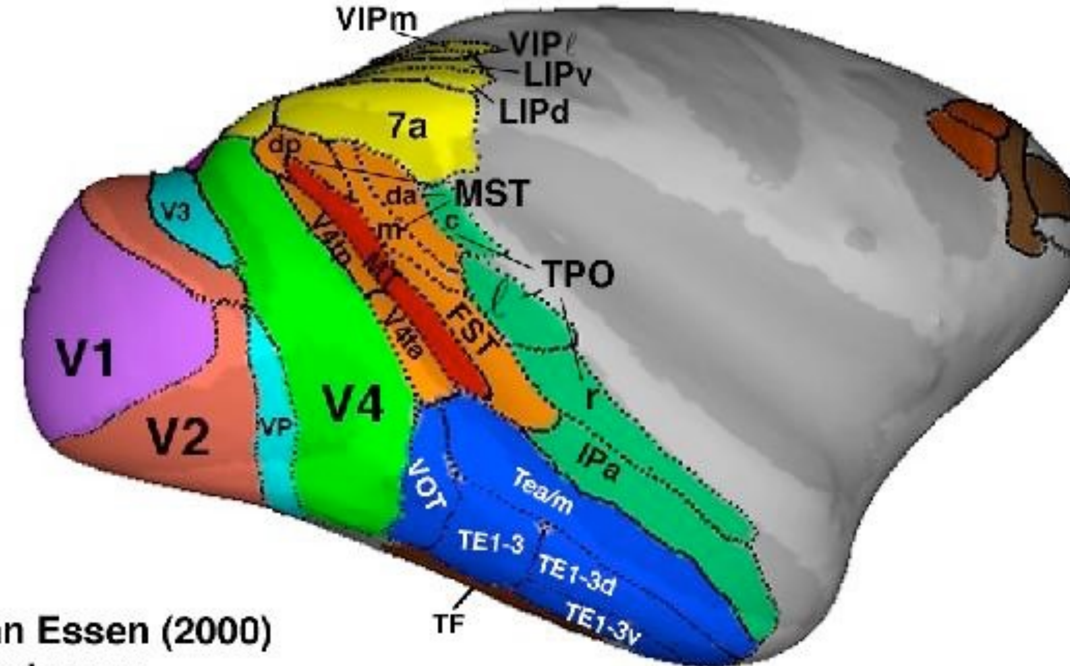
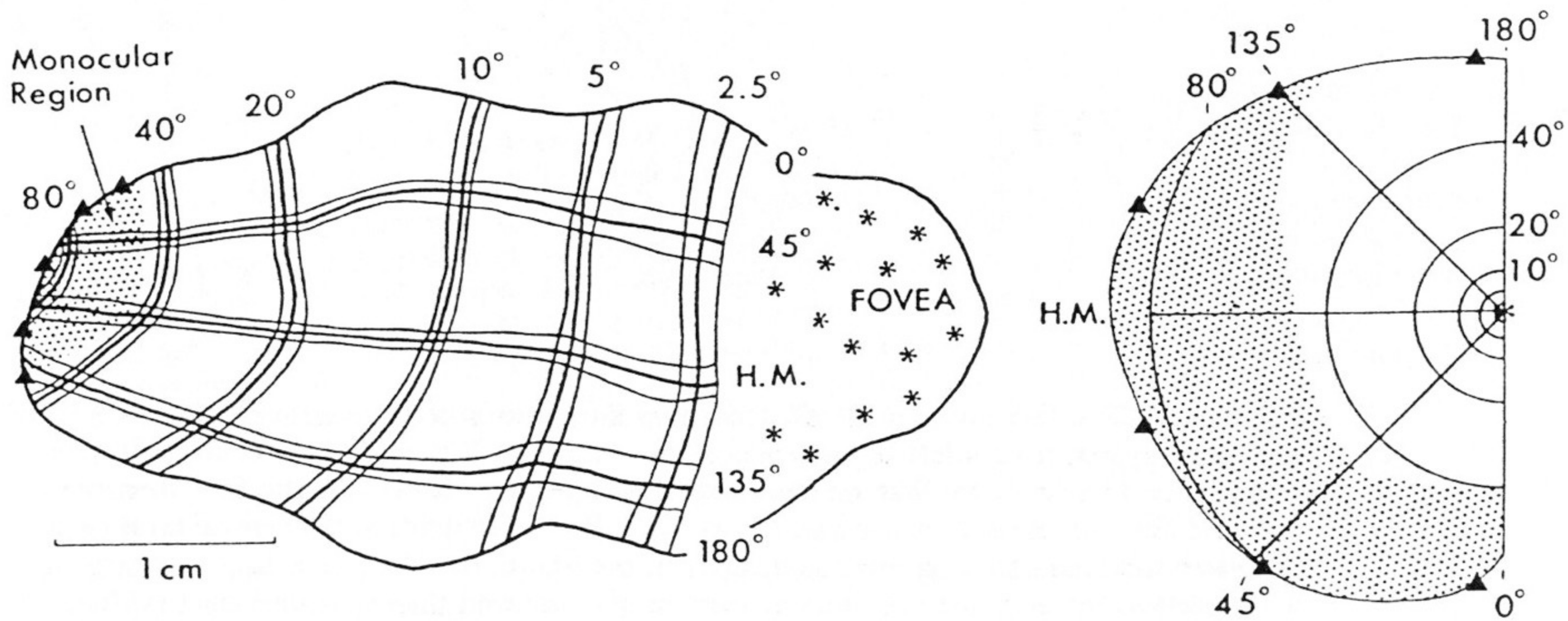
Visual cortex

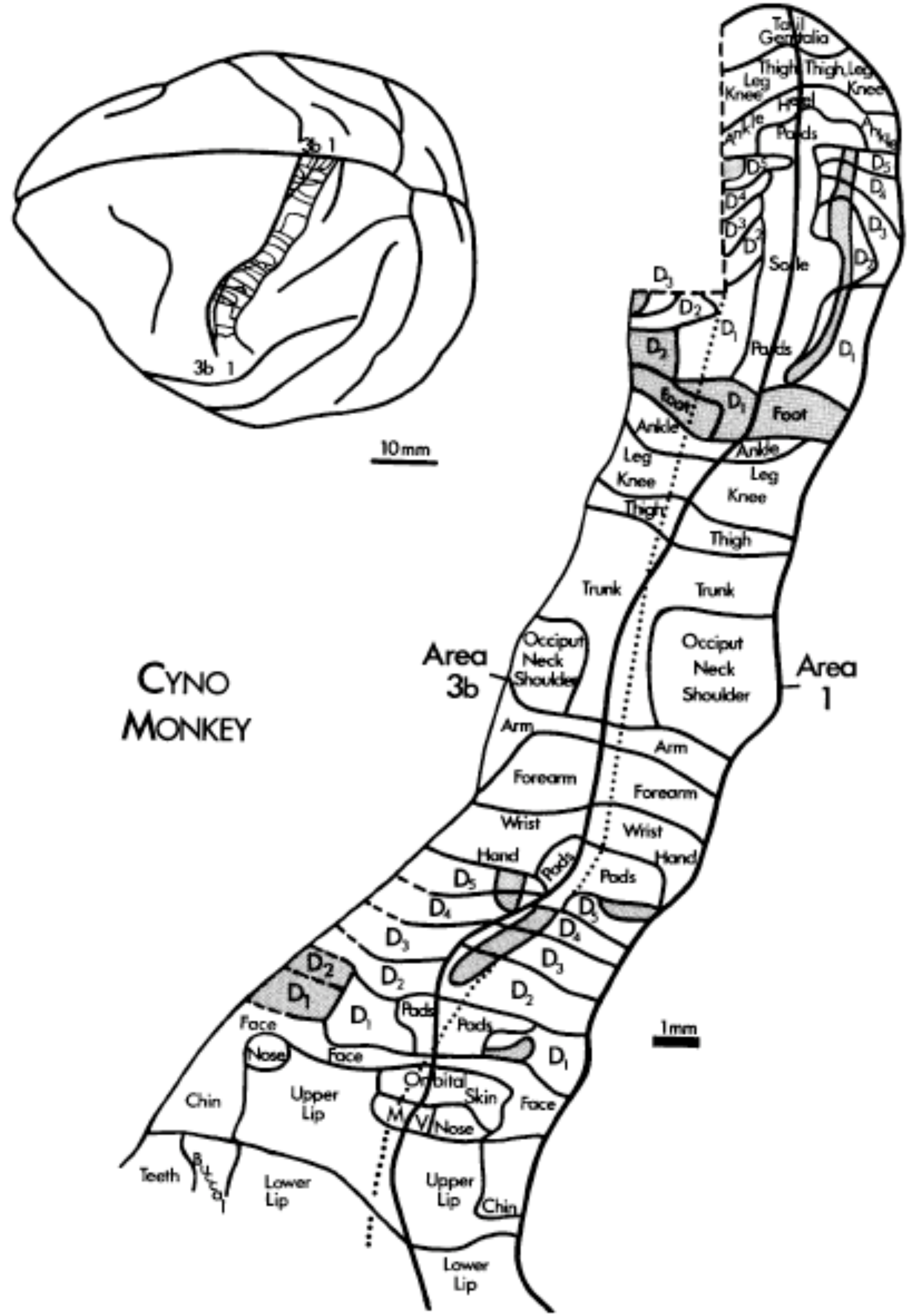


Lewis & Van Essen (2000)
Visual areas

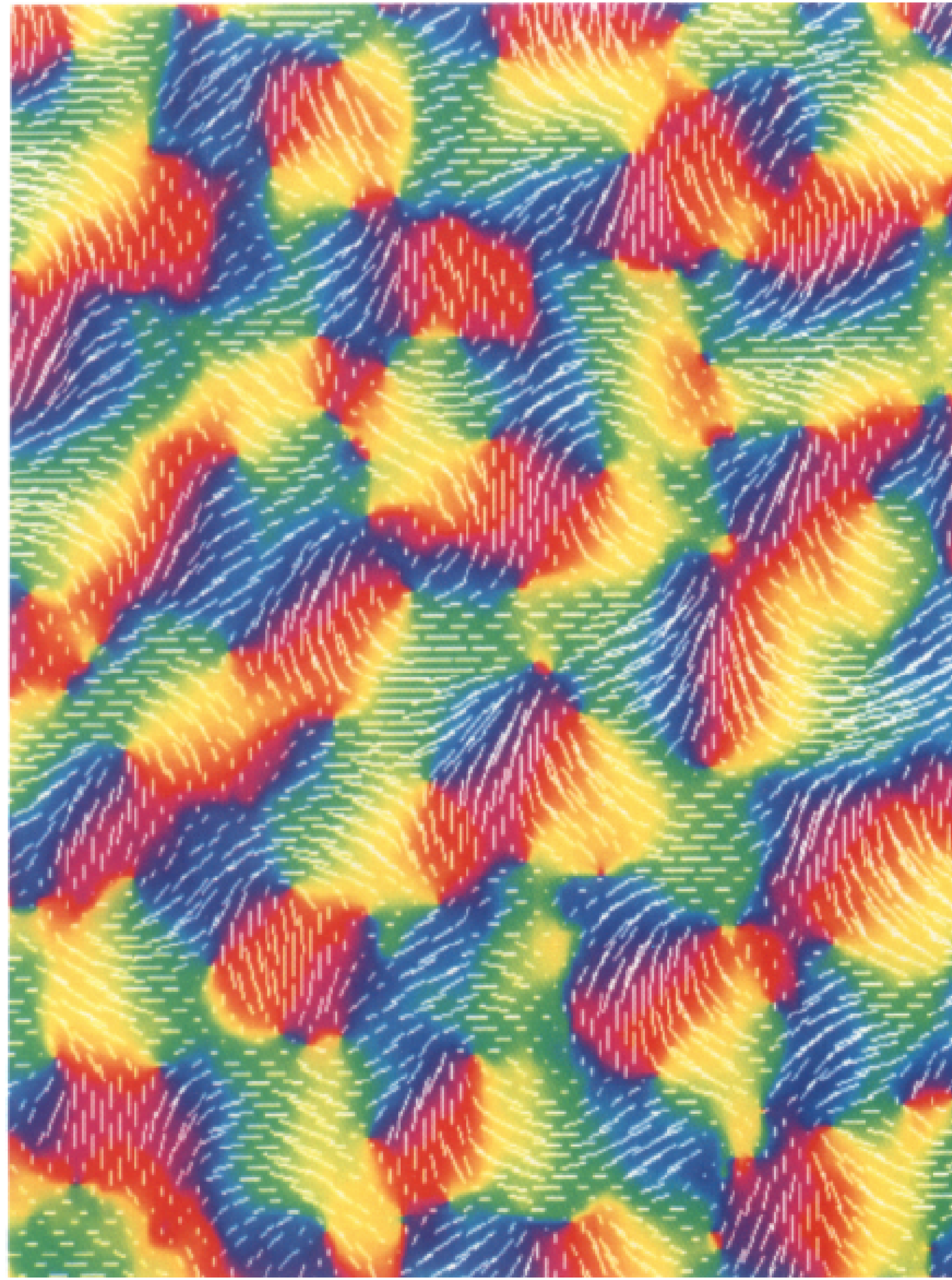
V1 topographic map



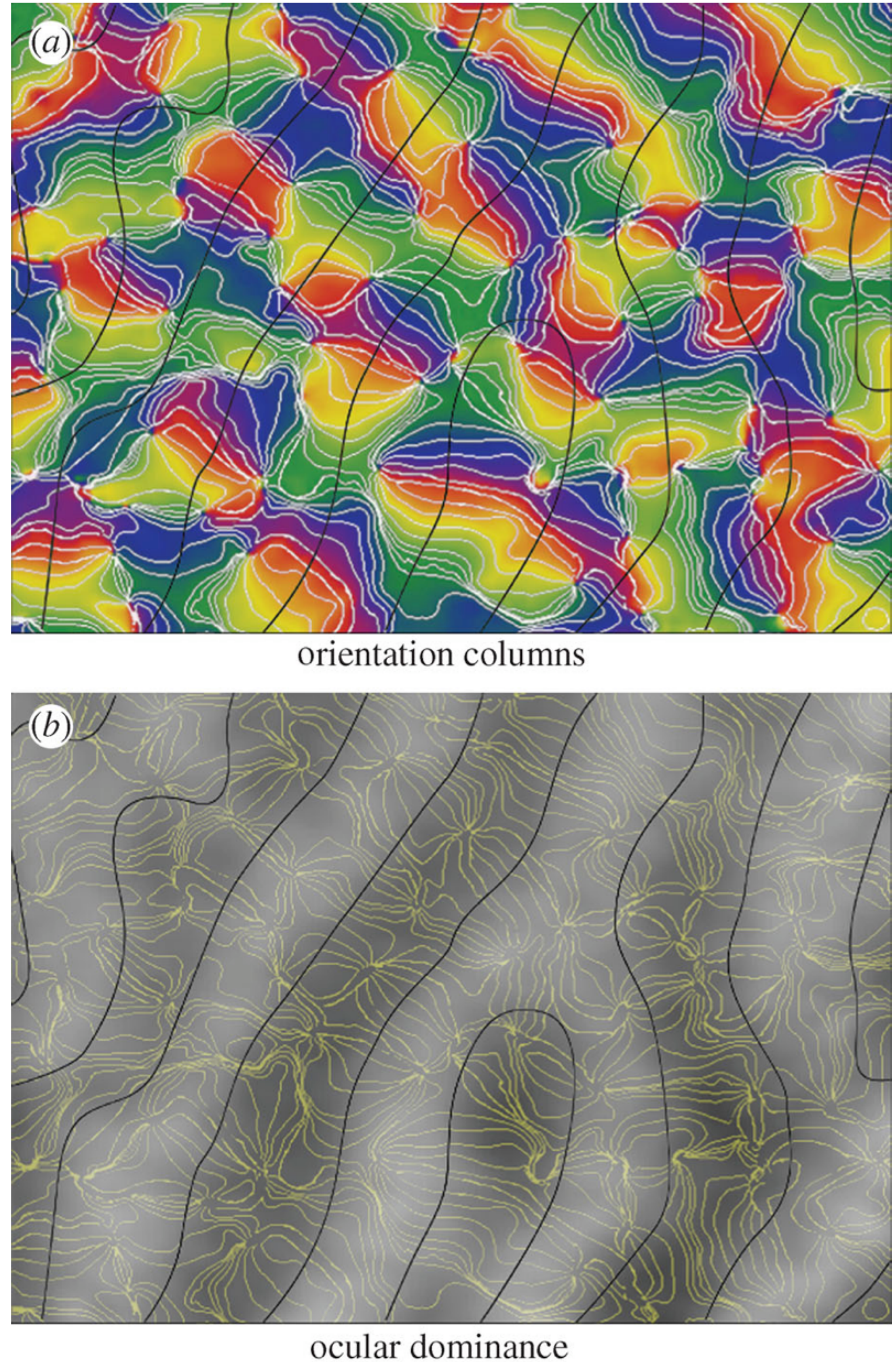
Somatosensory cortex



Orientation maps in V1 (Blasdel 1992)

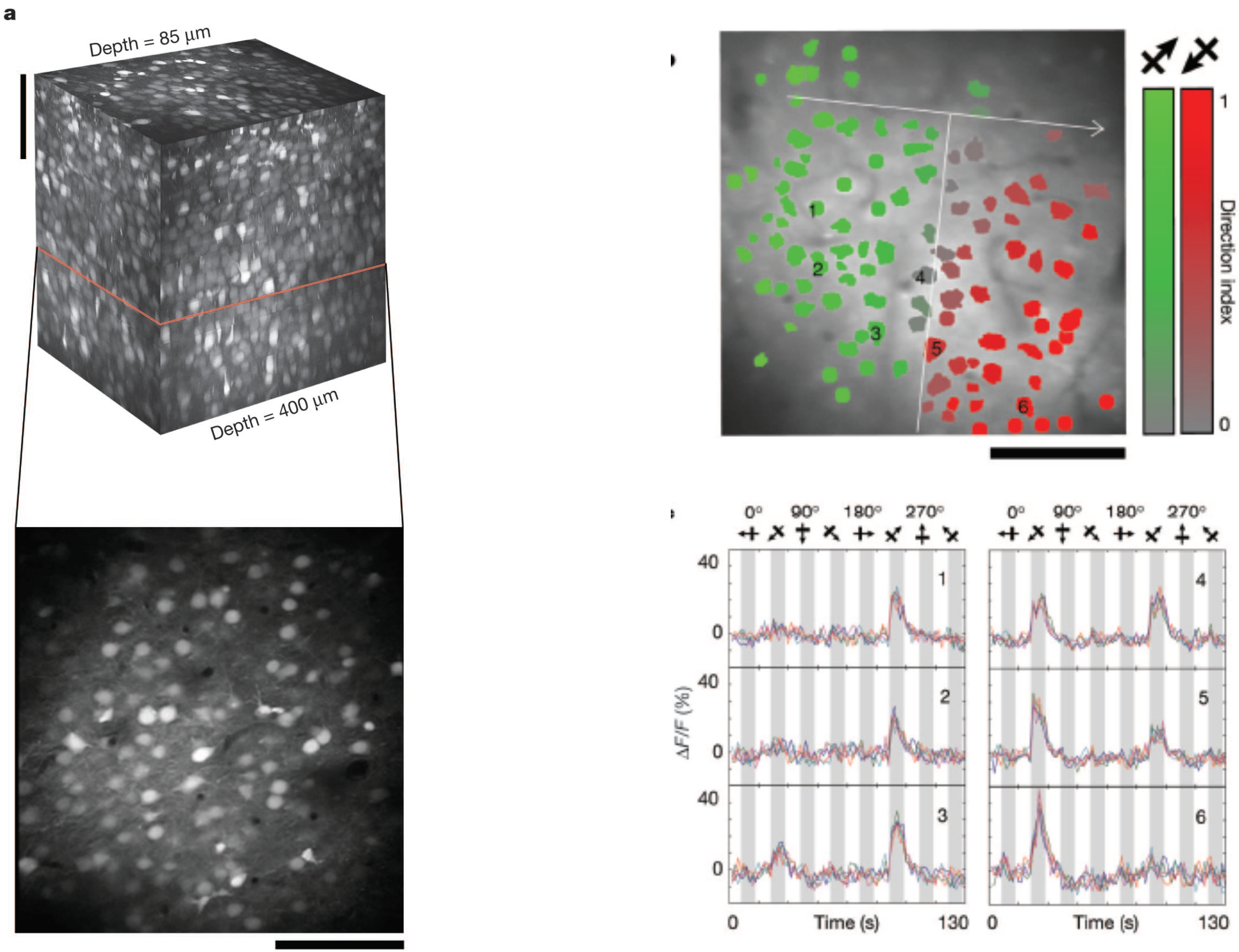


Joint map of orientation and ocular dominance demonstrates tradeoff between feature diversity vs. smoothness within each feature dimension (from Blasdel 1992)

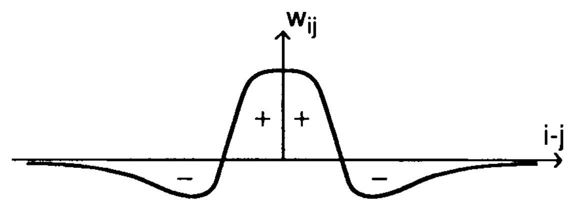


Direction selective cells in cat area 18

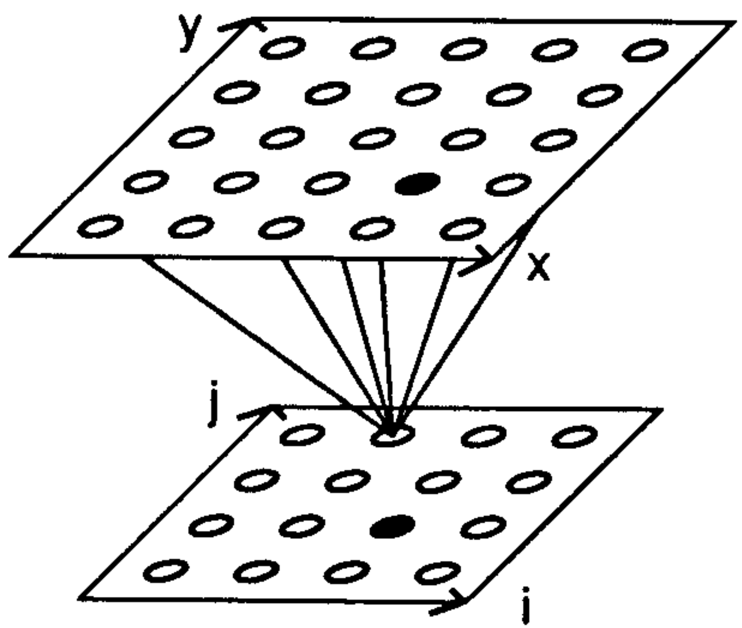
(Ca++ imaging - Ohki, Chung, Ch'ng, Kara, Reid, 2005)



Self-organizing maps



horizontal connections



output

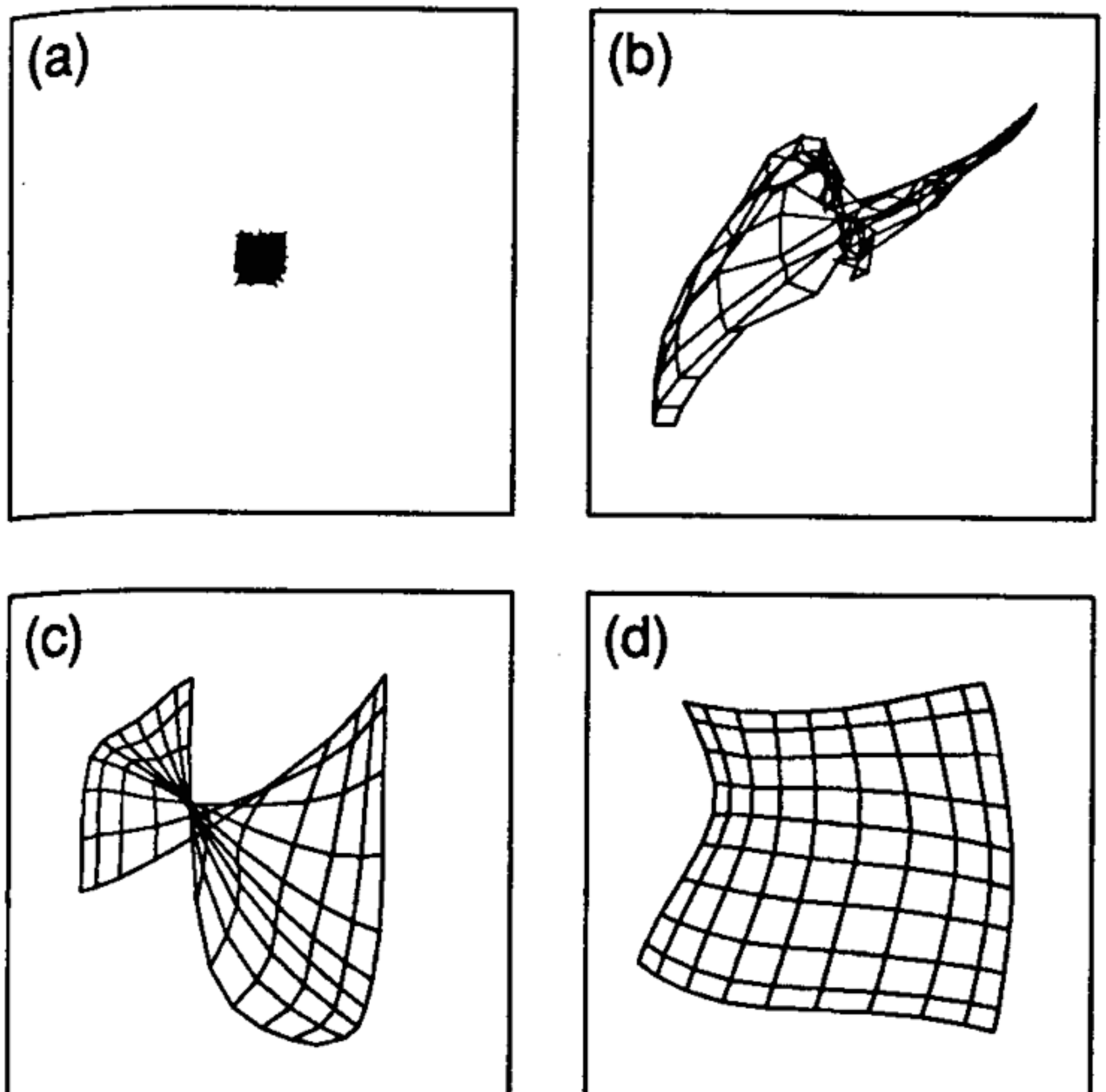
input

Kohonen's algorithm

$$\Delta \mathbf{w}_i = \eta \, \Lambda(i, i^*) \, (\mathbf{x} - \mathbf{w}_i)$$

$$\Lambda(i, i^*) = e^{-\frac{|\mathbf{r}_i - \mathbf{r}_{i*}|^2}{2\sigma^2}}$$

Kohonen's algorithm applied to a 2D input array



What is this organization good for?

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2005) **360**, 837–862
doi:10.1098/rstb.2005.1623
Published online 29 April 2005

The cortical column: a structure without a function

Jonathan C. Horton* and Daniel L. Adams

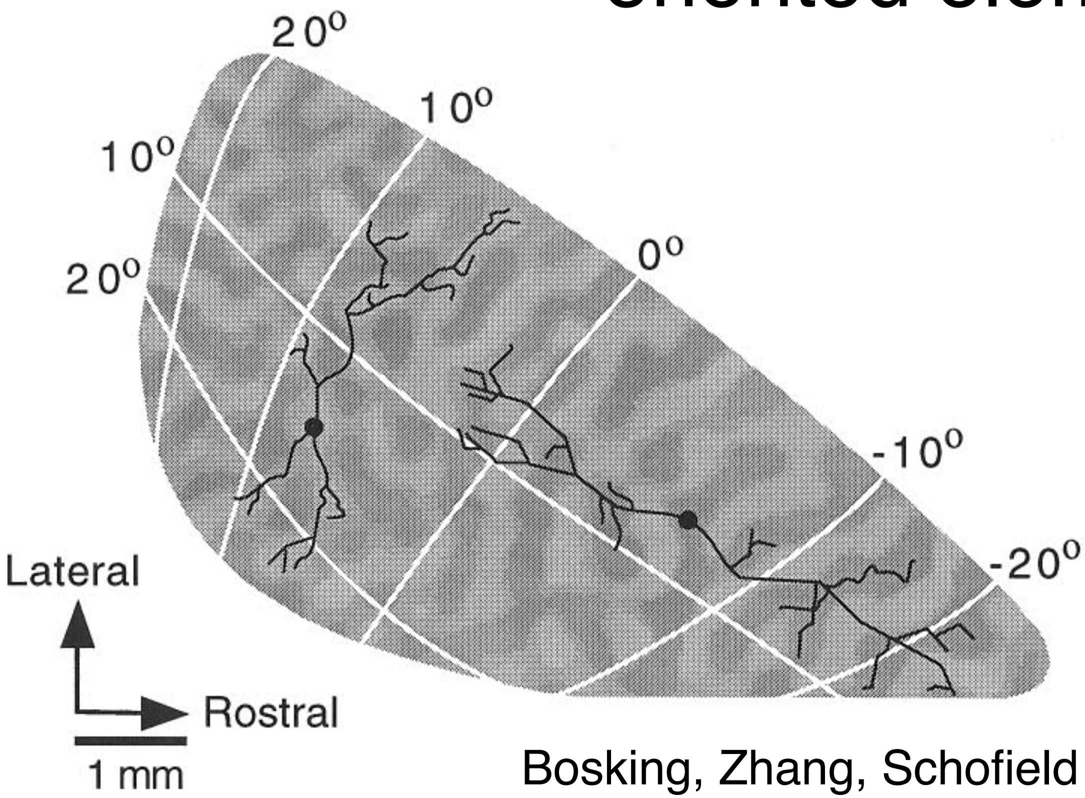
Beckman Vision Center, 10 Koret Way, University of California, San Francisco, CA 94143-0730, USA

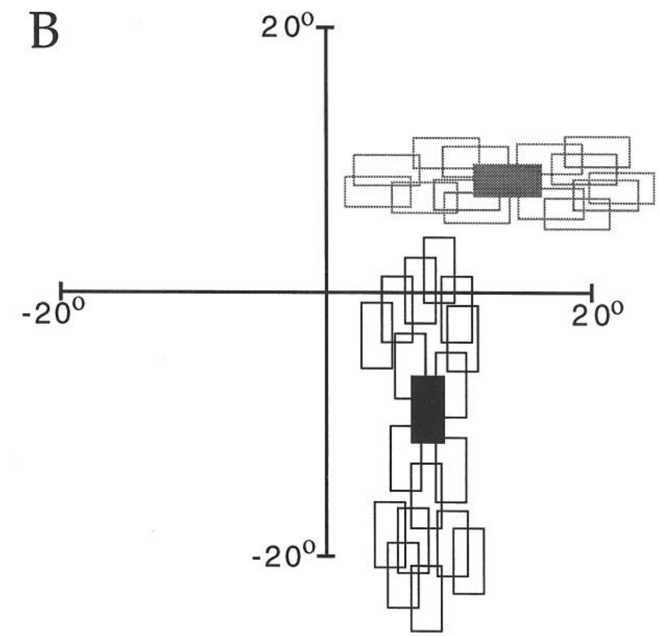
This year, the field of neuroscience celebrates the 50th anniversary of Mountcastle's discovery of the

Although the column is an attractive concept, it has failed as a unifying principle for understanding cortical function. *Unravelling the organization of the cerebral cortex will require a painstaking description of the circuits, projections and response properties peculiar to cells in each of its various areas*.

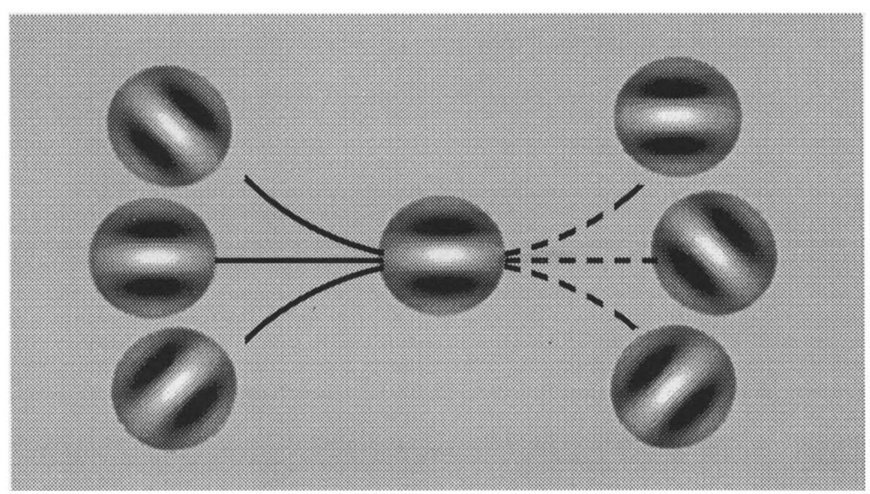
function. Unravelling the organization of the cerebral cortex will require a painstaking description of the circuits, projections and response properties peculiar to cells in each of its various areas.

Horizontal connections may enforce continuity among oriented elements



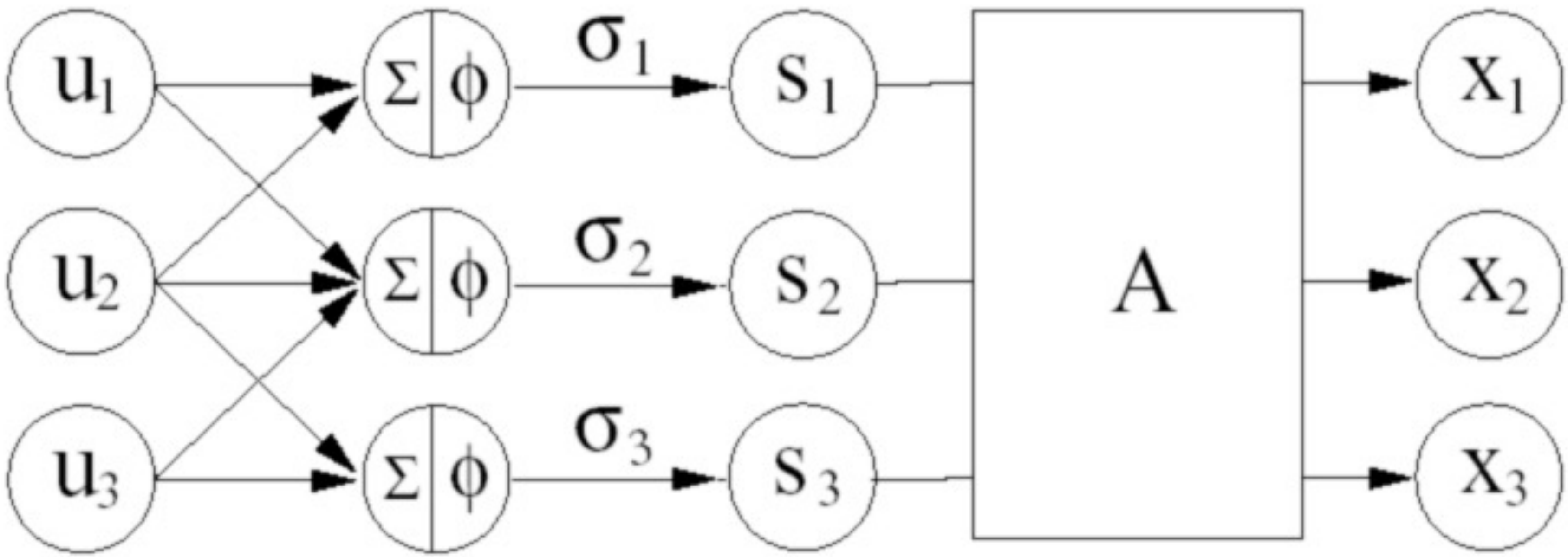


Bosking, Zhang, Schofield & Fitzpatrick (1993)

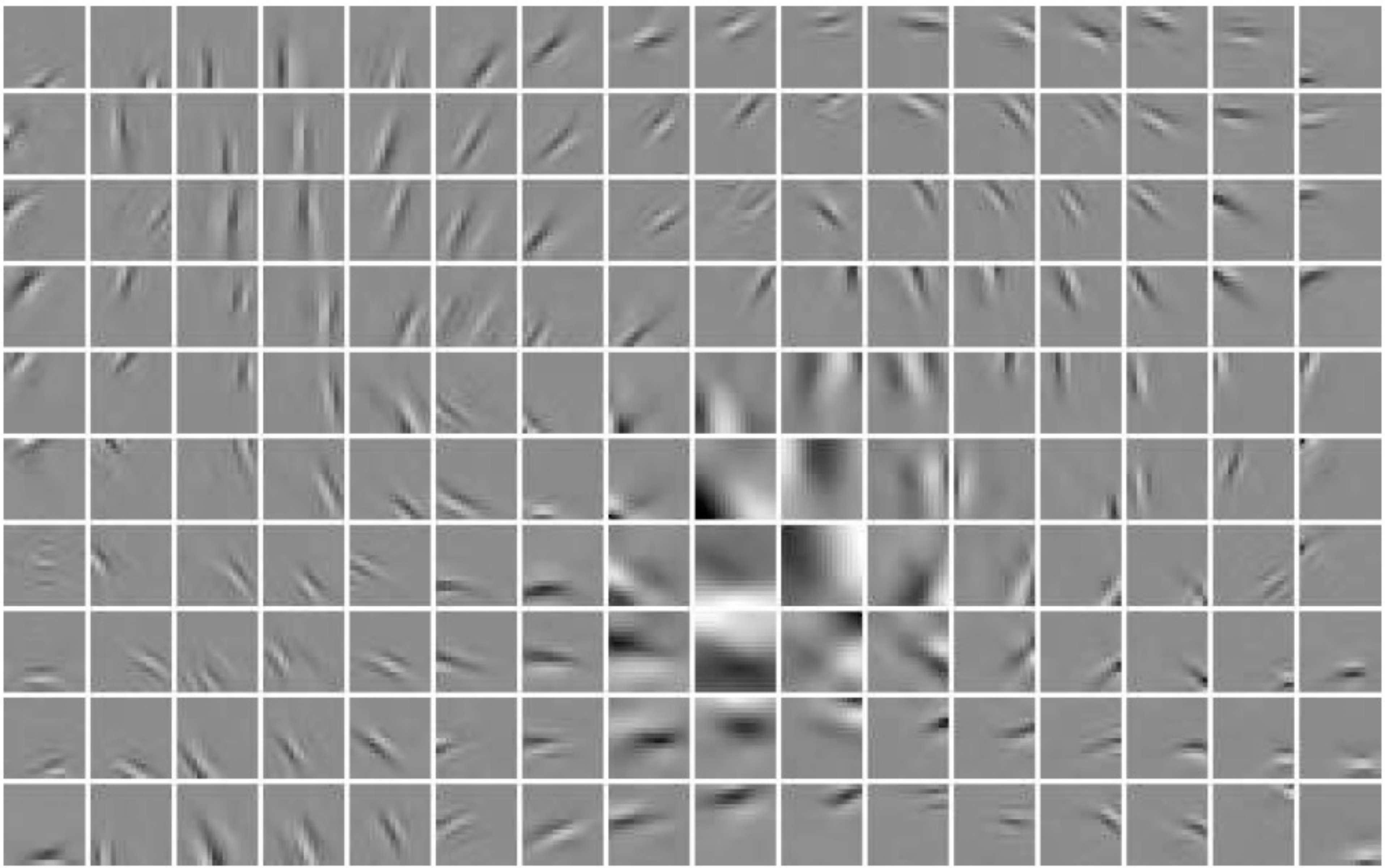


Field, Hayes & Hess (1993)

‘Topographic ICA’
(Hyvarinen & Hoyer)

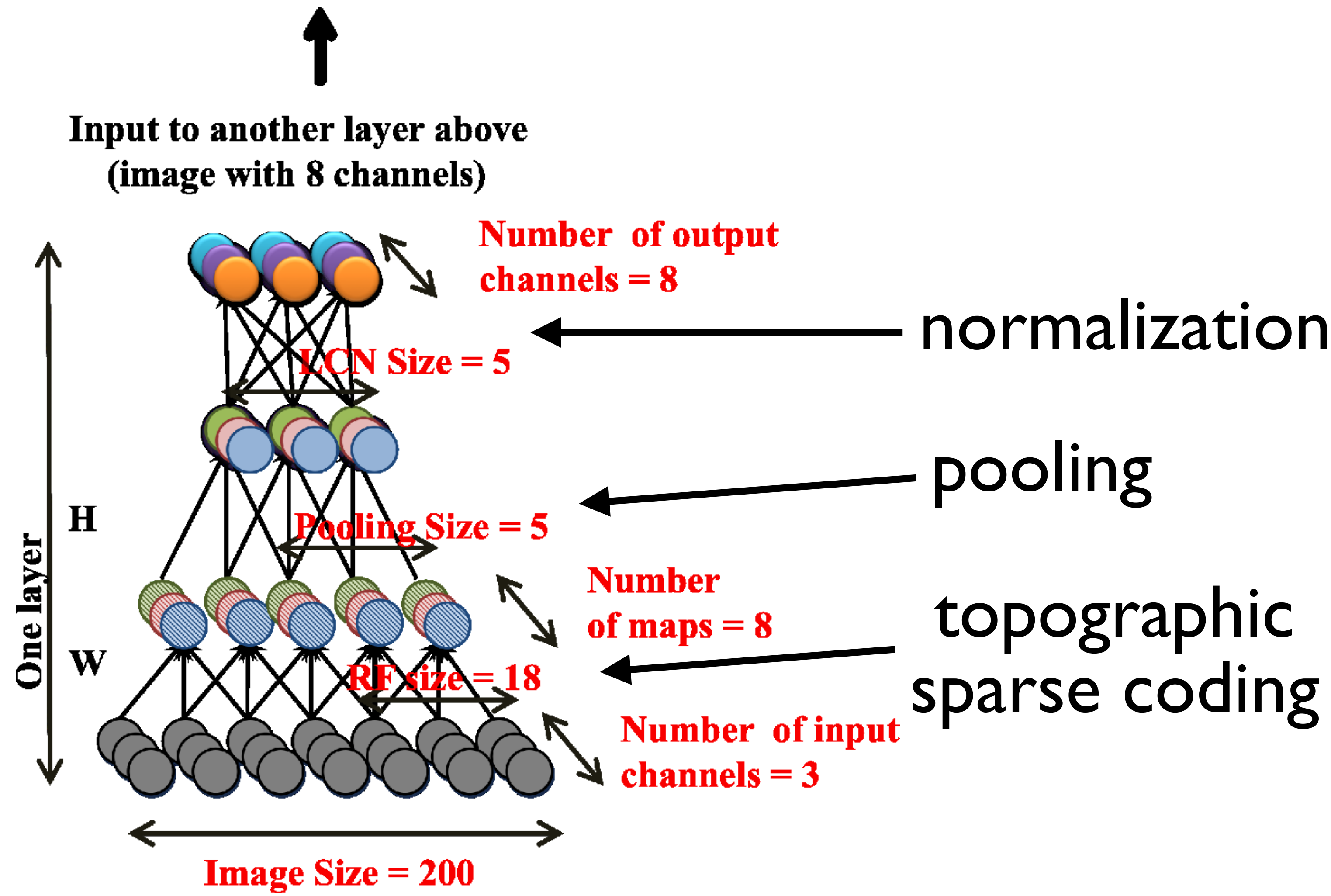


Topographic ICA/Sparse coding
(Hyvarinen & Hoyer)



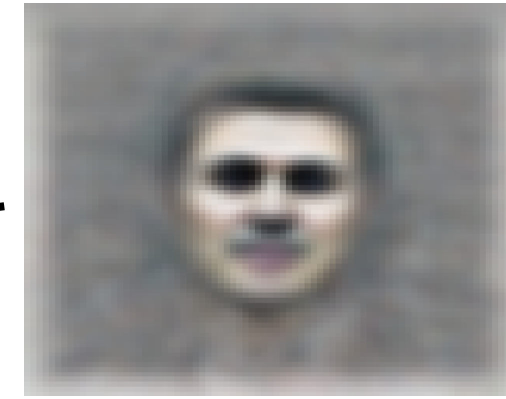
‘Google Brain’

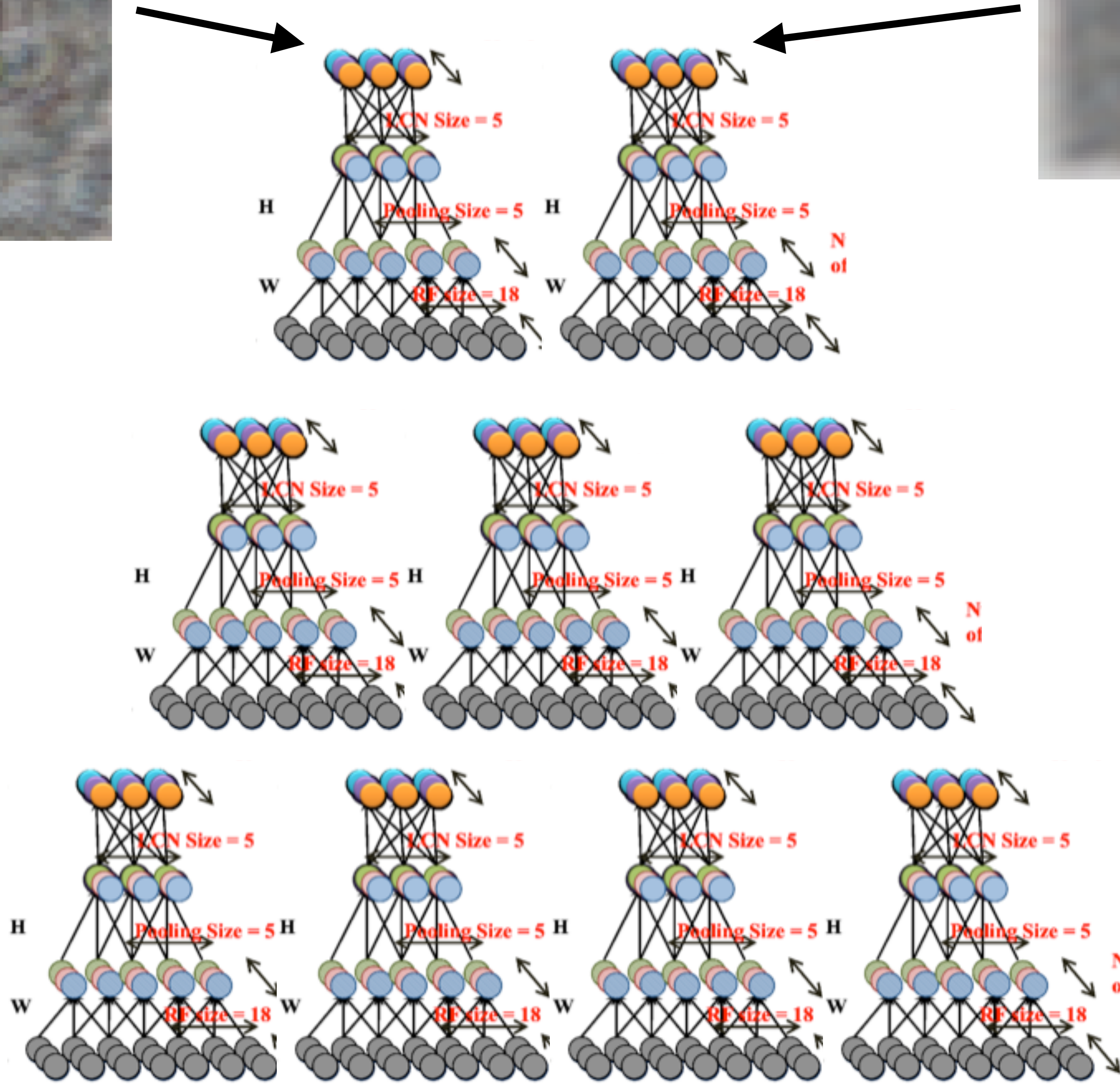
(Quoc Le et al. 2012)



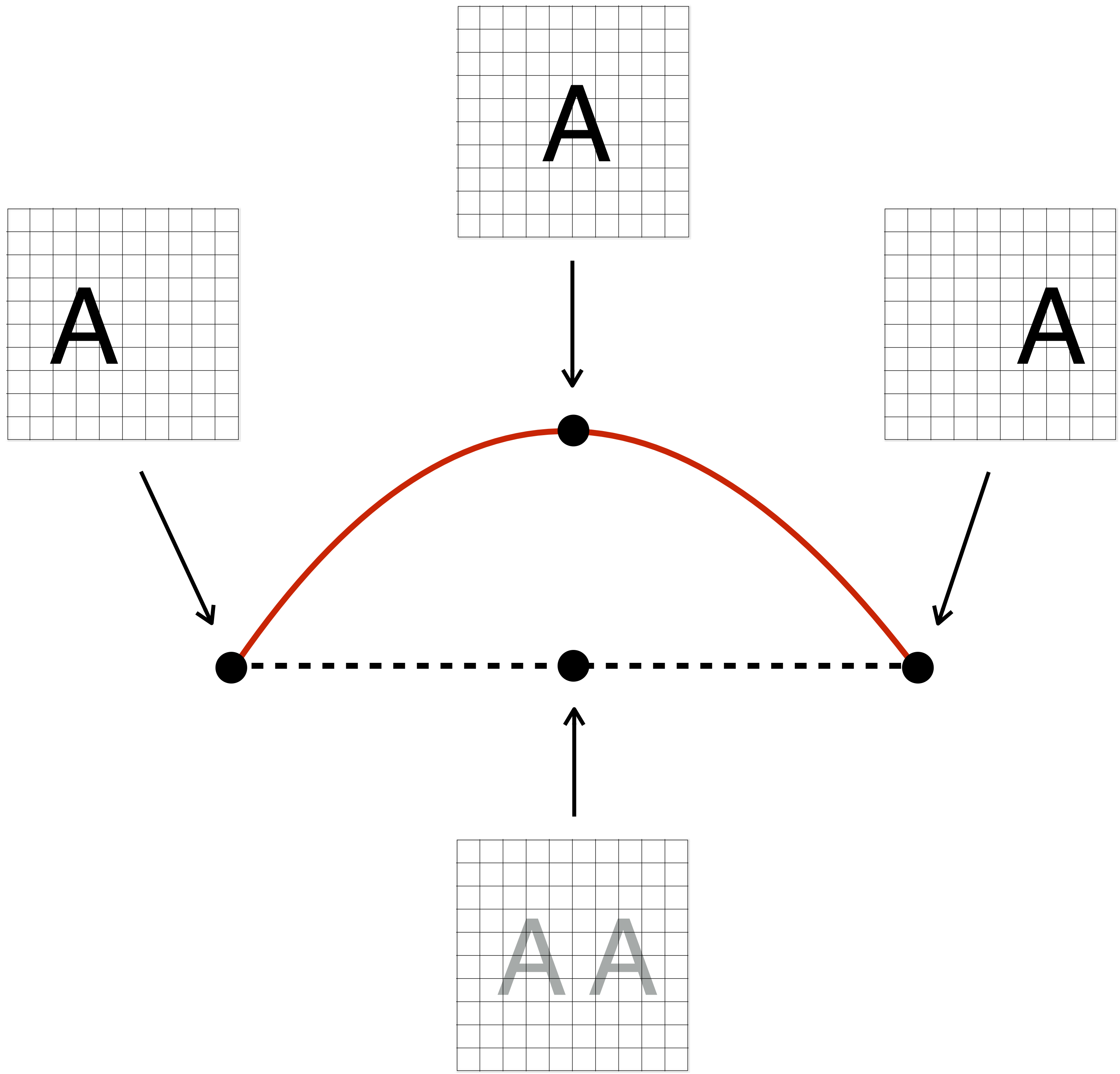
‘Google Brain’

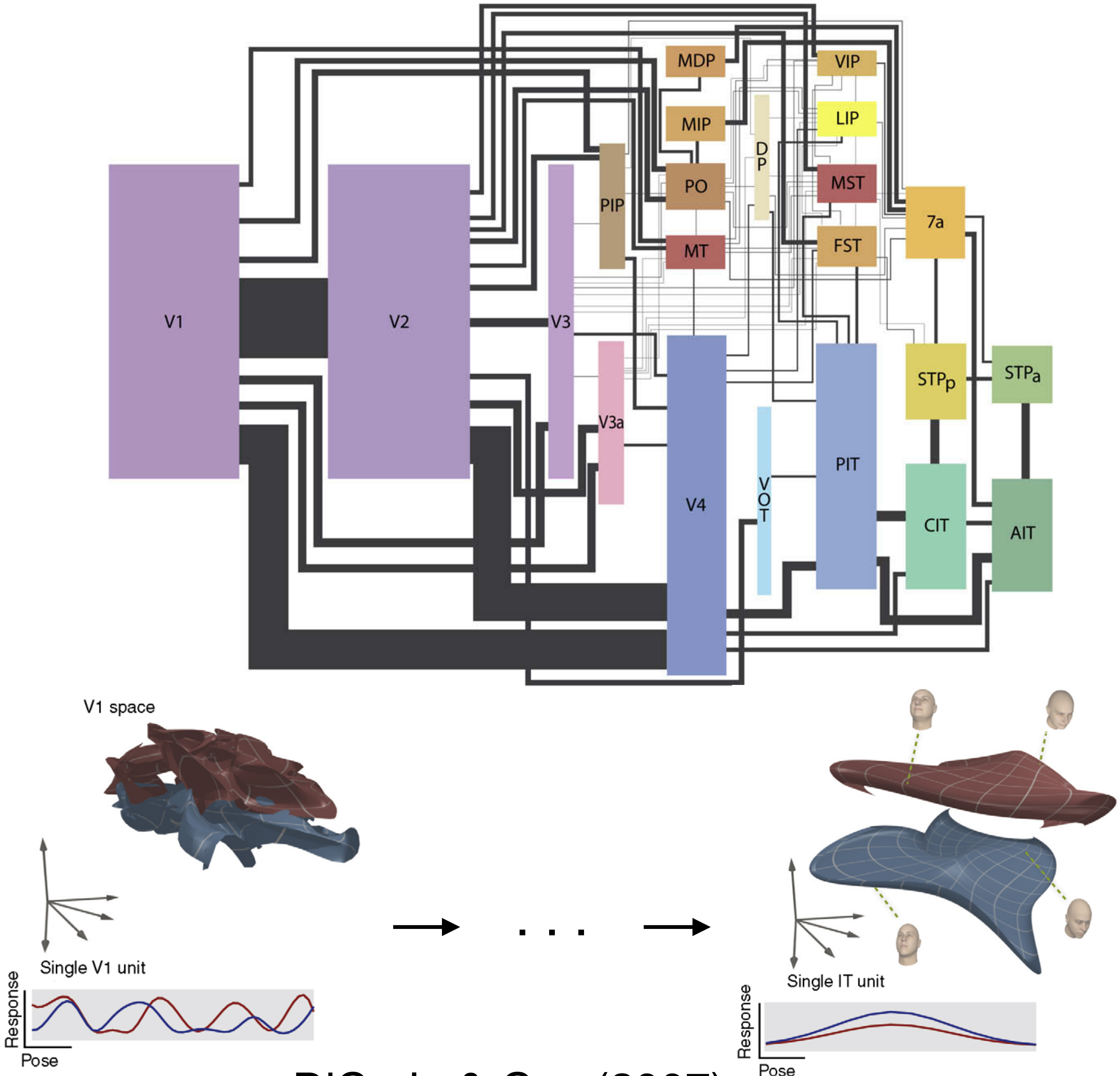
(Quoc Le et al. 2012)





Manifolds


A
A
A
A A



DiCarlo & Cox (2007)

(An aside on adaptation)

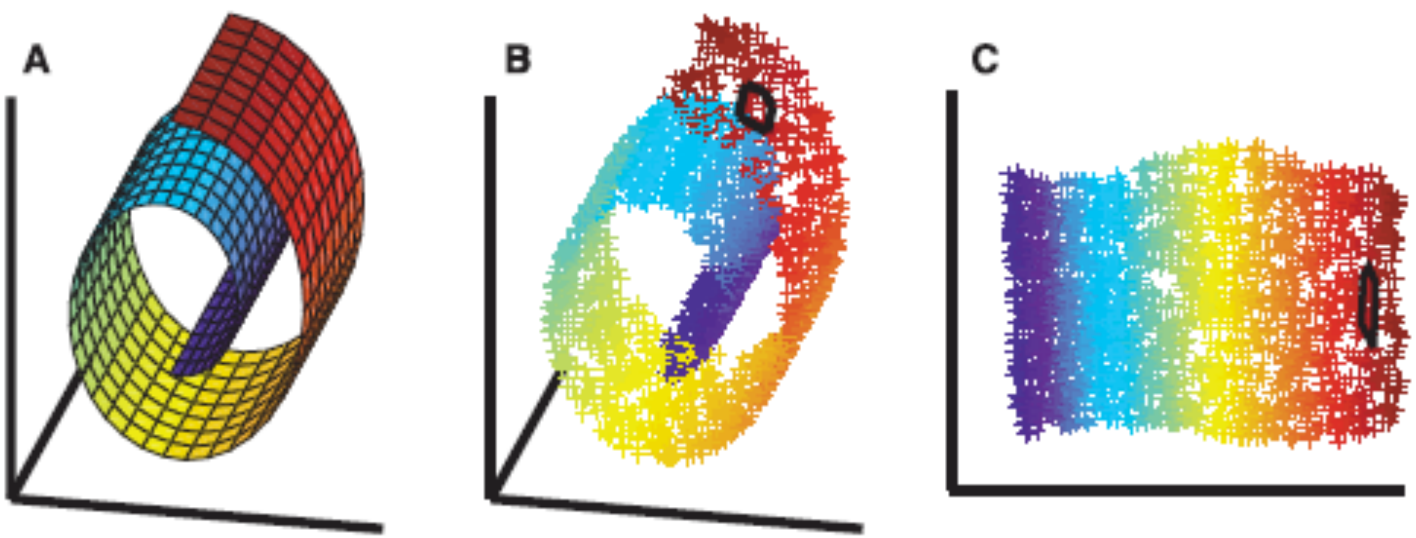
Nonlinear Dimensionality Reduction by Locally Linear Embedding

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

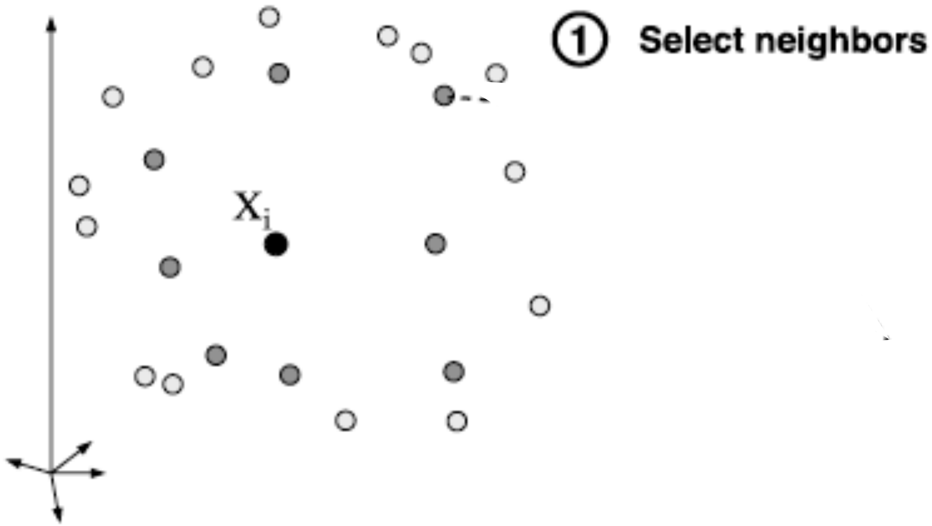
A Global Geometric Framework for Nonlinear Dimensionality Reduction

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

Science, 22 Dec. 2000



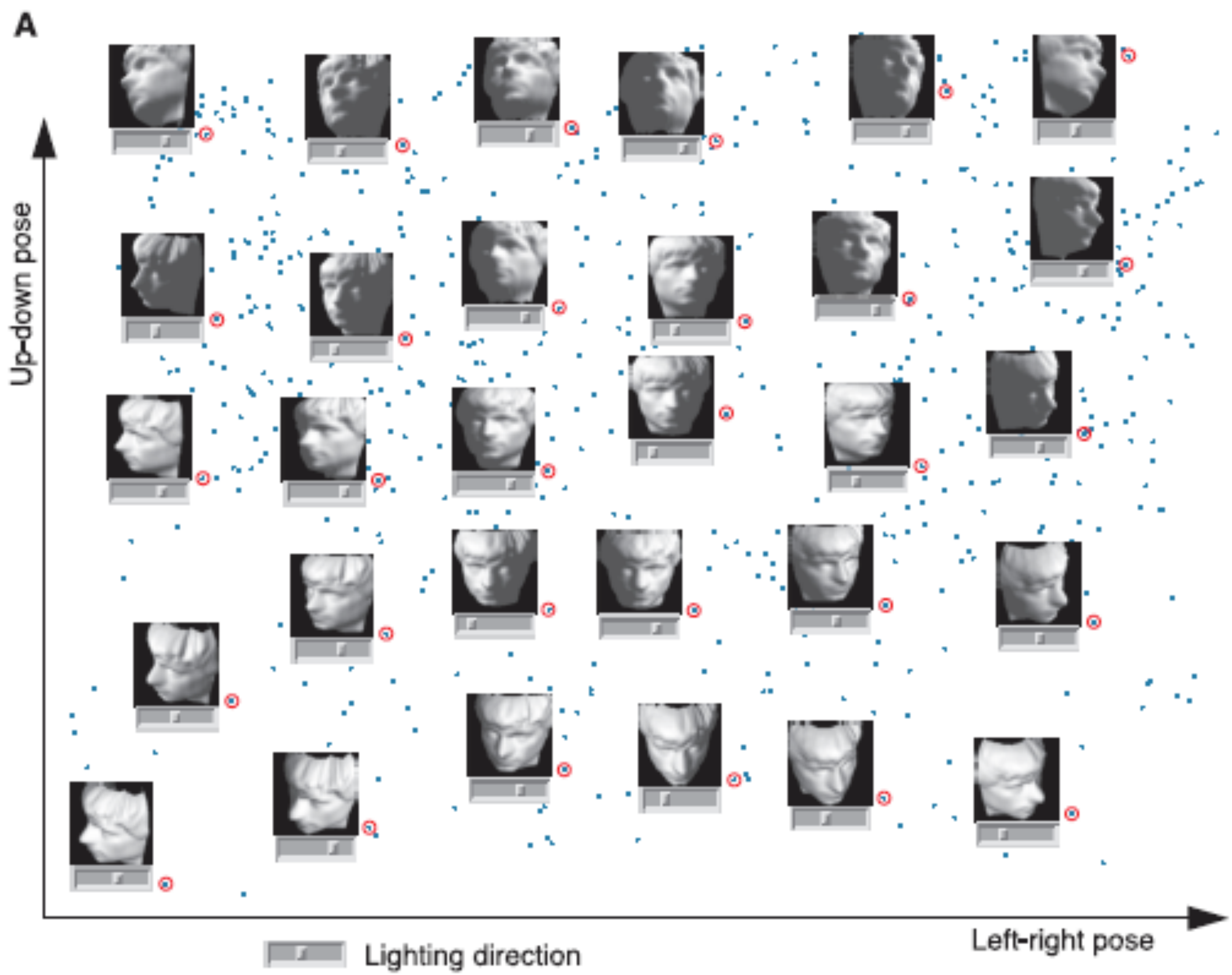
Local Linear Embedding (LLE)



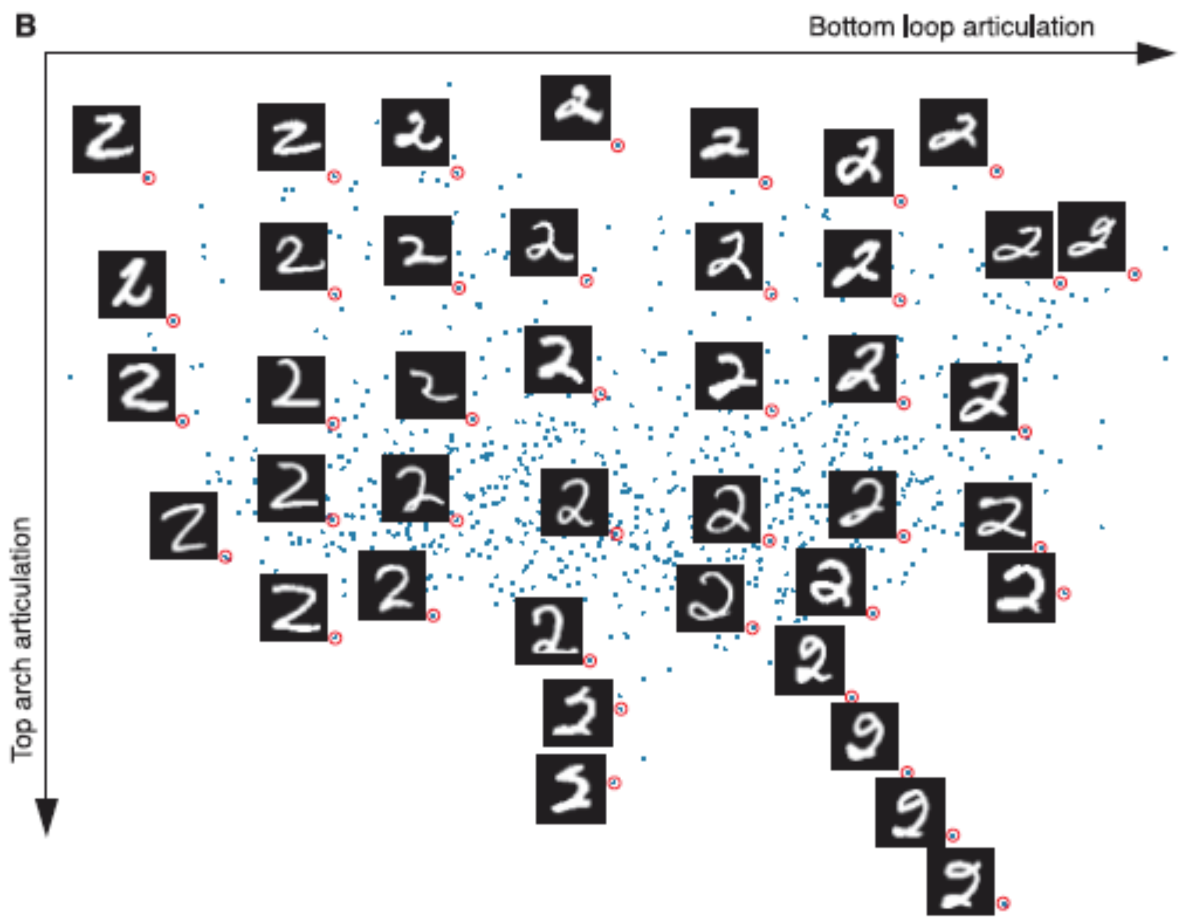
$$\varepsilon(W) = \sum_i \left| \vec{X}_i - \Sigma_j W_{ij} \vec{X}_j \right|^2$$

$$\Phi(Y) = \sum_i \left| \vec{Y}_i - \Sigma_j W_{ij} \vec{Y}_j \right|^2$$

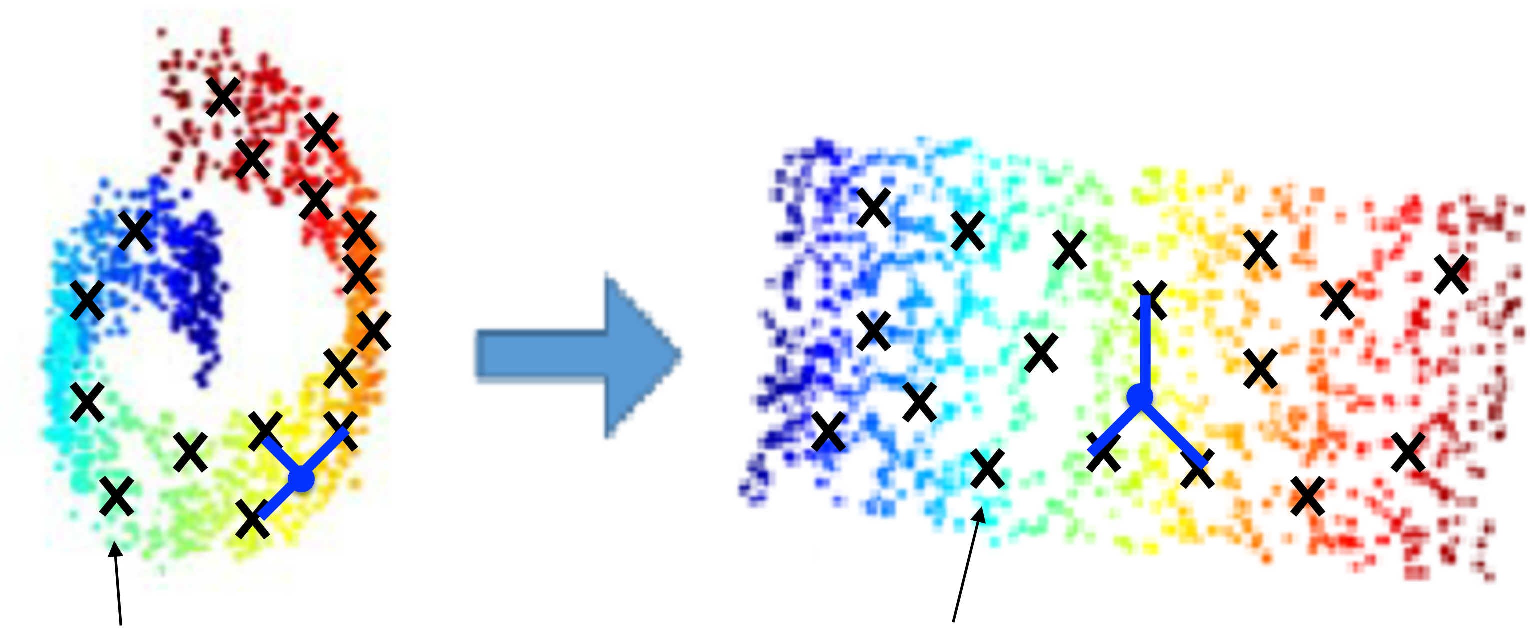
Manifold of facial pose and lighting



Hand-written digits



Local Linear Landmarks (LLL)
(Vladymyrov & Carreira-Perpinán, 2013)

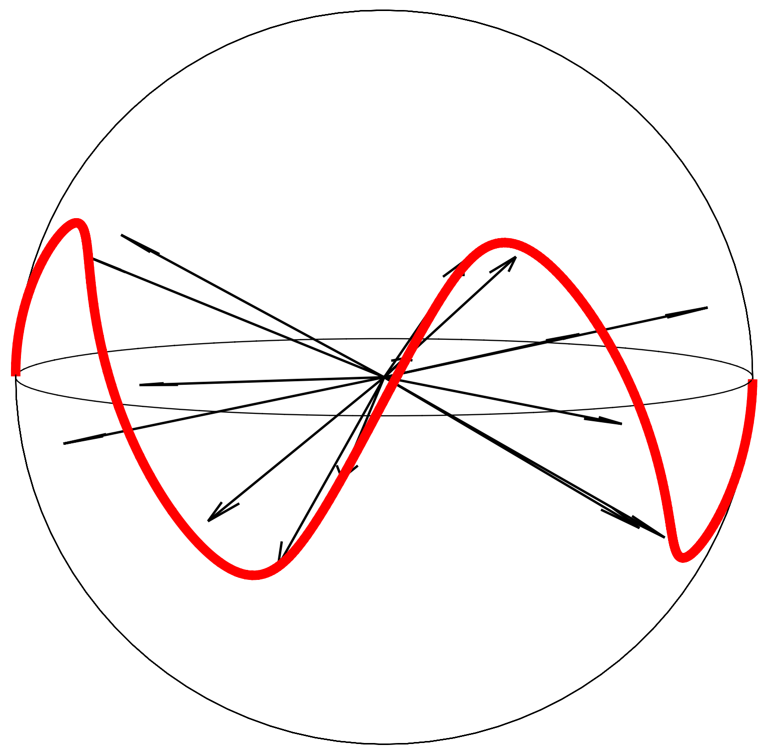


$$\Phi_i \longrightarrow P_i$$

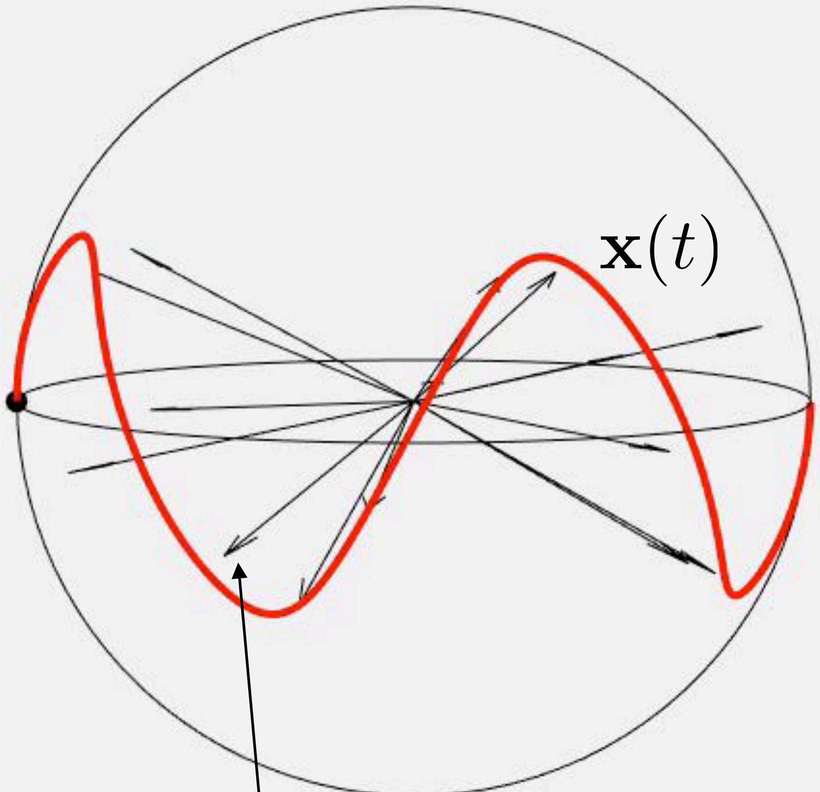
$$x = \Phi\,\alpha + n \longrightarrow y = P\,\alpha$$

local linear approximation — embedding

Basis functions learned by sparse coding form a locally linear approximation to the manifold of natural images



$$\mathbf{x}(t) = \Phi\,\alpha(t) + \mathbf{n}(t)$$

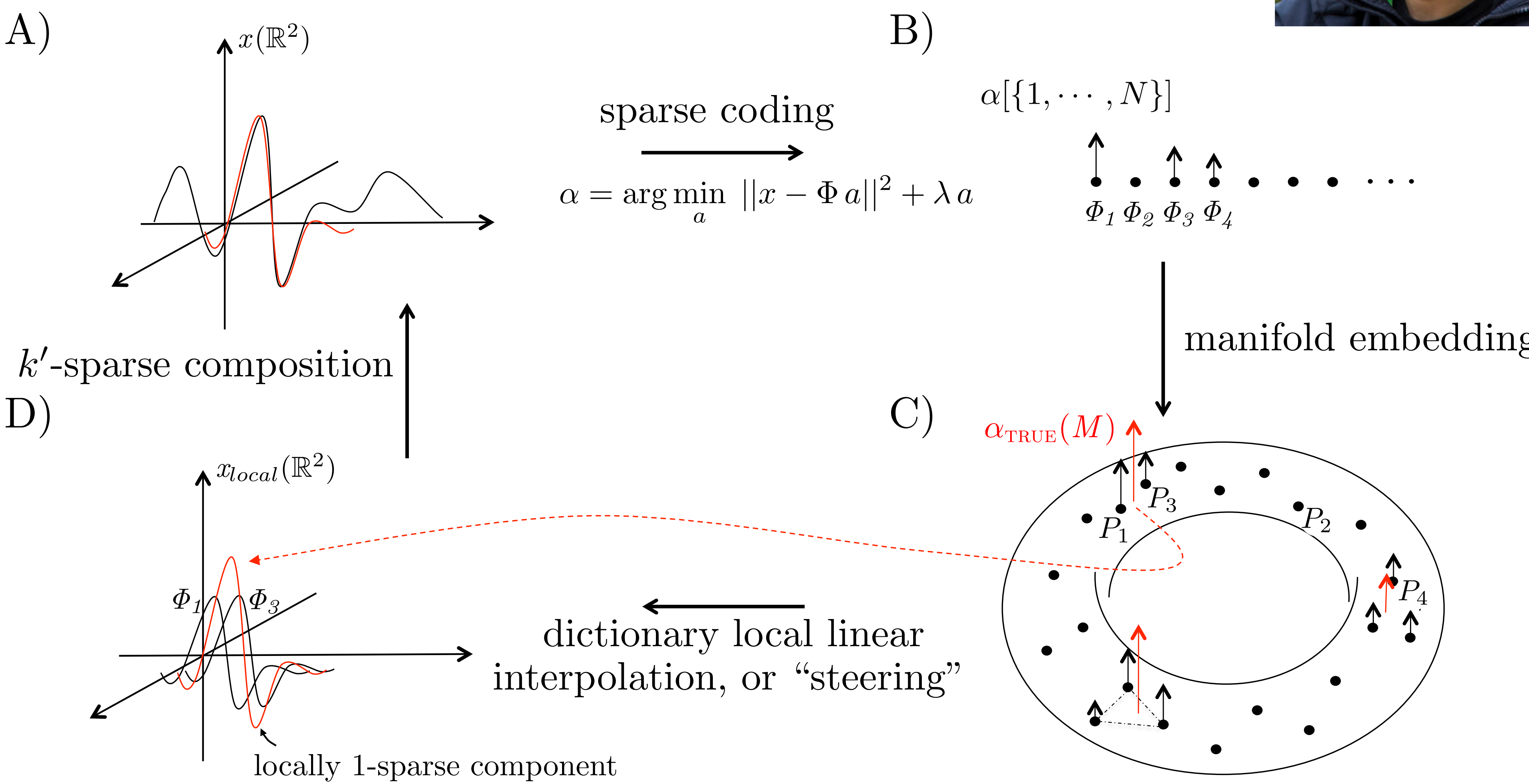


$$\beta(t) = P\,\alpha(t)$$

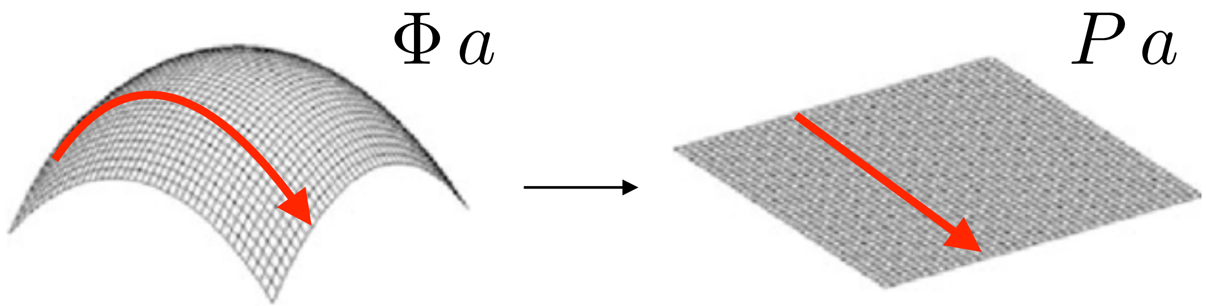
$$\Phi_i \longrightarrow P_i$$

Sparse Manifold Transform

(Yubei Chen, Ph.D. thesis; Neurips 2018)



We seek a geometric mapping $f : \Phi \rightarrow P$, s.t. each of the dictionary elements is mapped to a new vector, $P_j = f(\Phi_j)$. Continuous temporal transformations in the input should have a linear flow on M and also in the geometrical embedding space.



We desire: $$P a_t \approx \frac{1}{2} P a_{t-1} + \frac{1}{2} P a_{t+1}$$

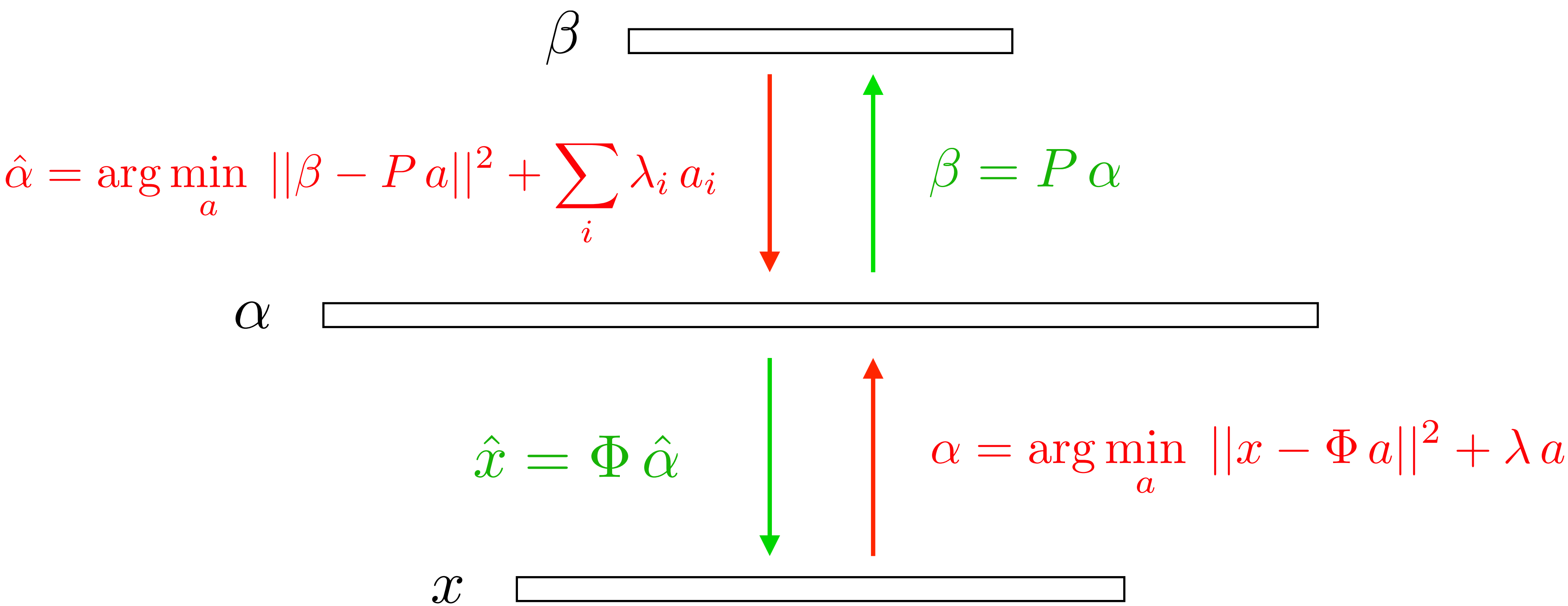
Objective function: $$\min_P ||P\,A\,D||_F^2 + \gamma\,||P\,V||_1$$

s.t. $$P\,V\,P^T = I$$

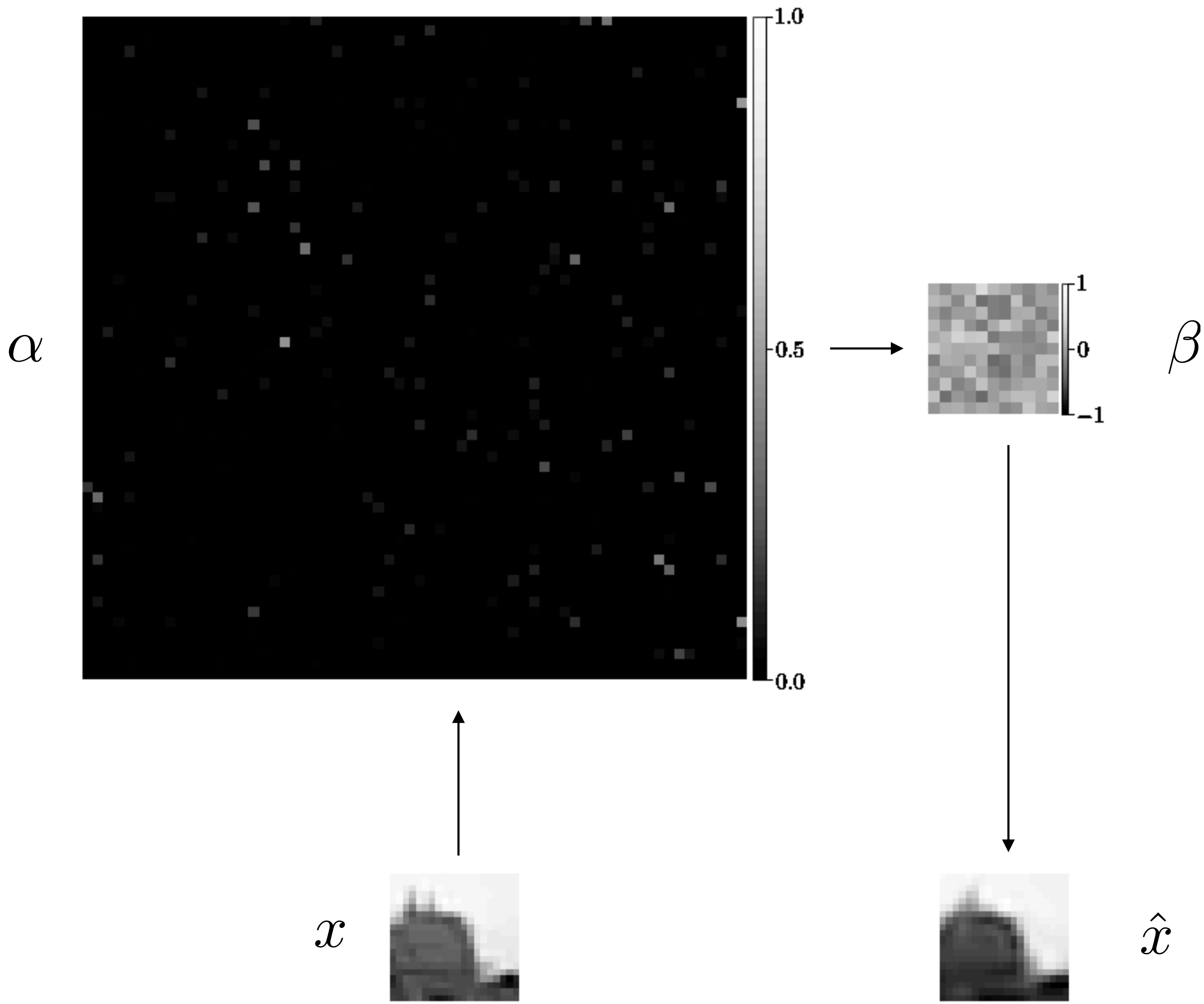
$$V = \mathrm{Cov}(a)$$

$$D = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 & \dots & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & \dots & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

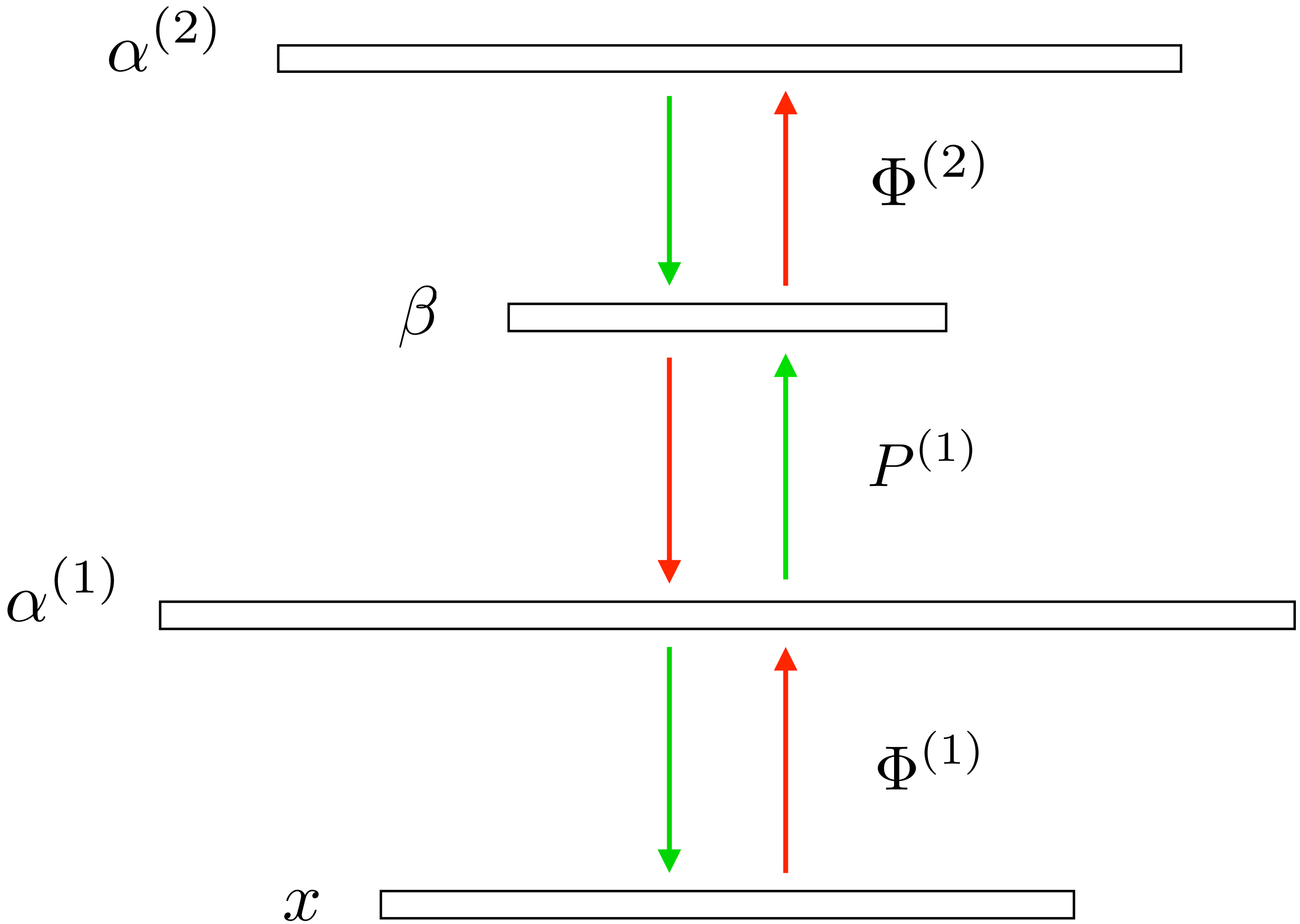
The sparse manifold transform



Encoding of a natural video sequence



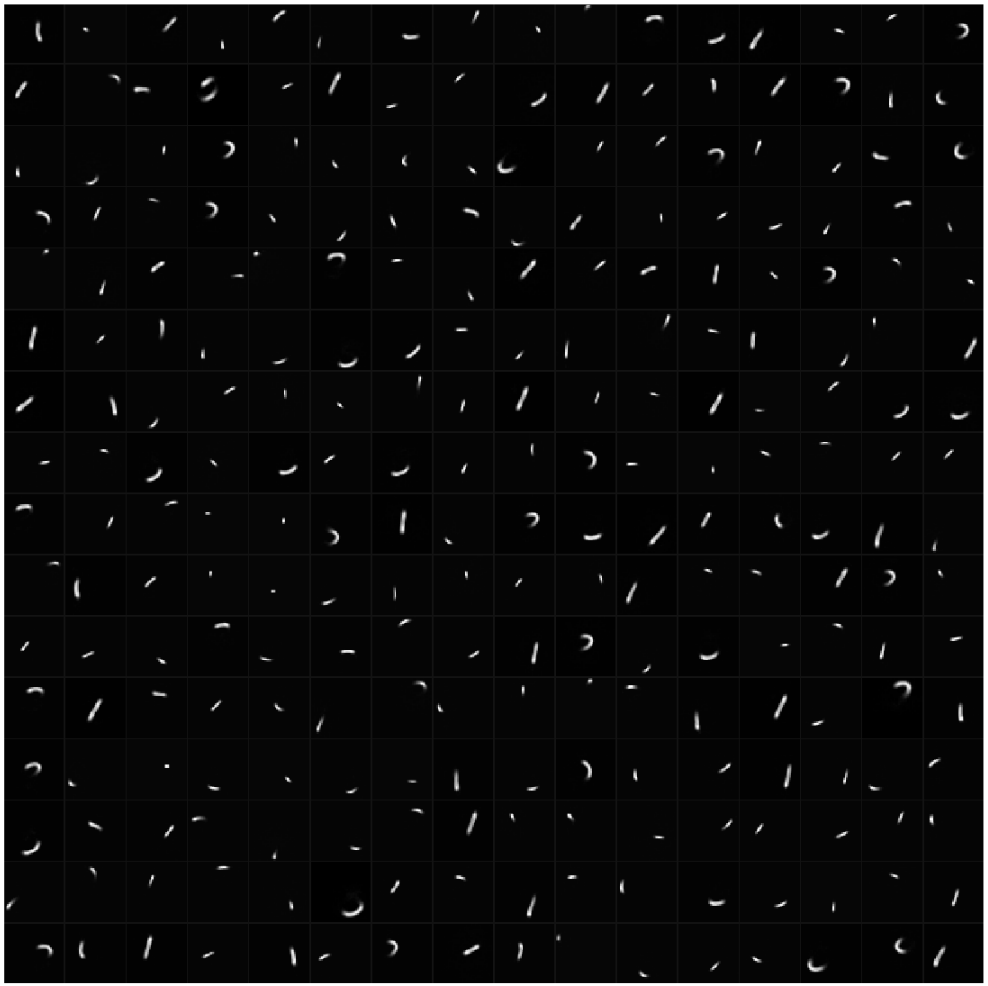
The *stacked* sparse manifold transform



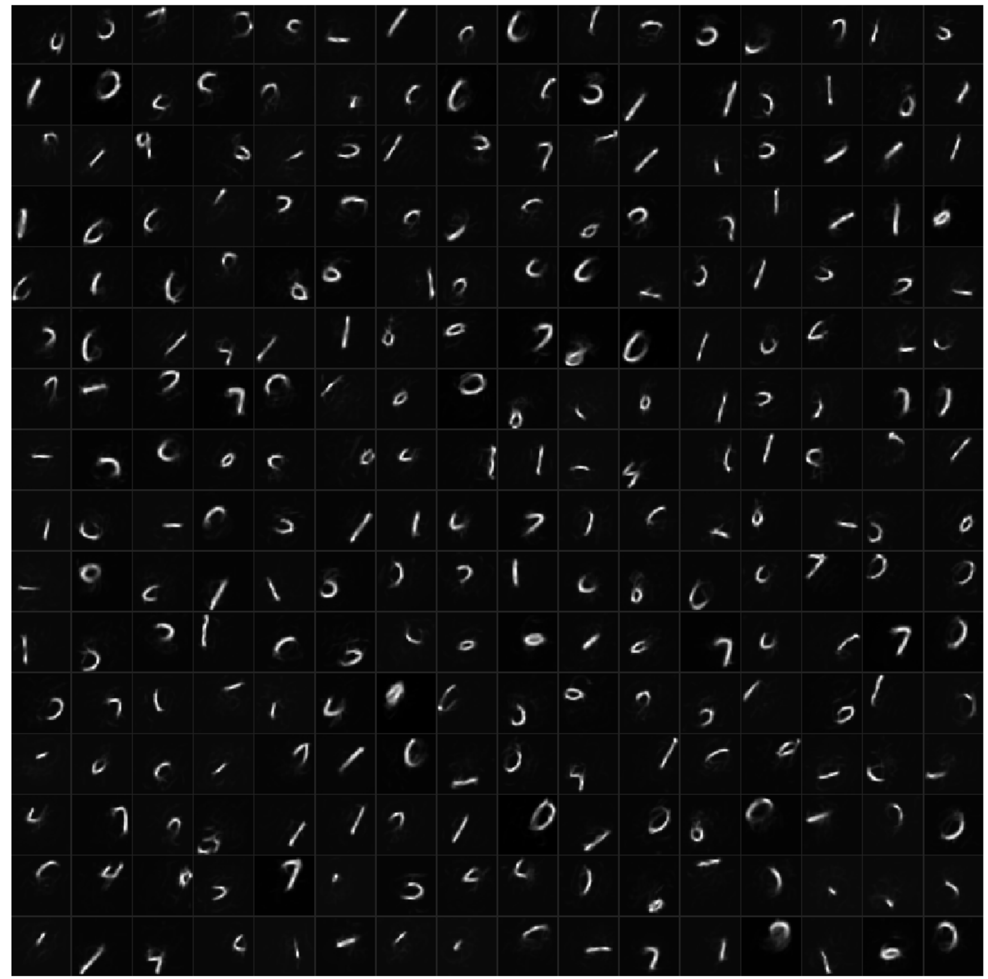
$$\min_{\alpha^{(1)},\alpha^{(2)}} \quad ||\Phi^{(2)}\alpha^{(2)} - P^{(1)}\alpha^{(1)}||^2 + ||x - \Phi^{(1)}\alpha^{(1)}||^2 + S(\alpha^{(1)}) + S(\alpha^{(2)})$$

Stacked Sparse Manifold Transform
(trained on MNIST)

Learned $\Phi^{(1)}$



Learned $\Phi^{(2)}$



Unsupervised learning principles

- Linear Hebbian learning → PCA
- Competitive Hebbian learning → clustering
- Sparse coding → feature learning
- Self-organizing map → topographic maps
- Sparse manifold transform → manifold learning
- Slow feature analysis → invariance