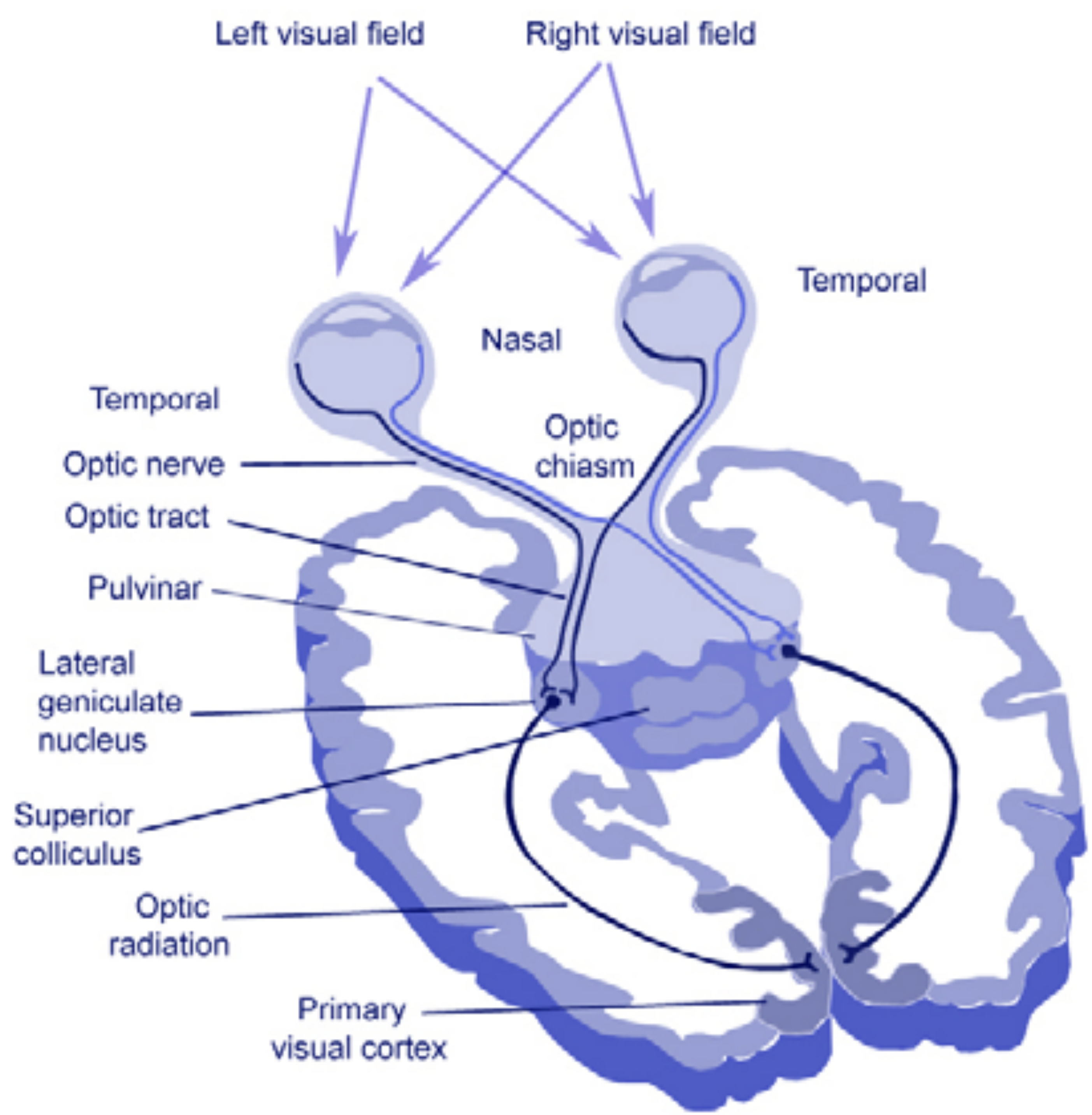
Emergence of simple-cell receptive field properties by learning a sparse code for natural images

Bruno A. Olshausen* & David J. Field

Department of Psychology, Uris Hall, Cornell University, Ithaca, New York 14853, USA

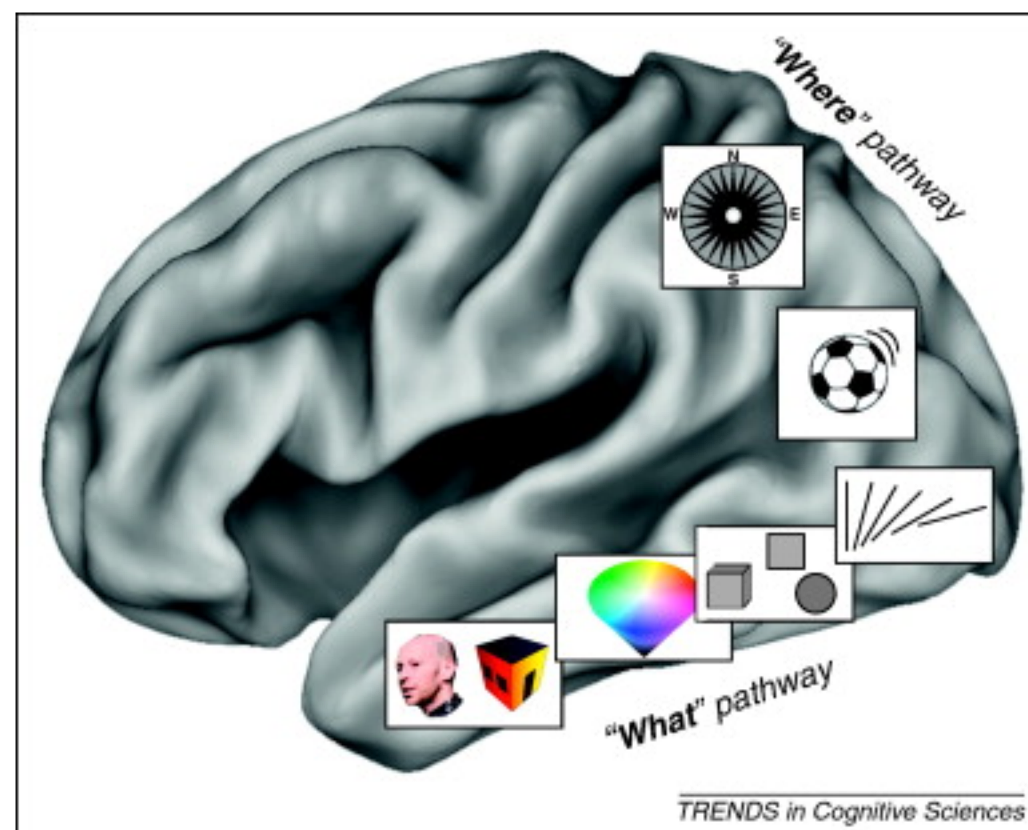
Presented by Dalin Guo
Oct. 16th

Visual Pathway



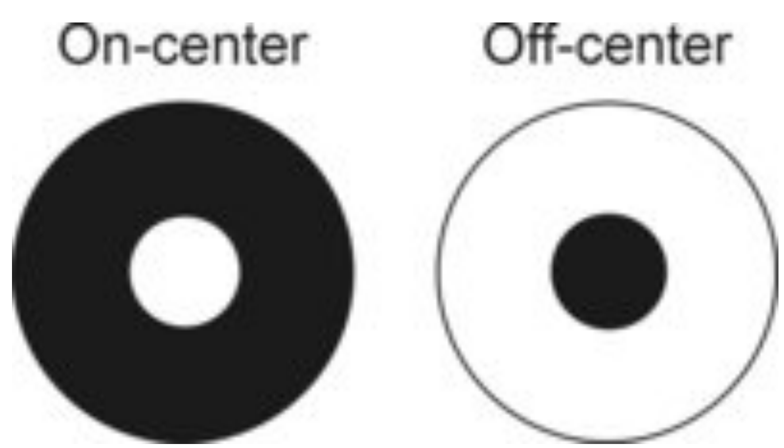
Primary Visual Cortex (V1)

- <u>First</u> stage of cortical processing of visual information
- Contains <u>a complete map</u> of the visual field
- Receive input from LGN, and output to subsequent cortical visual areas (~30)



The Receptive Fields

Definition: The particular region of the sensory space (e.g. the visual field) in which a stimulus will modify the firing of that neuron

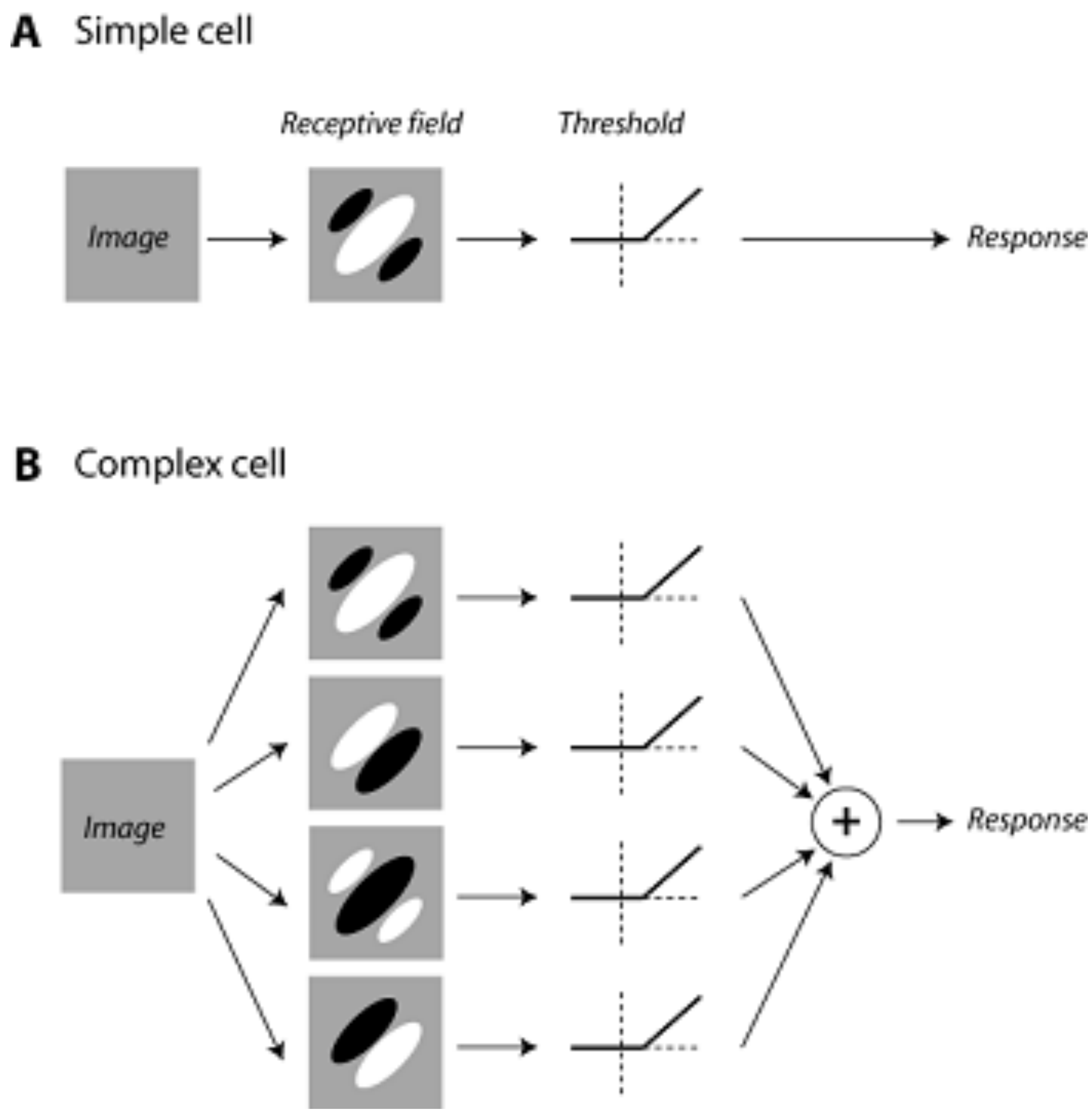


Ganglion Cell Responses: *On-center neurons respond to the presentation of a light spot on a dark background; Off-center neurons to the presentation of a dark spot on a light background.*

Primary Visual Cortex (V1)

- First stage of cortical processing of visual information
- Contains a complete map of the visual field
- Receive input from LGN, and output to subsequent cortical visual areas
- Based on the structure of receptive field, two types of neurons: simple and complex (Hubel and Wiesel, 1959)

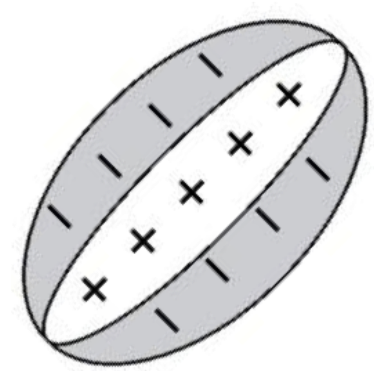
Primary Visual Cortex (V1)



Simple Cells

The receptive fields of simple cells in V1:

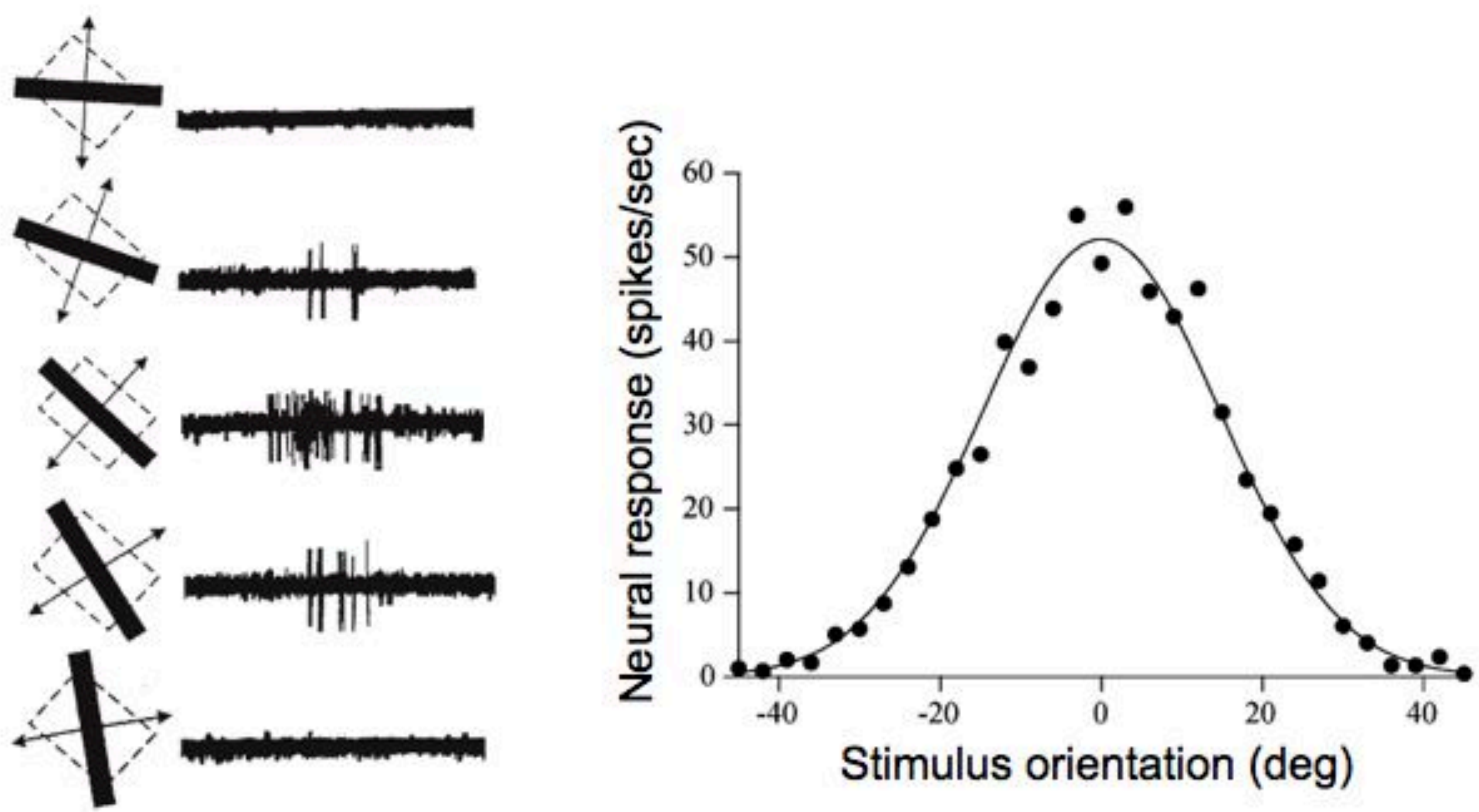
- Spatially localized
- Oriented
- Bandpass
    - Selective to structure at different spatial scales



Oriented receptive field of a neuron in primary visual cortex (V1)

Simple Cells

V1 physiology: orientation selectivity



Hubel & Wiesel, 1968

Complex Cells

The receptive fields of complex cells in V1:

- Not spatially localized
    - i.e. Moving the bar through the field produces a sustained response
- Direction-selectivity
    - Fire more when the bar moves in one direction, and are suppressed by motion in the opposite direction

Emergence of Receptive Field

Assumption: An image I(x,y) can be represented as:

$$I(x, y) = \sum_i \alpha_i \phi_i(x, y)$$

Efficient coding:

- Goal: a) a set of basis function that forms a complete code and b) results in the coefficient being statistically independent

- None has succeeded

Dataset

- **Ten** 512*512 images of natural surroundings in the American northwest
    - preprocessed by filtering with zero-phase whitening/ lowpass filter
- 16*16-pixel image patches extracted from natural scenes.

Emergence of Receptive Field

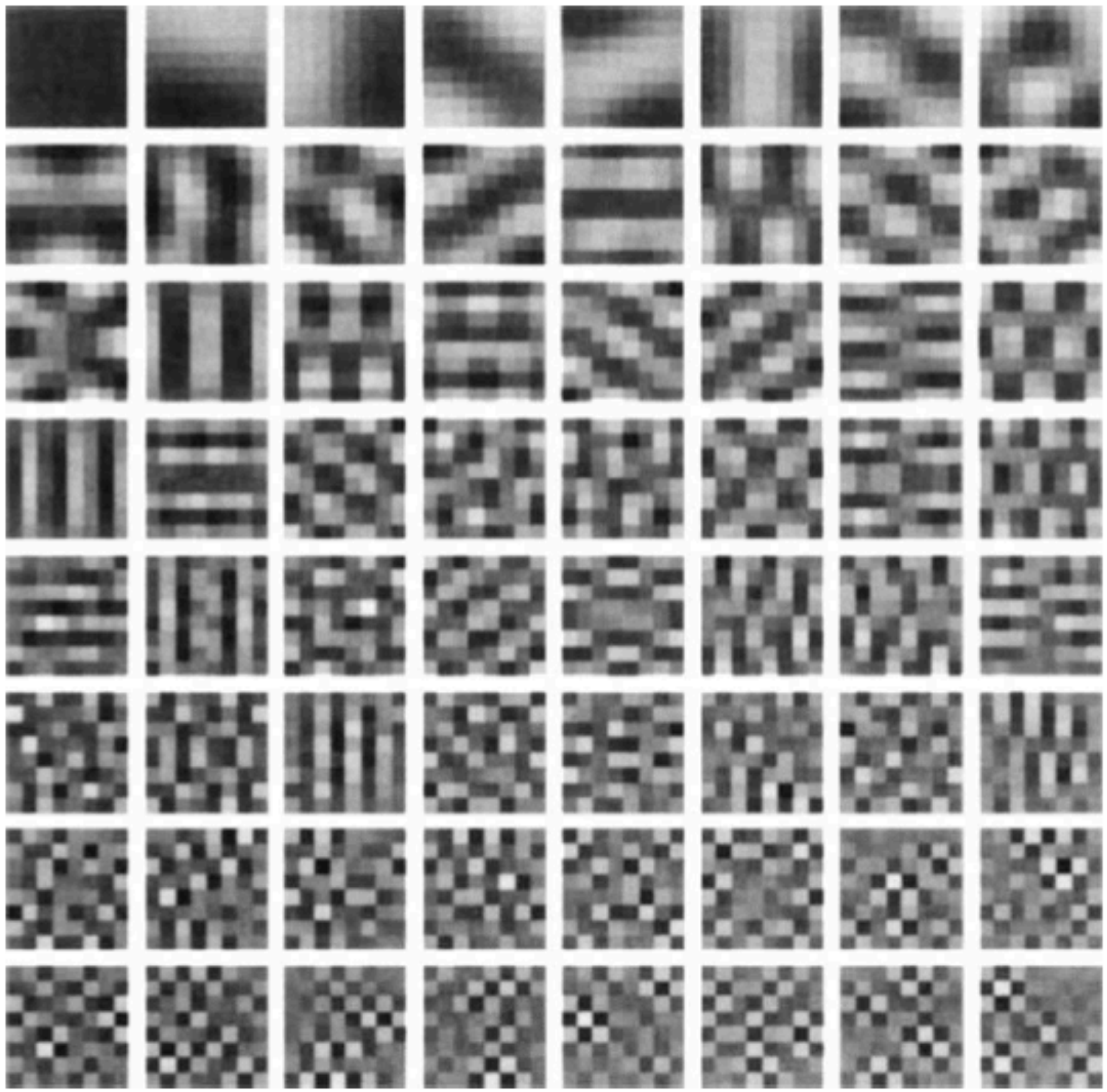
Principal Component Analysis (PCA)

<u>Goal:</u> a) a set of mutually orthogonal basis function and b) maximum variance in the data

<u>Property:</u> the coefficients are *pairwise decorrelated*

<u>Result:</u> non-localized receptive fields; majority do not at all resemble any known cortical receptive fields

PCA: Result



New Approach: Sparse Code

Algorithm: find ***sparse*** linear codes for natural scenes

Cost: E = -[preserve information] - λ[sparseness of α_i]

$$[\text{preserve information}] = -\sum_{xy} [I(x,y) - \sum_i \alpha_i \phi_i(x,y)]^2$$

$$[\text{sparseness of } \alpha_i] = -\sum_i S(\frac{\alpha_i}{\sigma})$$

Result: similar receptive fields with those three properties as in the primary visual cortex; higher degree of statistical independence among its outputs

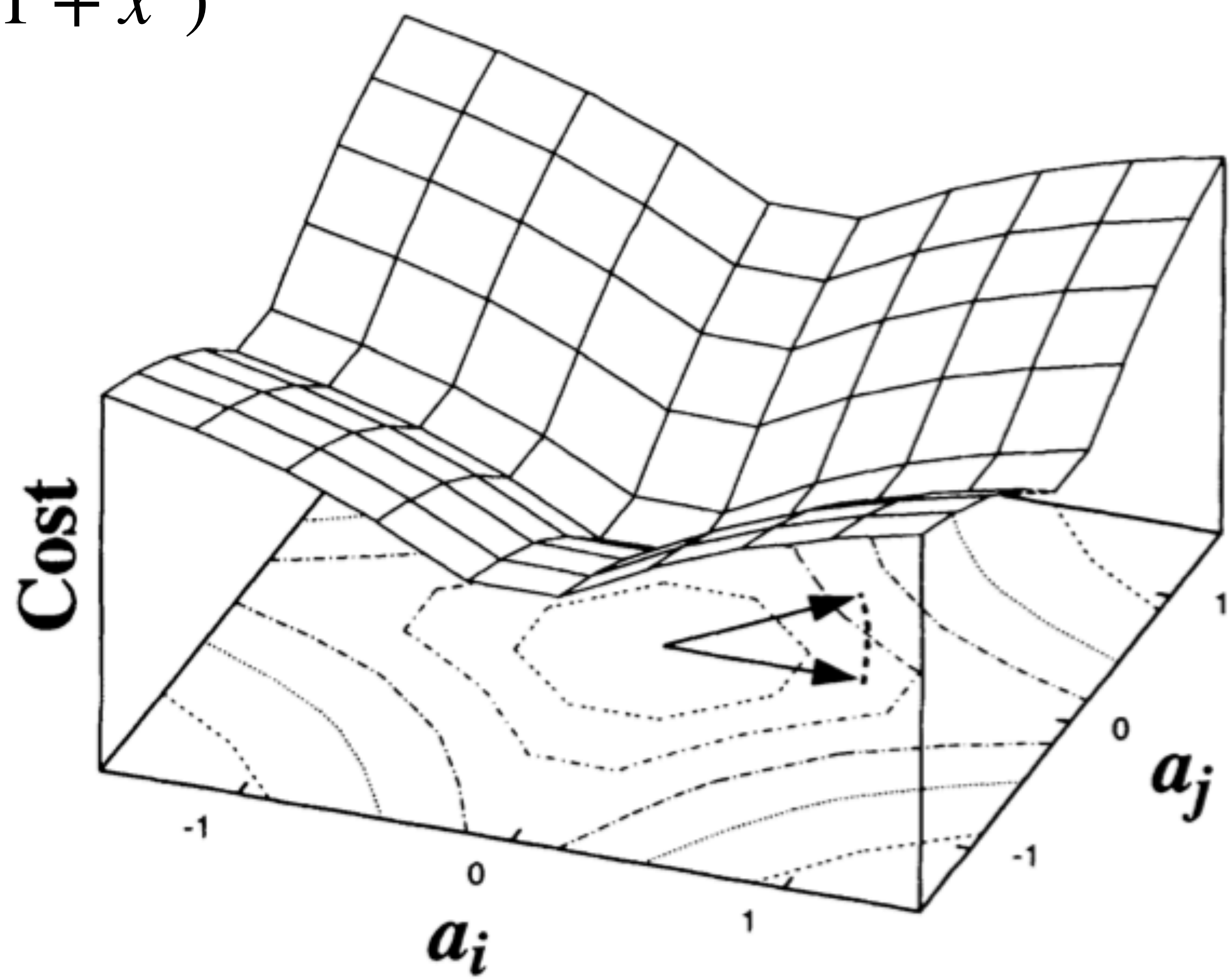
Cost for Sparseness

[sparseness of α_i] = $-\sum_i S(\frac{\alpha_i}{\sigma})$

$S(x) = \log(1 + x^2)$

$S(x) = -e^{-x^2}$

$S(x) = |x|$



Algorithm

Cost function: $E = \sum_{xy} [I(x, y) - \sum_i \alpha_i \phi_i(x, y)]^2 + \lambda \sum_i S(\frac{\alpha_i}{\sigma})$

α_i are determined from the equilibrium solution to the differential equation:

$$\dot{\alpha}_i = b_i - \sum_j C_{ij}\alpha_j - \frac{\lambda}{\sigma} S'(\frac{\alpha_i}{\sigma})$$

$$b_i = \sum_{x,y} \phi_i(x, y) I(x, y) \quad C_{ij} = \sum_{x,y} \phi_i(x, y)\phi_j(x, y)$$

The learning rule for ϕ

$$\Delta\phi_i(x_m \,.\, y_n) = \eta < \alpha_i[I(x_m, y_n) - \hat{I}(x_m, y_n)] >$$

$$\hat{I}(x_m, y_n) = \sum_i \alpha_i \phi_i(x_m, y_n)$$

Algorithm

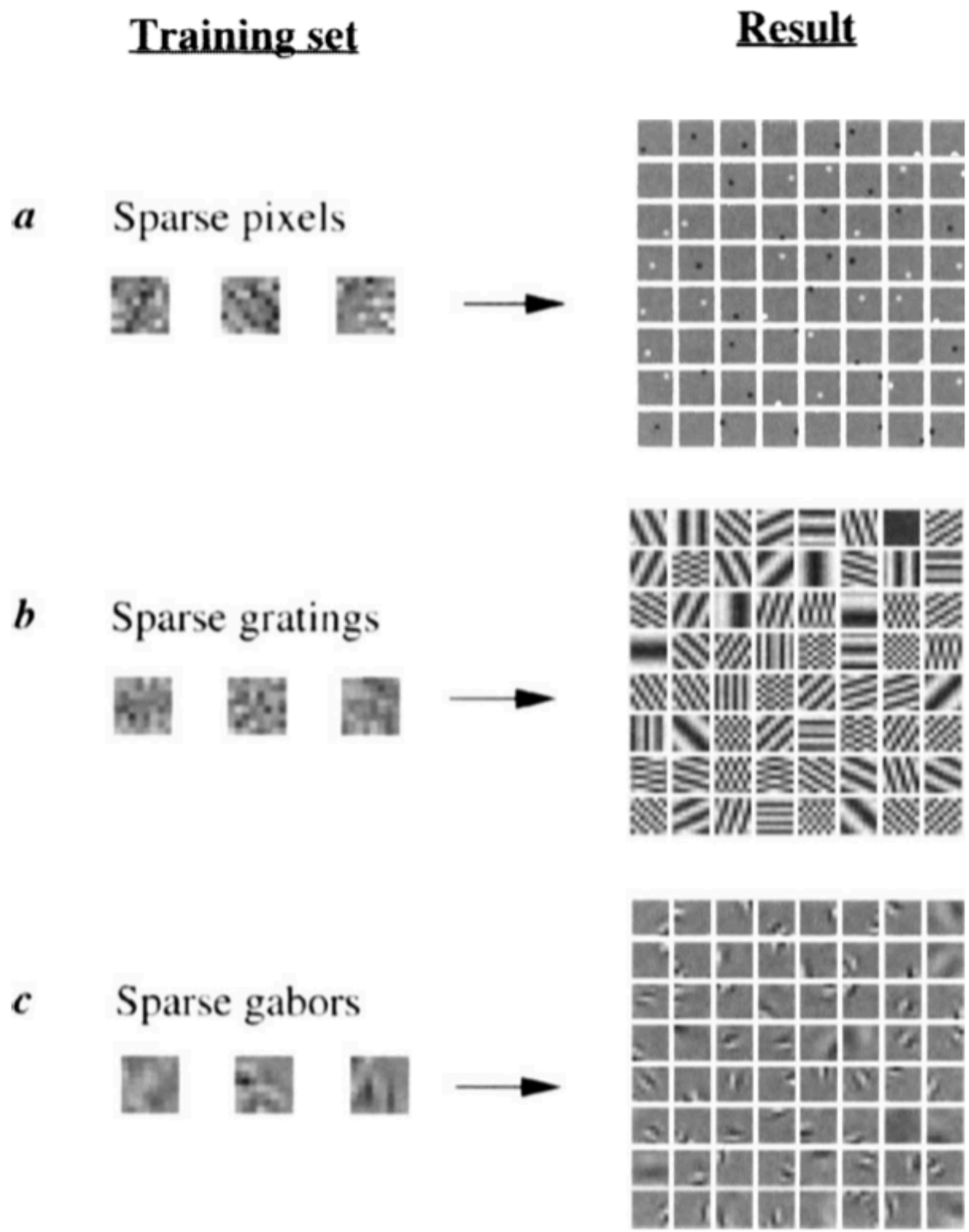
α_i are computed by the conjugate gradient method, halting when the change in the cost function is less than 1%

ϕ_i are initialized to random values and are updated every 100 images

The vector length (gain) of each basis function is adapted over time so as to maintain equal variance

$\lambda/\sigma = 0.14$, and σ^2 set to the variance of the images

Recovered Sparse Structure

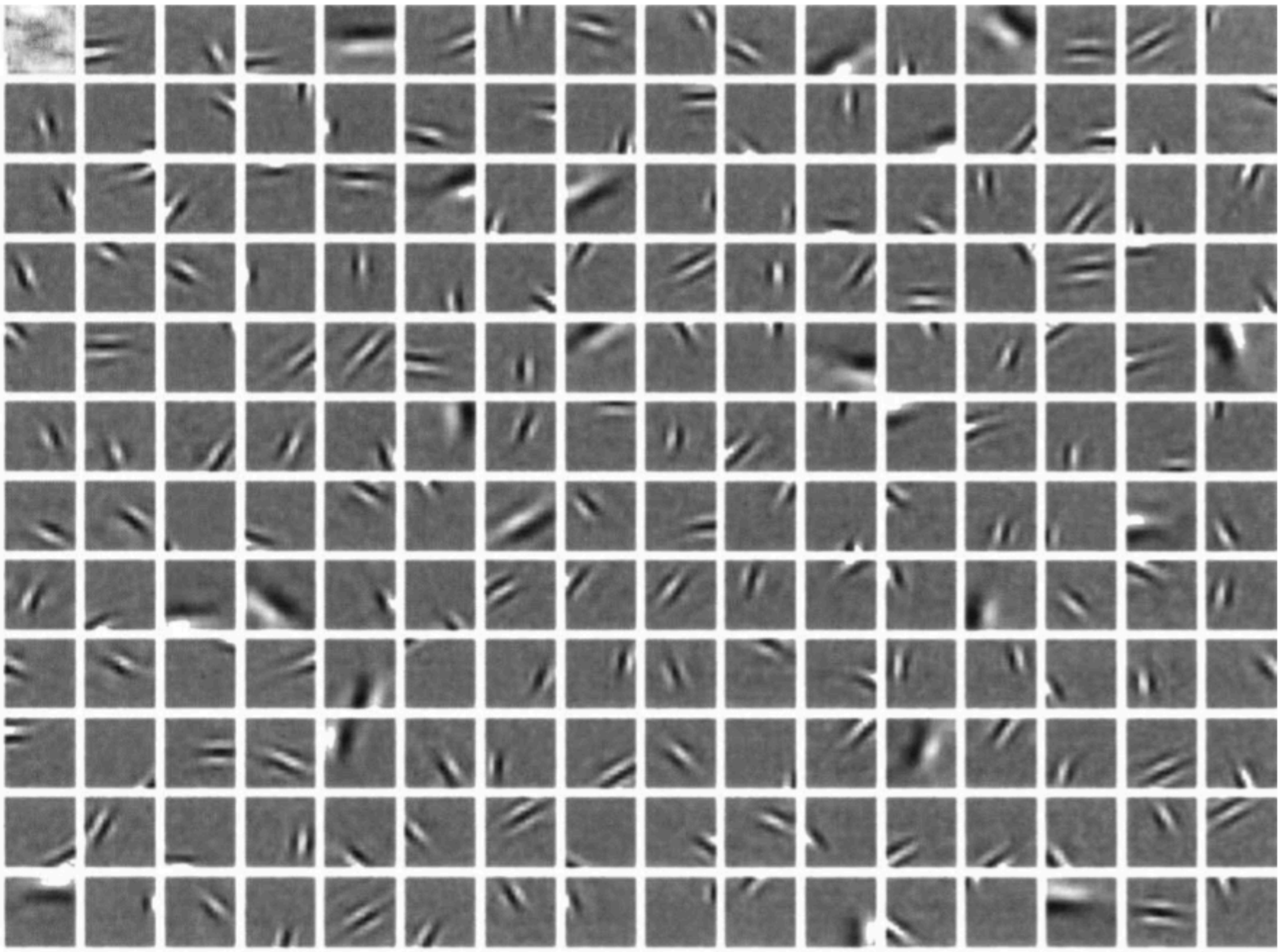


a. each pixel activated
$P(x) = e^{-|x|}/Z$

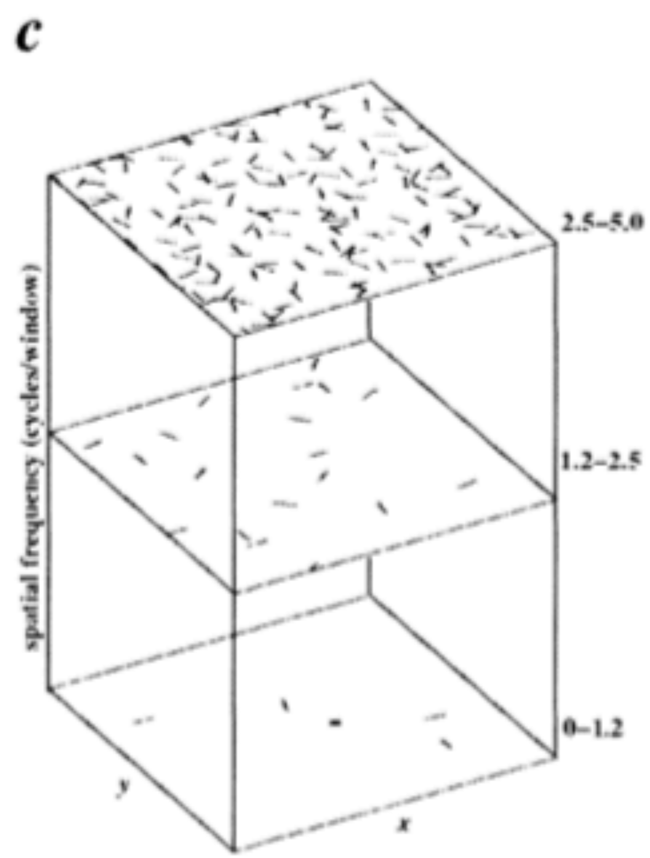
b. ‘Sparse pixels’ in Fourier domain

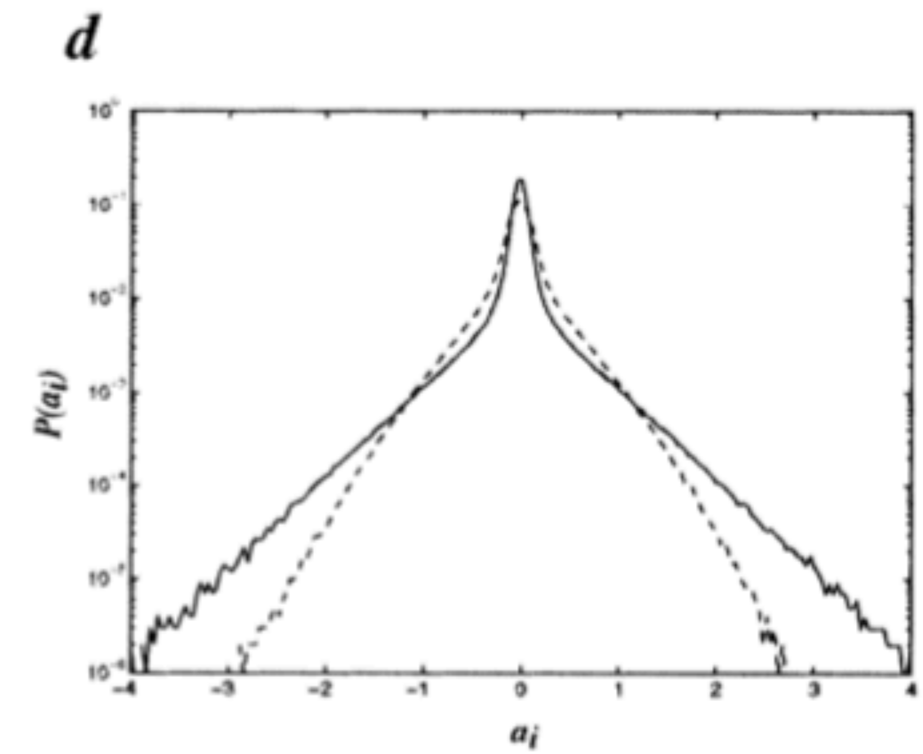
c. sparse, non-orthogonal Gabor functions

Recovery Result



Recovery Result





c. The distribution of the learned basis functions in space, orientation and scale

d. Histograms of all coefficients for the learned basis function (solid) and for random initial condition (broken)

Summary

- Localized, oriented, bandpass receptive fields emerge with two objectives:
  - Information preserved
  - Sparse representation
- Further challenge:
  - Other properties of simple cells (e.g. direction selectivity)
  - complex response at later stages (e.g. nonlinear)
  - By remaining statistical dependence in natural images?