Beyond Edges:
Multi-layer Sparse Coding

Jekanthan Thanga.

Space Robotics Group
Institute for Aerospace Studies
University of Toronto

Outline

- P. O. Hoyer and A. Hyvarinen. A Multi-Layer Sparse Coding Network Learns Contour Coding from Natural Images

  - Introduction
  - Information Theoretic vs Probabilistic Internal Model
  - Why Sparse Coding ?
  - Cell Models (Simple, Complex, Contour)
    - Overview, Experiments
  - Demos
  - Underlying Issues & Conclusion

Introduction

- Simulation of V1 cortical simple cells, complex cells and other higher order neurons*
- Result of evolution and neural learning
- Determine hierarchical role of the neurons in image processing
- Emergence of higher order functionality through unsupervised training

Information Theoretic vs Probabilistic Internal Model

- Role of Lower level neurons maybe understood from an information theoretic view
- But this may not hold for higher neurons.
- Is the goal to capture/store image data efficiently ?

  or
- Efficient modeling of data for specific task, such as determining shape, direction

Why Sparse Coding ?

- Conjecture: Natural image may be described using few primitives (Field, 1994)
- Evidence of sparse coding using Gabor filters (shows high kurtosis)
- Potential advantages in storage and retrieval using associative memory (Field,1997)
  - Increased Capacity (Baum, Moody & Wilczek, 1988)
  - Wiring length (Foldiak, 1995)
  - Metabolic Efficiency (Baddeley, 1996)

Cortical Simple Cell Model

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i s_i + \mathbf{n}$$

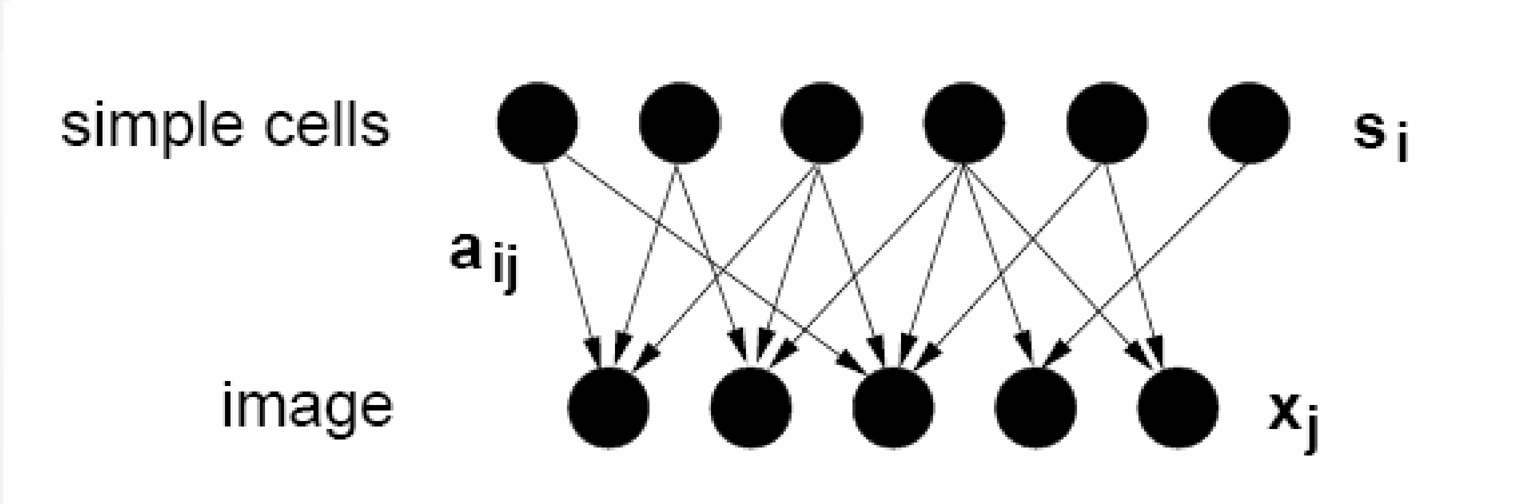
$$a_{ij} = \frac{1}{2\pi\sigma_x\sigma_y} exp[-\frac{\tilde{x}^2}{2\sigma_x^2} - \frac{\tilde{y}^2}{2\sigma_y^2}] \cos(k\tilde{x} - \phi)$$

$$\tilde{x} = x\cos(\theta) + y\sin(\theta)$$

$$\tilde{y} = -x\sin(\theta) + y\cos(\theta)$$

- Position, direction, phase sensitive
- Mutually independent and s_i sparsely activated
- a_{ij} Gabor functions
- x_j image data
- n_j Gaussian Noise

Simple Cell Model



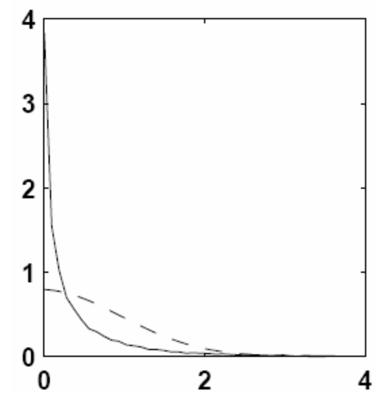
Simple Cell Training

- Unsupervised
- Gradient descent learning
- Minimize:

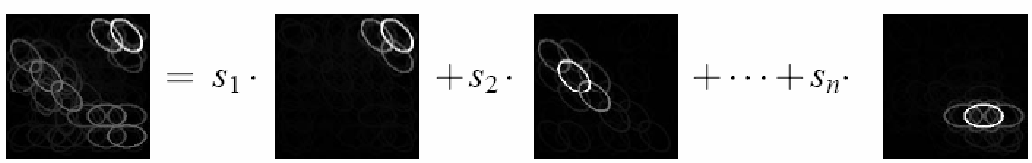
$$C(\mathbf{A}, \mathbf{s}^{(n)}, n = 1 \ldots N) = \sum_{n} \left[\|\mathbf{x}^{(n)} - \mathbf{A}\mathbf{s}^{(n)}\|^2 + \lambda \sum_{i} s_i^{(n)} \right]$$

- Learned on patches of 100,000 natural images
- Short term Learning objective:
  - Sparse config of s_i
- Long term
  - generative weight a_{ij} pattern match data with max prob.

Sparse Coding



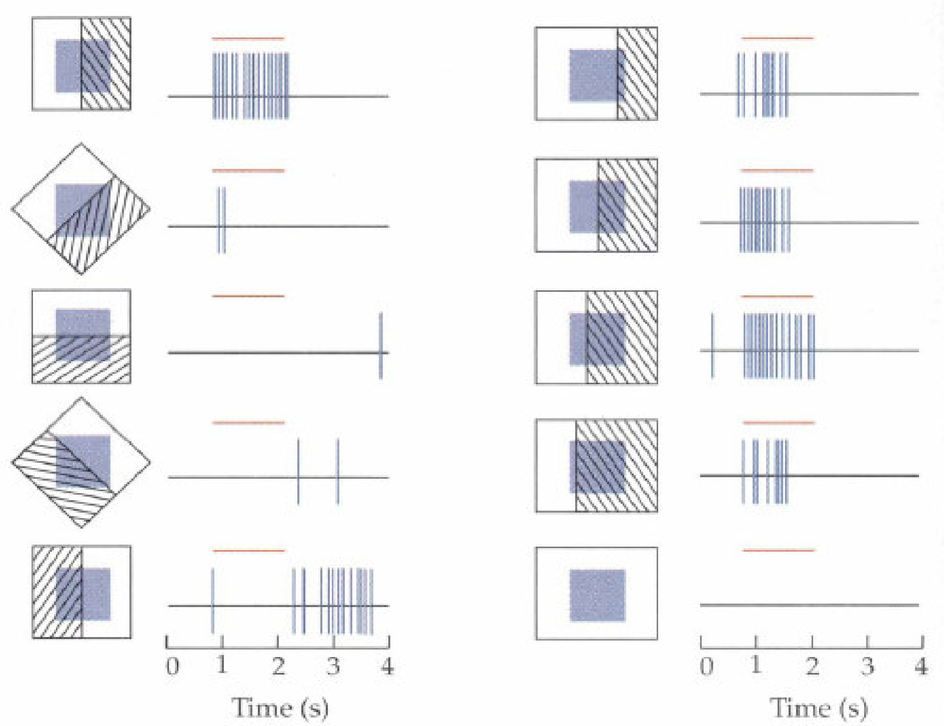
- Property of scale invariance
- Highly peaked at zero, with heavy tail
- Change in response histogram represented by Kurtosis
- Implies data described with few active s_i



- Still need all the basis pattern a_{ij} to select from

Cortical Complex Cells

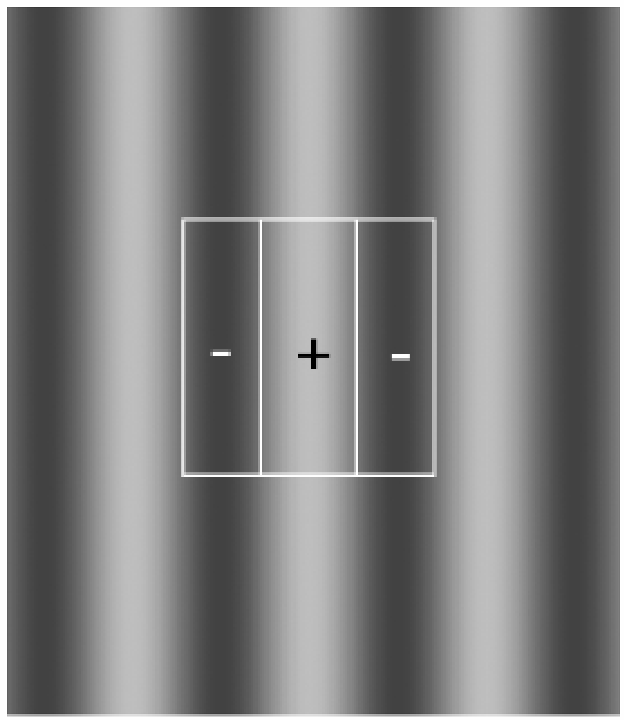
- Not sensitive to position or phase
- Can't use linear model
- Simple energy model:



Hubel, D. H. and T. N. Wiesel (1962).

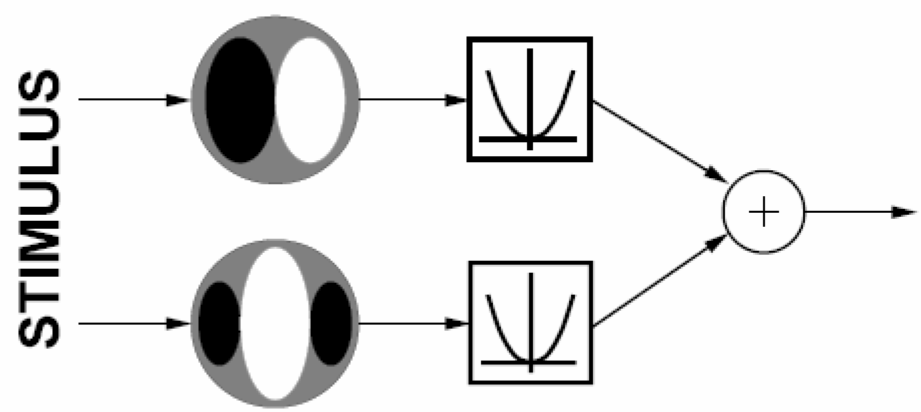
Cortical Complex Cells

- Not sensitive to position or phase
- Can't use linear model
- Simple energy model:



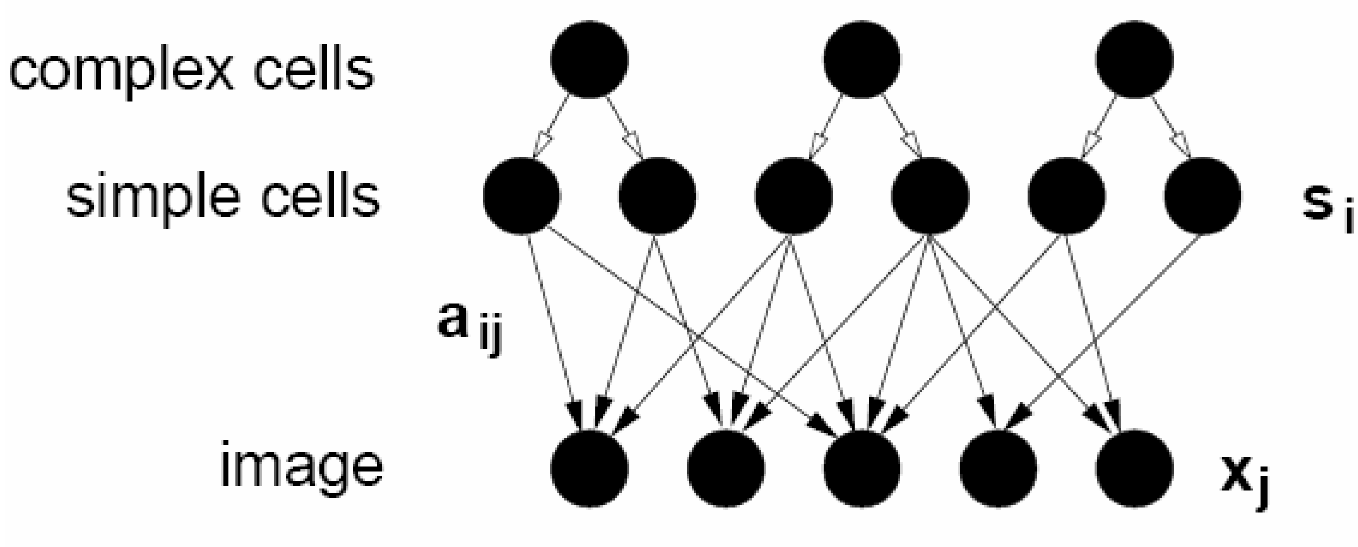
Cortical Complex Cells

- Not sensitive to position or phase
- Can't use linear model
- Simple energy model:



Cortical Complex Cells

- Not sensitive to position or phase
- Can't use linear model
- Simple energy model:



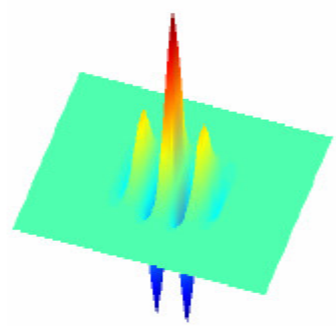
2-D Spatial Gabor Function

$$D_e(x,y,\theta) = \frac{1}{2\pi\sigma_x\sigma_y} exp[-\frac{\tilde{x}^2}{2\sigma_x^2} - \frac{\tilde{y}^2}{2\sigma_y^2}] \cos(k\tilde{x} - \phi)$$

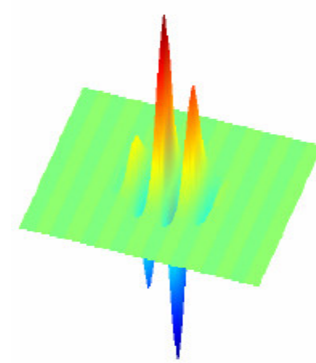
$$D_o(x,y,\theta) = \frac{1}{2\pi\sigma_x\sigma_y} exp[-\frac{\tilde{x}^2}{2\sigma_x^2} - \frac{\tilde{y}^2}{2\sigma_y^2}] \sin(k\tilde{x} - \phi)$$

$$\tilde{x} = x\cos(\theta) + y\sin(\theta)$$
$$\tilde{y} = -x\sin(\theta) + y\cos(\theta)$$

- θ Orientation
- k frequency along x-axis
- ϕ Phase

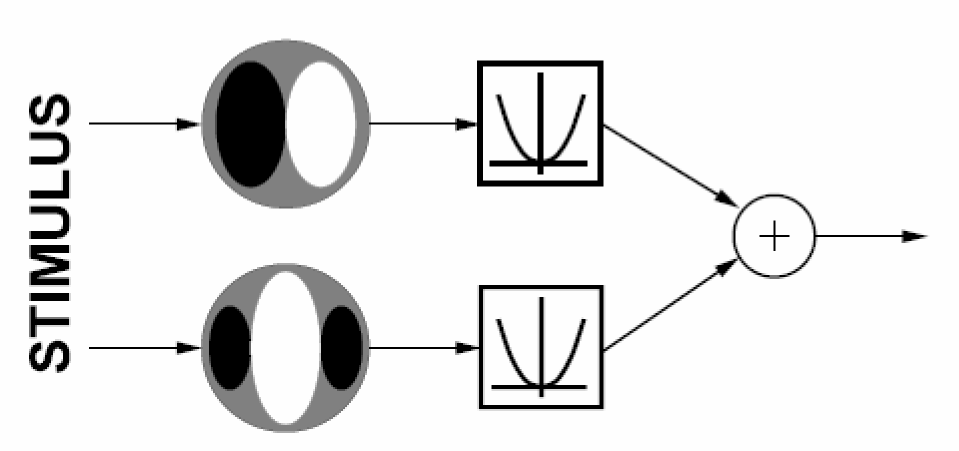


Even Symmetric
Cosine Gabor



Odd Symmetric
Sine Gabor

Complex Cell Model



$$C(x_c, y_c, \theta) = \left(\sum_{x_i, y_i} (D_{e,x_i,y_i}(x_c, y_c, \theta) I_{x_i,y_i}) \right)^2 + \left(\sum_{x_i, y_i} (D_{o,x_i,y_i}(x_c, y_c, \theta) I_{x_i,y_i}) \right)^2$$

- x_c, y_c Center of filter
- x_i, y_i Image patch coordinates
- D_e – Even symmetric spatial Gabor filter
- D_o – Odd symmetric spatial Gabor filter
- θ Orientation

Phase Invariance

$$D_e(x,y,\theta) = G(x,y,\theta)\cos(k\tilde{x}-\phi)_+$$
$$D_o(x,y,\theta) = G(x,y,\theta)\sin(k\tilde{x}-\phi)_+$$

$$C(x_c,y_c,\theta) = \left(\sum_{x_i,y_i}(D_{e,x_i,y_i}(x_c,y_c,\theta)I_{x_i,y_i})\right)^2 + \left(\sum_{x_i,y_i}(D_{o,x_i,y_i}(x_c,y_c,\theta)I_{x_i,y_i})\right)^2$$

$$C(x_c,y_c,\theta) = \sum_{x_{i1},y_{i1}}(D_{e,x_{i1},y_{i1}}(x,y,\theta)I_{x_{i1},y_{i1}})\cdot\sum_{x_{i2},y_{i2}}(D_{e,x_{i2},y_{i2}}(x,y,\theta))I_{x_{i2},y_{i2}})_+$$
$$+\sum_{x_{i1},y_{i1}}(D_{o,x_{i1},y_{i1}}(x_c,y_c,\theta)I_{x_{i1},y_{i1}})\cdot\sum_{x_{i2},y_{i2}}(D_{o,x_{i2},y_{i2}}(x_c,y_c,\theta)I_{x_{i2},y_{i2}})_+$$

$$C(x_c,y_c,\theta) = \sum_{x_{i1},y_{i1}}\sum_{x_{i2},y_{i2}}\left((G_{x_{i1},y_{i1}}(x,y,\theta)I_{x_{i1},y_{i1}})\cdot(G_{x_{i2},y_{i2}}(x,y,\theta)I_{x_{i2},y_{i2}})\cos^2(k\tilde{x}-\phi)\right)_+$$
$$+\sum_{x_{i1},y_{i1}}\sum_{x_{i2},y_{i2}}\left((G_{x_{i1},y_{i1}}(x,y,\theta)I_{x_{i1},y_{i1}})\cdot(G_{x_{i2},y_{i2}}(x,y,\theta)I_{x_{i2},y_{i2}})\sin^2(k\tilde{x}-\phi)\right)_+$$

$$C(x_c,y_c,\theta) = \sum_{x_{i1},y_{i1}}\sum_{x_{i2},y_{i2}}((G_{x_{i1},y_{i1}}(x,y,\theta)I_{x_{i1},y_{i1}})\cdot(G_{x_{i2},y_{i2}}(x,y,\theta)I_{x_{i2},y_{i2}})$$

Invariant to spatial phase, sensitive to frequency and orientation but not within RF

Complex Cell Training

Learning Goal:

- Few active complex cells
  - sparseness
- Adaptation of a_{ij} so that prob. of data maximized
- a_{ij} and s_i non-negative values

Details

Complex Cell Training Details

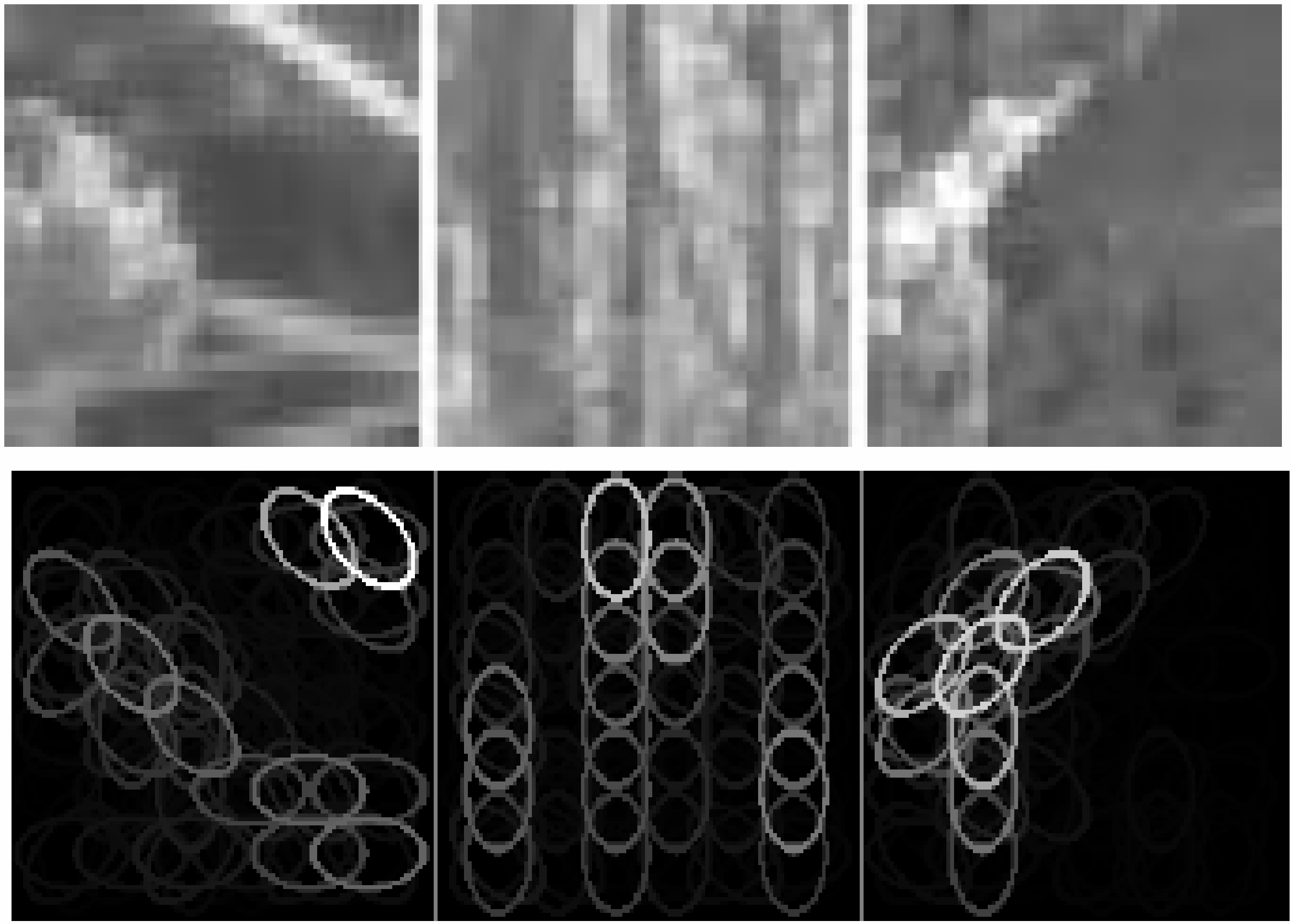
- Gradient descent learning
- Minimize:

$$C(\mathbf{A}, \mathbf{s}^{(n)}, n = 1 \ldots N) = \sum_n \left[\|\mathbf{x}^{(n)} - \mathbf{A}\mathbf{s}^{(n)}\|^2 + \lambda \sum_i s_i^{(n)} \right]$$

- Iteratively seeking w_i that minimize:

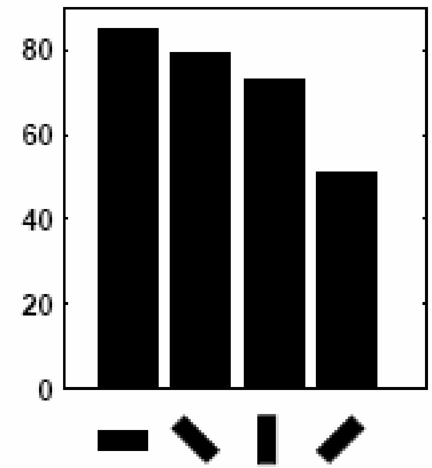
$$\sum_n \{ (s_i^{(n)} - \mathrm{rect}(\mathbf{w}_i^T \bar{\mathbf{x}}^{(n)}))^2 \}$$

Results

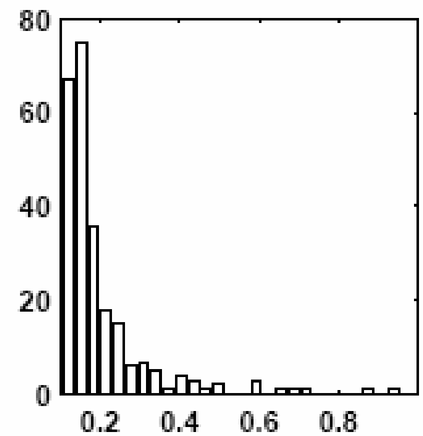


Results

- Co-linearity
- Characteristics of basis pattern
- Distribution of Pattern Lengths
  - Quadratic drop in frequency for increased length size
  - Captures scale invariance of natural images
- Lack of co-circularity
  - May need even higher order functionality
  - Could also be the dataset



Orientation Histogram.



Length Distribution

Results

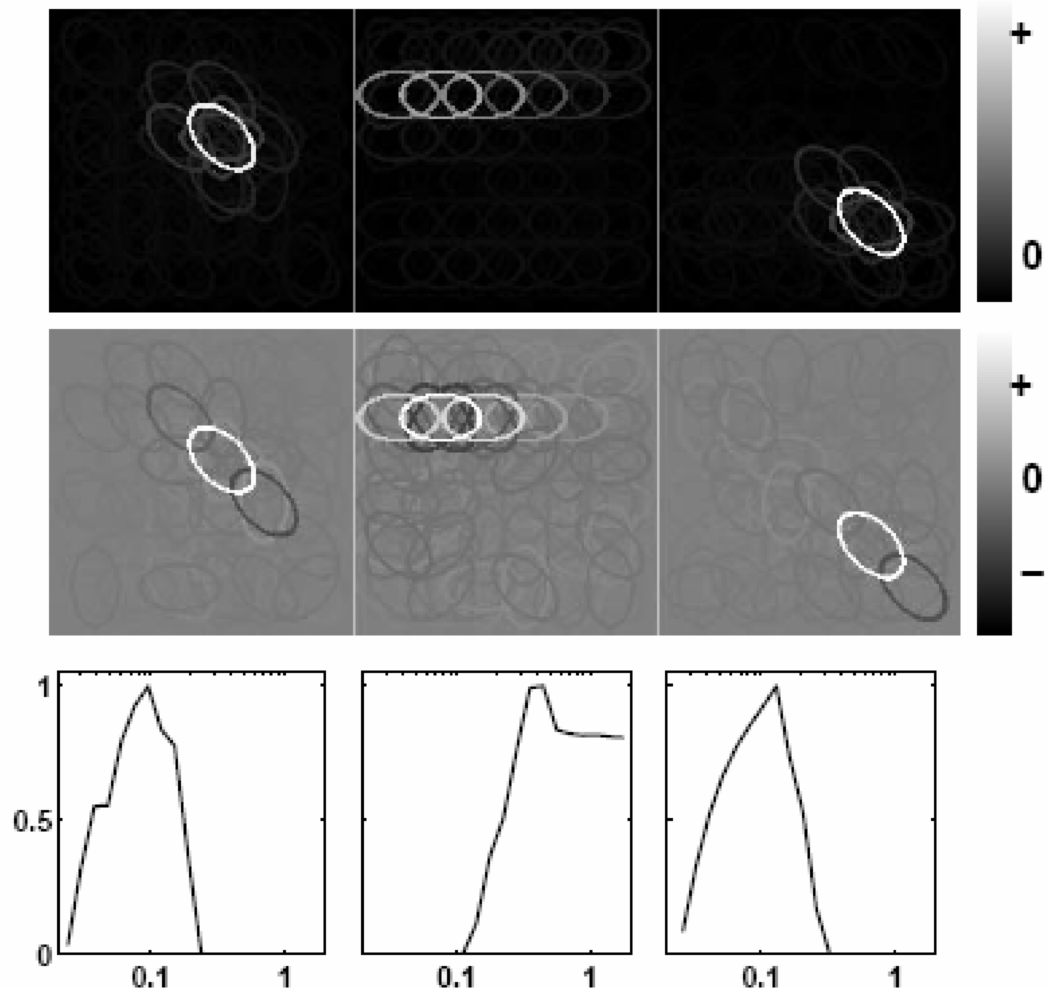
- Sparseness results in competition for better representation of short and long contours
- Basis patterns selective for contour length
- Long don't respond to short, vice versa
- Short detectors develop limited end stopping characteristics
- Long detectors develop cross inhibitory characteristics

Results

Basis Patterns

Min. mean square error, half-rectification

Stimulus



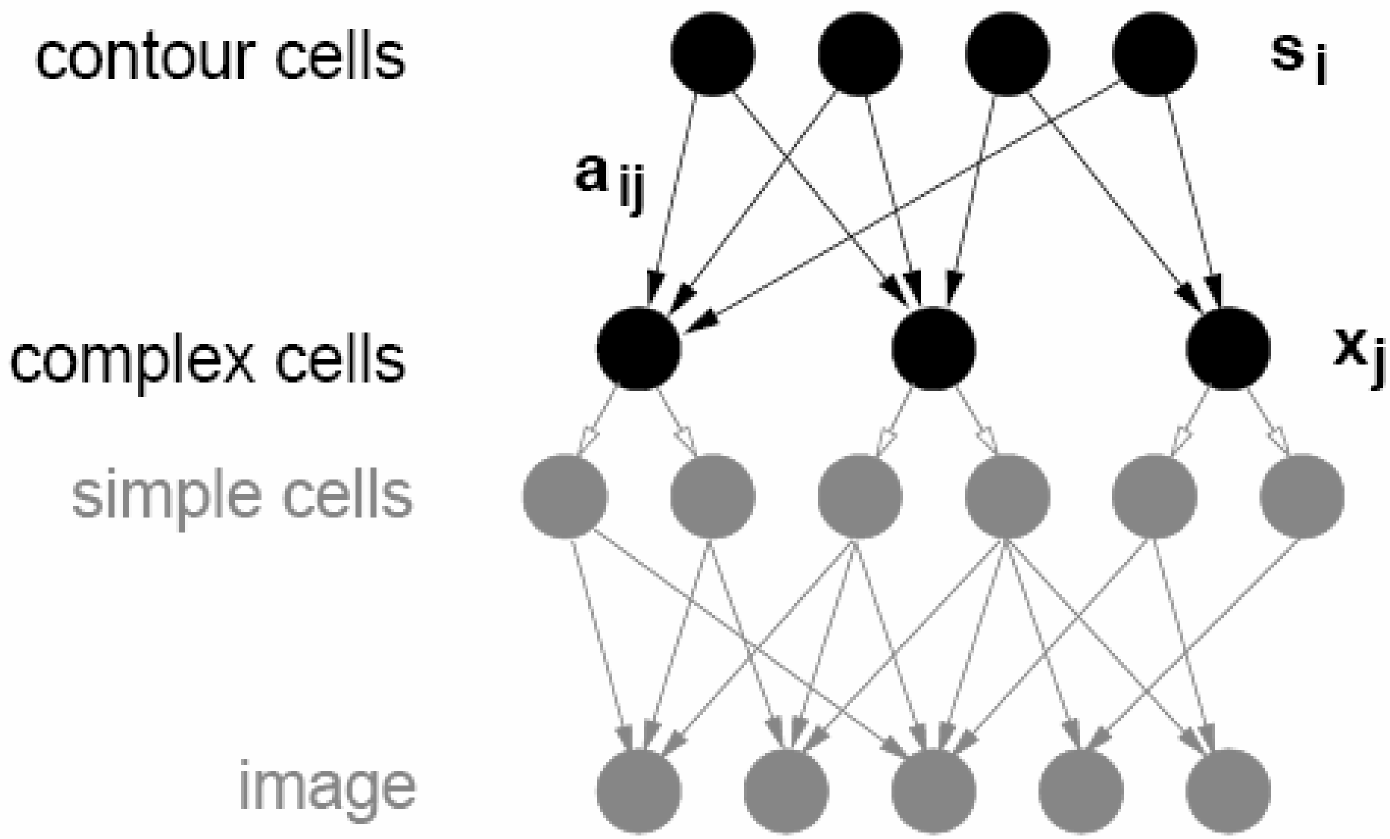
Contour Integration

- Extend architecture to include feedback
- Feedback
  - When signal faint/contradictory
  - Resolve to mostly likely interpretation
  - Bottom up sensory information, generative feedback response
- Feedback used as a form of noise reduction

Contour Cell Model

Simple Cell Model

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i s_i + \mathbf{n}$$



Contour Cell Training Details

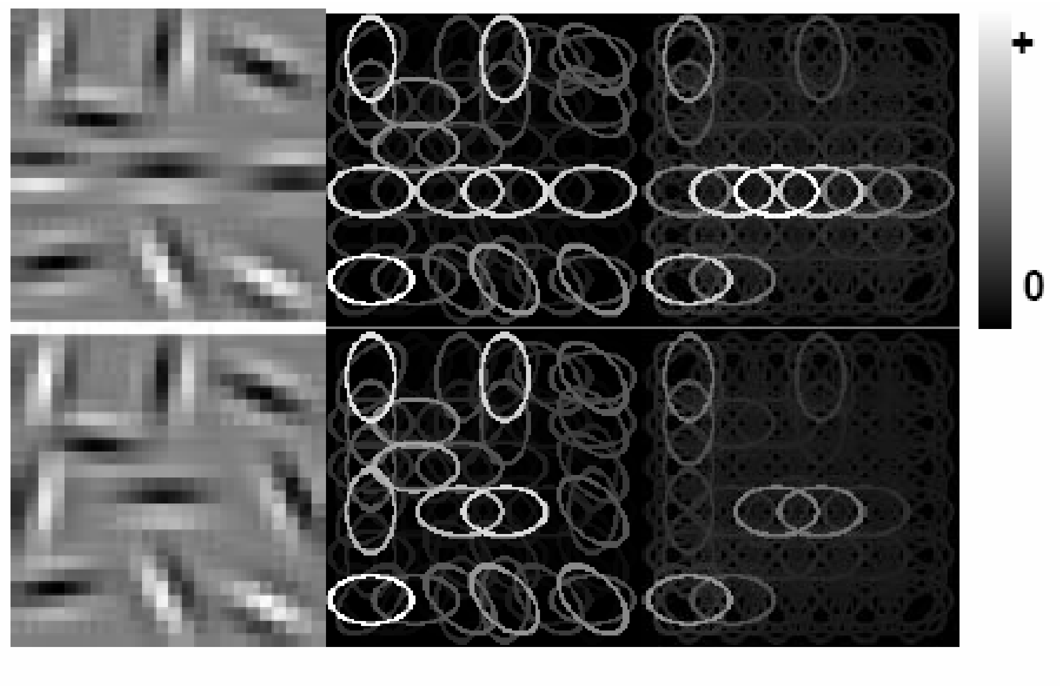
- Minimize with respect to s:

$$\|\mathbf{x} - \mathbf{A}\mathbf{s}\|^2 + \lambda \sum_i s_i$$

- λ - based on representation error and sparseness – noise level fudge factor
- Fixed simple cell layer.

Contour Integration

- Captures smooth contours
- End stopping characteristics
- Suppression of spurious edges



Complex Contour

Demos

- Simple Cell Network
- Complex Cells
- Contour Cells (Feedback model)

Underlying Issues

- Co-linearity dominant, lack of co-circularity
- Role of Sparse coding framework
- Role/location of contour coding units
  - End stopping evident at two different layers
  - Feedback could be handled by horizontal connections
  - Need for extra layer ?

Questions & Comments