Sparse and Overcomplete Representation: Finding Statistical Orders in Natural Images

Amir Reza Saffari Azar

Institute for Theoretical Computer Science,
Graz University of Technology

amir@igi.tugraz.at

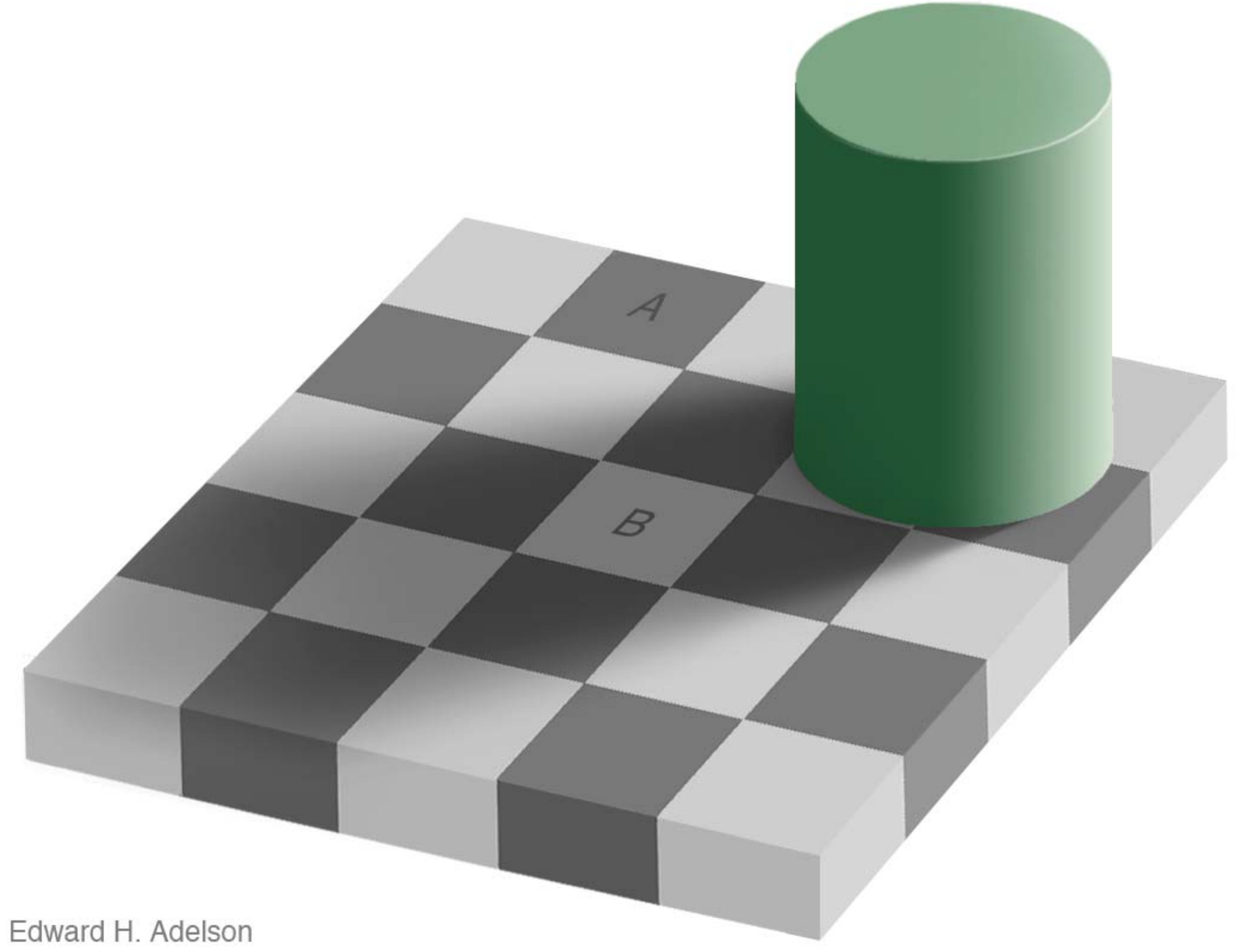
Outline

- Visual Cortex.
- Sparse and Overcomplete Representation.
- Nonlinear Hierarchial Model for Modeling Higher-order Structures.
- Discussion.

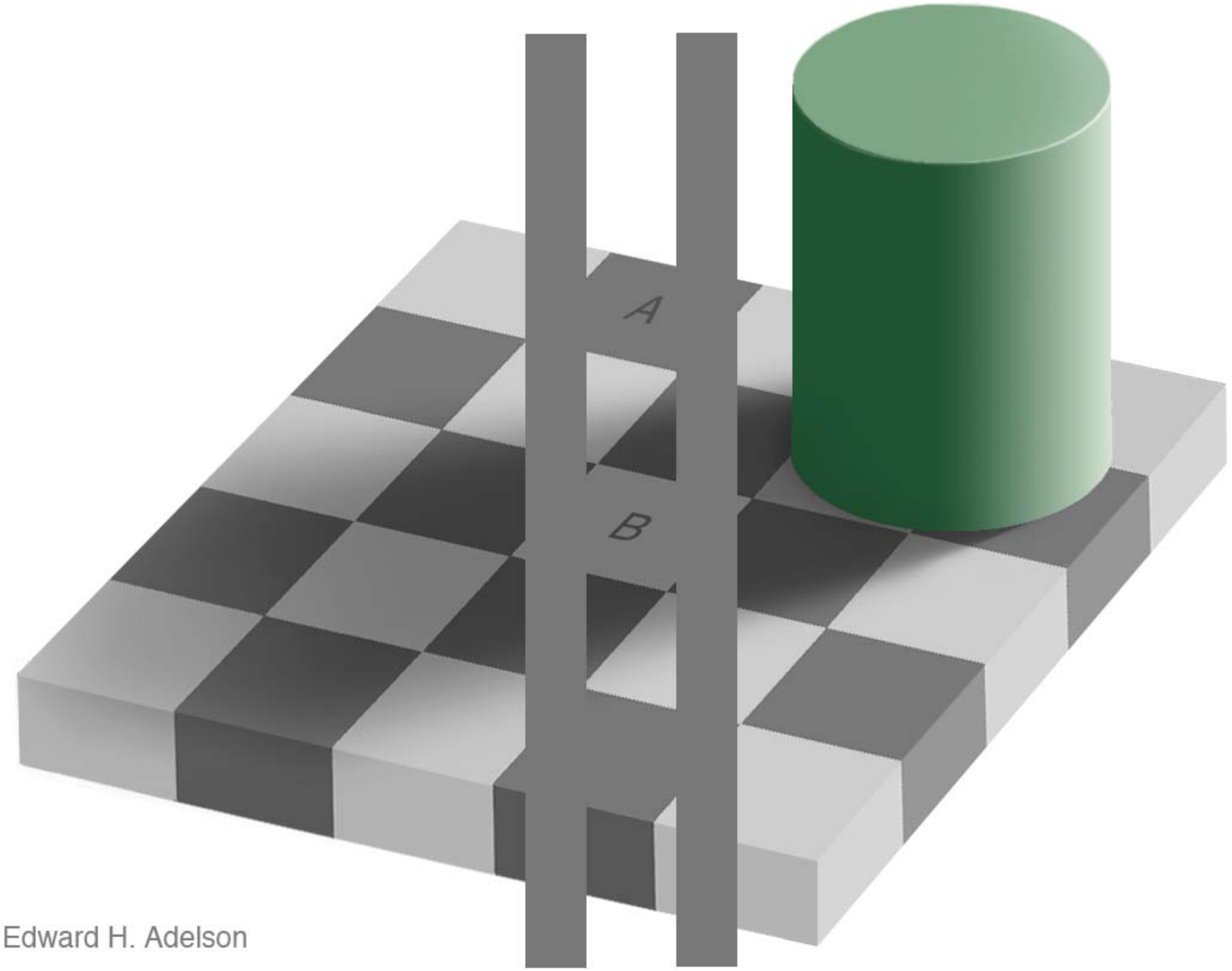
Visual Cortex

- How visual cortex (V1) represents images?
- Visual cortex as an hierarchical inference machine.
- Environment (E), Observed data (D): $P(D|E)$, Priors: $P(E)$.
- Inference: $P(E|D) = P(D|E)P(E)/Z$.

Illusion



Illusion



Visual Cortex

- Perception as probabilistic inference (Helmoltz 1867/1962).
- Redundancy reduction (Barlow 1961).
- Theoretical models for retina (Srinivasan et al. 1982) and LGN (van Hateren 1993) response properties.

Visual Cortex

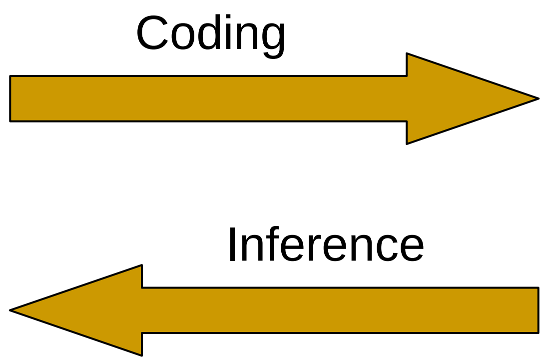
- Anatomical convergence of about 100 million photoreceptors onto 1 million ganglion cells.
- In cat V1, there are 25 times as many output fibers as there are input fibers from LGN.
- In macaque this ratio is on the order of 50:1.
- Increased redundancy.
- Overcomplete representation.

Visual Cortex

- A meaningful representation can be achieved by finding a code that has less active units at each time.
- Sparse representation.
- Evidences for sparse activity in V1.
- The average activity in primate cortex is less than 1 Hz (Lennie 2003).

Linear Image Model

$$I(\mathbf{x}) = \sum_i a_i \Phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

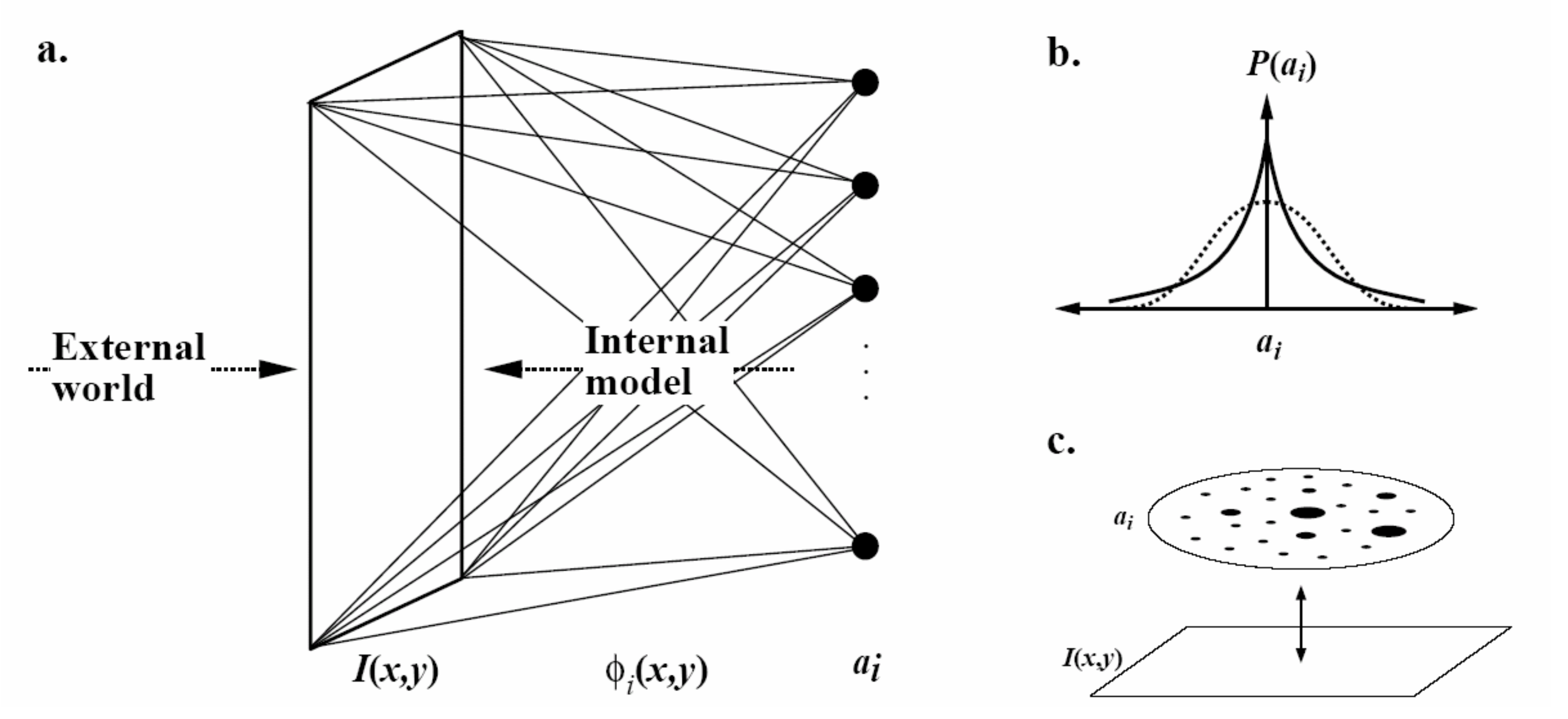


$$E = \sum_{\mathbf{x}} \left(I(\mathbf{x}) - \sum_i a_i \Phi_i(\mathbf{x}) \right)^2$$

$$E = \sum_{\mathbf{x}} \left(I(\mathbf{x}) - \sum_i a_i \Phi_i(\mathbf{x}) \right)^2 + \beta \sum_i S(a_i)$$

- (Olshausen 2003)

Linear Image Model



- (Olshausen 2002)

Probabilistic Formulation

$$I(\mathbf{x}) = \sum_i a_i \Phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

$$P(\mathbf{a}) = \prod_i P(a_i)$$

$$P(a_i) = \frac{1}{Z_S} e^{-S(a_i)}$$

$$P(\mathbf{a} \mid \mathbf{I}, \theta) \propto P(\mathbf{I} \mid \mathbf{a}, \theta) P(\mathbf{a} \mid \theta)$$

$$P(\mathbf{I} \mid \mathbf{a}, \theta) = \frac{1}{Z_{\lambda_N}} e^{-\frac{\lambda_N}{2}|\mathbf{I} - \Phi \mathbf{a}|^2}$$

$$P(\mathbf{a} \mid \theta) = \prod_i \frac{1}{Z_S} e^{-S(a_i)}$$

- (Olshausen 2002)

Finding Sparse Codes

$$I(\mathbf{x}) = \sum_i a_i \Phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}} P(\mathbf{a} \mid \mathbf{I}, \theta)$$

$$\dot{\mathbf{a}} \propto \nabla_{\mathbf{a}} \log P(\mathbf{a} \mid \mathbf{I}, \theta)$$
$$= -\nabla_{\mathbf{a}} \left[\frac{\lambda_N}{2} \left| \mathbf{I} - \Phi \mathbf{a} \right|^2 + \sum_i S(a_i) \right]$$

$$\tau \dot{a}_i = b_i - \sum_j C_{ij} a_j - S'(a_i)$$

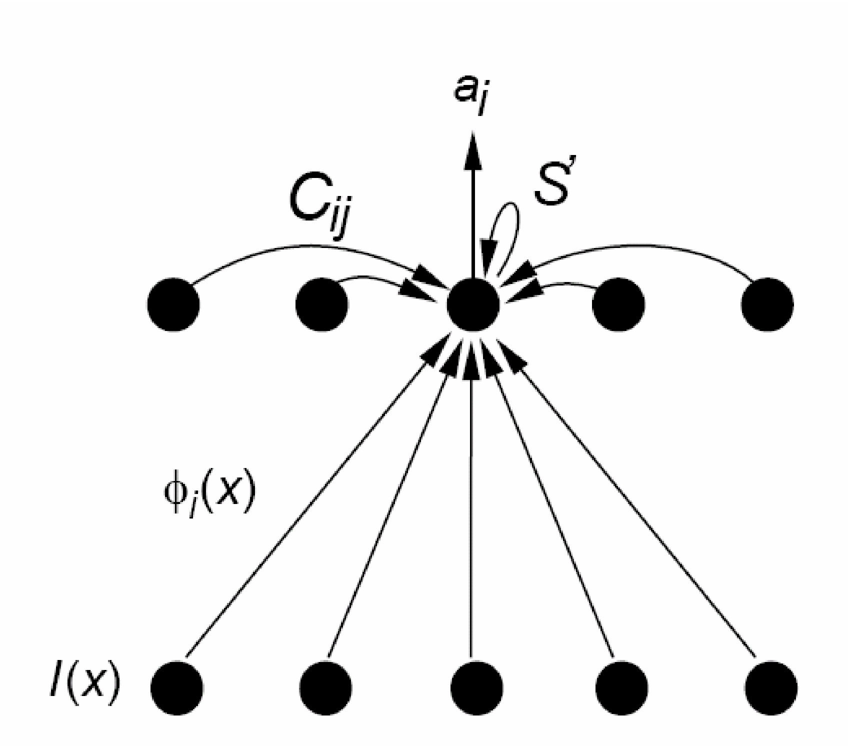
$$b_i = \lambda_N \sum_{\mathbf{x}} \Phi_i(\mathbf{x}) I(\mathbf{x})$$

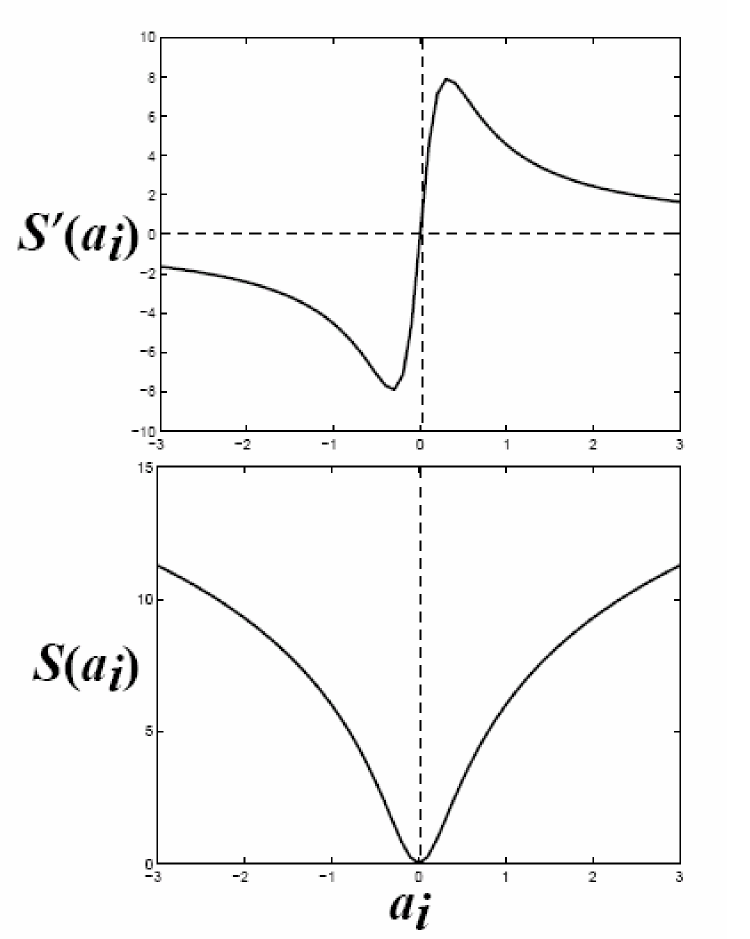
$$C_{ij} = \lambda_N \sum_{\mathbf{x}} \Phi_i(\mathbf{x}) \Phi_j(\mathbf{x})$$

$$S(a) = \gamma \log(1 + (a/\sigma)^2)$$

- (Olshausen 2003)

Neural Interpretation





- (Olshausen 2002, 2003)

Learning Basis Functions

$$I(\mathbf{x}) = \sum_i a_i \Phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

$$\hat{\theta} = \arg\max_\theta \langle P(\mathbf{I} \mid \theta) \rangle$$

$$P(\mathbf{I} \mid \theta) = \int P(\mathbf{I}, \mathbf{a} \mid \theta) d\mathbf{a} \propto P(\mathbf{I} \mid \hat{\mathbf{a}}, \theta)$$

$$\mathbf{e} = \mathbf{I} - \mathbf{\Phi a}$$

$$\Delta \mathbf{\Phi} \propto \frac{\partial \langle \log P(\mathbf{I} \mid \theta) \rangle}{\partial \mathbf{\Phi}}$$

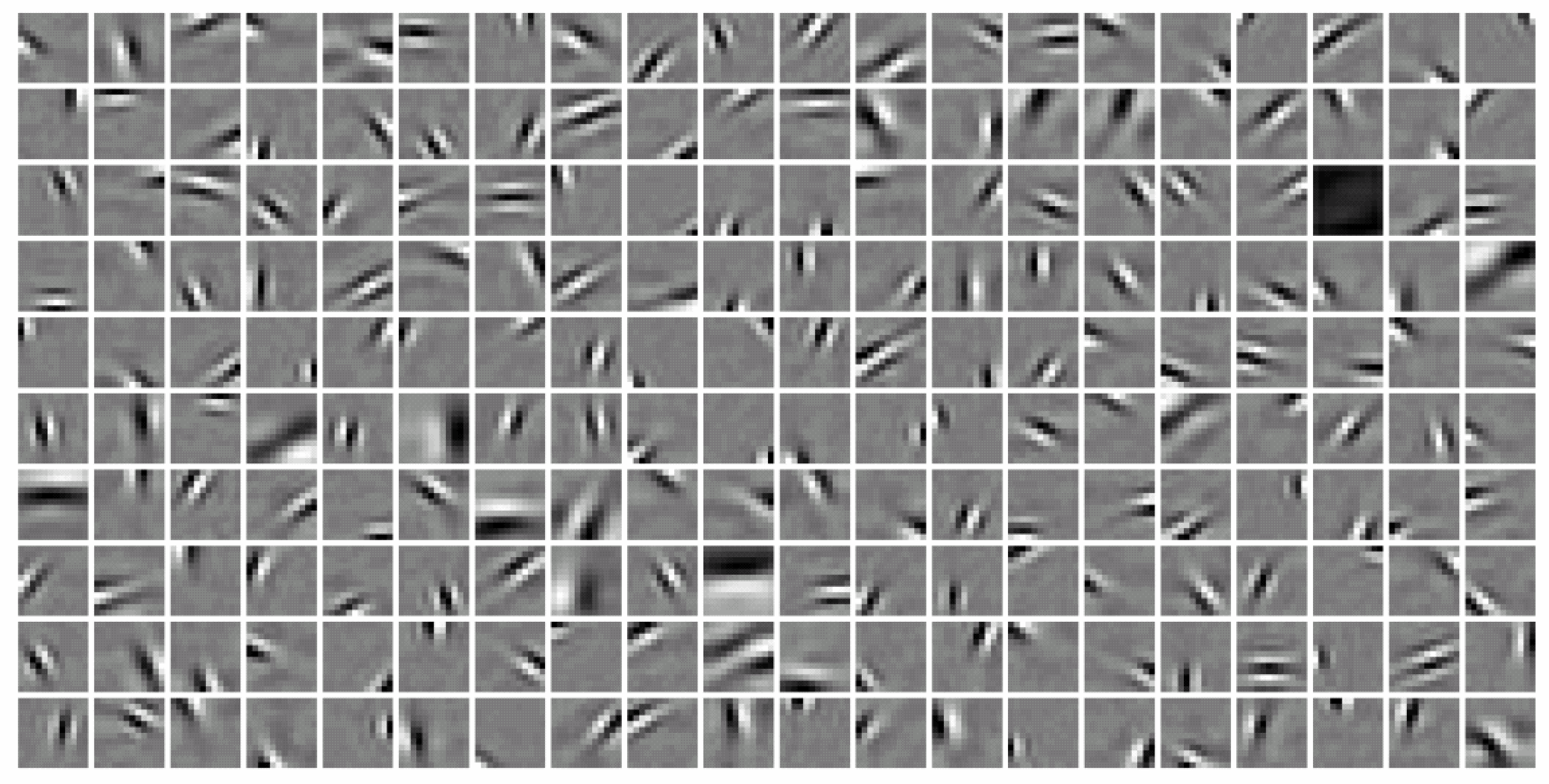
$$= \lambda_N \langle \mathbf{e} \hat{\mathbf{a}}^T \rangle$$

$$g_i = \left| \Phi_i \right|_{L2}$$

$$g_i^{new} = g_i^{old} \left[\frac{\langle a_i^2 \rangle}{\sigma^2} \right]^\alpha$$

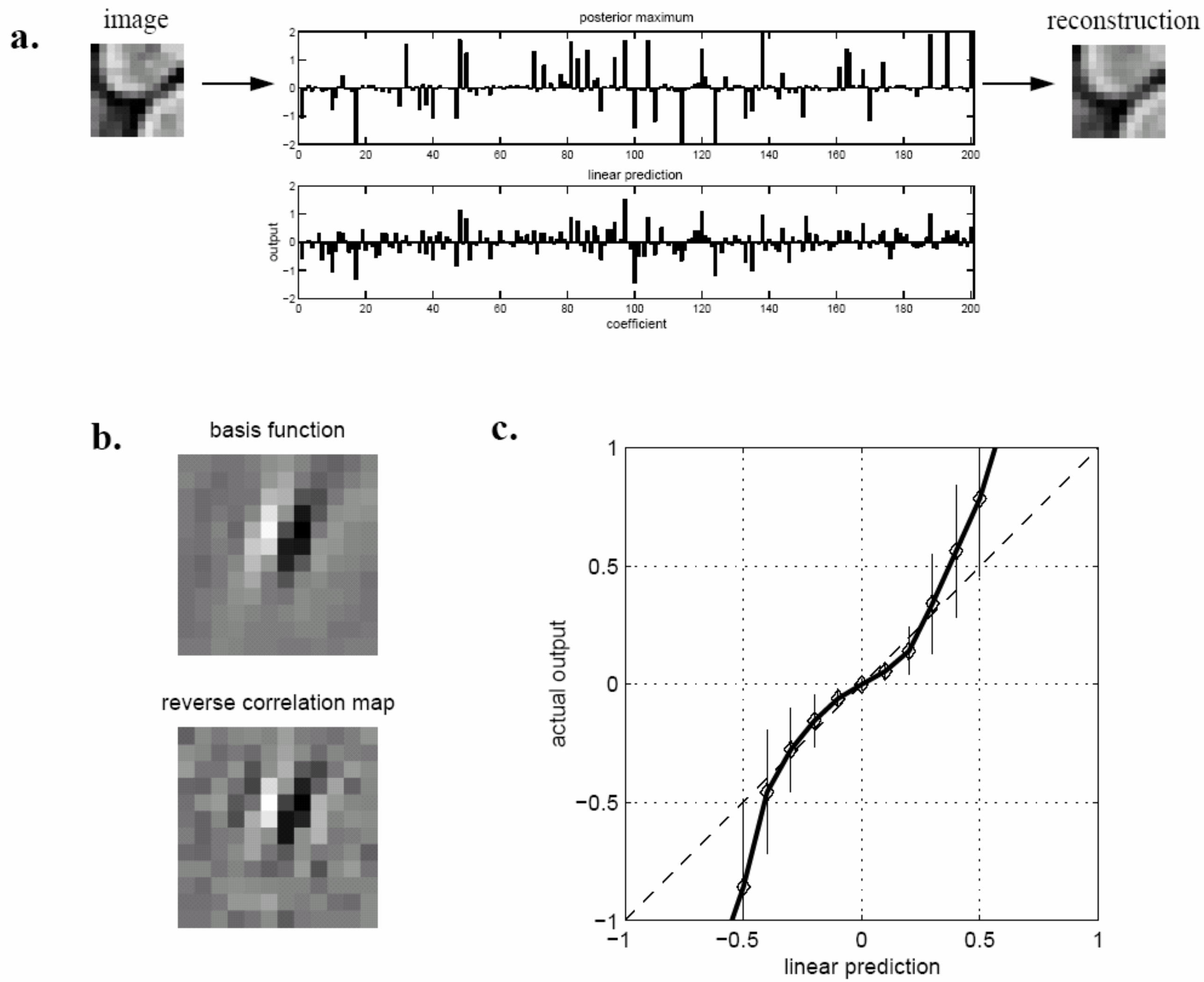
- (Olshausen 2002, 2003)

Results



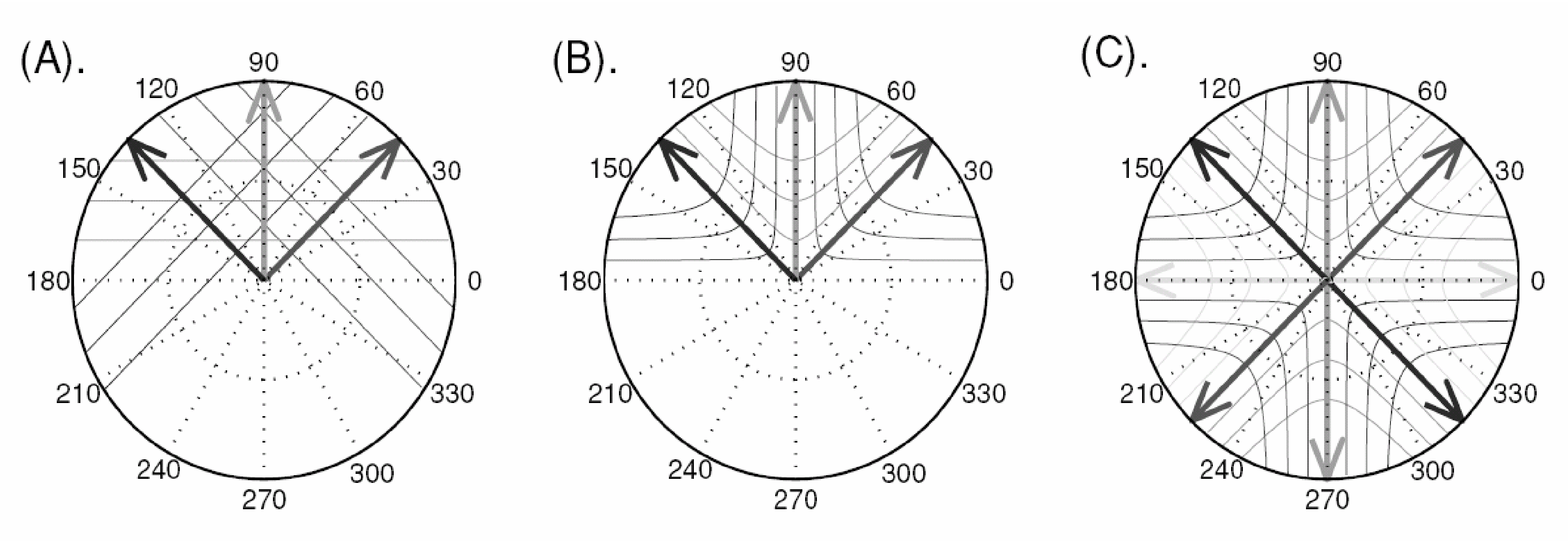
- (Olshausen 2002, 2003)

Sparsification



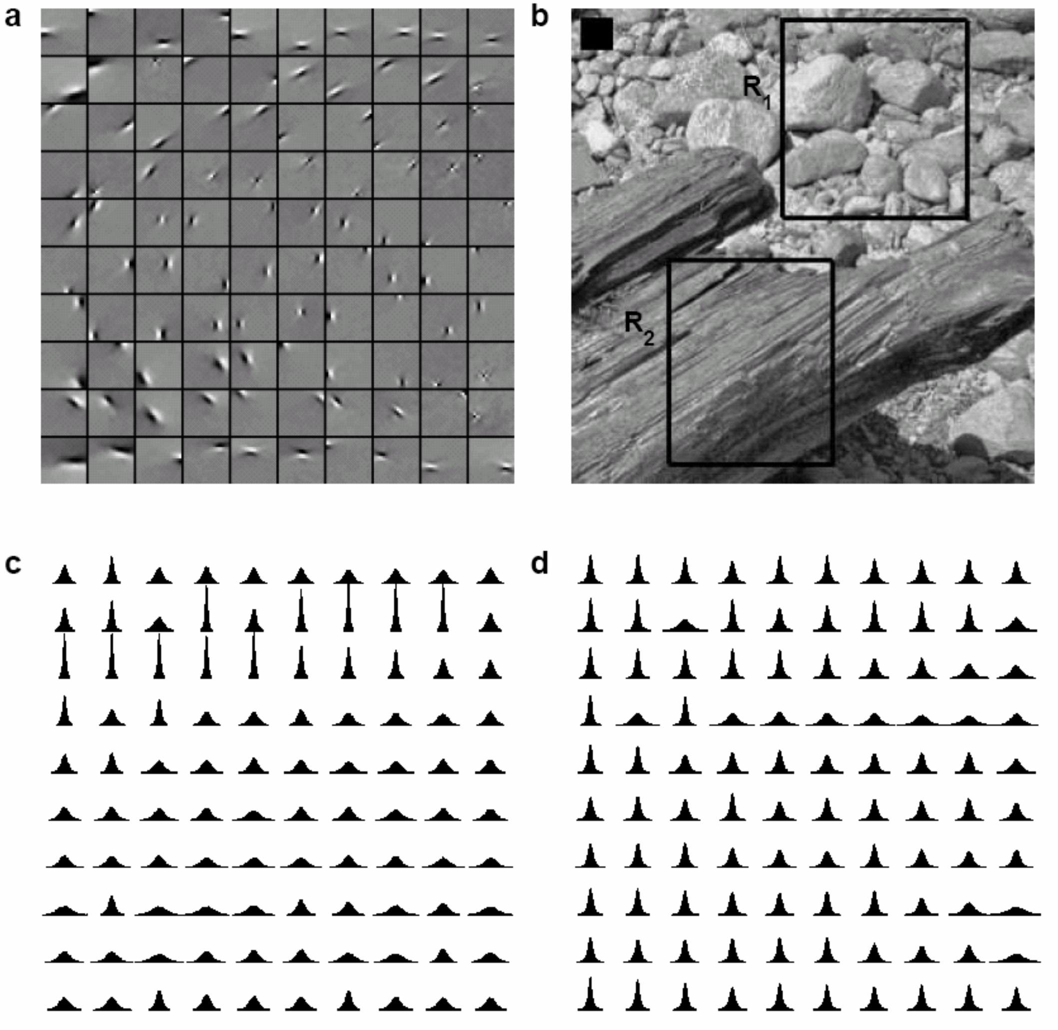
- (Olshausen 2002)

Overcomplete Sparse Representation



- (Olshausen 2005)

Non-Stationary Statistics in Natural Images



- (Karklin and Lewicki 2005)

Probabilistic Formulation

$$I(\mathbf{x}) = \sum_i a_i \Phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

$$P(\mathbf{a}) = \prod_i P(a_i)$$

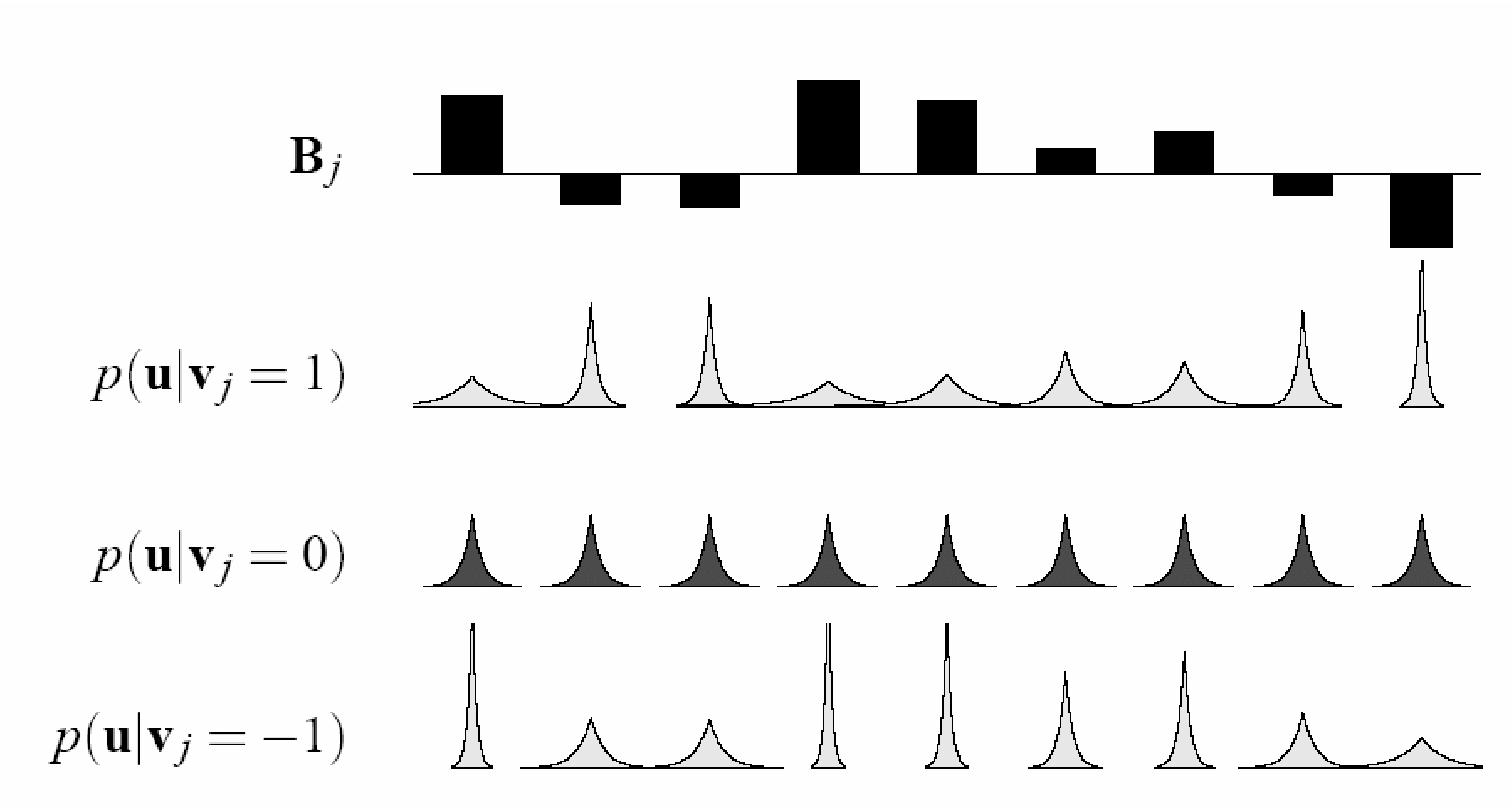
$$P(a_i) = N(0, \lambda_i, q_i) = z_i \exp\left(-\left|\frac{a_i}{\lambda_i}\right|^{q_i}\right)$$

$$\log(\boldsymbol{\lambda}) = \mathbf{B}\mathbf{v}$$

$$-\log P(\mathbf{a} \mid \mathbf{B}, \mathbf{v}) \propto \sum_i \left|\frac{a_i}{\exp(\sum_j B_{ij} v_j)}\right|^{q_i}$$

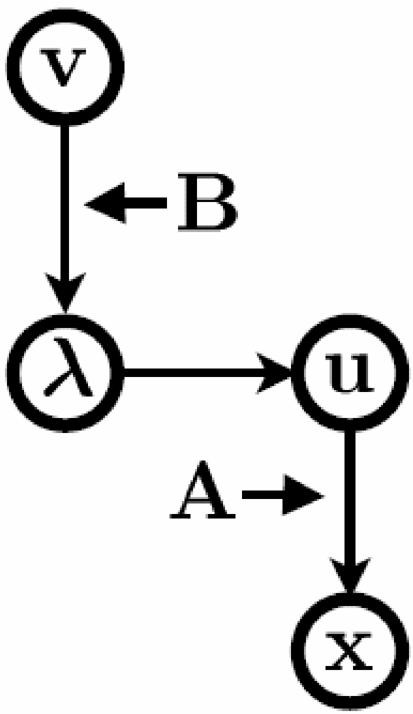
- (Karklin and Lewicki 2003)

Example



- (Karklin and Lewicki 2005)

Hierarchical Structure



$$
\begin{aligned}
v_i &\sim \mathcal{N}(0, 1, q_i) \\
\lambda_j &= c \exp^{[\mathbf{Bv}]_j} \\
u_j | \lambda_j &\sim \mathcal{N}(0, \lambda_j, q_j) \\
x_k &= \sum_j \mathbf{A}_{kj} u_j
\end{aligned}
$$

- (Karklin and Lewicki 2005)

Encoding Variance Coefficients

$$\hat{\mathbf{v}} = \arg\max_{\mathbf{v}} P(\mathbf{v} \mid \mathbf{a}, \mathbf{B}) = \arg\max_{\mathbf{v}} P(\mathbf{a} \mid \mathbf{B}, \mathbf{v}) P(\mathbf{v})$$

$$P(\mathbf{v}) = \prod_i P(v_i)$$

$$P(v_i) = N(0, 1, r_i)$$

$$\dot{v}_j = \sum_{i=1}^{N} \left[B_{ij} + q_i B_{ij} \left| \frac{u_i}{e^{[\mathbf{Bv}]_i}} \right|^{q_i} \right] - \mathrm{sign}(v_j) r_j |v_j|^{r_j - 1}$$

- (Karklin and Lewicki 2003)

Learning Variance Basis Functions

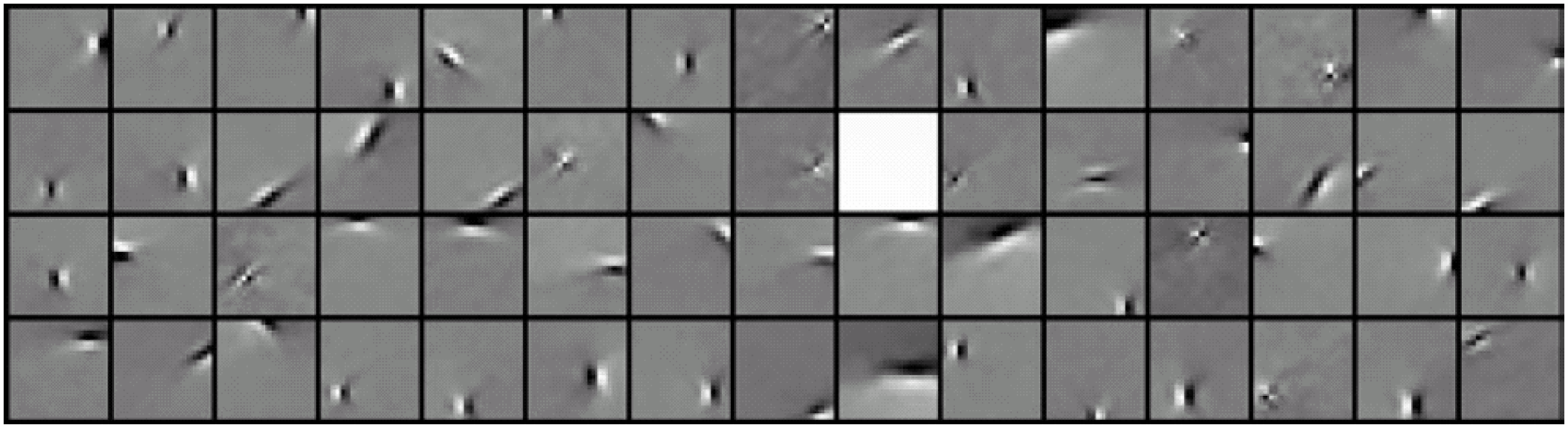
$$\hat{\mathbf{B}} = \arg\max_{\mathbf{B}} \left\langle P(\mathbf{B} \mid \mathbf{I}, \mathbf{\Phi}) \right\rangle = \arg\max_{\mathbf{B}} \left\langle P(\mathbf{I} \mid \mathbf{\Phi}, \mathbf{B}) P(\mathbf{B}) \right\rangle$$

$$\log P(\mathbf{I} \mid \mathbf{\Phi}, \mathbf{B}) \propto \log P(\mathbf{a} \mid \mathbf{B}, \hat{\mathbf{v}}) P(\hat{\mathbf{v}}) \big/ \left|\det \mathbf{\Phi}\right|$$

$$\Delta B_{ij} = \left\langle -v_j + v_j q_i \left| \frac{a_i}{e^{[\mathbf{Bv}]_i}} \right|^{q_i} - B_{ij} \right\rangle$$

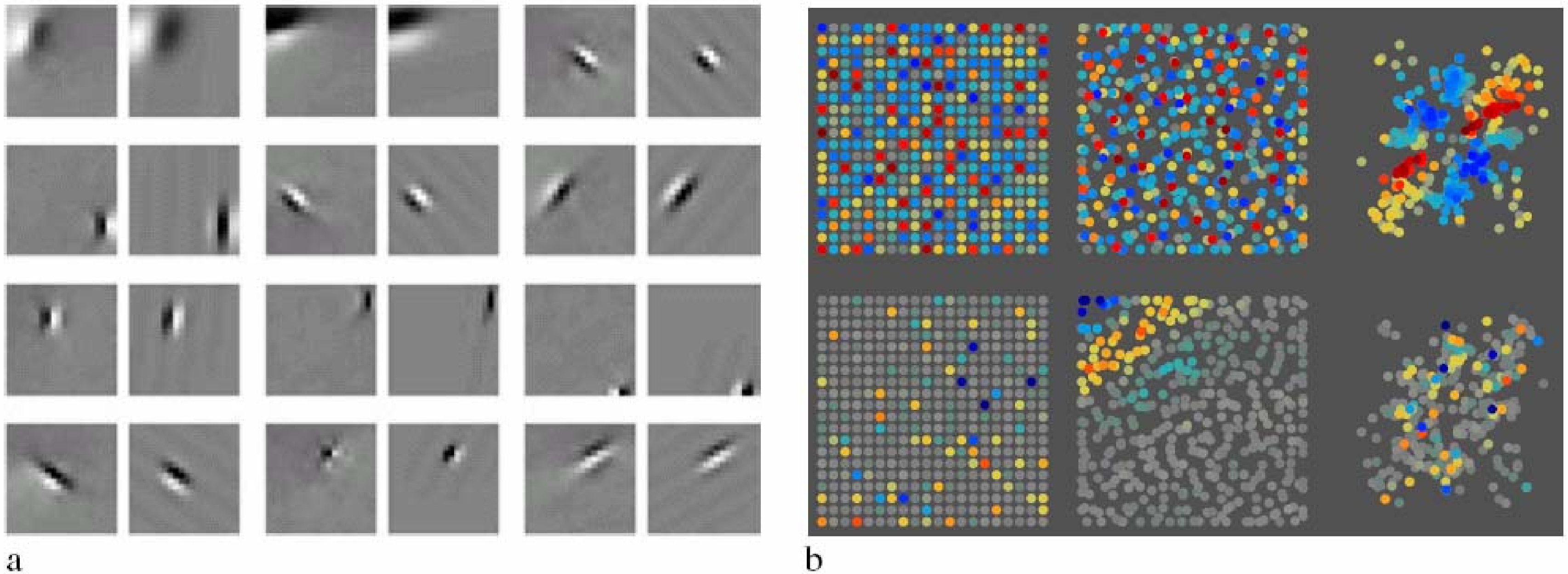
- (Karklin and Lewicki 2003)

Results: Basis Functions



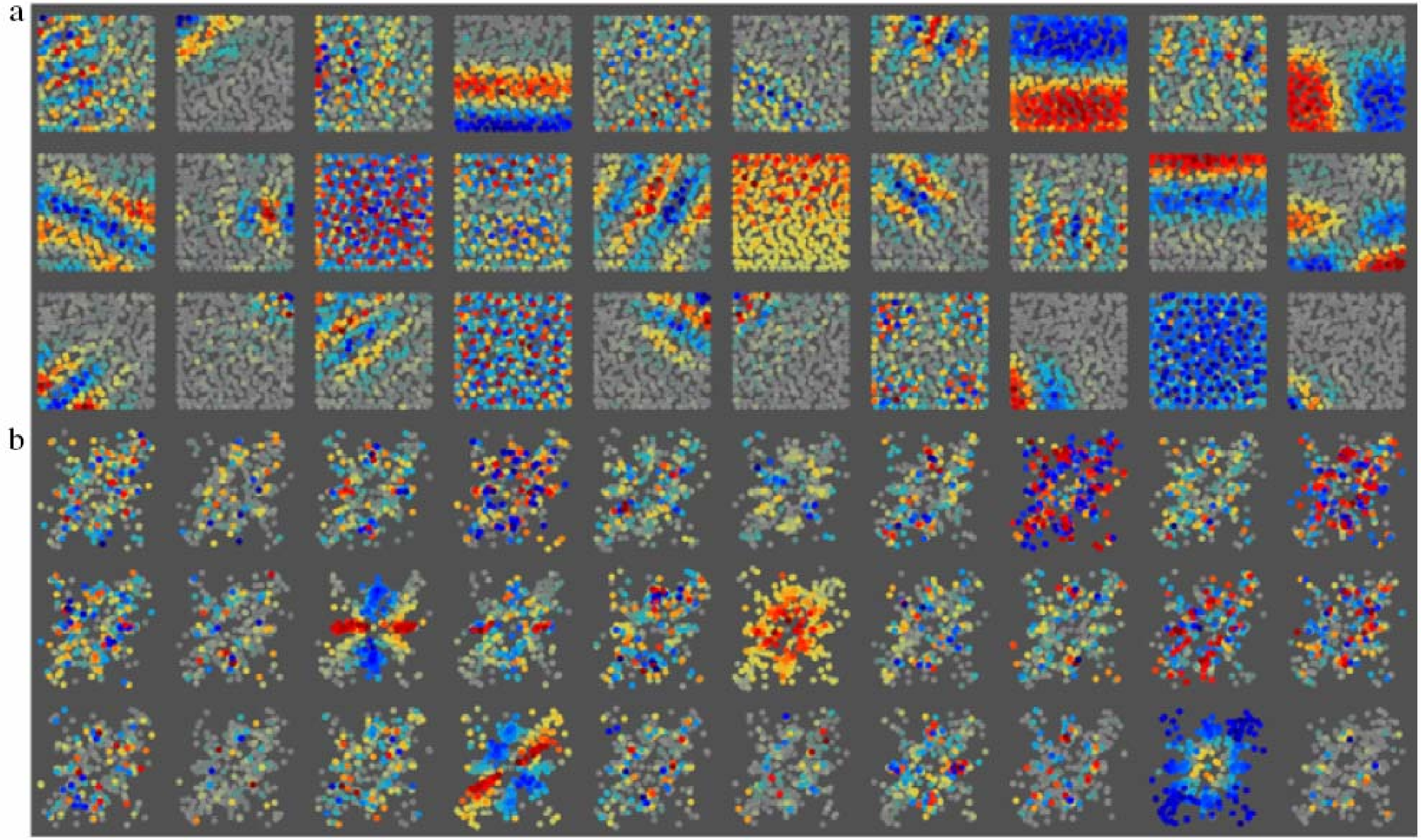
- (Karklin and Lewicki 2003)

Results: Basis Functions



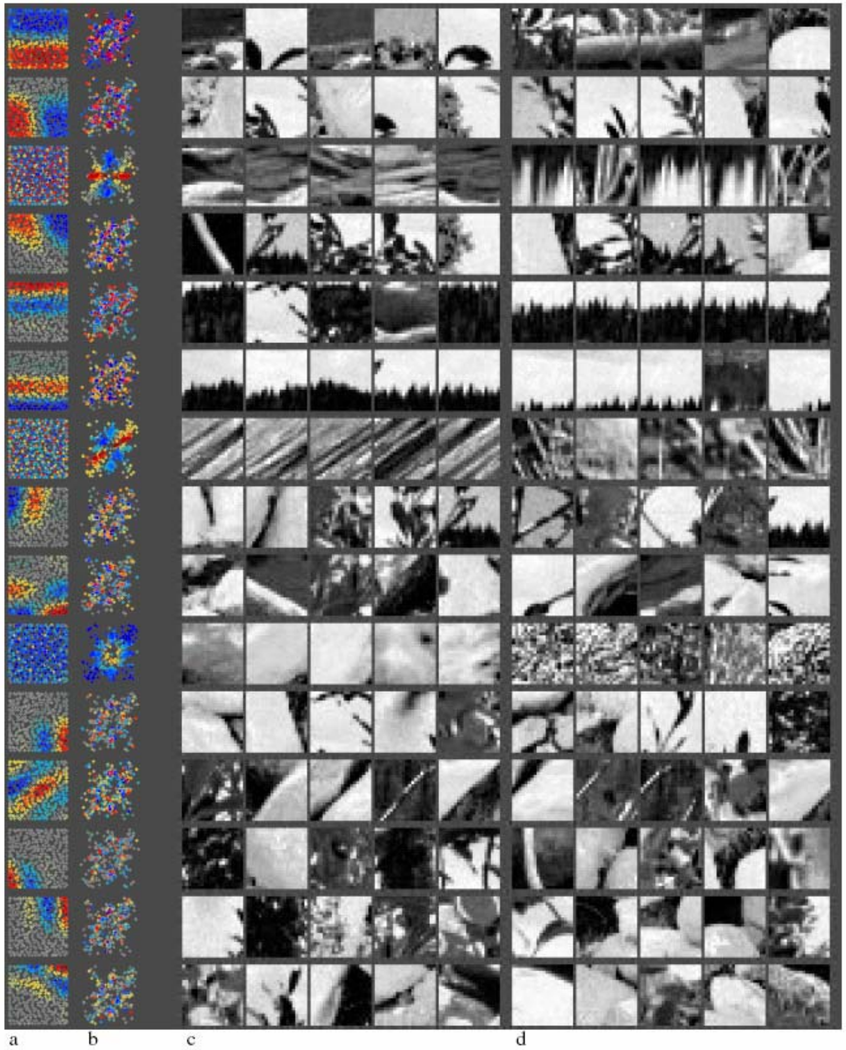
- (Karklin and Lewicki 2003)

Results: Variance Basis Functions



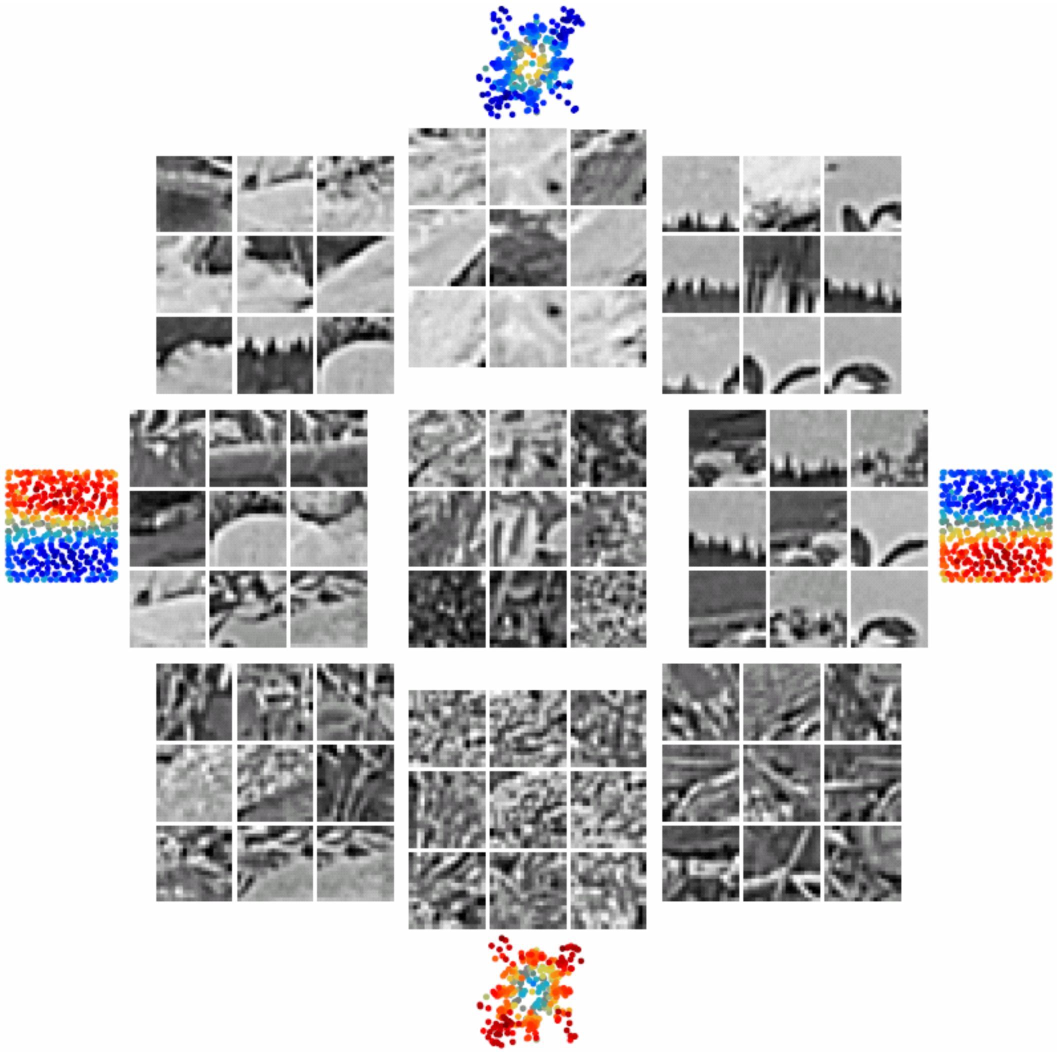
- (Karklin and Lewicki 2003)

Results: Variance Basis Functions



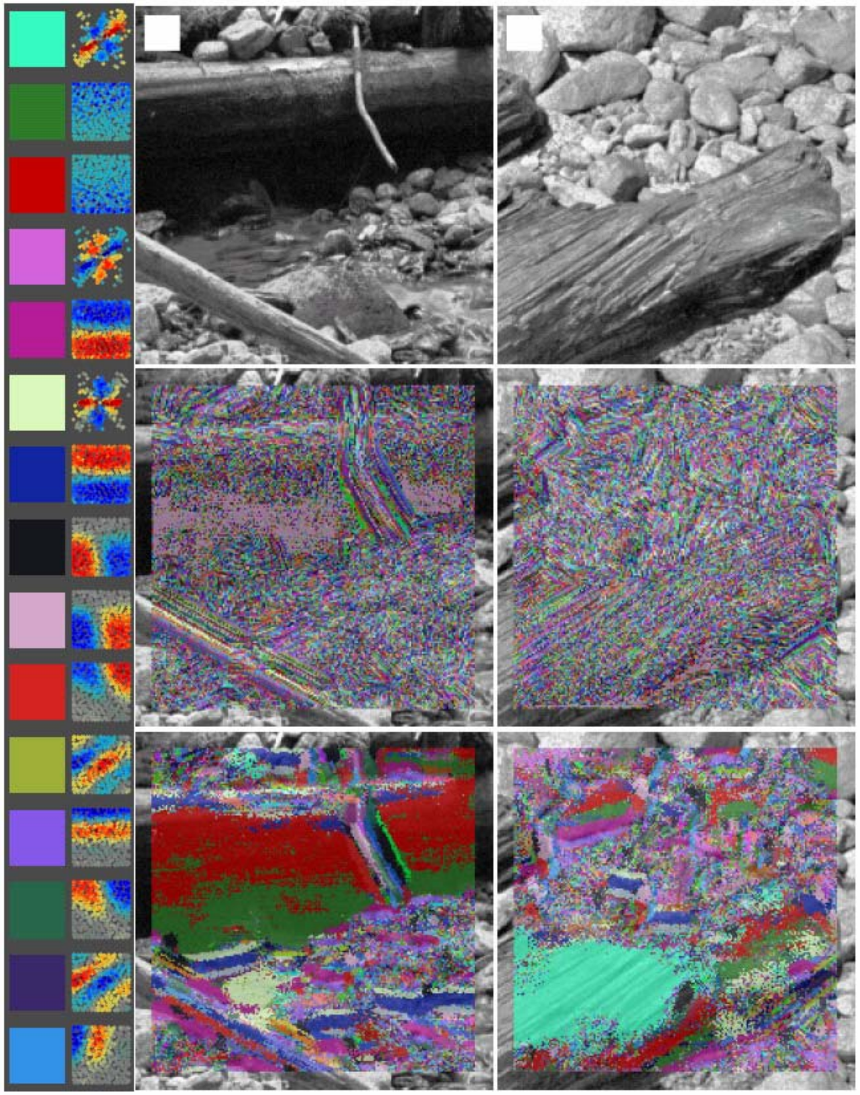
- (Karklin and Lewicki 2003)

Results: Variance Basis Functions



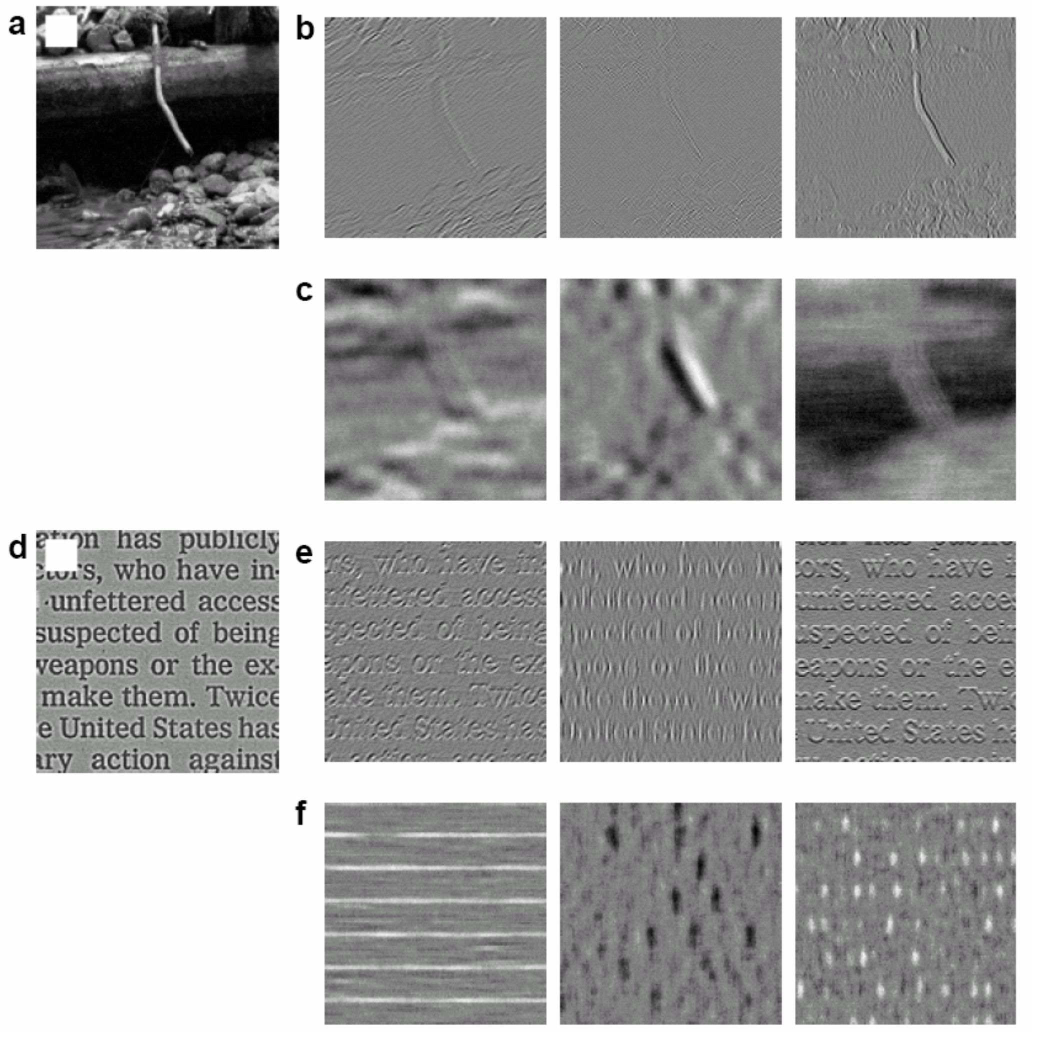
- (Karklin and Lewicki 2003)

Results: Comparison of Basis Functions



- (Karklin and Lewicki 2003)

Results: Comparison of Basis Functions



- (Karklin and Lewicki 2005)

References

- Olshausen, B. A. (2003) **Principles of Image Representation in Visual Cortex**, In: *The Visual Neuroscience,* Eds: Chalupa, L. M., Werner, J. S., MIT Press, 1603-1615.
- Hyvärinen, A. , Hoyer, P. O., Hurri, J., Gutmann, M. (2005) **Statistical models of images and early vision**, *Proceedings of the Int. Symposium on Adaptive Knowledge Representation and Reasoning (AKRR2005),* Espoo, Finland.
- Lennie, P. (2003) **The cost of cortical computation**, *Curr. Biol. 13,* 493-497.

References

- Barlow, H. P. (1961), **Possible principles underlying the transformations of sensory messages**, In: *Sensory Communication,* Ed: W. A. Rosenblith, MIT Press, 217-234.
- Srinivasan, M. V., Laughlin, S. B., Dubs, A. (1982), **Predictive coding: a fresh view of inhibition in the retina**, *Proc R Soc Lond, B, 216,* 427-459.
- van Hateren (1993), **Spatiotemporal contrast sensitivity of early vision**, *Vision Research, 33,* 257-267.

References

- Olshausen, B. A. (2002), **Sparse Codes and Spikes**, In: *Probabilistic Models of the Brain: Perception and Neural Function,* Eds: R. P. N. Rao, B. A. Olshausen, M. S. Lewicki, MIT Press, pp. 257-272.
- Olshausen B. A., Field D. J. (1996), **Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images**, *Nature, 381,* pp. 607-609.
- Olshausen B. A., Field D. J. (2005) **How close are we to understanding V1?** *Neural Computation*, in press.

References

- Karklin Y., Lewicki M. S., (2005) **A hierarchical Bayesian model for learning non-linear statistical regularities in non-stationary natural signals**, *Neural Computation, 17* (2), pp. 397-423.
- Karklin Y., Lewicki M. S., (2003) **Learning higher-order structures in natural images**, *Network: Computation in Neural Systems, 14*, pp. 483-499.