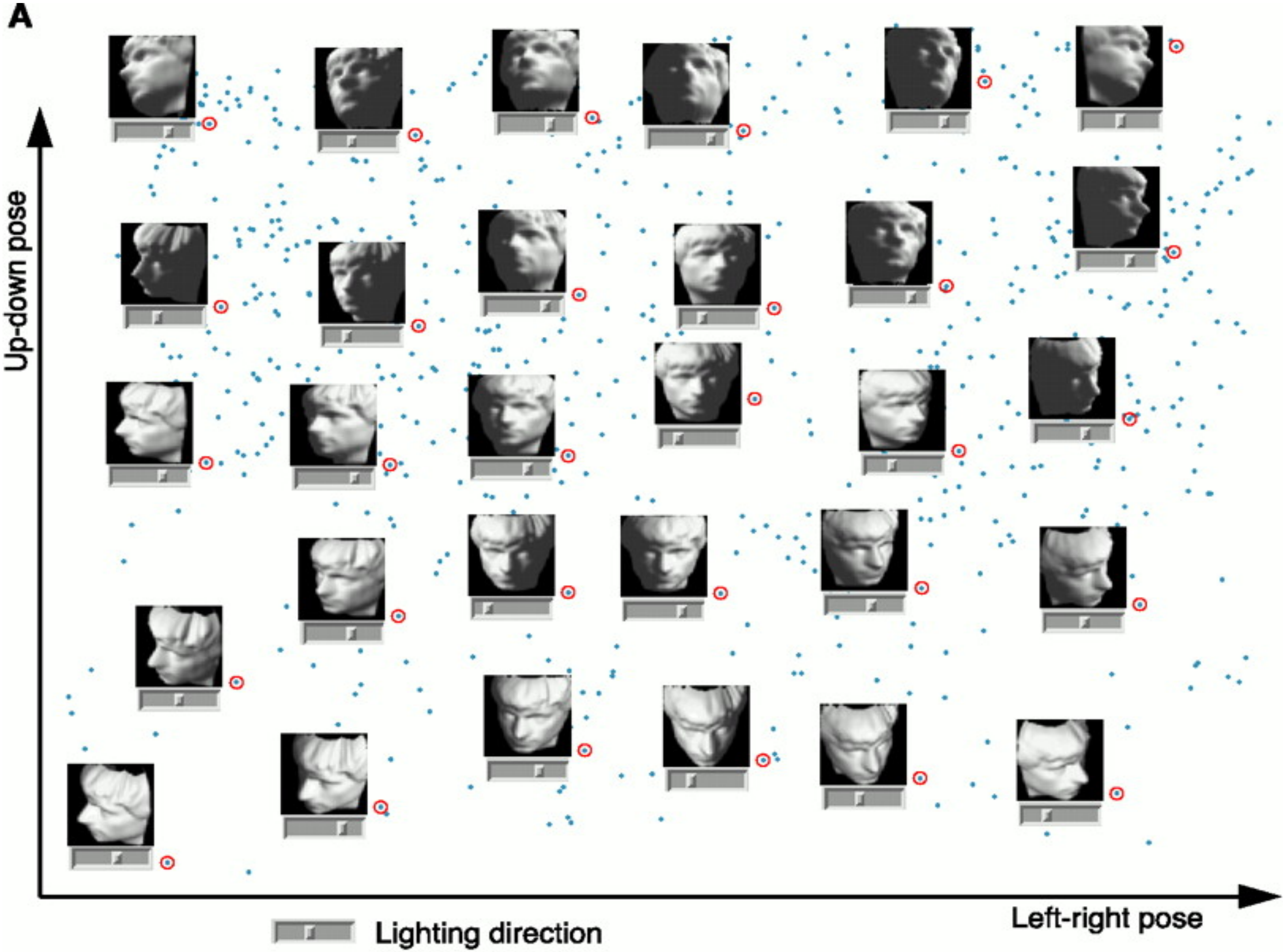
Unfolding Manifolds: ISOMAP vs. LLE
--几何数据处理初步

姚 远

高等统计选讲

2009.10.14


A
Up-down pose
Left-right pose
Lighting direction

Dimensionality Reduction

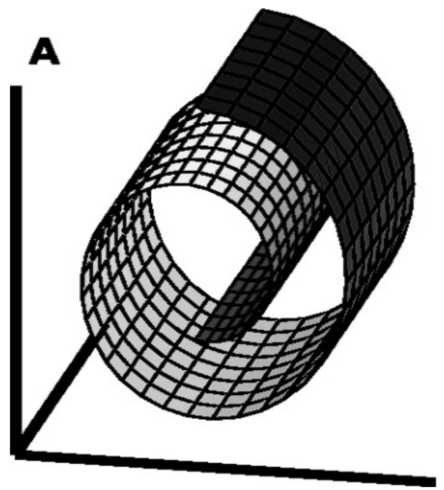
- Need to analyze large amounts multivariate data.
  - Human Faces.
  - Speech Waveforms.
  - Global Climate patterns.
  - Gene Distributions.

- Difficult to visualize data in dimensions just greater than three.

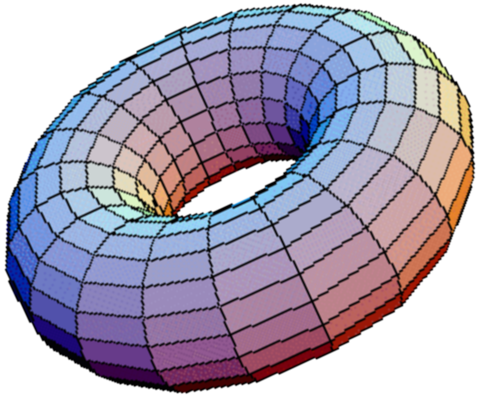
- Discover compact representations of high dimensional data.
  - Visualization.
  - Compression.
  - Better Recognition.
  - Probably meaningful dimensions.

Types of structure in multivariate data..

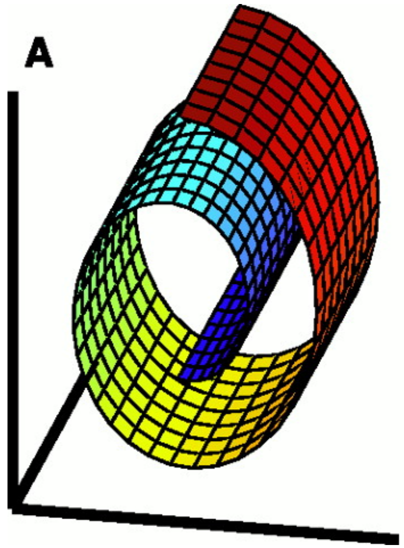
- Clusters.
  - Density Estimation Techniques.
- On or around low Dimensional Manifolds
  - Linear： Principal Component Analysis
  - NonLinear: ISOMAP, LLE, Laplacian Eigenmap, Diffusion Map, etc.



Concept of Manifolds



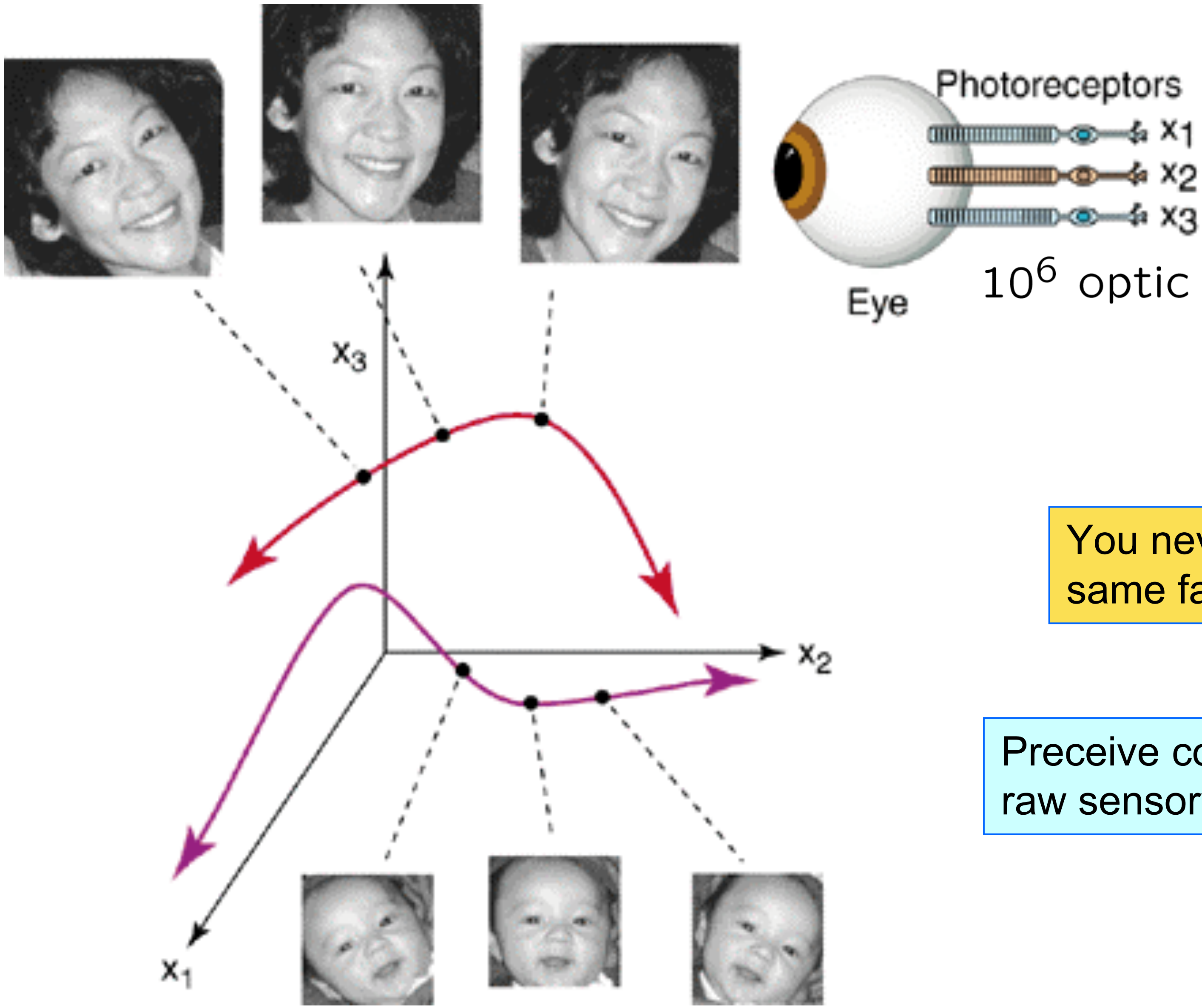
- *“A manifold is a topological space which is locally Euclidean.”*
- In general, any object which is nearly "flat" on small scales is a manifold.
- Euclidean space is a simplest example of a manifold.
- Concept of submanifold.
- Manifolds arise naturally whenever there is a smooth variation of parameters [like pose of the face in previous example]
- The dimension of a manifold is the minimum integer number of co-ordinates necessary to identify each point in that manifold.



Concept of Dimensionality Reduction:

Embed data in a higher dimensional space to a lower dimensional manifold

Manifolds of Perception..Human Visual System

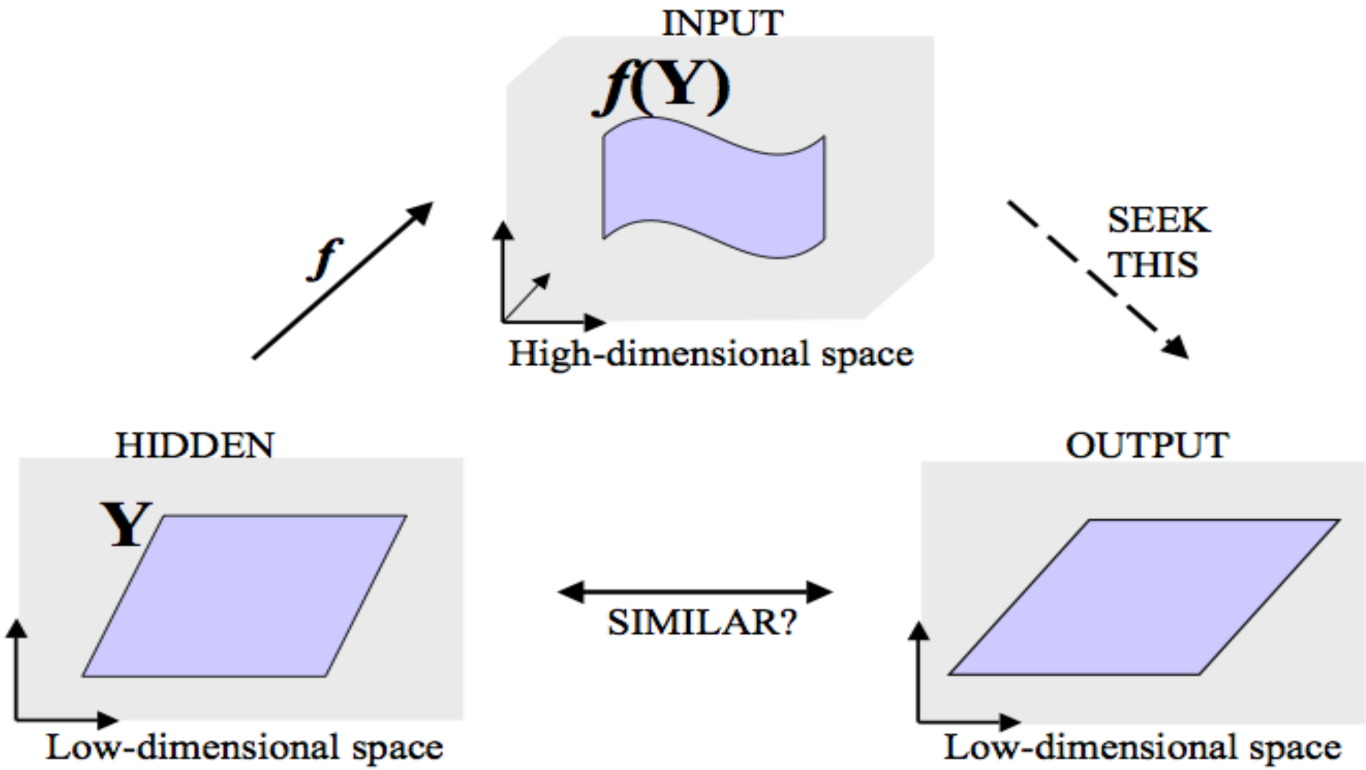


10^6 optic nerve fibres

You never see the same face twice.

Preceive constancy when raw sensory inputs are in flux..

Generative Models in Manifold Learning

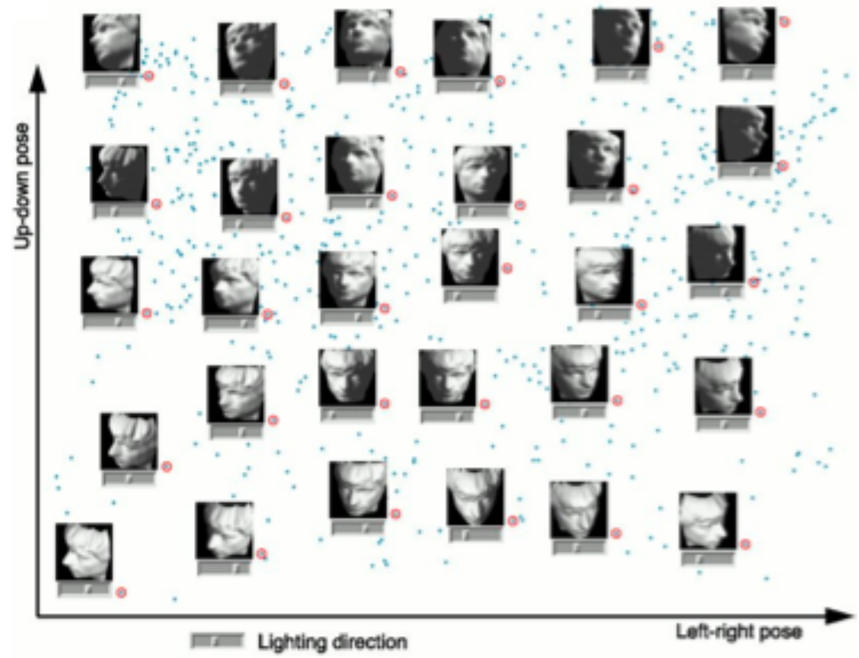


Example: faces

- Given input:
  - randomly ordered sequence of images

  - varied in pose and lighting

- Desired output:
  - Intrinsic dimensionality: 3
  - Low-dimensional representation:

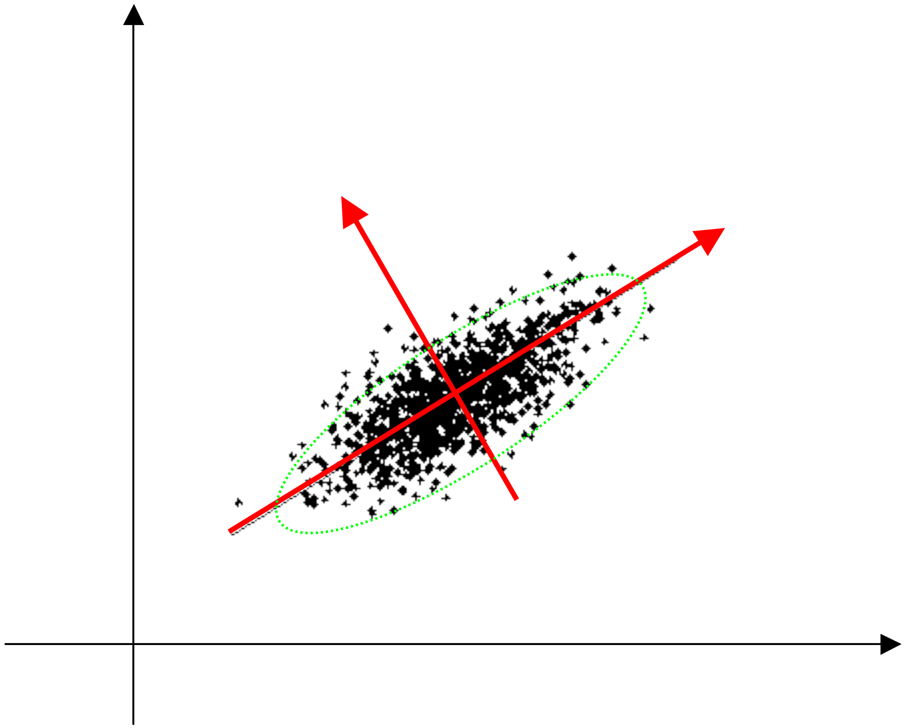


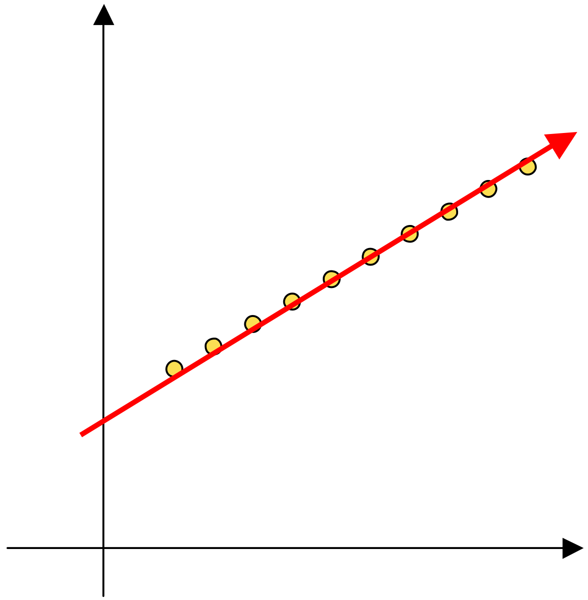
Linear methods..

- Principal Component Analysis (PCA)

$$X_{D\times N} = [X_1|X_2|....|X_N]$$

EigenValue Decomposition of XX^T

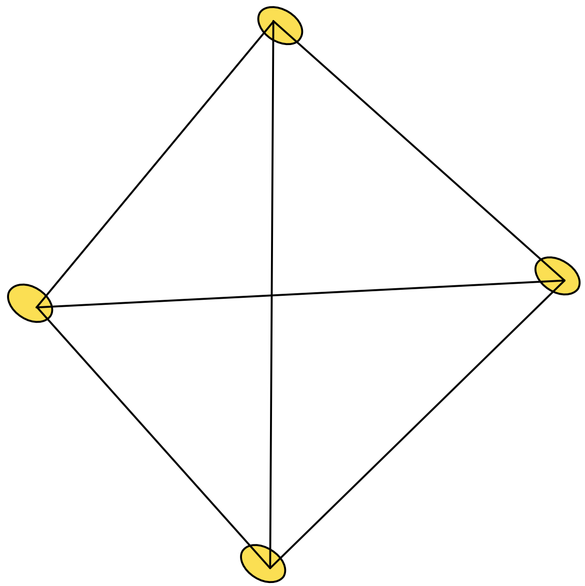




One Dimensional Manifold

MultiDimensional Scaling..

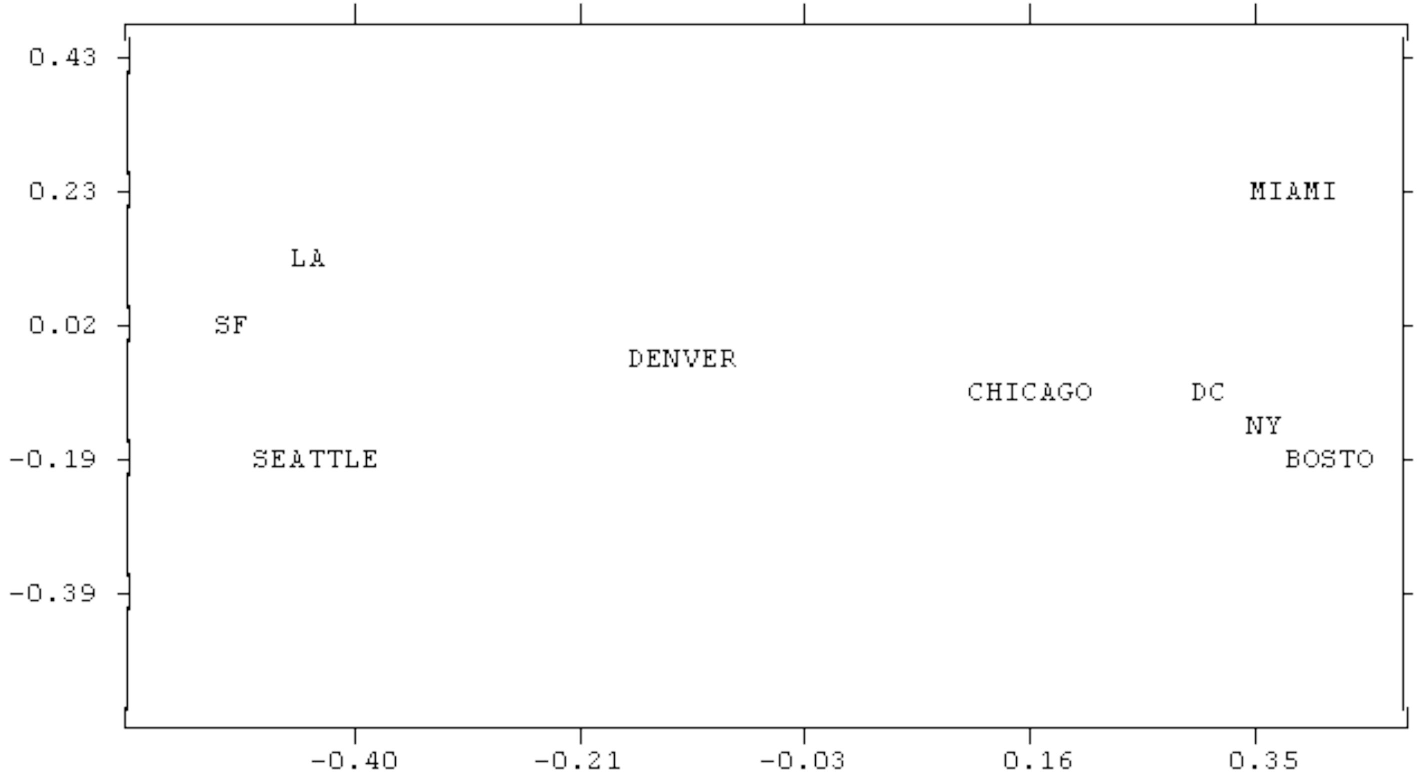
- Here we are given pairwise distances instead of the actual data points.
  - First convert the pairwise distance matrix into the dot product matrix XX^T
  - After that same as PCA.



If we preserve the pairwise distances do we preserve the structure??

Example of MDS…

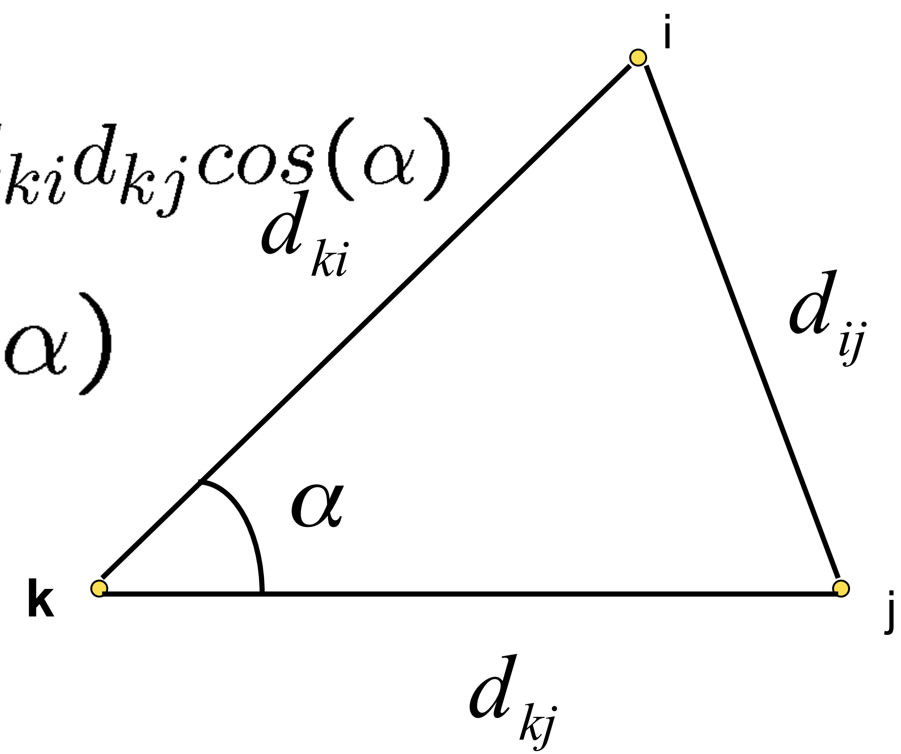
		1 BOST	2 NY	3 DC	4 MIAM	5 CHIC	6 SEAT	7 SF	8 LA	9 DENV
1	BOSTON	0	206	429	1504	963	2976	3095	2979	1949
2	NY	206	0	233	1308	802	2815	2934	2786	1771
3	DC	429	233	0	1075	671	2684	2799	2631	1616
4	MIAMI	1504	1308	1075	0	1329	3273	3053	2687	2037
5	CHICAGO	963	802	671	1329	0	2013	2142	2054	996
6	SEATTLE	2976	2815	2684	3273	2013	0	808	1131	1307
7	SF	3095	2934	2799	3053	2142	808	0	379	1235
8	LA	2979	2786	2631	2687	2054	1131	379	0	1059
9	DENVER	1949	1771	1616	2037	996	1307	1235	1059	0



How to get dot product matrix from pairwise distance matrix?

$$d_{ij}^2 = d_{ki}^2 + d_{kj}^2 - 2d_{ki}d_{kj}cos(\alpha)$$

$$b_{ij} = d_{ki}d_{kj}cos(\alpha)$$



$$b_{ij} = \frac{1}{2}(d_{ki}^2 + d_{kj}^2 - d_{ij}^2)$$

MDS..

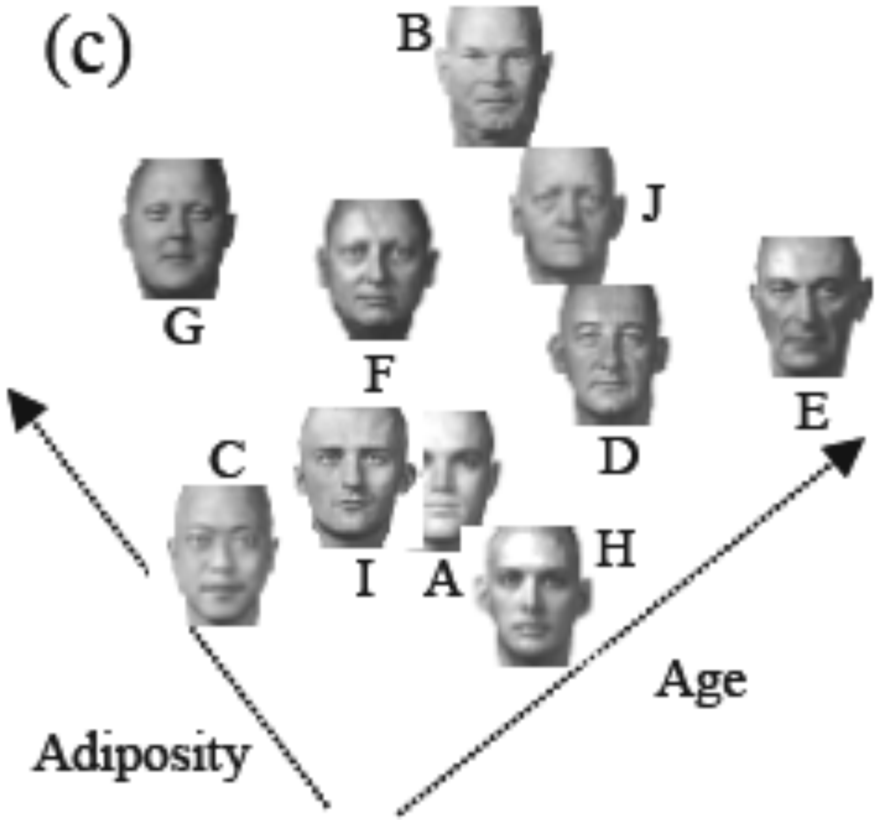
- MDS—origin as one of the points and orientation arbitrary.

Centroid as origin

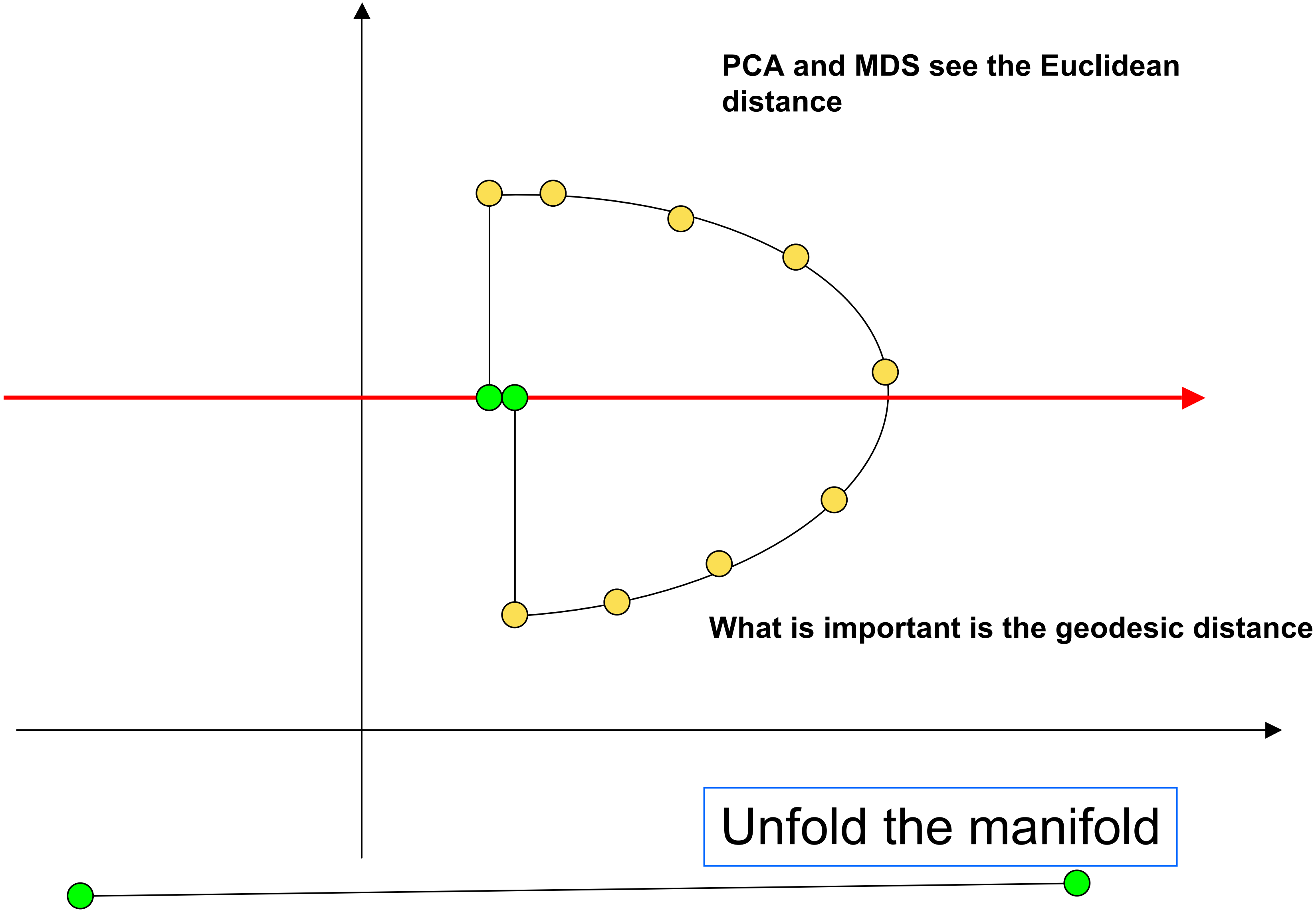
$$b_{ij}^{*} = -\frac{1}{2}\left[d_{ij}^{2} - \frac{1}{N}\sum_{l=1}^{N} d_{il}^{2} - \frac{1}{N}\sum_{m=1}^{N} d_{mj}^{2} + \frac{1}{N^{2}}\sum_{o=1}^{N}\sum_{p=1}^{N} d_{op}^{2}\right]$$

MDS is more general..

- Instead of pairwise distances we can use paiwise "dissimilarities".
- When the distances are Euclidean MDS is equivalent to PCA.
- Eg. Face recognition, wine tasting
- Can get the significant cognitive dimensions.

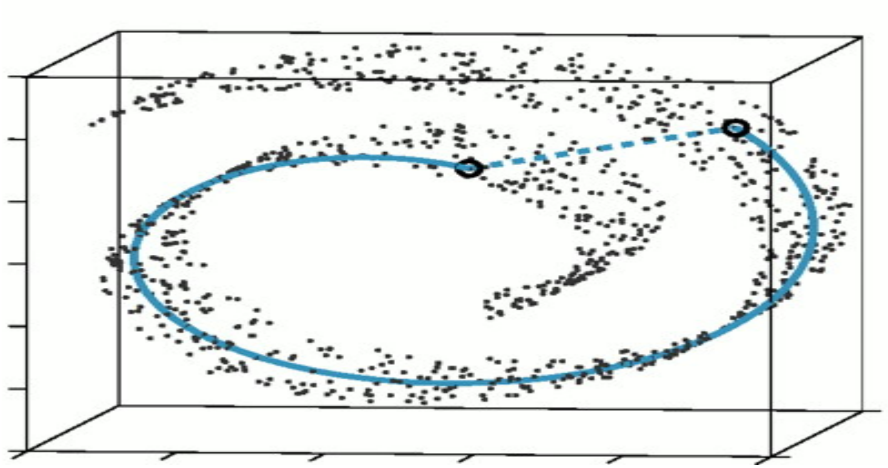


Nonlinear Manifolds..



Intrinsic Description..

- ***To preserve structure, preserve the geodesic distance and not the euclidean distance.***



Two Basic Geometric Embedding Methods

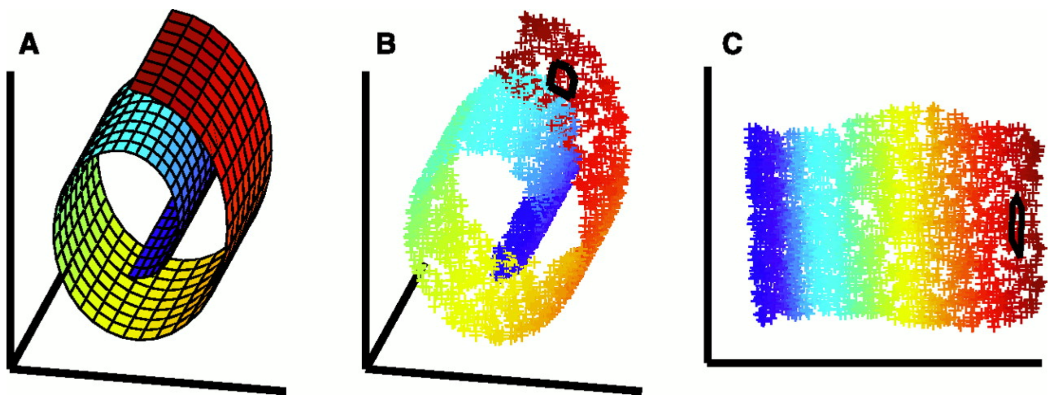
- Tenenbaum et.al's Isomap Algorithm
  - Global approach.
  - On a low dimensional embedding
    - Nearby points should be nearby.
    - Farway points should be faraway.

- Roweis and Saul's Locally Linear Embedding Algorithm
  - Local approach
    - Nearby points nearby

Isomap

- **Estimate the geodesic distance between faraway points.**
- For neighboring points Euclidean distance is a good approximation to the geodesic distance.
- For farway points estimate the distance by a series of short hops between neighboring points.
  - Find shortest paths in a graph with edges connecting neighboring data points

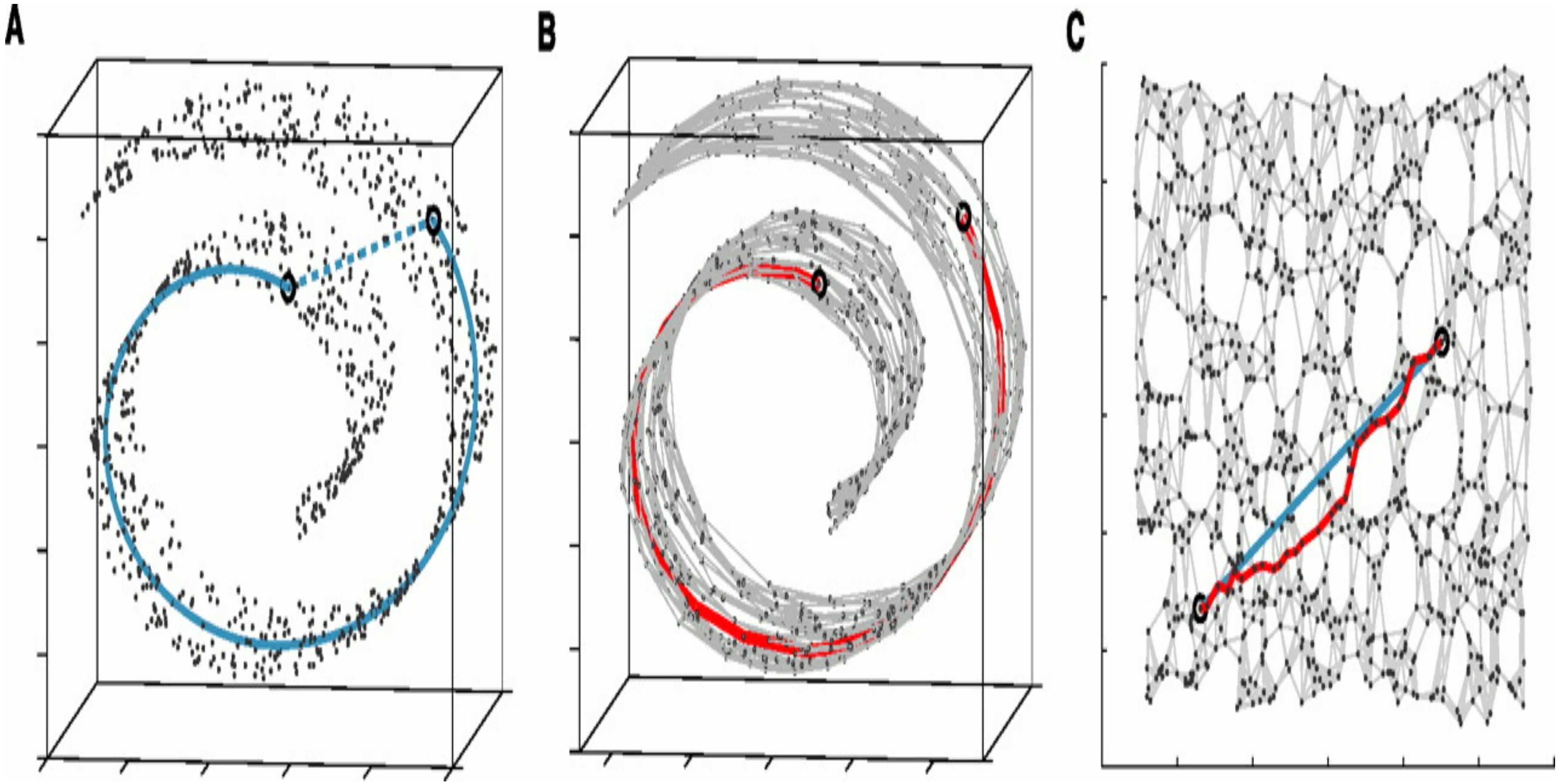
Once we have all pairwise geodesic distances use classical metric MDS

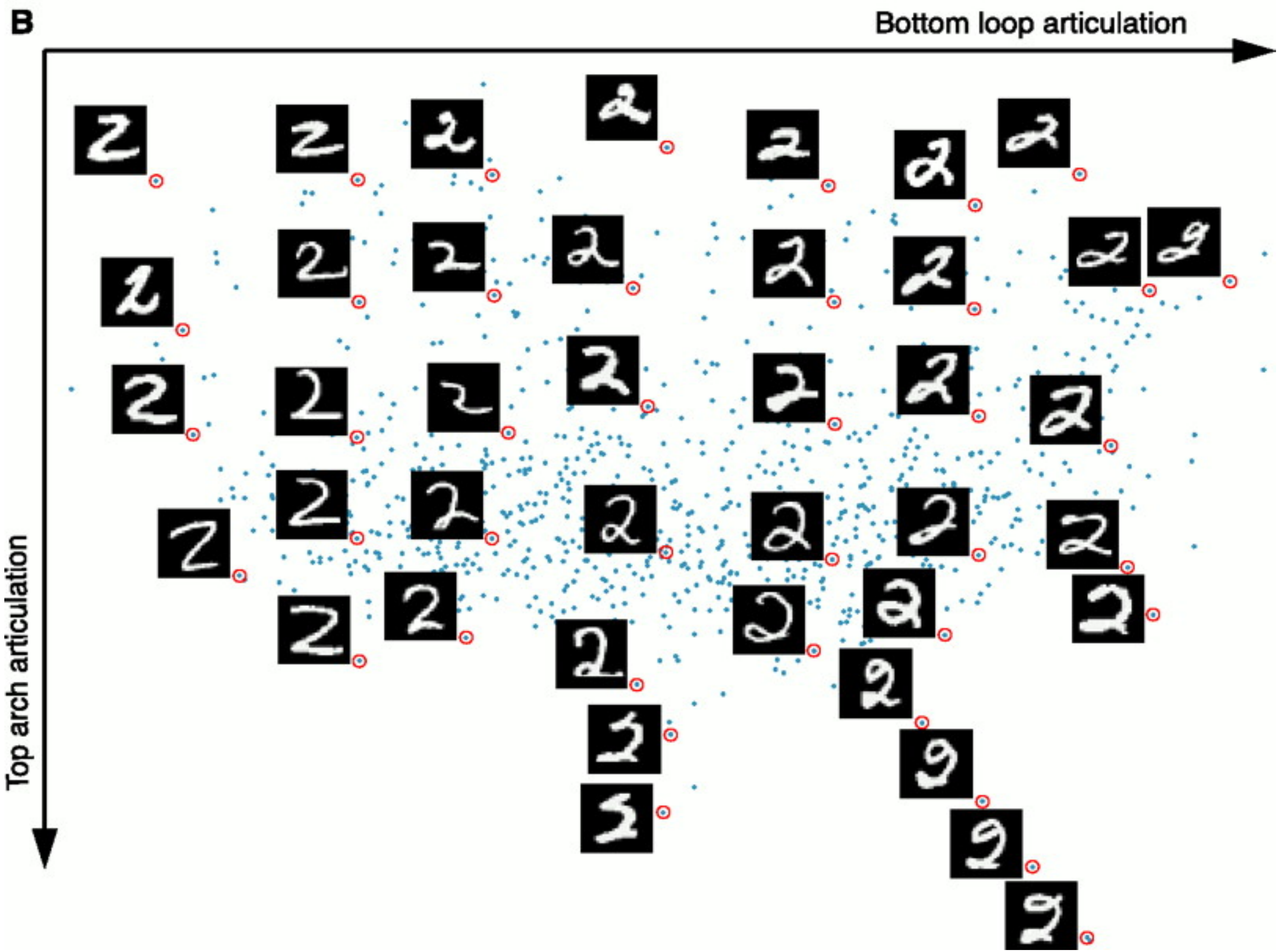


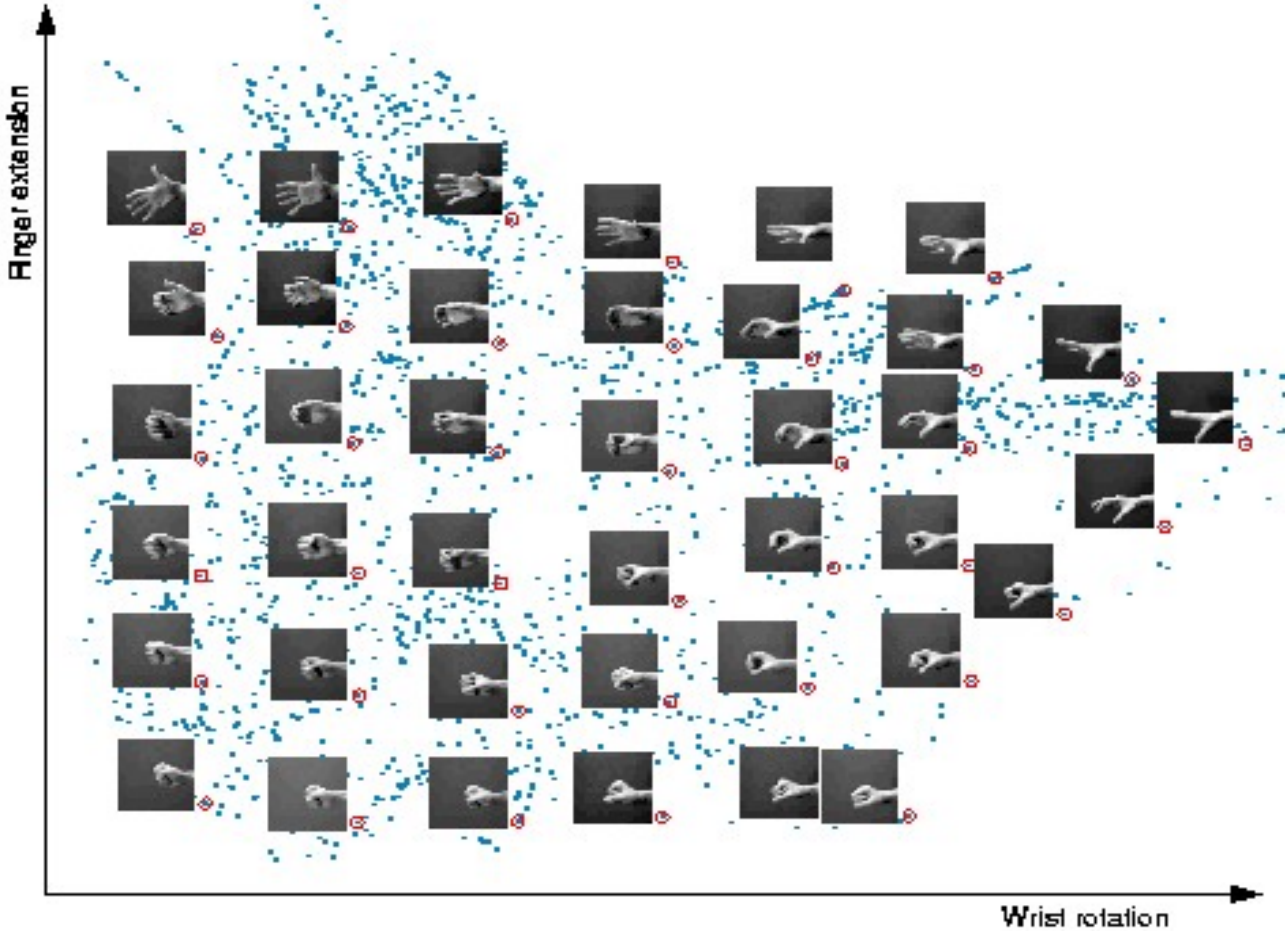
Isomap - Algorithm

- Determine the neighbors.
  - All points in a fixed radius.
  - K nearest neighbors
- Construct a neighborhood graph.
  - Each point is connected to the other if it is a K nearest neighbor.
  - Edge Length equals the Euclidean distance
- Compute the shortest paths between two nodes
  - Floyd's Algorithm
  - Djkastra's ALgorithm
- Construct a lower dimensional embedding.
  - Classical MDS

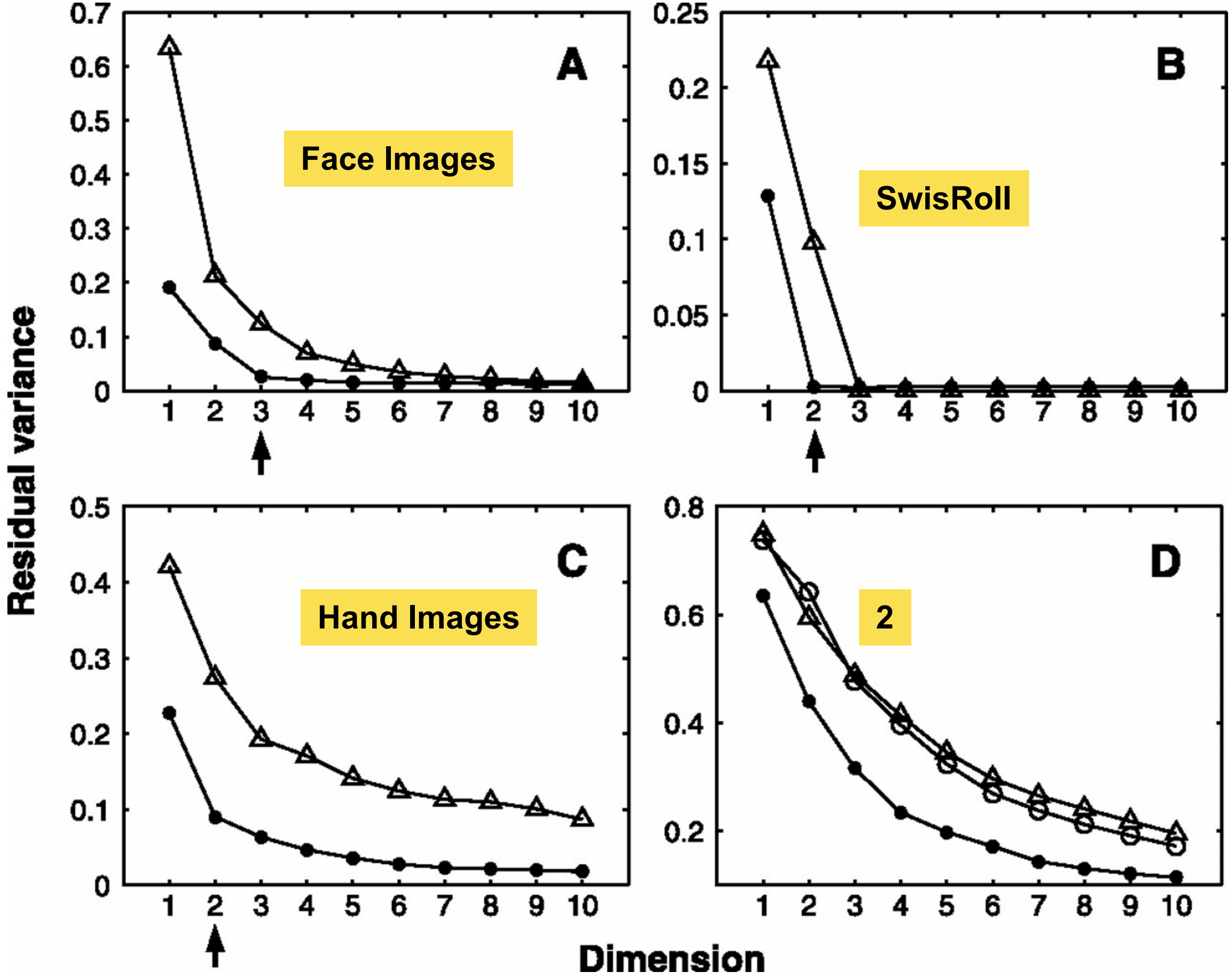
Isomap




B
Bottom loop articulation
Top arch articulation


Finger extension
Wrist rotation

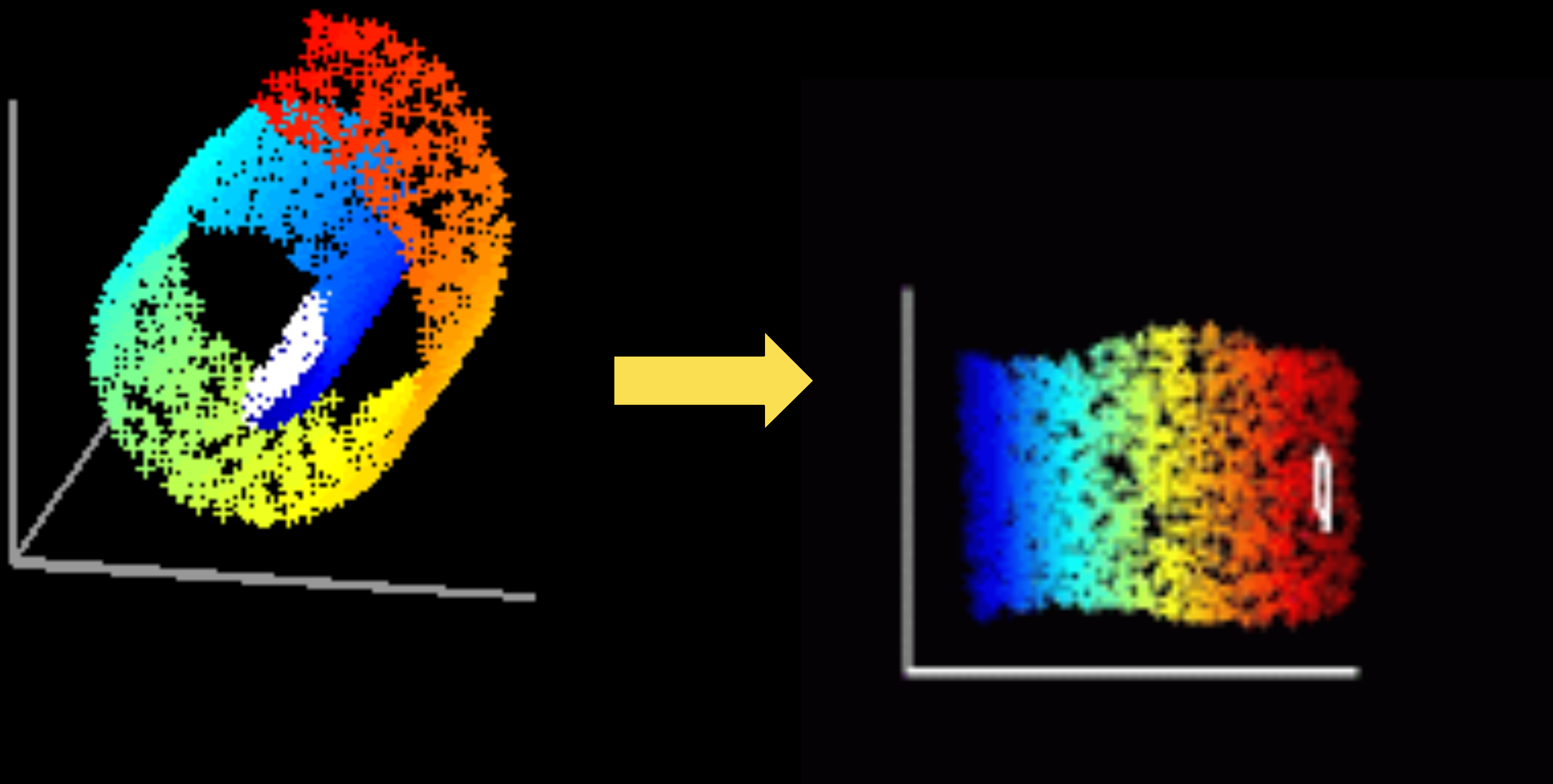
Residual Variance



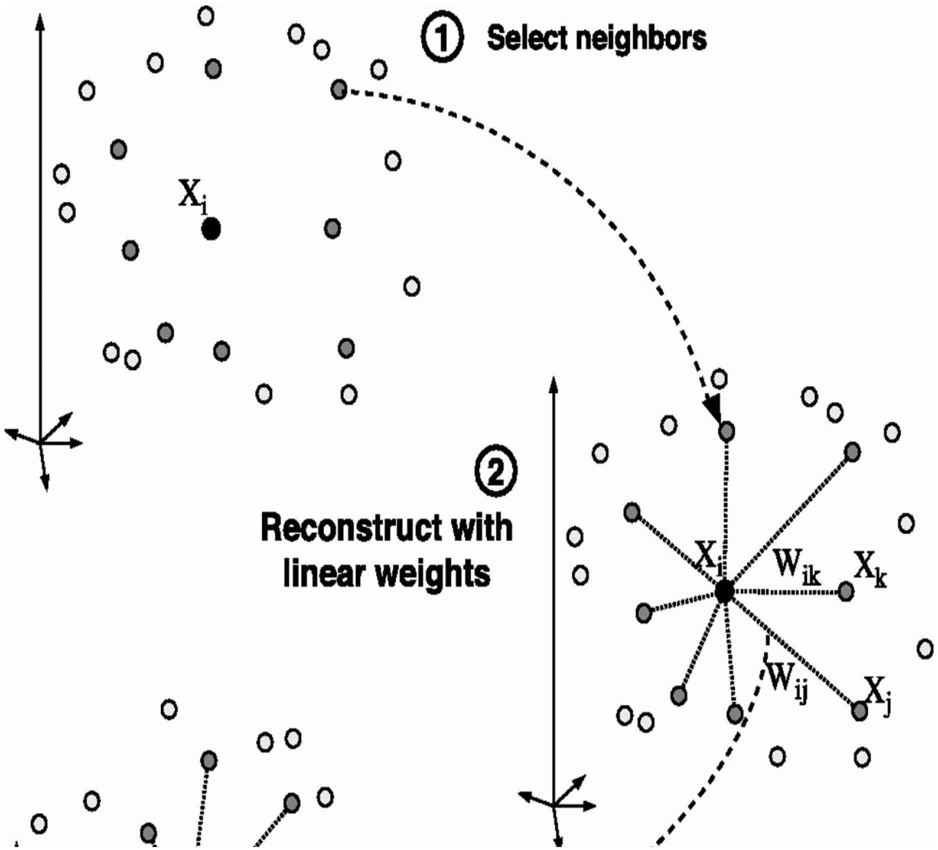
Locally Linear Embedding

manifold is a topological space which is locally Euclidean

Fit Locally , Think Globally



Fit Locally…



We expect each data point and its neighbours to lie on or close to a locally linear patch of the manifold.

Each point can be written as a linear combination of its neighbors.
The weights choosen to minimize the reconstruction Error.

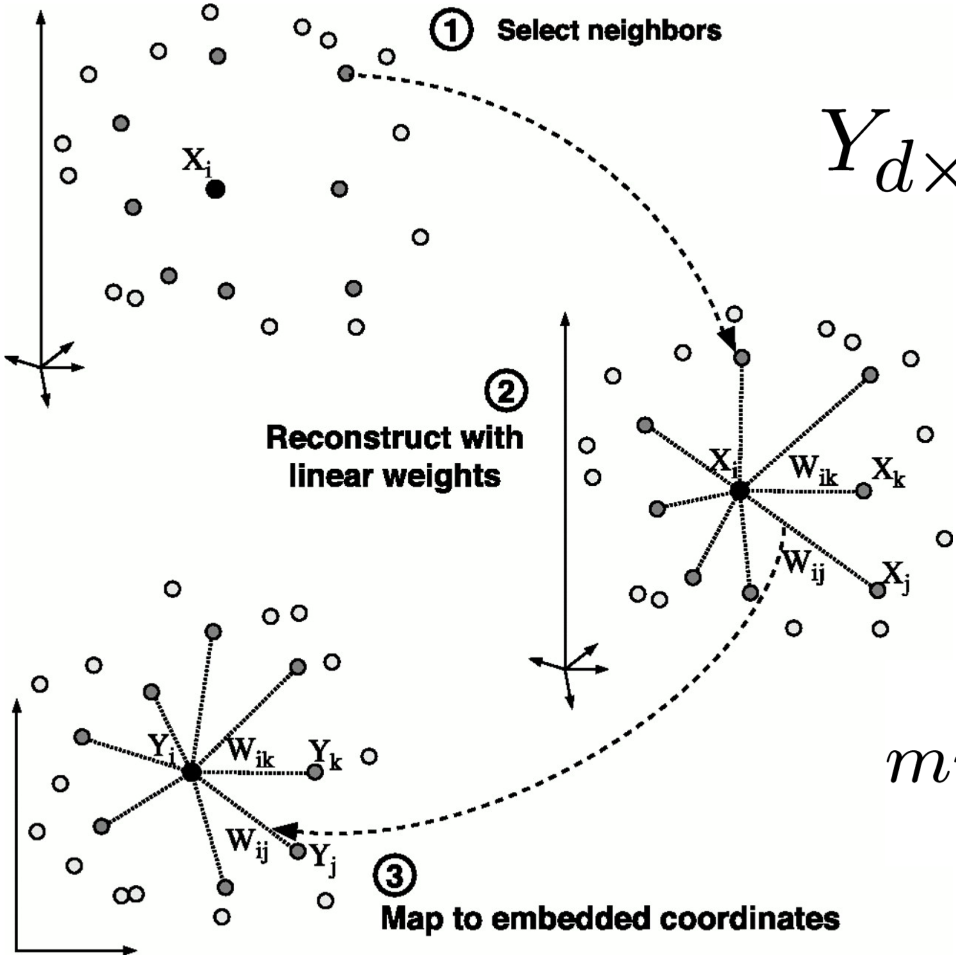
$$min_W \parallel X_i - \sum_{j=1}^{K} W_{ij} X_j \parallel^2 \qquad (1)$$

Derivation on board

Important property...

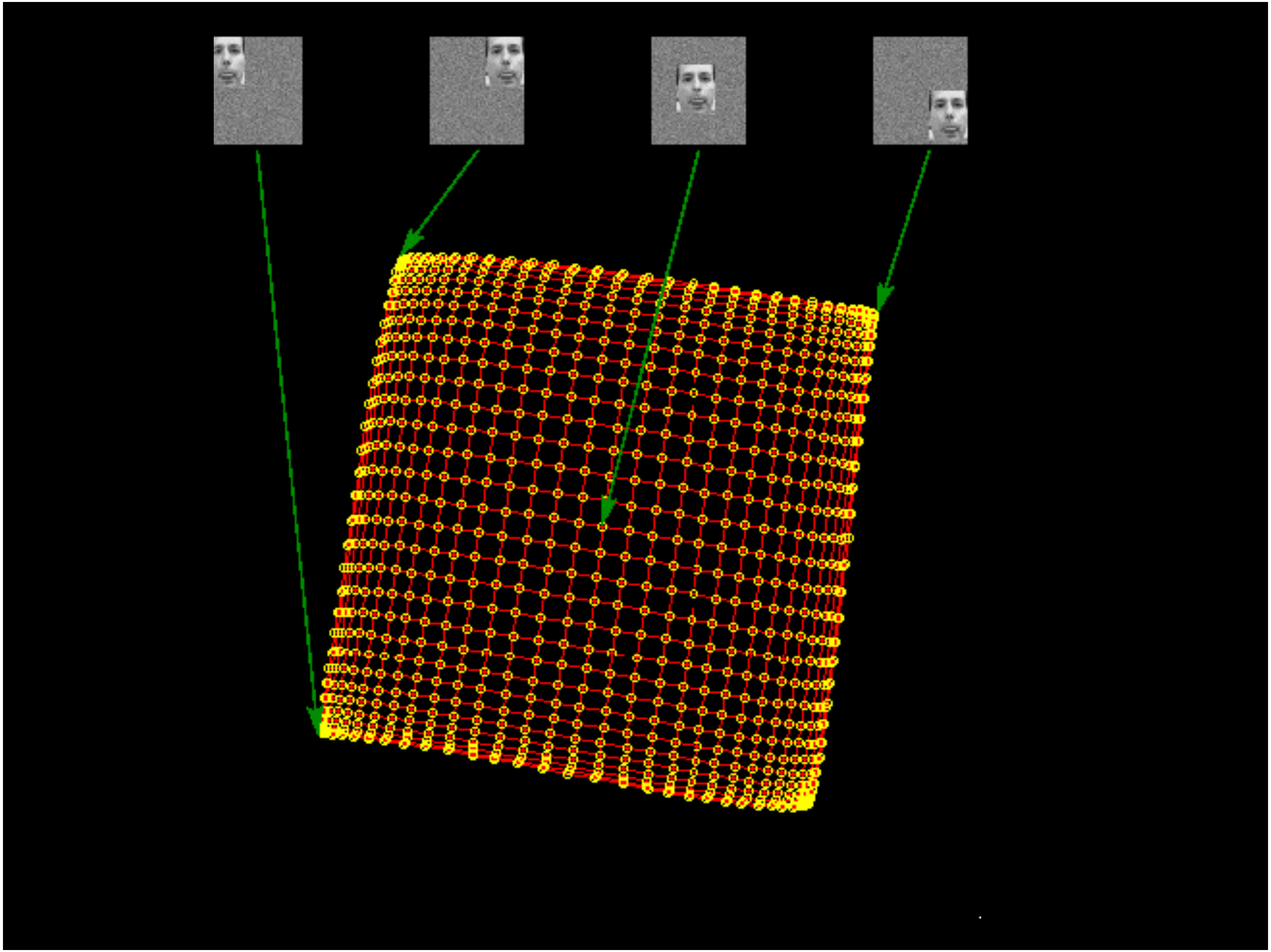
- The weights that minimize the reconstruction errors are invariant to rotation, rescaling and translation of the data points.
  - Invariance to translation is enforced by adding the constraint that the weights sum to one.
- **The same weights that reconstruct the datapoints in D dimensions should reconstruct it in the manifold in d dimensions.**
  - The weights characterize the intrinsic geometric properties of each neighborhood.

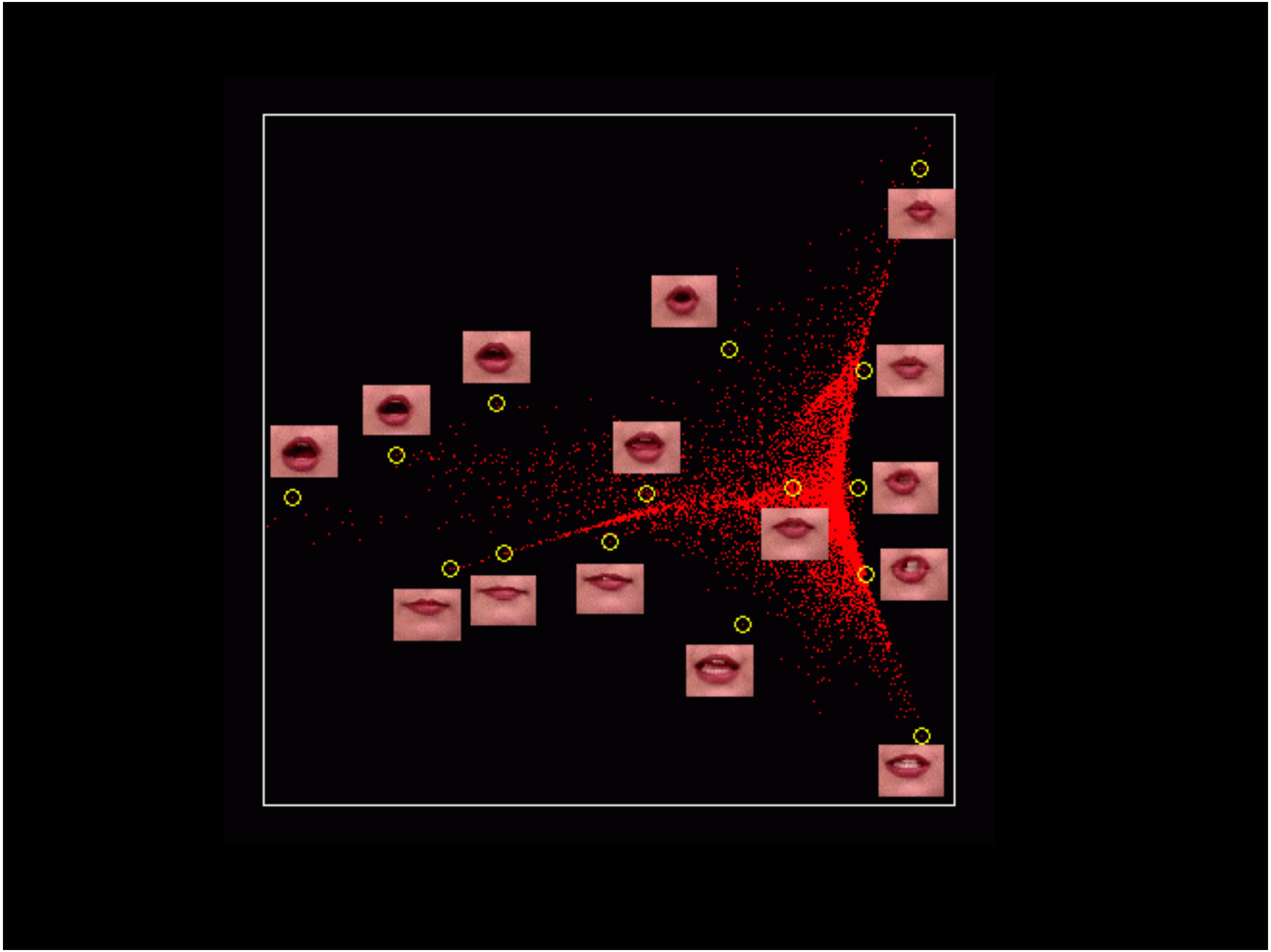
Think Globally...

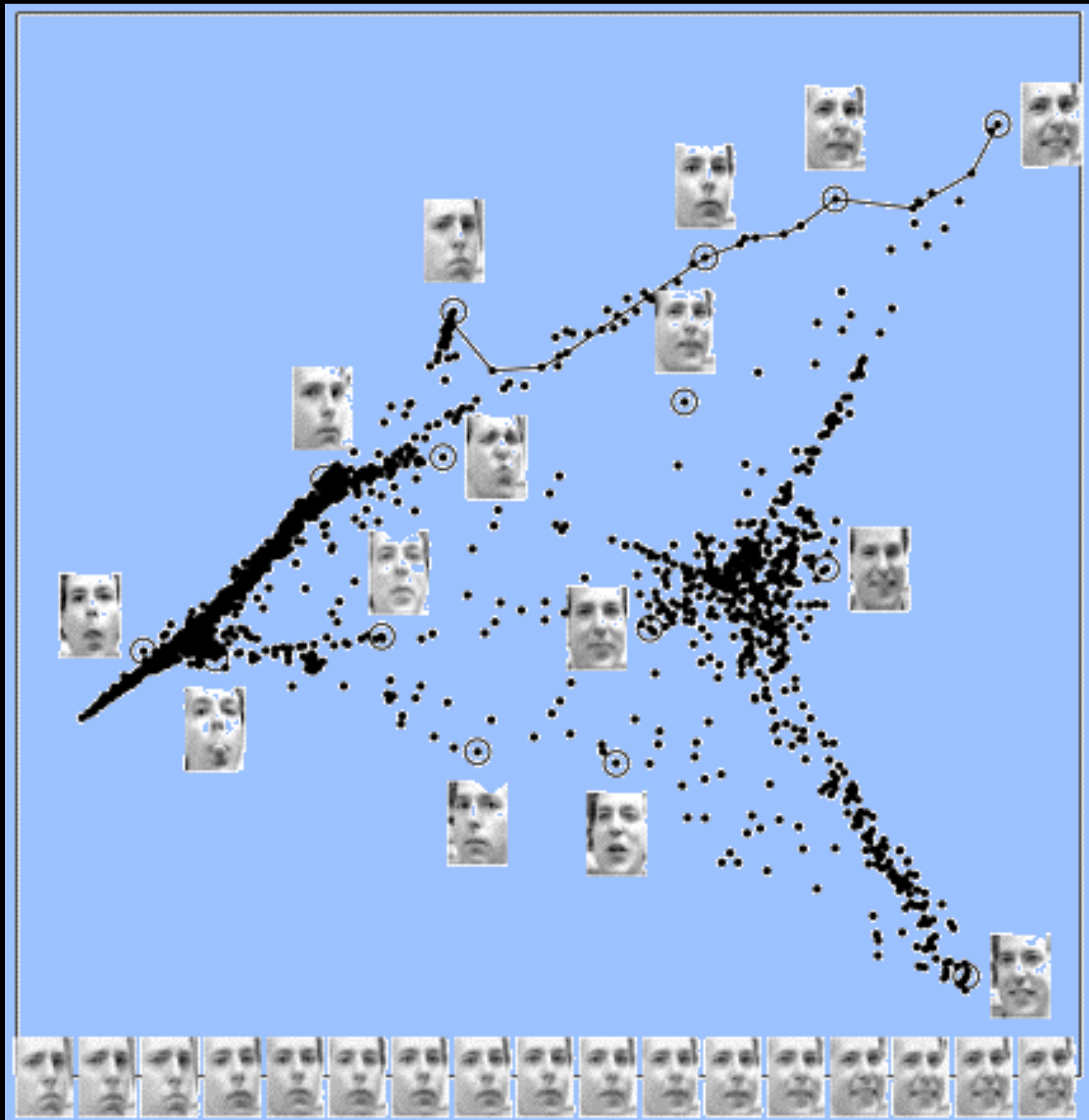


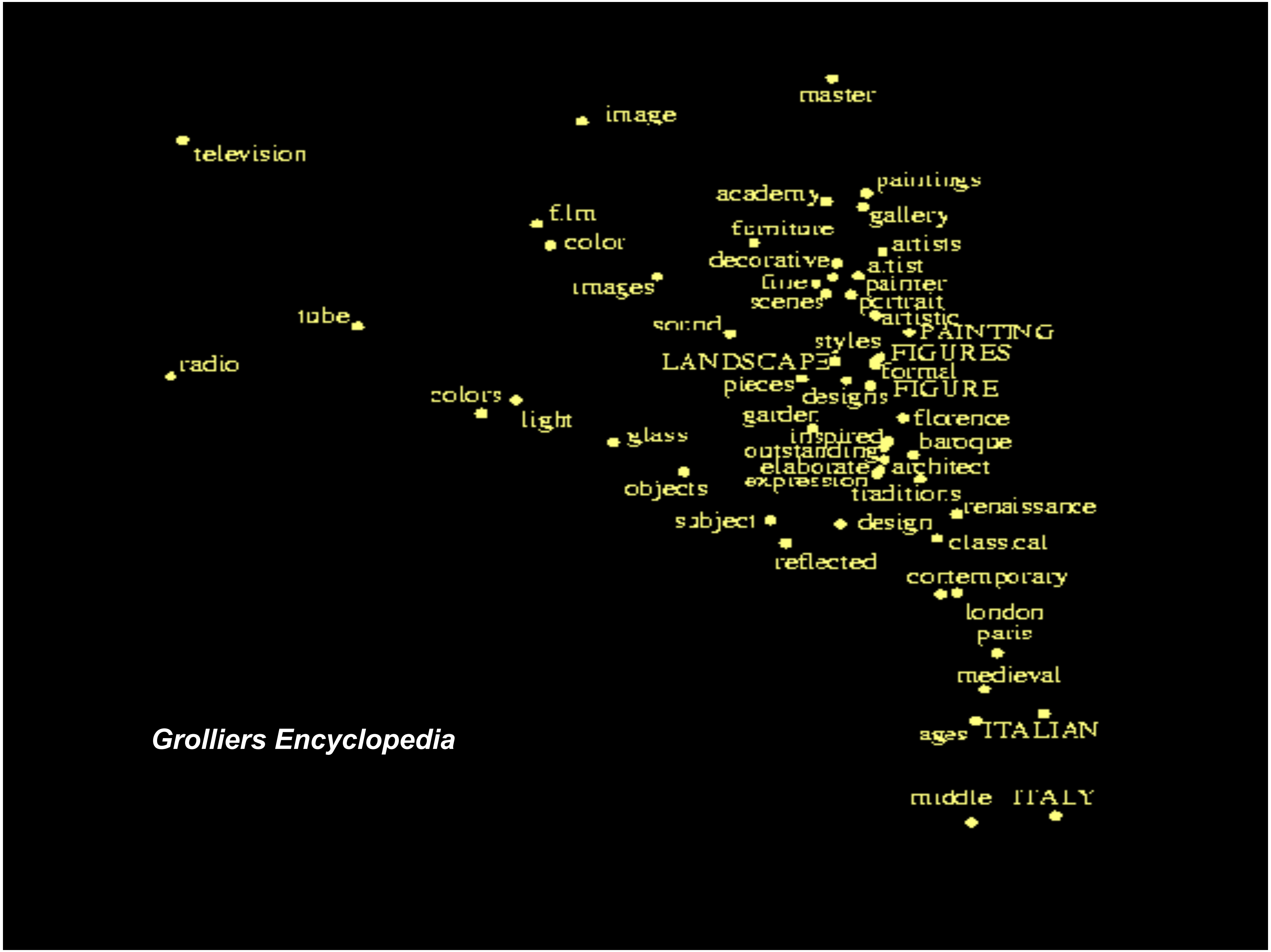
$$Y_{d \times N} = [Y_1 | Y_2 | ... | Y_N]$$

$$min_Y \sum_{i=1}^{N} \parallel Y_i - Y W_i \parallel^2$$








master
image
television
paintings
academy
gallery
film
color
furniture
artists
decorative
artist
fine
painter
images
scenes
portrait
artistic
tube
sound
PAINTING
styles
LANDSCAPE
FIGURES
formal
radio
pieces
designs
FIGURE
colors
light
garden
florence
glass
inspired
baroque
outstanding
elaborate
architect
expression
objects
traditions
renaissance
subject
design
classical
reflected
contemporary
london
paris
medieval
ages
ITALIAN
middle
ITALY
Grolliers Encyclopedia

Summary..

ISOMAP	LLE
Do MDS on the geodesic distance matrix.	Model local neighborhoods as linear a patches and then embed in a lower dimensional manifold.
Global approach	Local approach
Might not work for nonconvex manifolds with holes	Nonconvex manifolds with holes
Extensions: Conformal & Isometric ISOMAP	Extensions: Hessian Eigenmaps, Laplacian Eigenmaps etc.

Both needs manifold finely sampled.

Conformal & Isometric Embedding

Y d-dimensional domain in Euclidean space R^D

$f : Y -> R^D$ smooth embedding

Recover Y and f based on a given set of x_i in R^D.

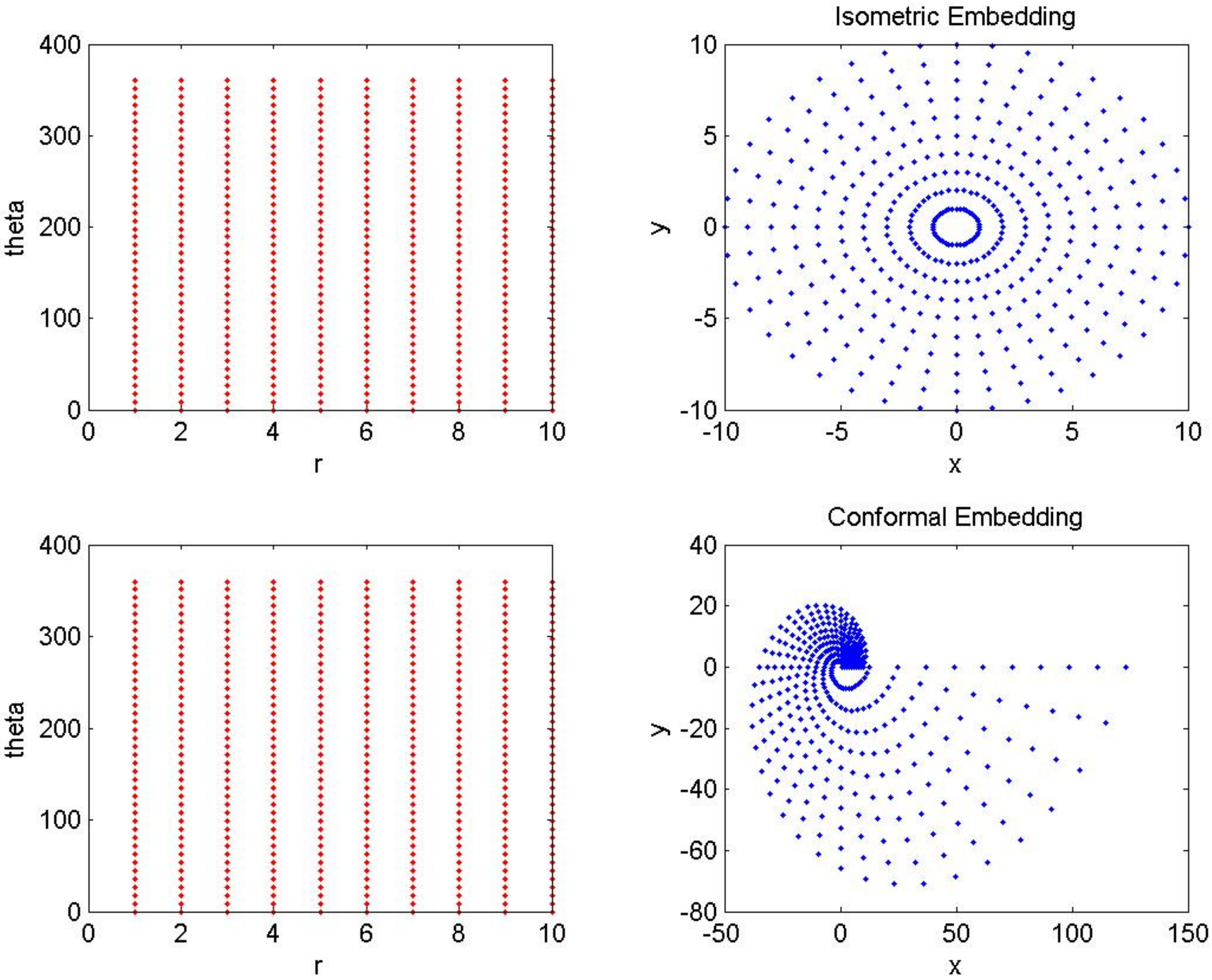
f is an isometric embedding if f preserves infinitesimal lengths and angles.

f is a conformal embedding if f preserves infinitesimal angles.

At every point y there is a scalar $s(y) > 0$ such that the infintesimal vectors at y get magnified in length by a factor $s(y)$.

Isometric Mapping and C-Isomap

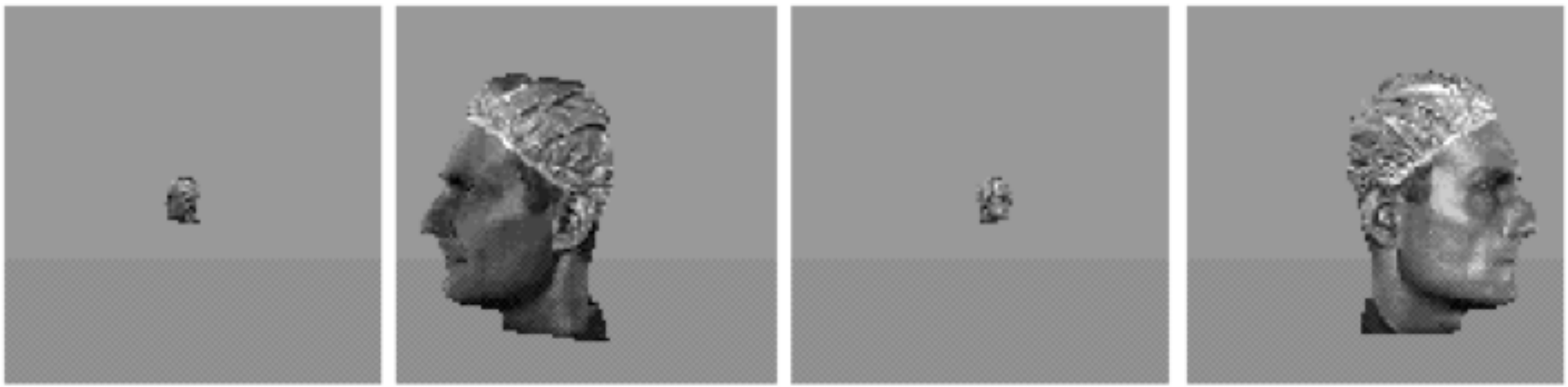
- Isometric mapping
  - Intrinsically flat manifold
  - Invariants??
    - Geodesic distances are reserved.
    - Metric space under geodesic distance.
- Conformal Embedding
  - Locally isometric upto a scale factor s(y)
  - Estimate s(y) and rescale.
  - C-Isomap
  - Original data should be uniformly dense



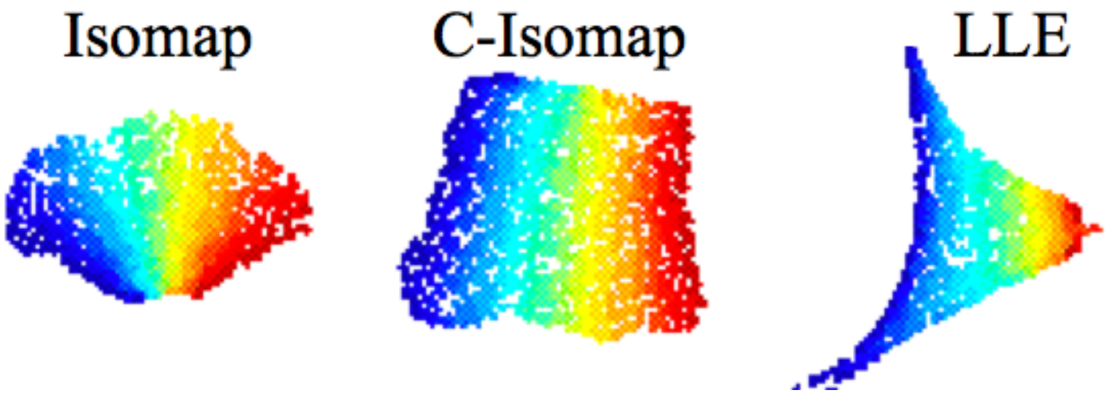
Isometric Embedding
Conformal Embedding
theta
r
x
y

C-Isomap

- C-Isomap is similar to Isomap, but the graph weights are renormalised.
- Suitable when observed effect of parameter variation is not constant over the manifold.



More Example: C-Isomap



- C-Isomap succeeds when
  - Y is a convex subsets of Euclidean space
  - Data are densely sampled, uniformly over Y
  - F is a conformal embedding

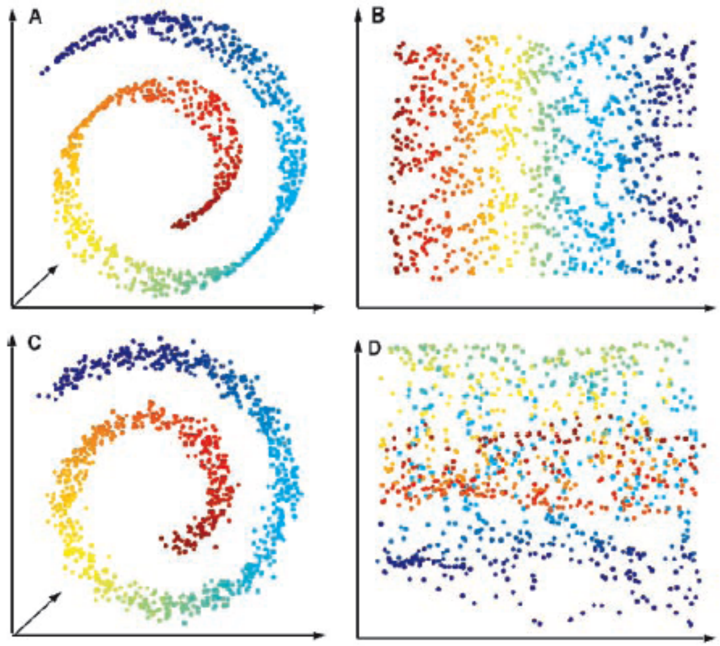
Short Circuit Problem???

Unstable?

Only free parameter is

How many neighbours?

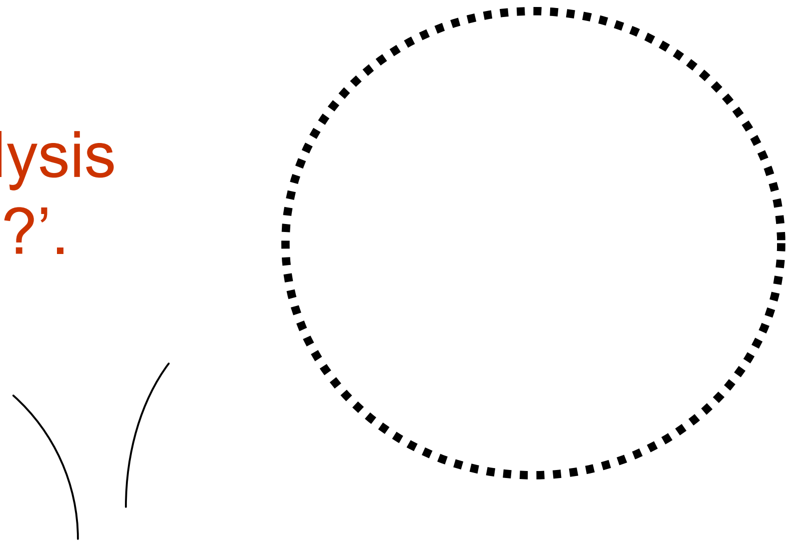
– How to choose neighborhoods.
    - Susceptible to short-circuit errors if neighborhood is larger than the folds in the manifold.
    - If small we get isolated patches.



???

- Isomap might not work on closed manifold, manifolds with holes?
- Noisy Data?
- Sparse Data?

Develop Multiscale analysis to solve some of those '?'.



Acknowledgement

- Slides stolen from Vikas C. Raykar, Vin de Silva, et al.