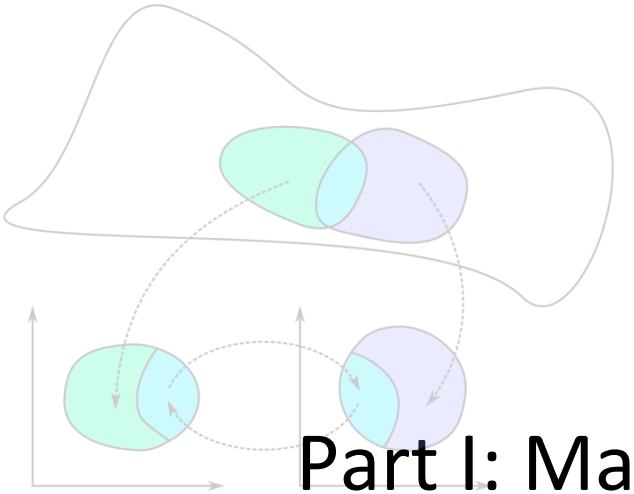
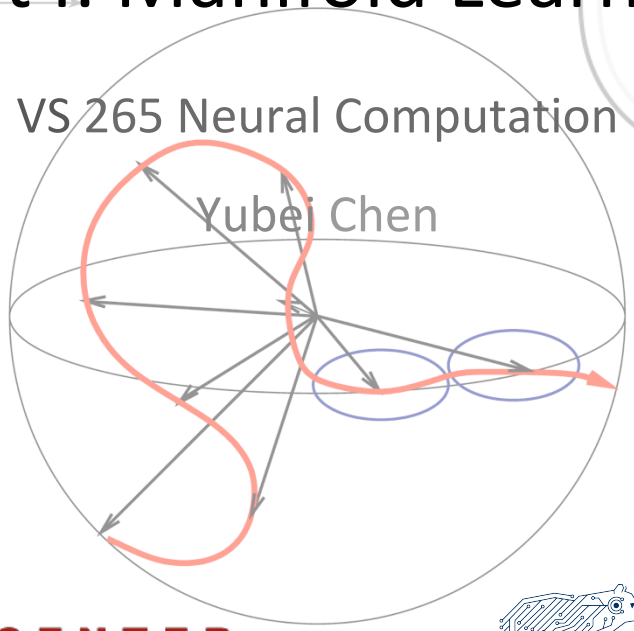
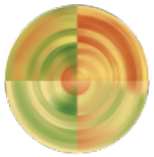


Part I: Manifold Learning

VS 265 Neural Computation

Yubei Chen

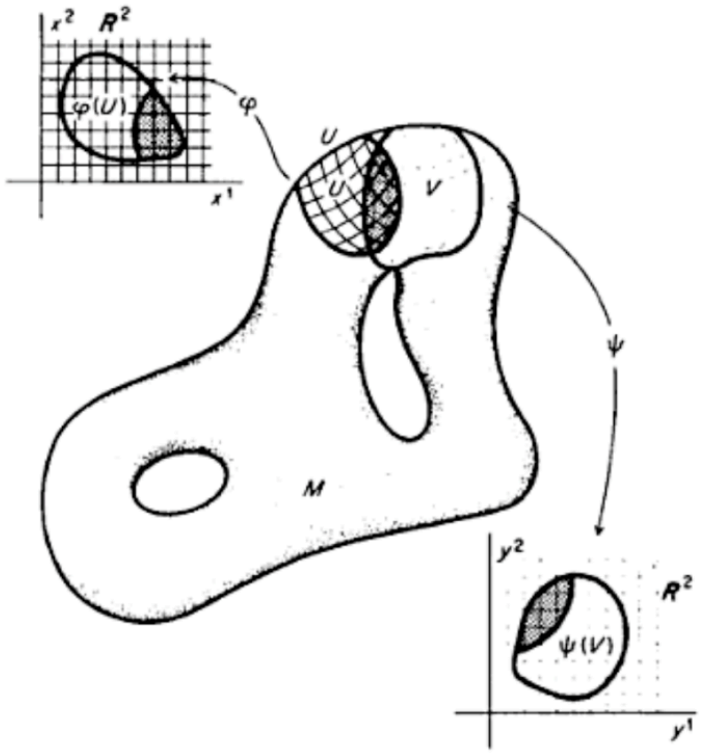


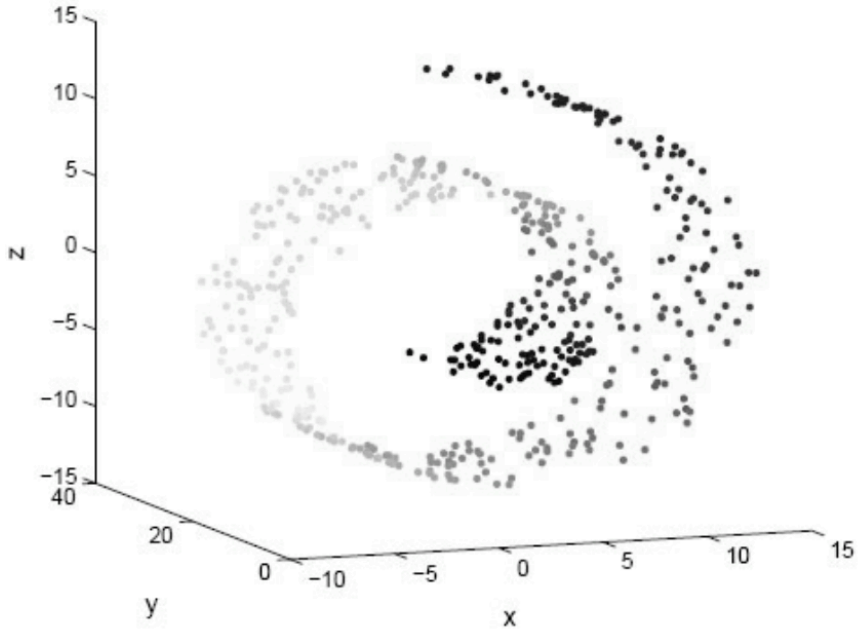


REDWOOD CENTER
for Theoretical Neuroscience

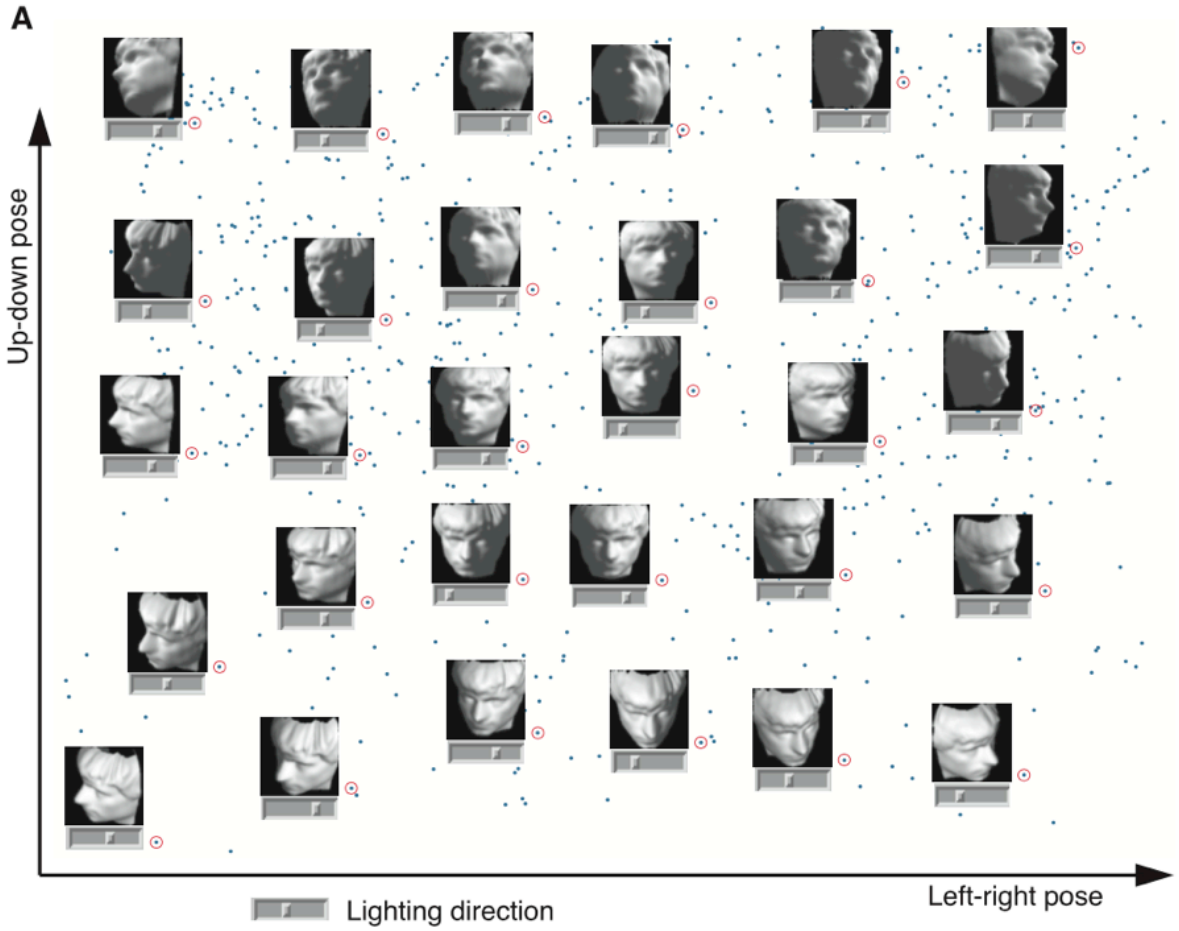
BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Manifold Hypothesis





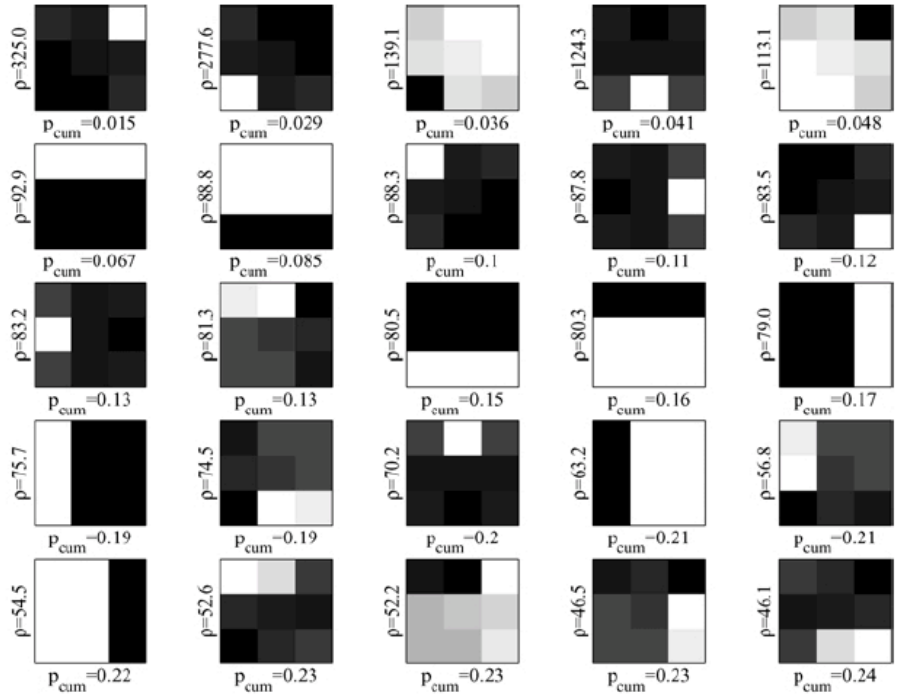
Manifold of facial pose and lighting



(Tenenbaum et al. ‘00)

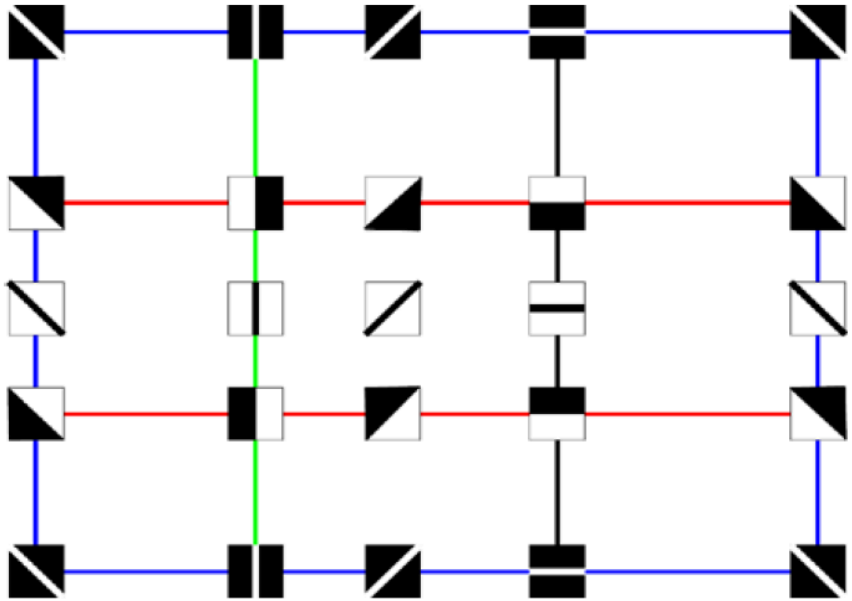
Adapted From Bruno Olshausen’s slides

High−Contrast 3x3 Image Patches



(Lee,Kim & Mumford, ’03)

High−Contrast 3x3 Image Patches



Klein bottle

(Carlsson et al., ’09)

Motion Sequence

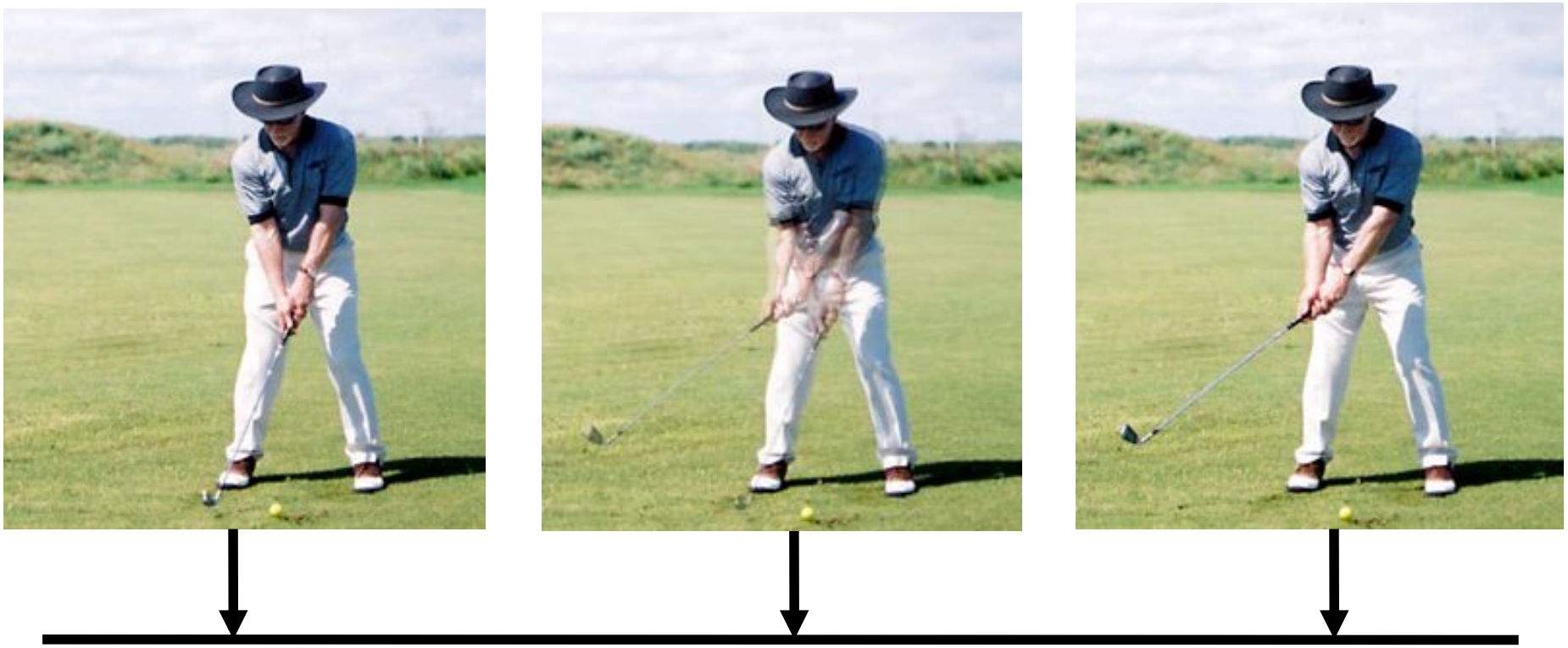
Images from www.golfswingphotos.com, and adapted from Alyosha Efros's Slides

reasonable distance metrics

?

Images from www.golfswingphotos.com, and adapted from Alyosha Efros's Slides

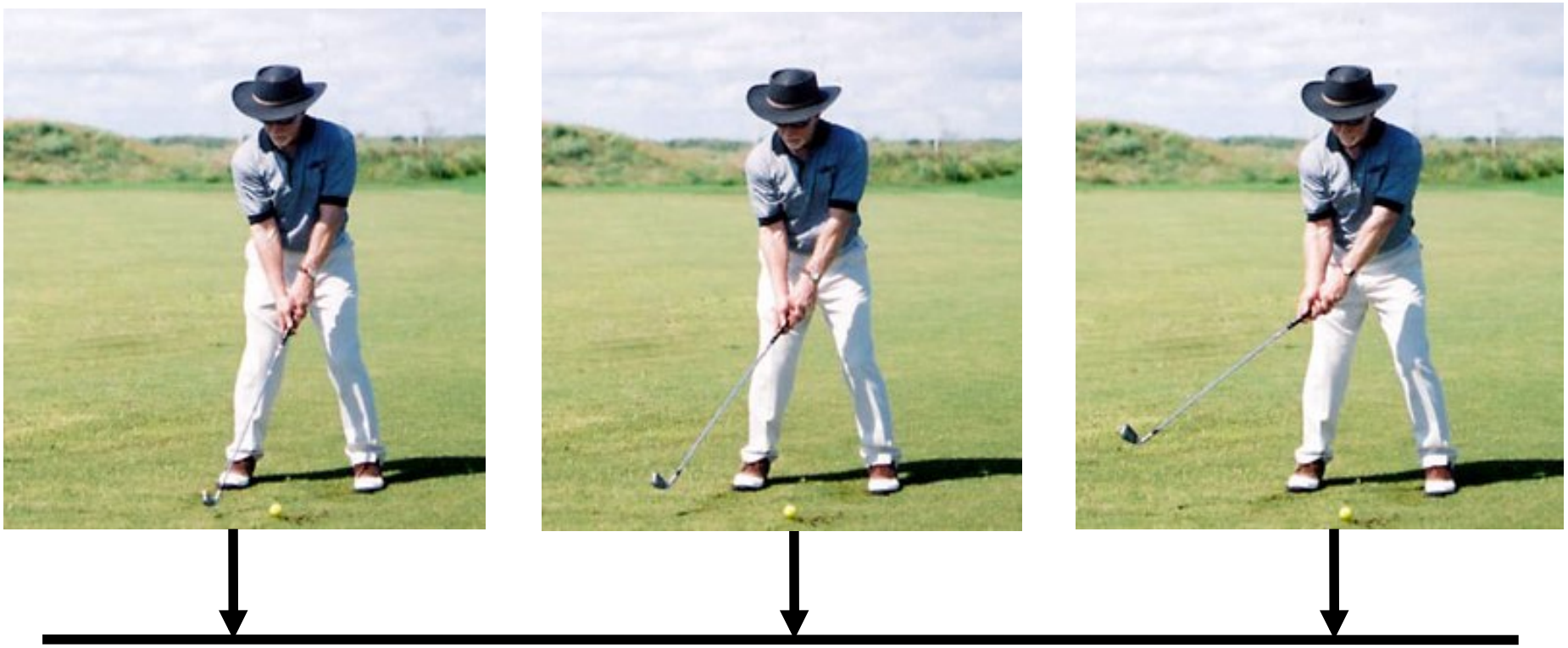
reasonable distance metrics



linear interpolation

Images from www.golfswingphotos.com, and adapted from Alyosha Efros's Slides

reasonable distance metrics

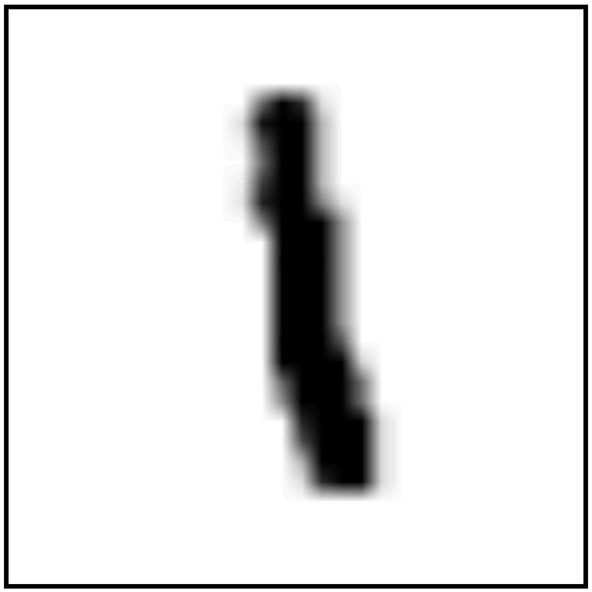


manifold interpolation

Images from www.golfswingphotos.com, and adapted from Alyosha Efros's Slides

Manifold Interpolation

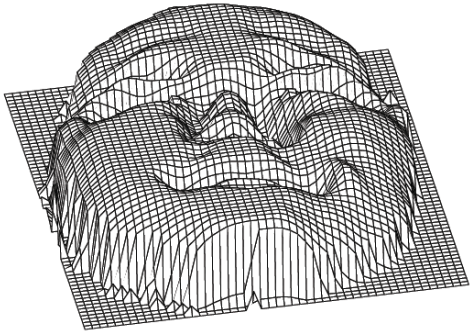
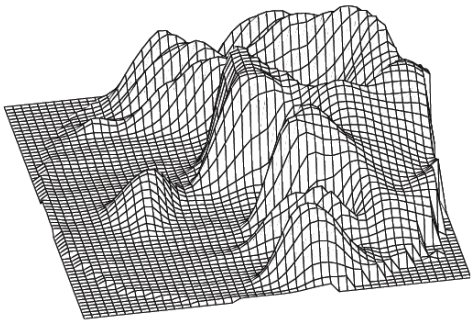
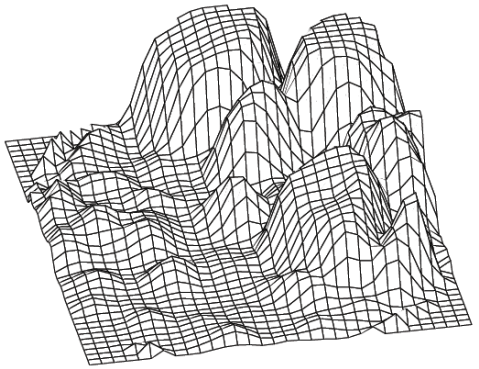
Digits Images, MNIST

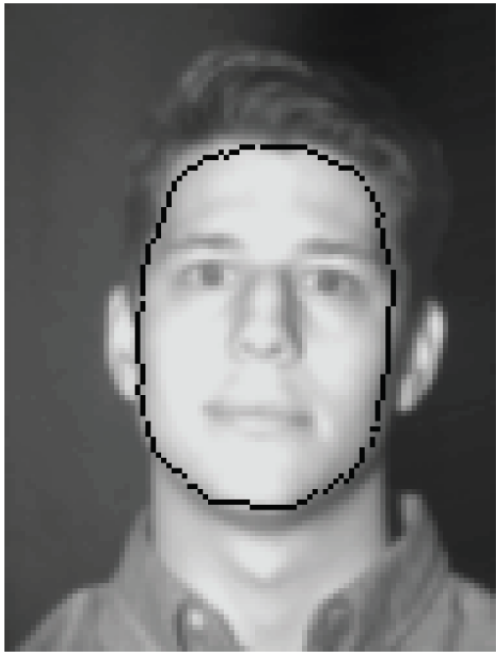


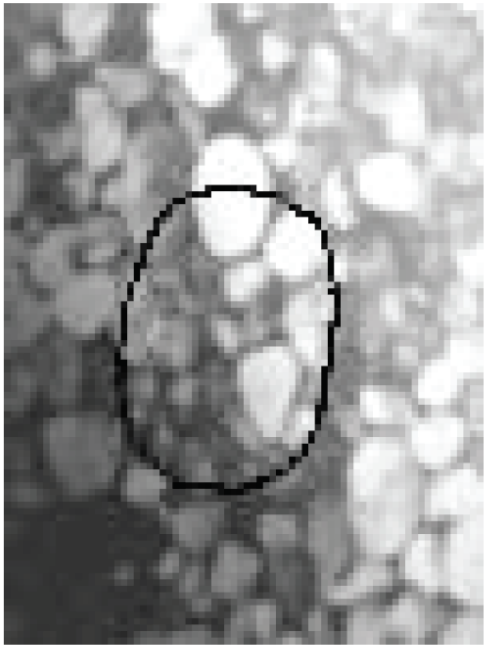
$$\mathbb{R}^{14\times14}$$

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .6 & .8 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .5 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .9 & 1 & .1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .3 & 1 & .1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Which is Close to the Middle Image?





Images from Pattern Theory, Mumford & Desolneux '12

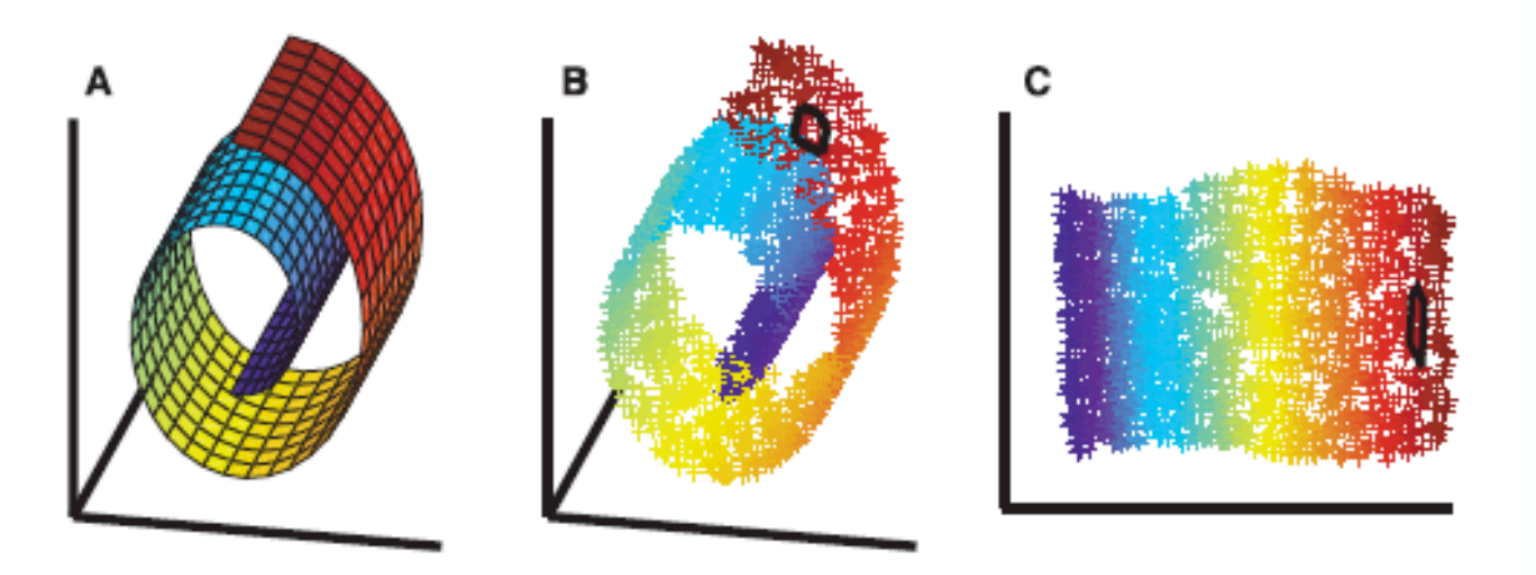
Nonlinear Dimensionality Reduction by Locally Linear Embedding

Sam T. Roweis[1] and Lawrence K. Saul[2]

A Global Geometric Framework for Nonlinear Dimensionality Reduction

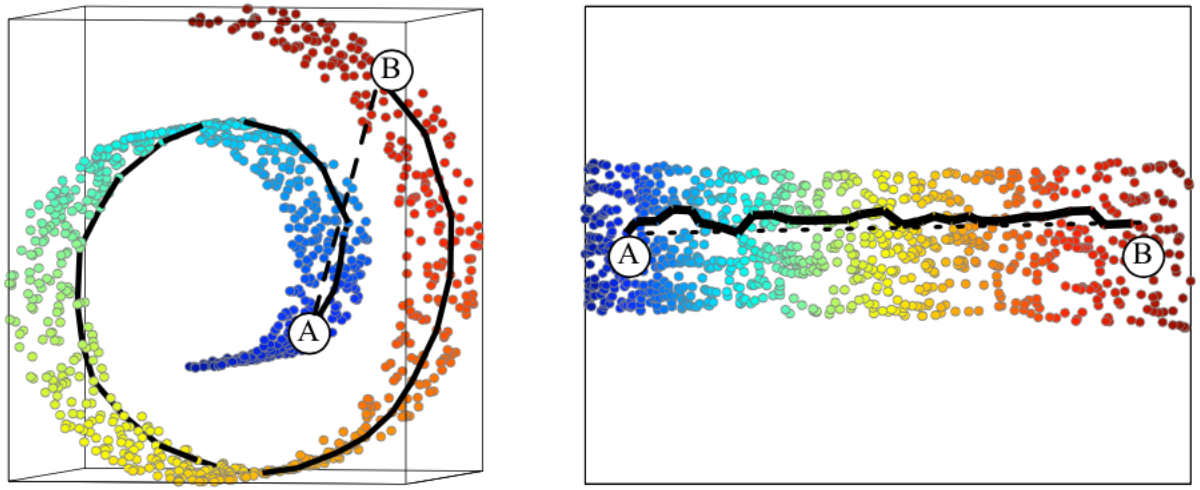
Joshua B. Tenenbaum,[1]* Vin de Silva,[2] John C. Langford[3]

Science, 22 Dec. 2000



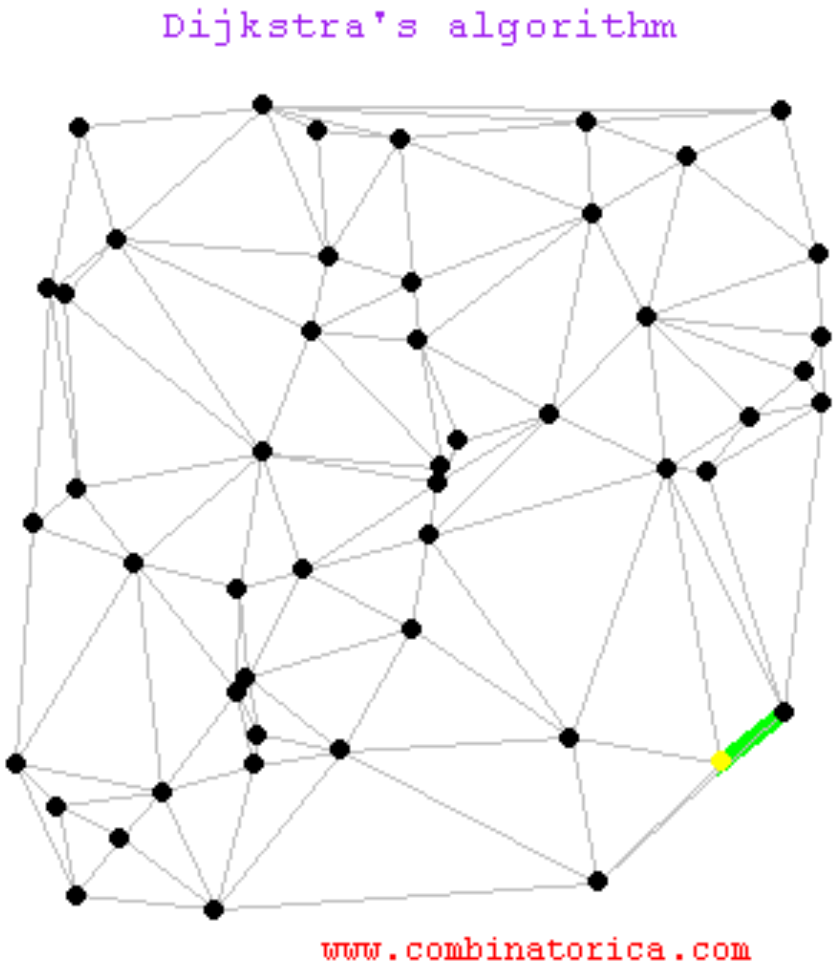
Adapted From Bruno Olshausen's slides

ISOMAP: To Preserve the Pairwise Geodesic Distance



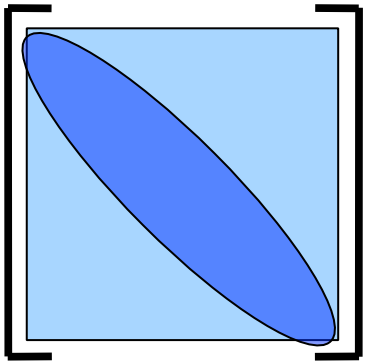
(Tenenbaum et al. ‘00)

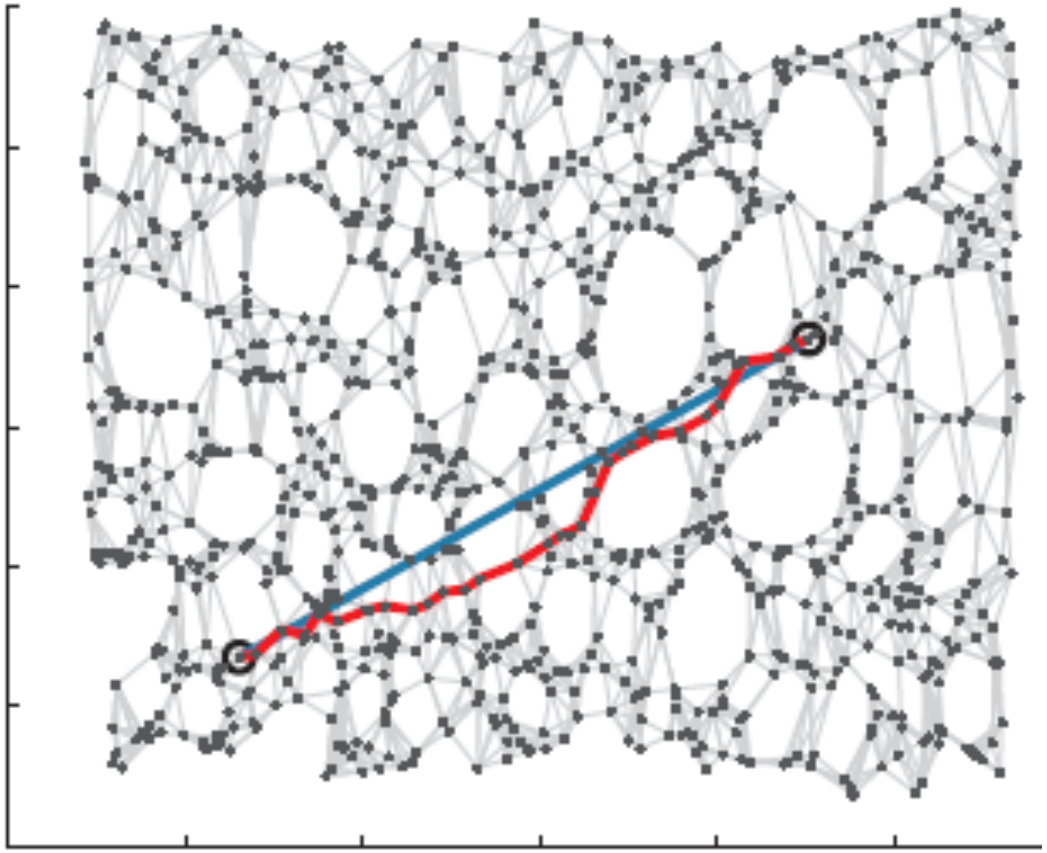
ISOMAP: Geodesic Distance Approximation



Isomap

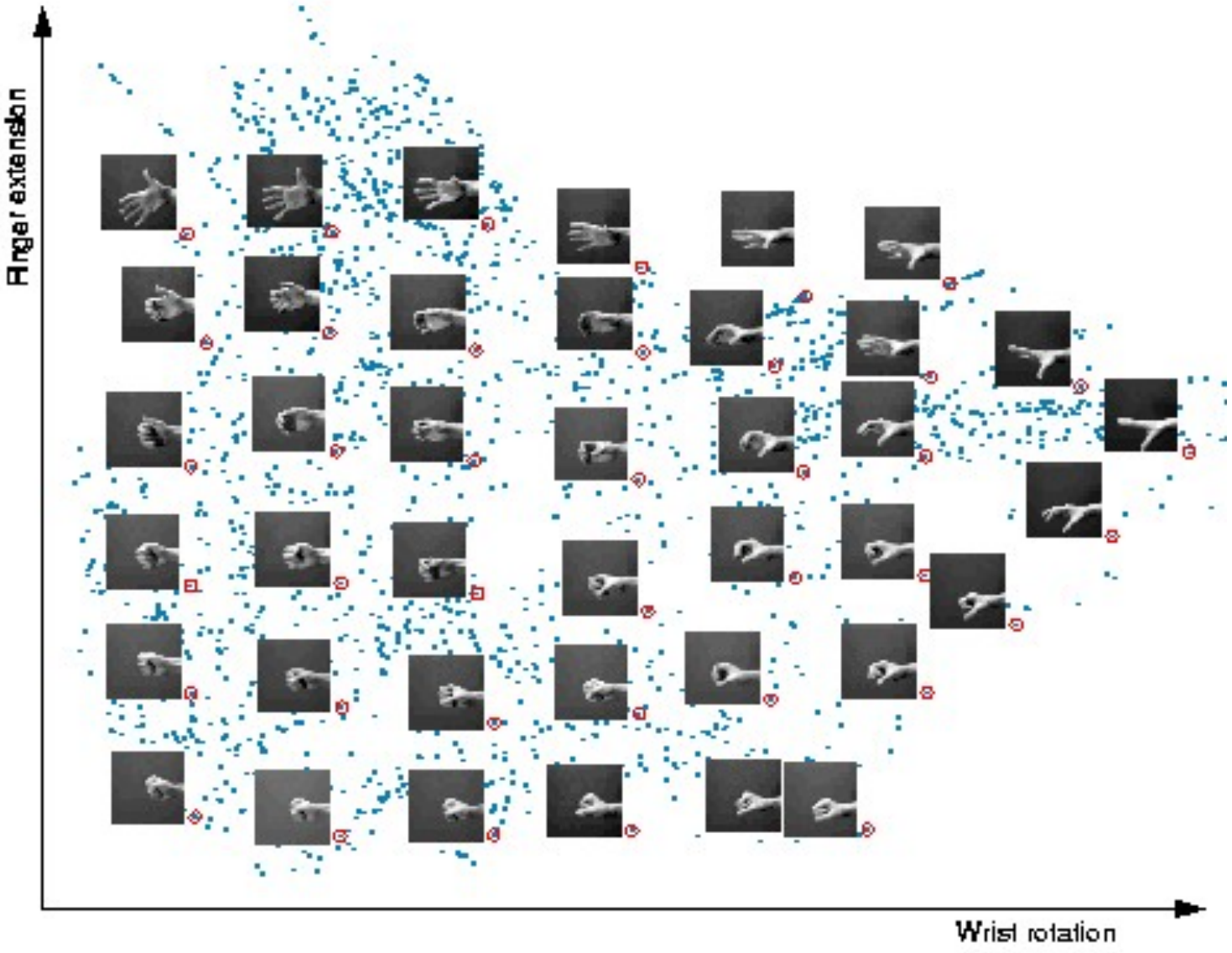
2. Infer other interpoint distances by finding shortest paths on the graph (Dijkstra's algorithm).

$D_g =$ 

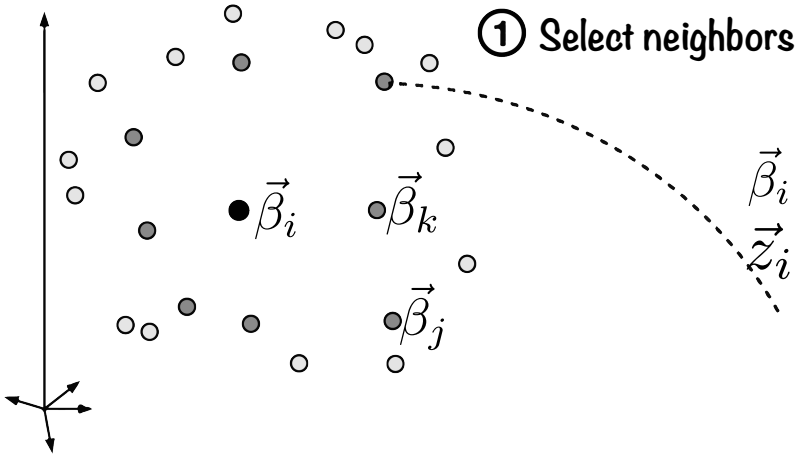


Adapted from Alyosha Efros's Slides

ISOMAP Result: Hands Gestures



Manifold Learning, LLE

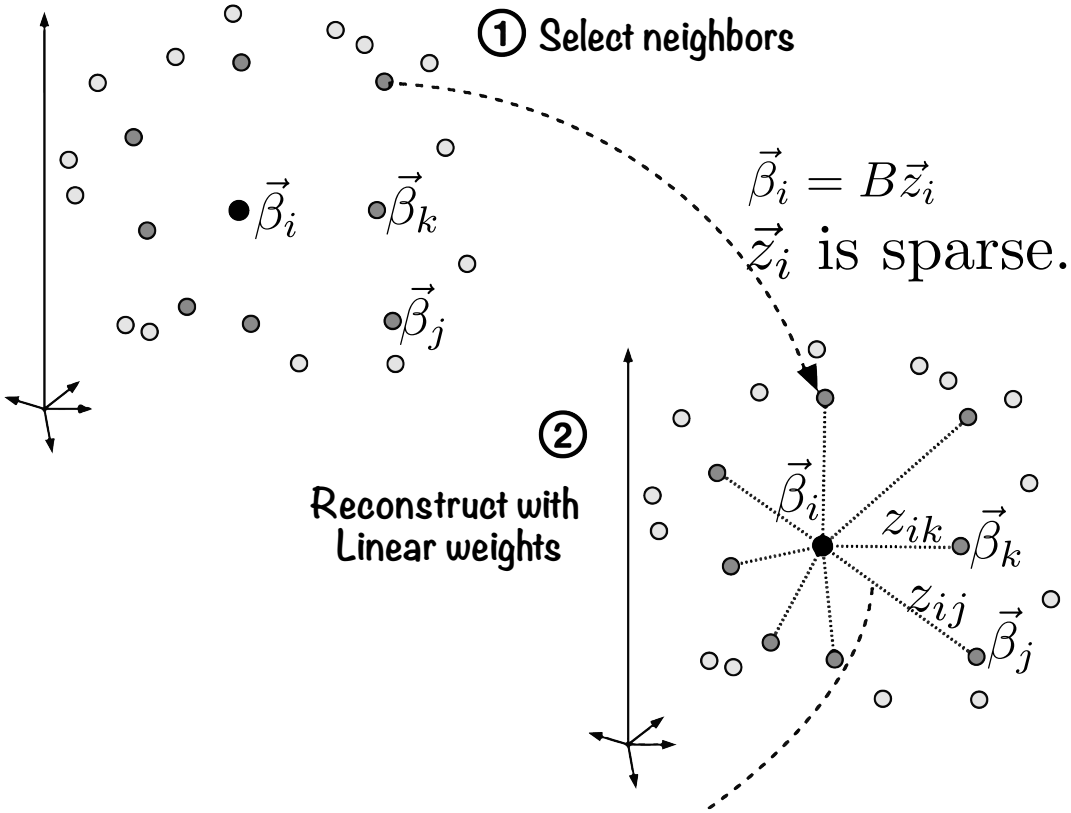


$$\vec{\beta}_i = B\vec{z}_i$$

$\vec{z}_i$ is sparse. $\|\vec{z}_i\|_1 = 1$

(Roweis & Saul, '00)

Manifold Learning, LLE



$$\vec{\beta}_i = B\vec{z}_i$$

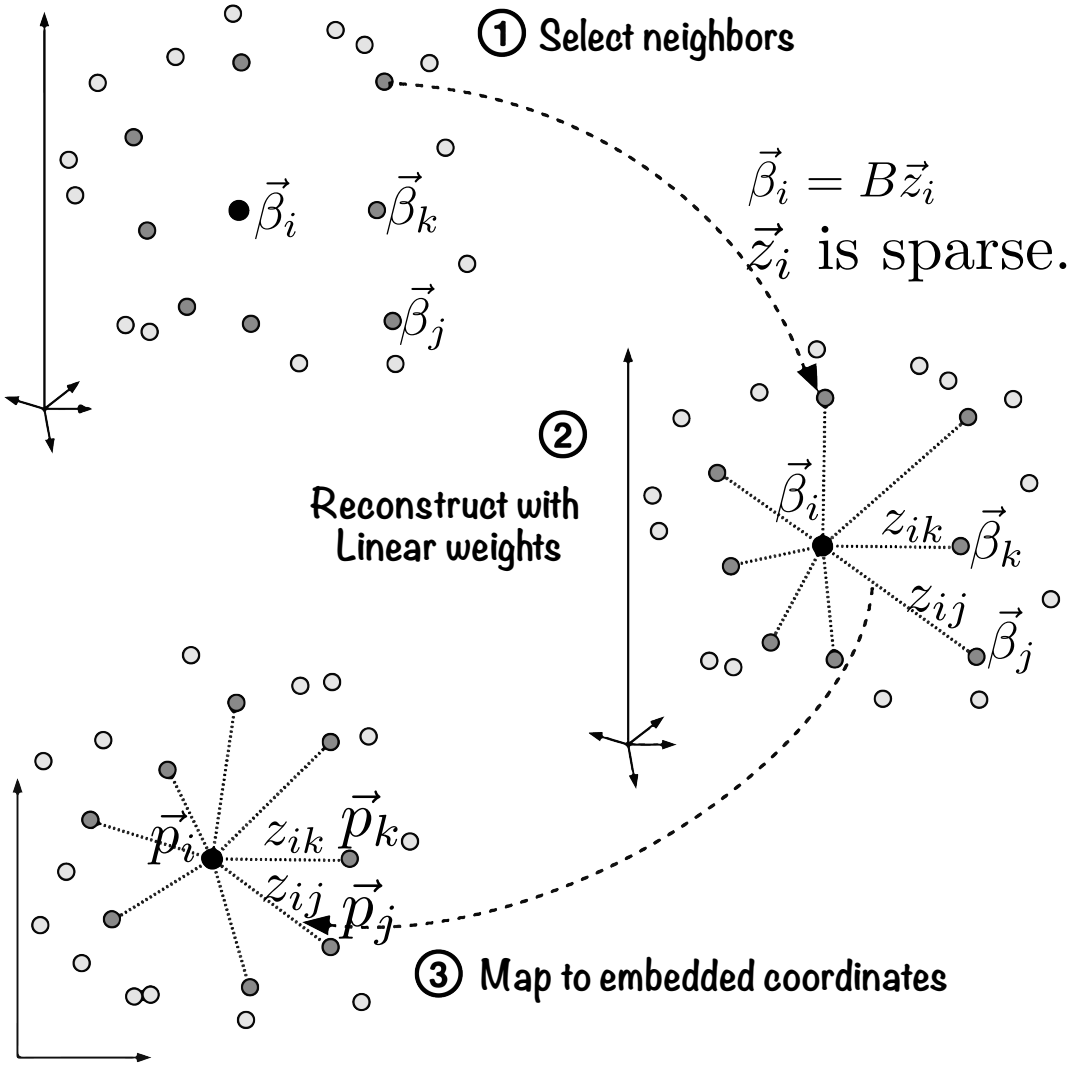
$\vec{z}_i$ is sparse. $\|\vec{z}_i\|_1 = 1$

$$B = BZ \rightarrow B(I - Z) = 0$$

Goal: $P(I - Z) \approx 0$

(Roweis & Saul, '00)

Manifold Learning, LLE



$$\vec{\beta}_i = B\vec{z}_i$$

$\vec{z}_i$ is sparse. $\|\vec{z}_i\|_1 = 1$

$$B = BZ \rightarrow B(I - Z) = 0$$

Goal: $P(I - Z) \approx 0$

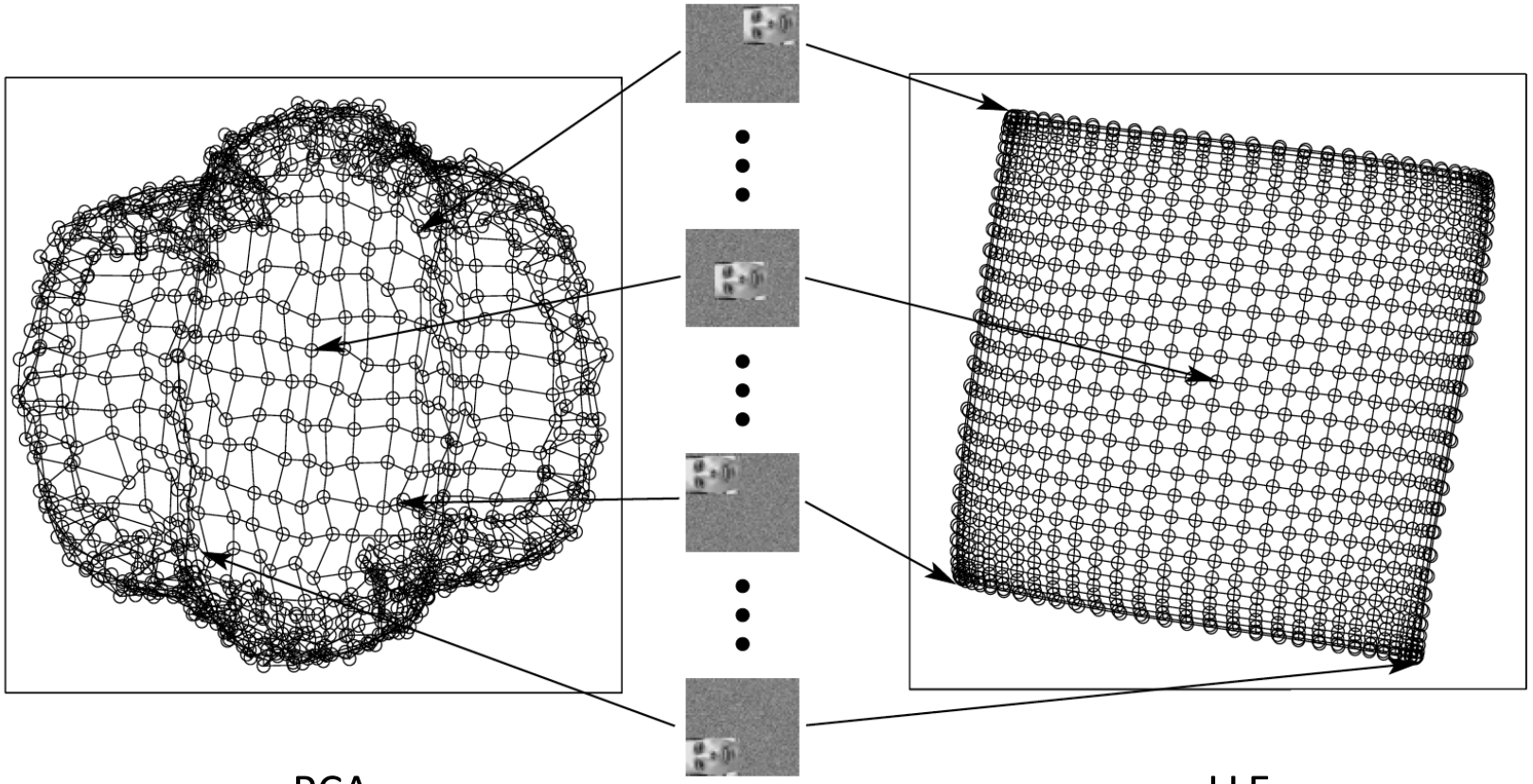
Opt:

$$\min_P \quad tr\ P(I-Z)(I-Z)^T P^T$$

$$\text{s.t.} \quad PUP^T = I$$

(Roweis & Saul, '00)

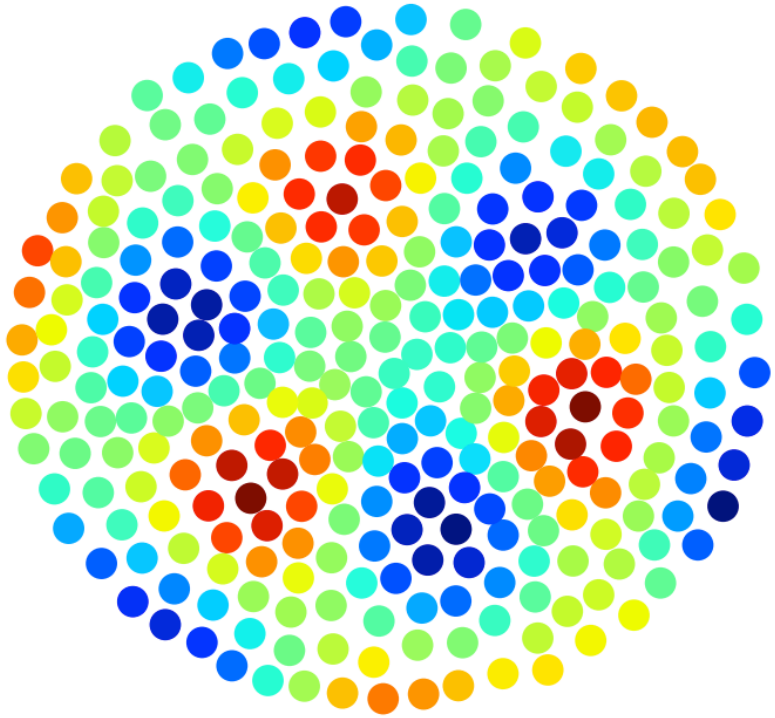
LLE Result: Translated Face patches



PCA

LLE

(Roweis & Saul, '00)



Part II: The Sparse Manifold Transform

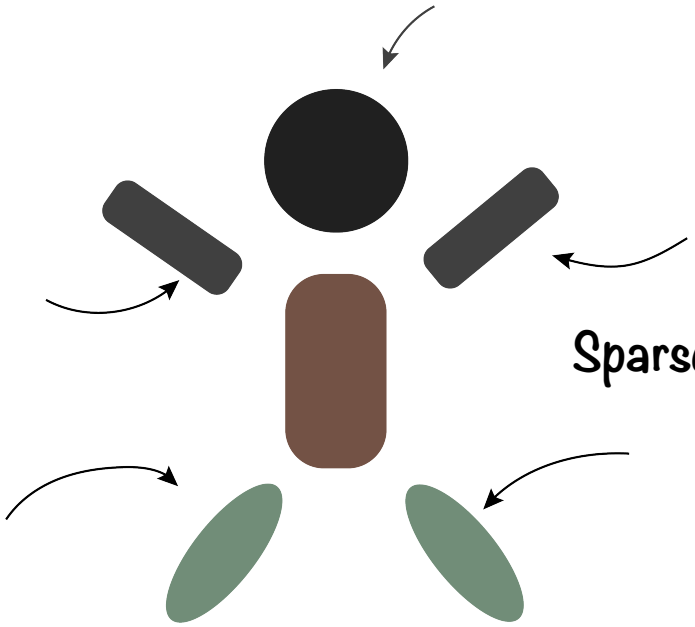
Unifying Discreteness and Continuity in Signal Representation

VS 265 Neural Computation

Yubei Chen

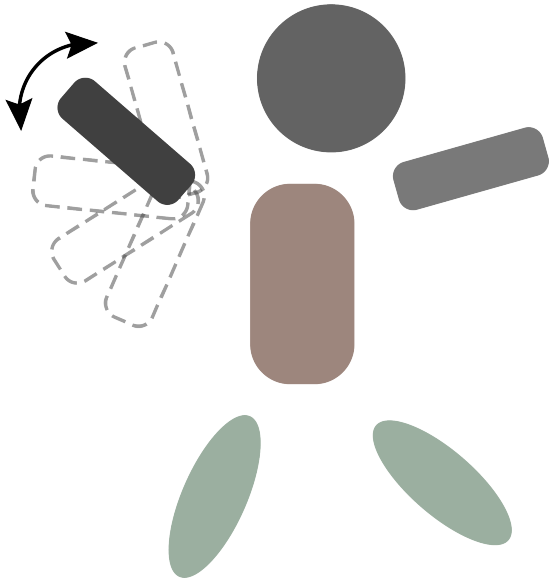
REDWOOD CENTER
for Theoretical Neuroscience

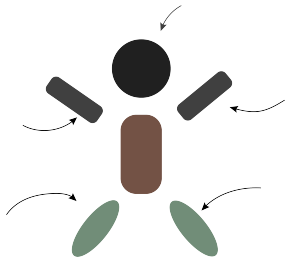
BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH



Sparse and discrete elements

Simple geometric transformations





Sparse Coding

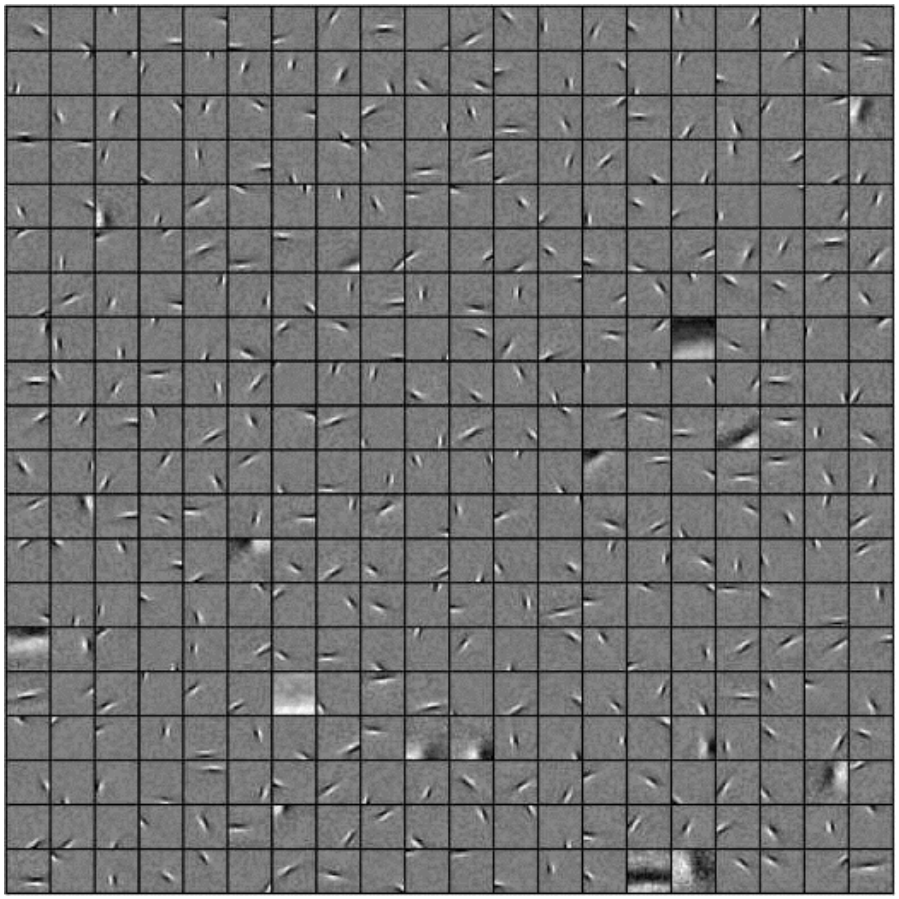
$\vec{\beta}_i$ is the input signal

$$\vec{\beta}_i = \Phi \vec{\alpha}_i$$

$\vec{\alpha}_i$ is sparse

Opt:

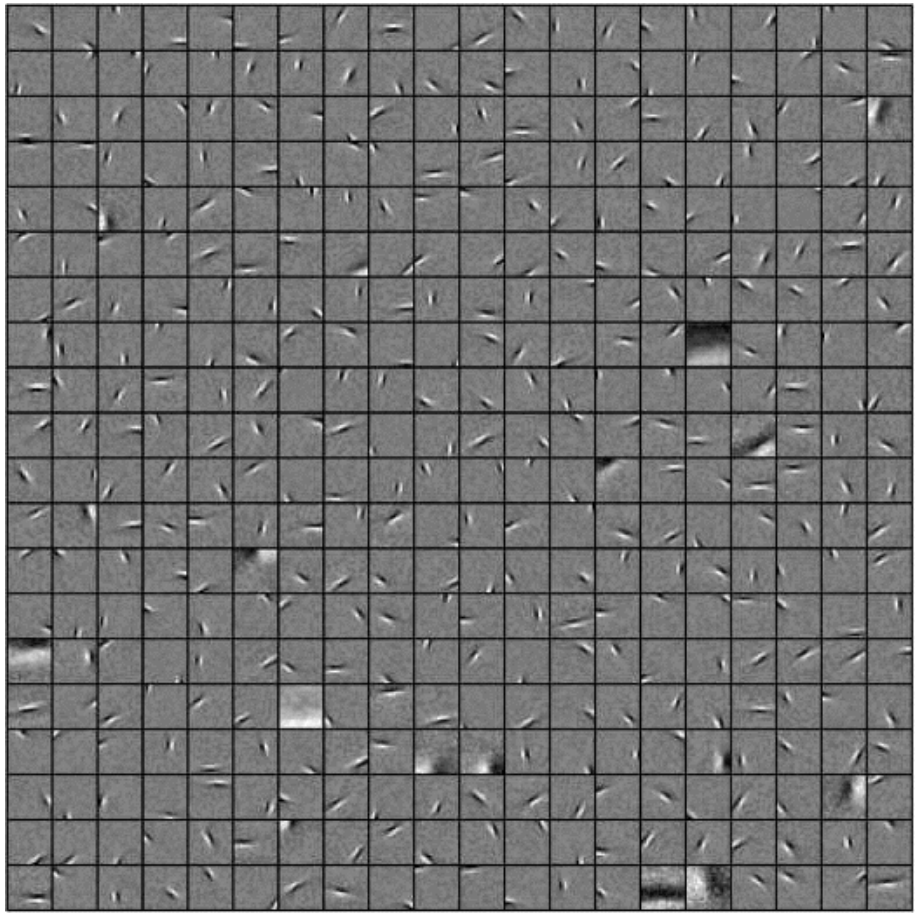
$$\min_{\Phi, \vec{\alpha}_i, i \in 1 \ldots N} \quad \frac{1}{N} \sum_{i=1}^{N} \|\vec{\beta}_i - \Phi \vec{\alpha}_i\|_2^2 + S(\vec{\alpha}_i)$$



(Olshausen & Field '96)

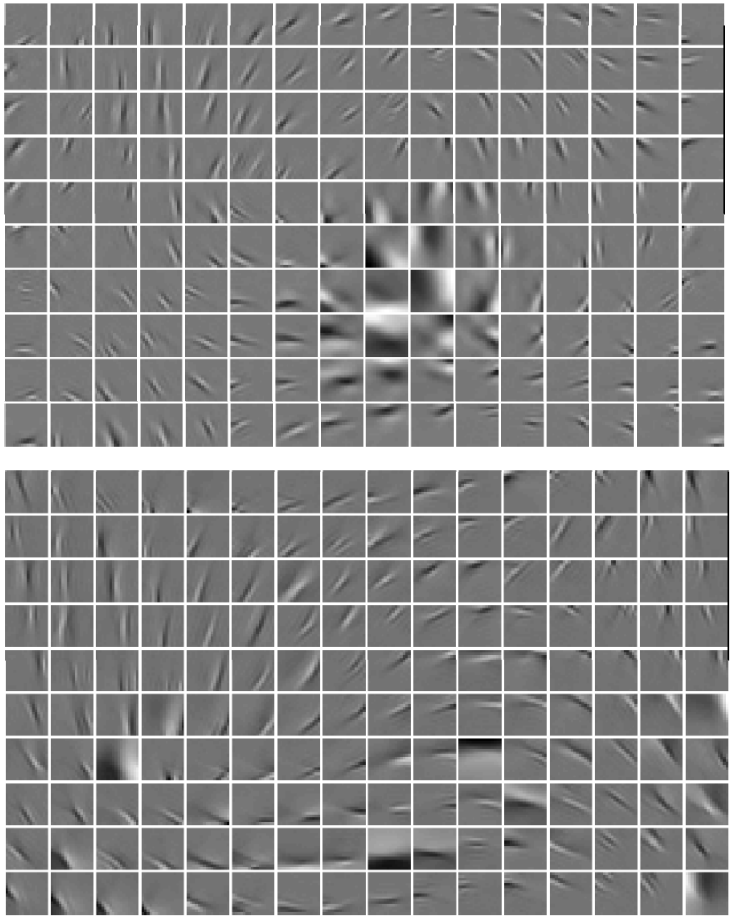
Topography and Subspace

Sparse Coding



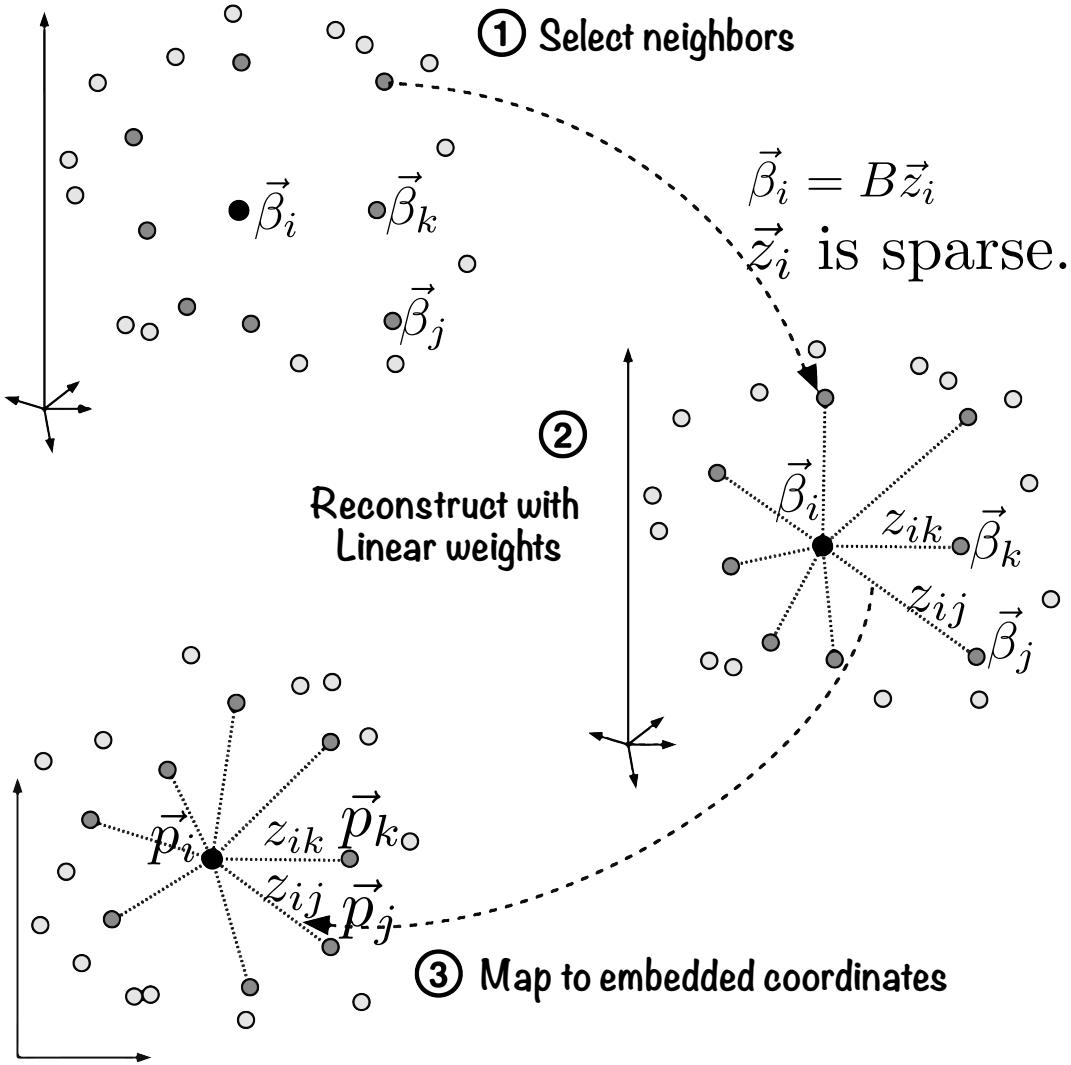
(Olshausen & Field ’96)

Topographic ICA



(Hyvarinen et al. , ‘01)

Manifold Learning, LLE



$$\vec{\beta}_i = B\vec{z}_i$$

$$\vec{z}_i \text{ is sparse. } \|\vec{z}_i\|_1 = 1$$

$$B = BZ \rightarrow B(I - Z) = 0$$

Goal: $P(I - Z) \approx 0$

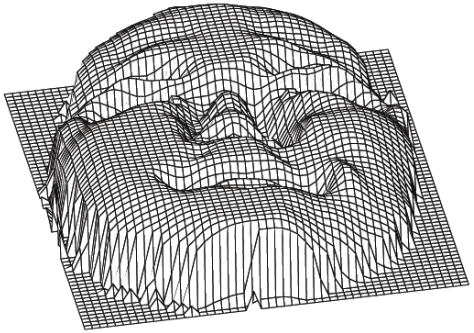
Opt:

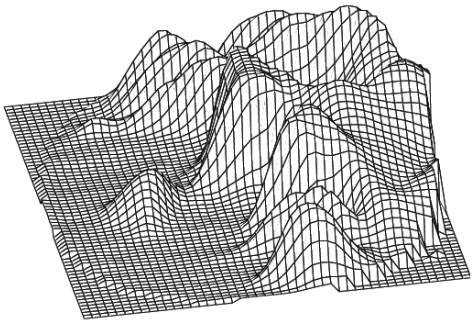
$$\min_P \quad tr\ P(I-Z)(I-Z)^T P^T$$

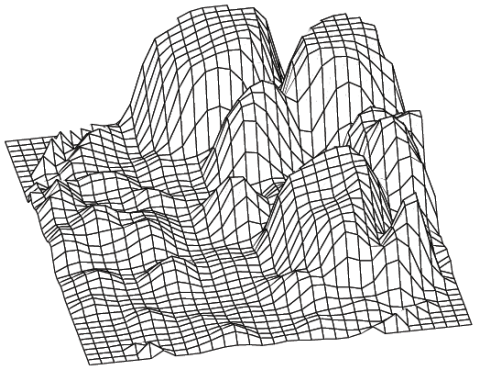
$$\text{s.t.} \quad PUP^T = I$$

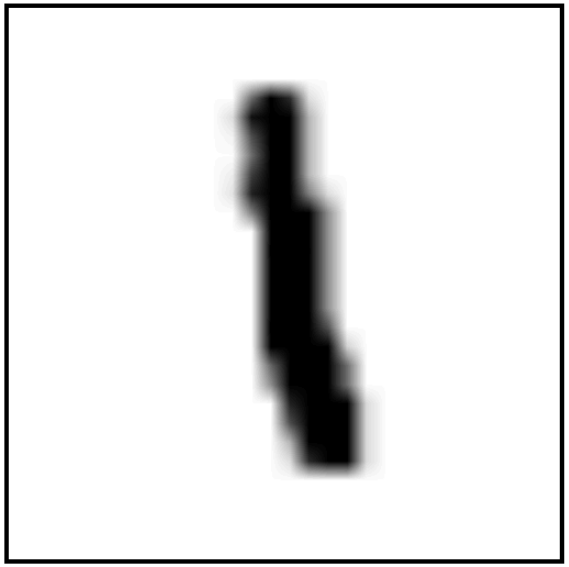
(Roweis & Saul, '00)

Images as Functions on $\mathbb{R}^2$



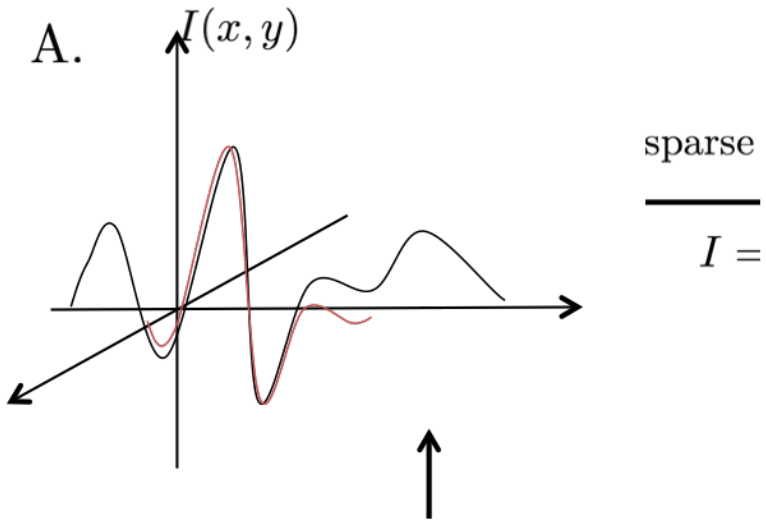






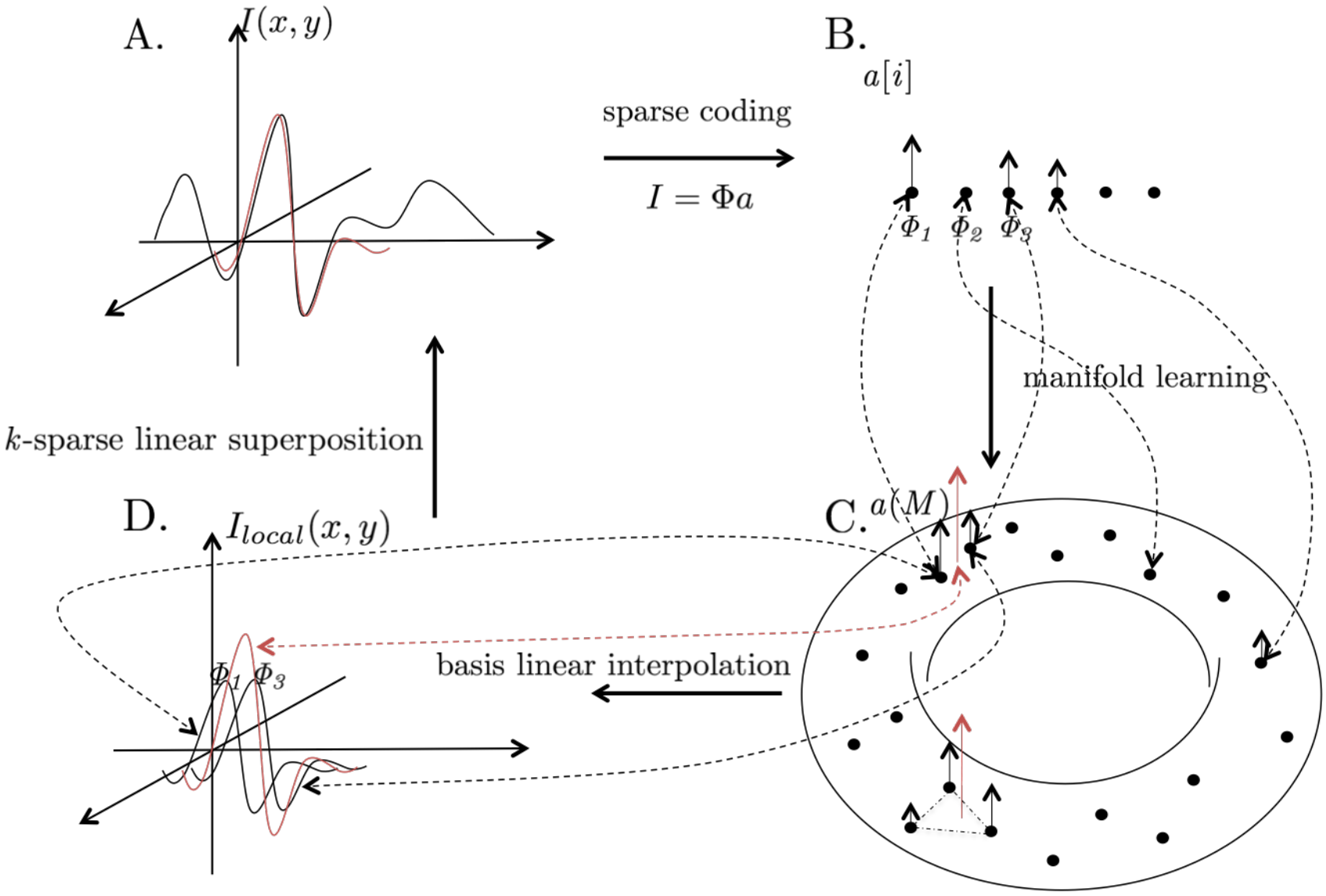
$$I \supseteq \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .6 & .8 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .5 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & .7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .9 & 1 & .1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .3 & 1 & .1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Dictionary Elements as New ‘Pixels’

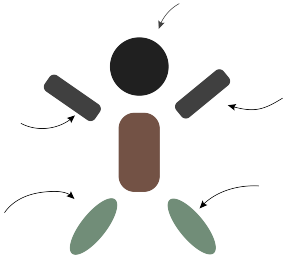


(The Sparse Manifold Transform, NIPS ‘18)

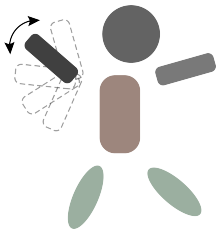
Dictionary Elements as New ‘Pixels’

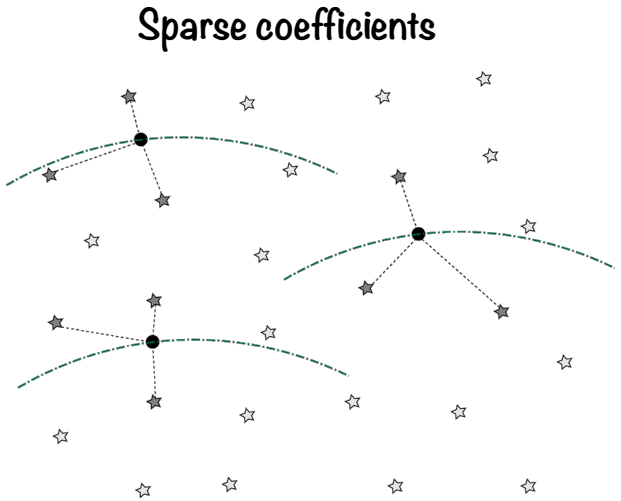


(The Sparse Manifold Transform, NIPS ‘18)



The Sparse Manifold Transform





$$\vec{\beta}_t = \Phi \vec{\alpha}_t$$

$$P\vec{\alpha}_t = \tfrac{1}{2} P(\vec{\alpha}_{t-1} + \vec{\alpha}_{t+1})$$

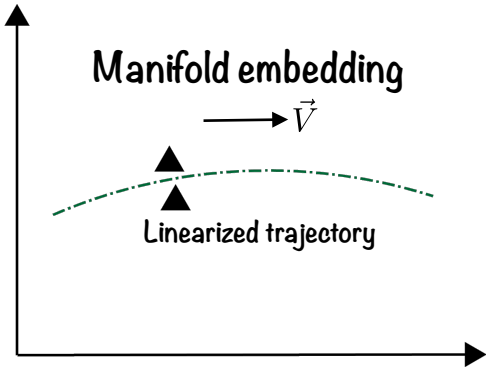
$$\text{Def: } \ddot{\vec{\alpha}}_t = \vec{\alpha}_t - \tfrac{1}{2}(\vec{\alpha}_{t-1} + \vec{\alpha}_{t+1})$$

Opt:

$$\min_P \quad tr\ P \sum_{t=1}^{T} (\ddot{\vec{\alpha}}_t \ddot{\vec{\alpha}}_t^T) P^T$$

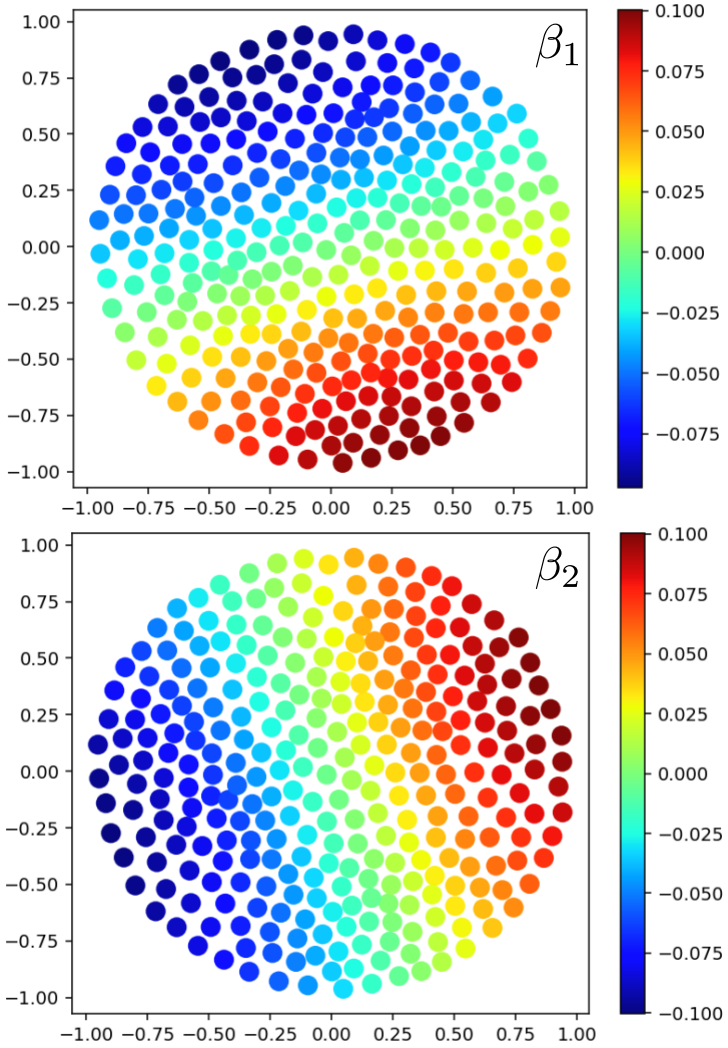
$$\text{s.t.} \quad PVP^T = I, \ (unit\ variance\ \&\ decorrelation)$$

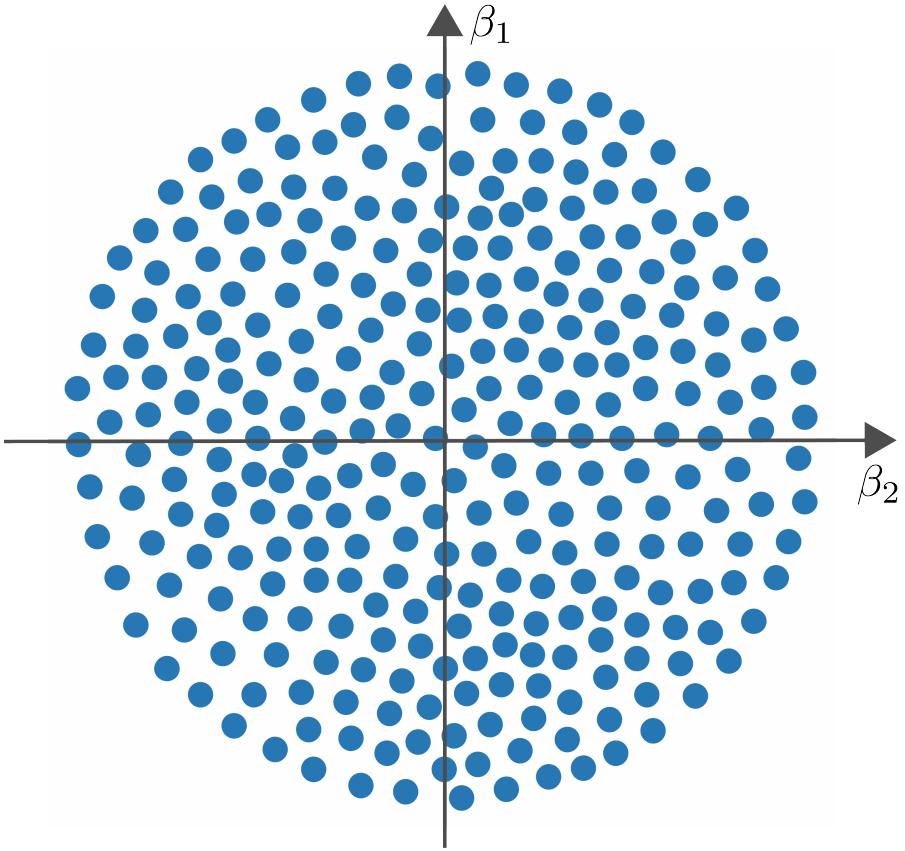
$$V\ is\ covariance\ matrix\ of\ \vec{\alpha}$$



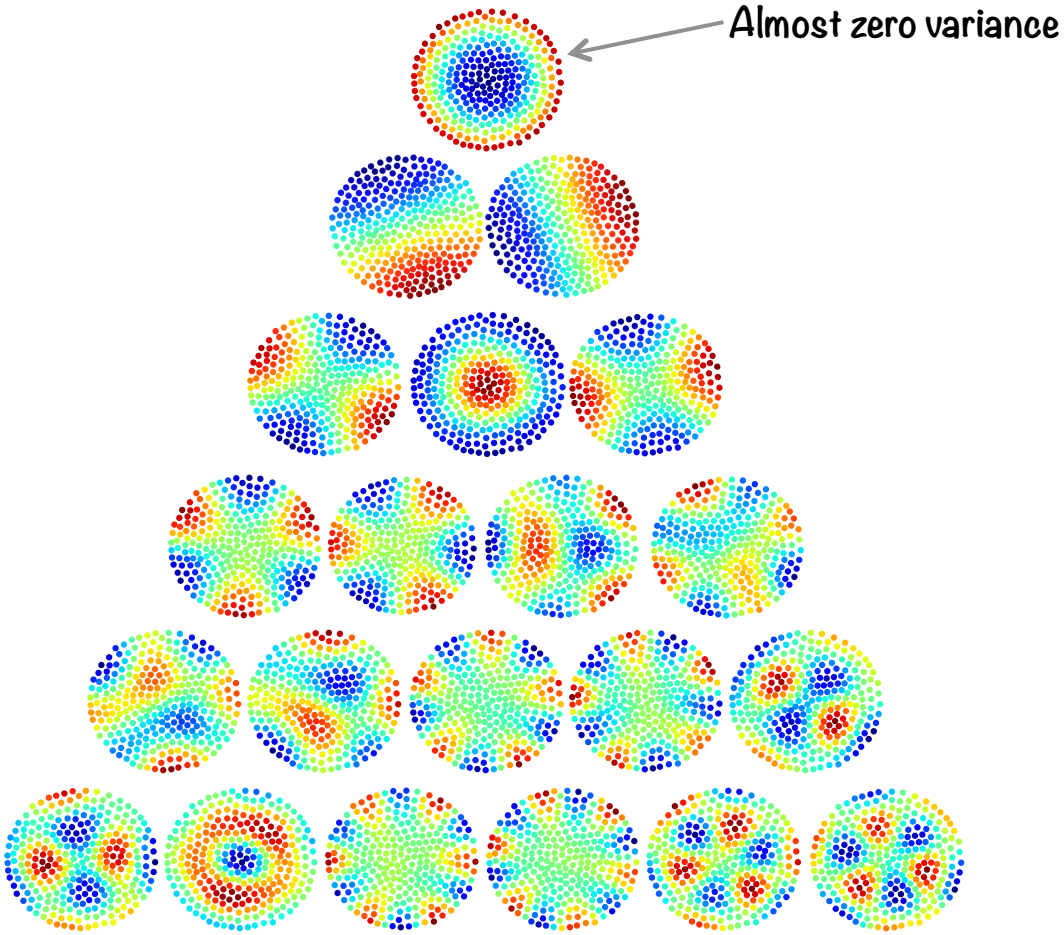
(The Sparse Manifold Transform, NIPS '18)

A Classical View of Manifold Learning



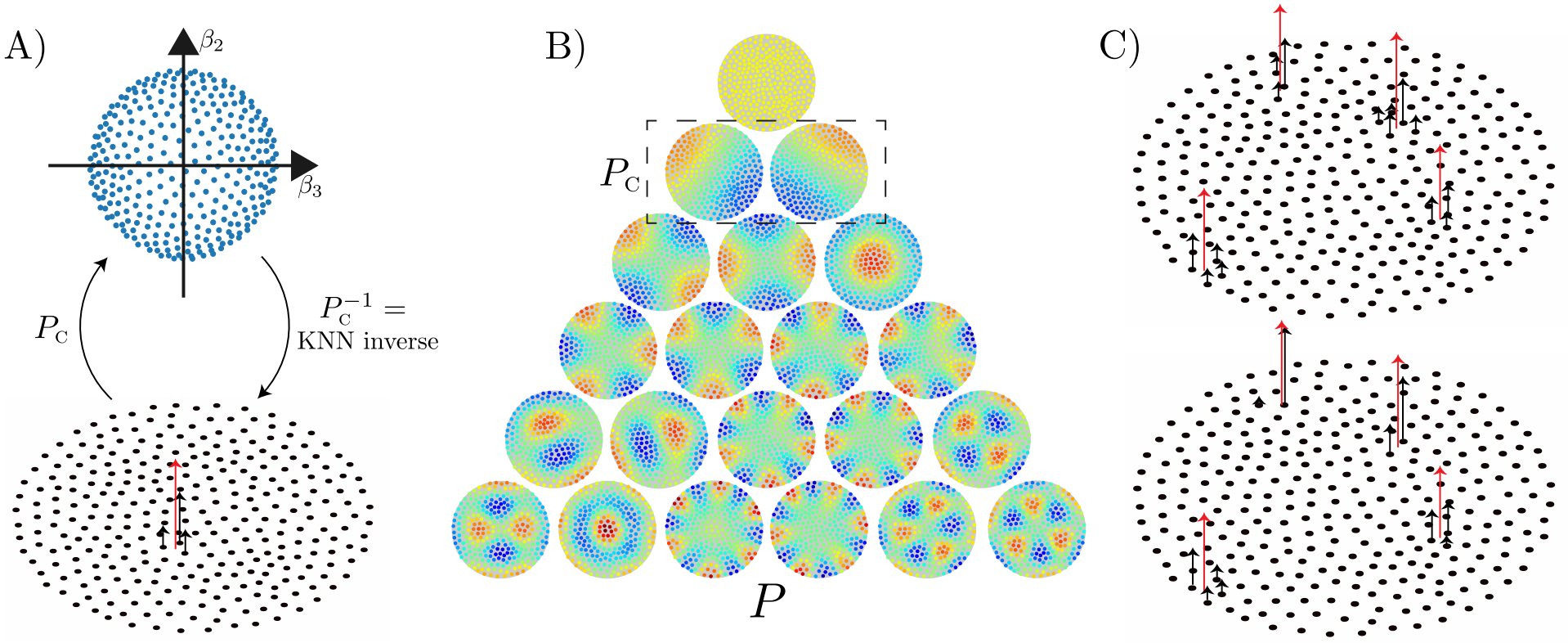


What if we keep going?



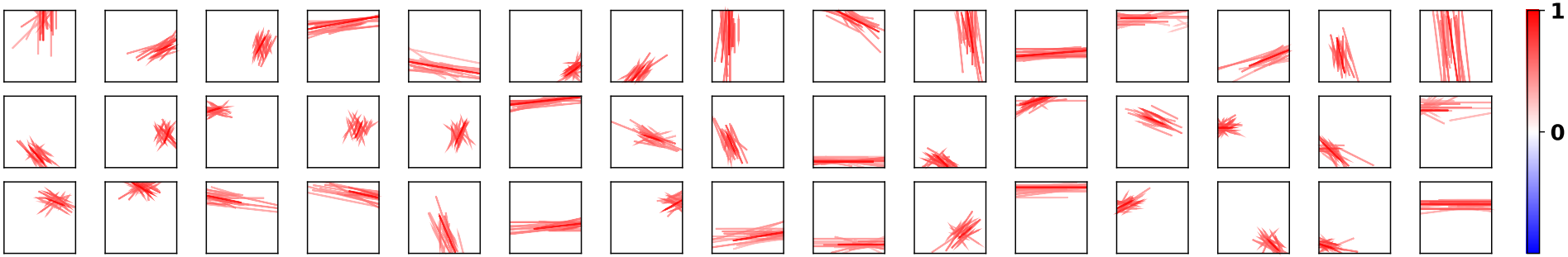
(The Sparse Manifold Transform, NIPS ‘18)

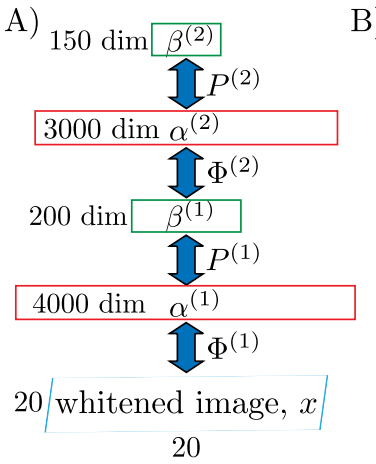
From Embedding to Sensing

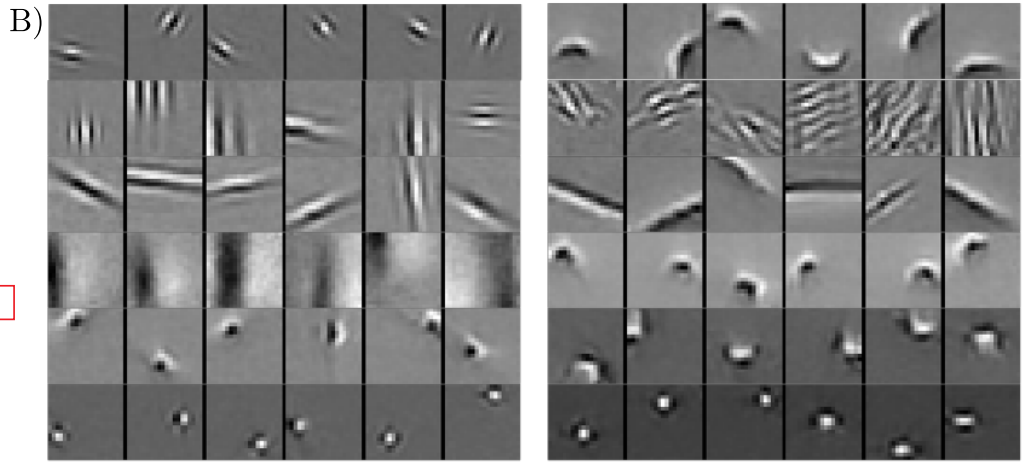


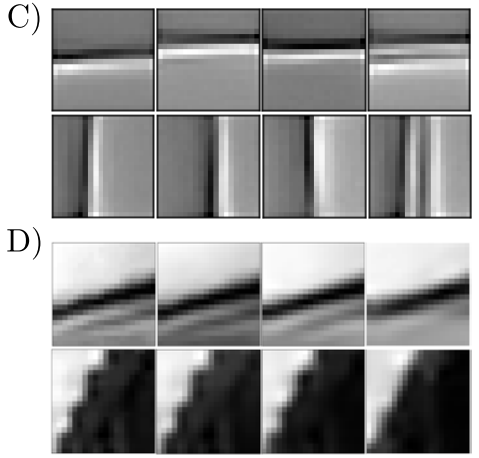
(The Sparse Manifold Transform, NIPS ‘18)

Groups and Hierarchy









(The Sparse Manifold Transform, NIPS ‘18)

Thanks!