Scene Statistics
Part 1b

Odelia Schwartz
University of Miami
Summer 2016

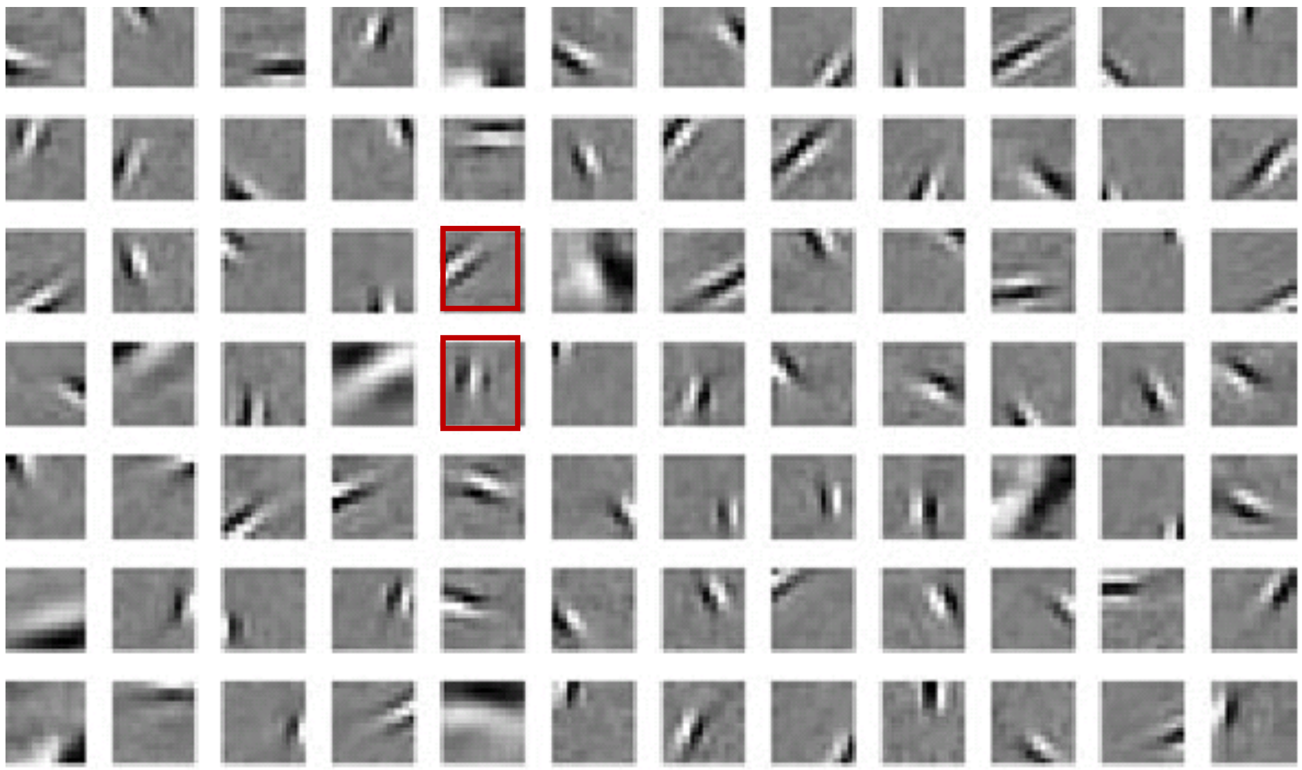
Summary

- We've considered bottom-up scene statistics, efficient coding, and relation of linear transforms to visual filters

- This class: nonlinearities

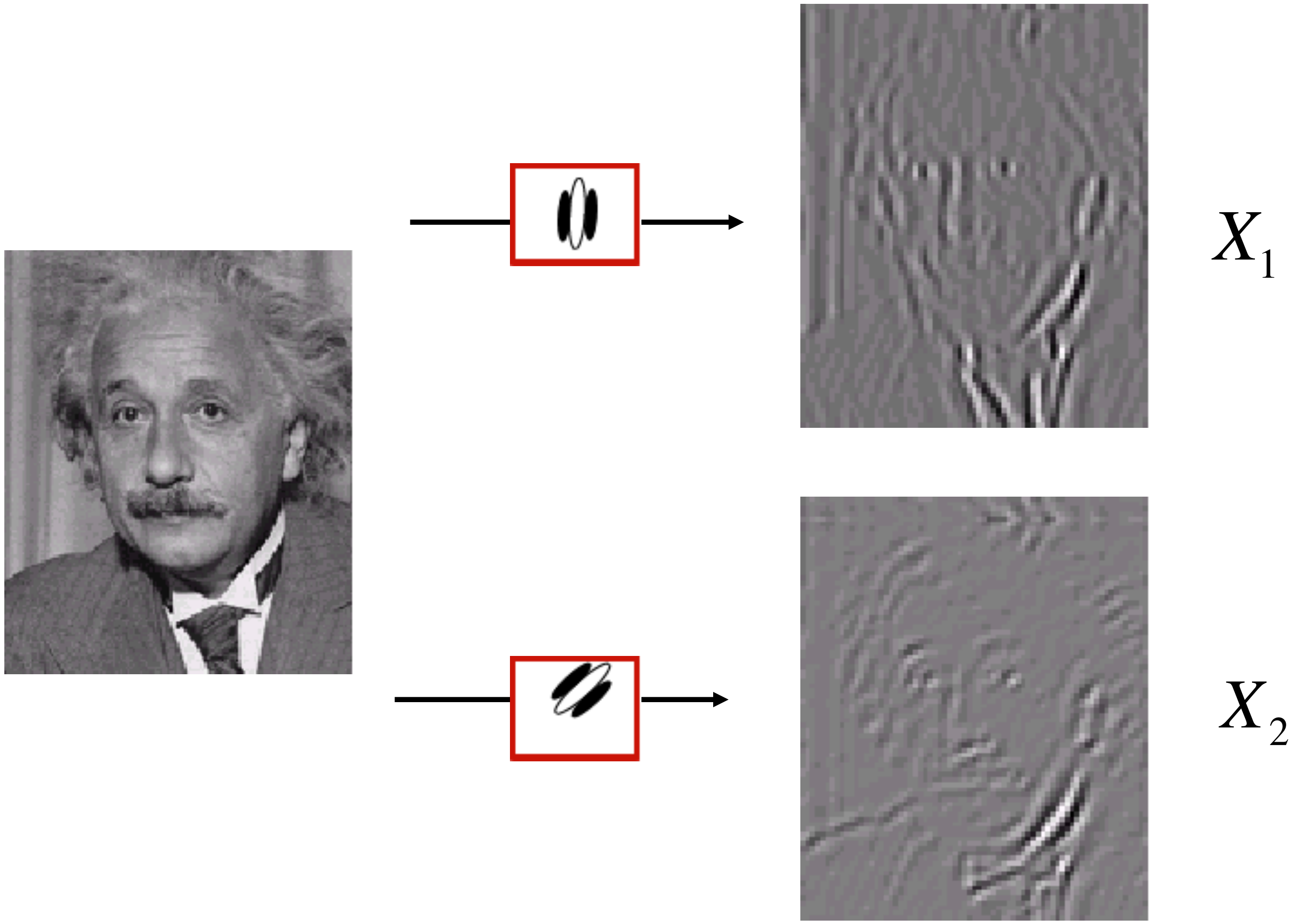
- This class: generative, top-down, perspective

Beyond linear

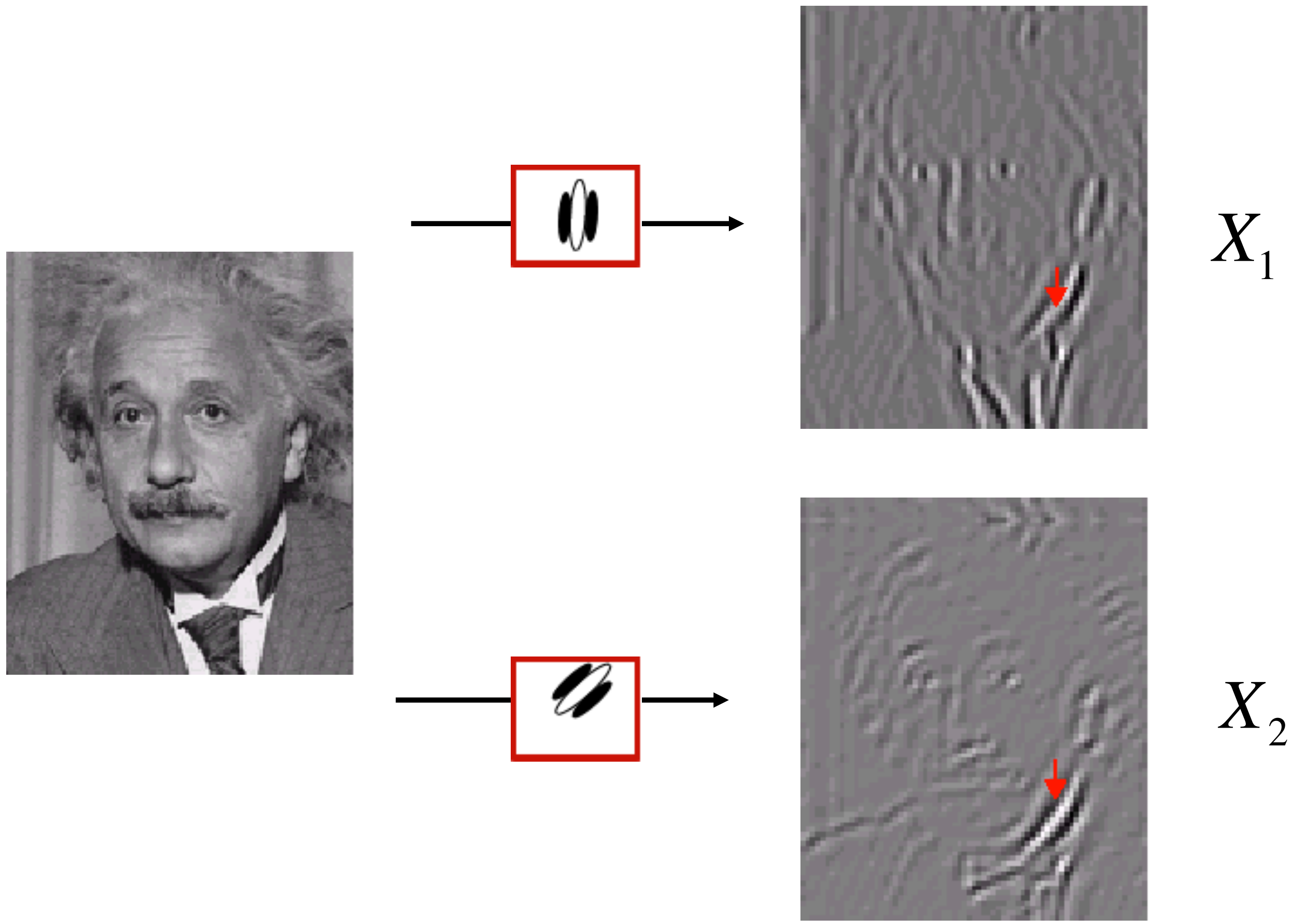


- Filter responses as independent as possible assuming a linear transform
- But are they independent?

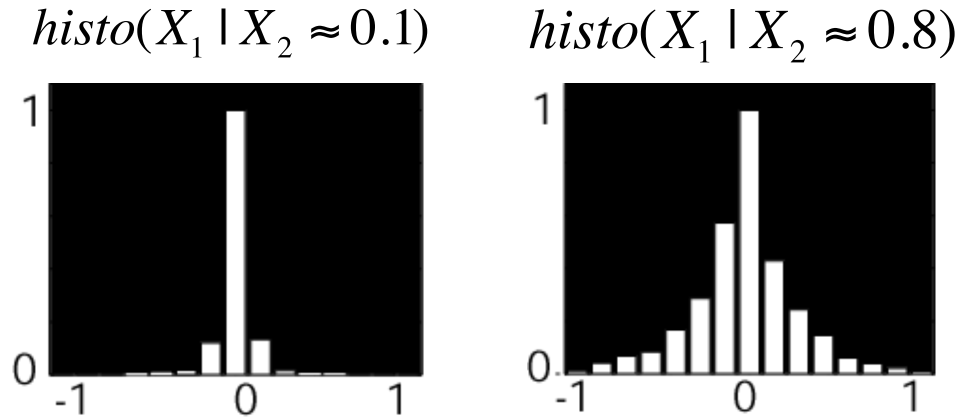
Bottom-up Joint Statistics

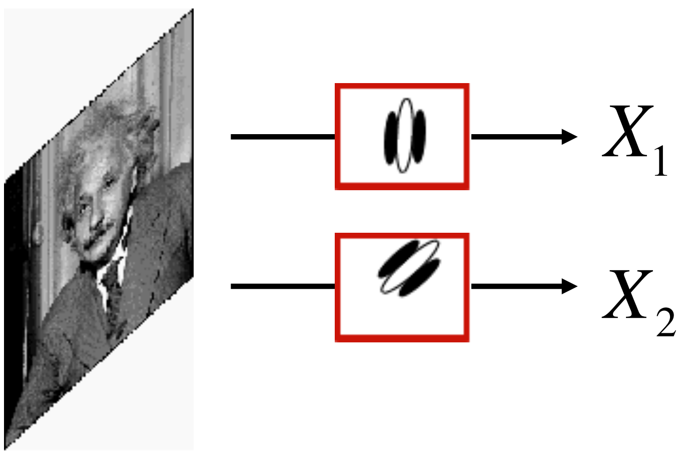


Bottom-up Joint Statistics



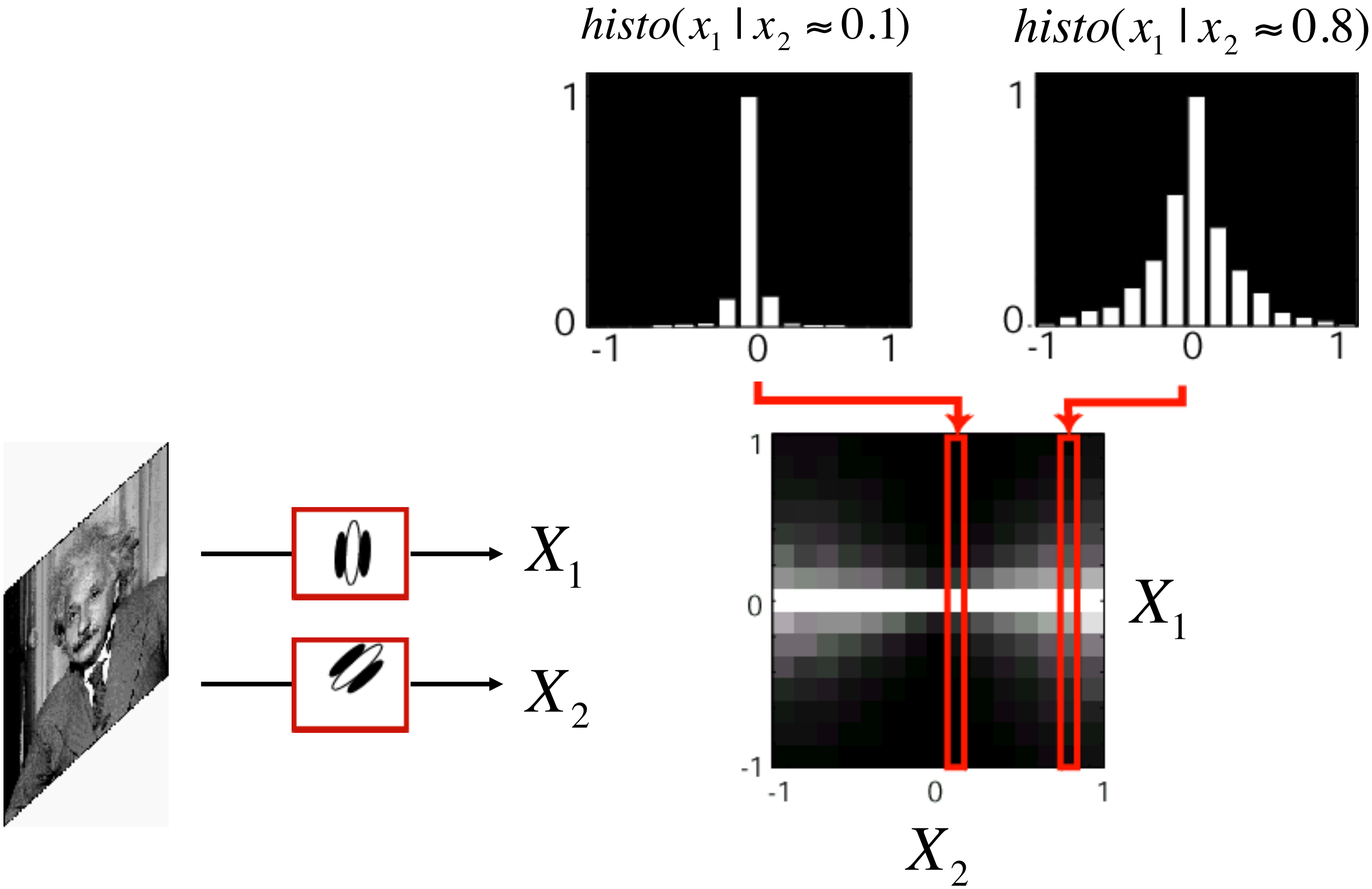
Bottom-up Joint Statistics





Are X_1 and X_2 statistically independent?

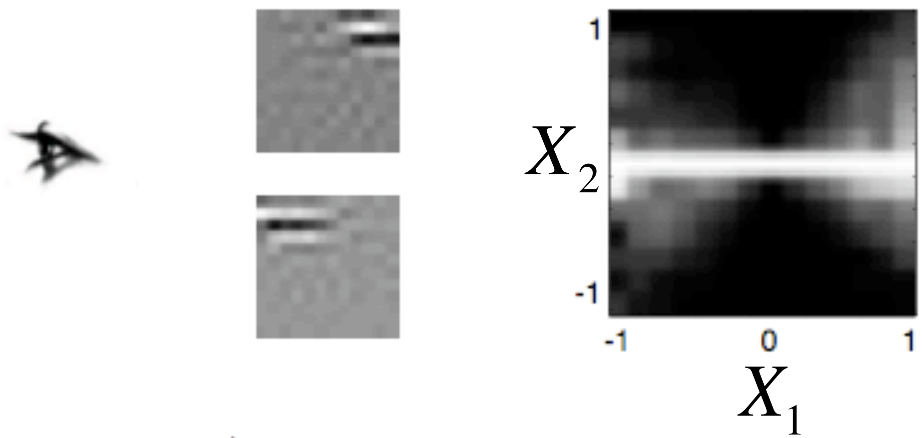
Bottom-up Joint Statistics

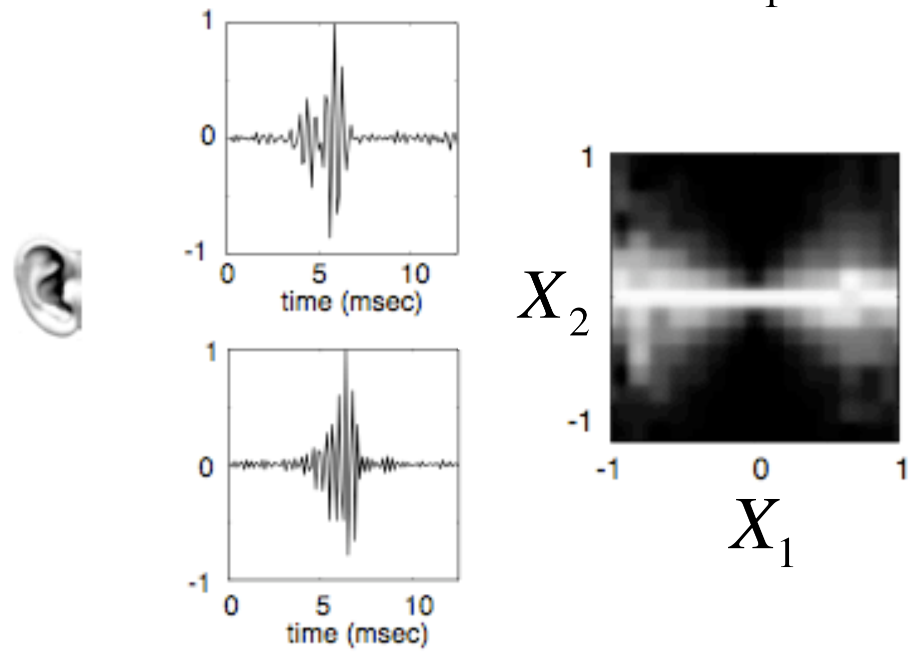


X_1 and X_2 are not statistically independent

Schwartz and Simoncelli, 2001

Bottom-up Joint Statistics



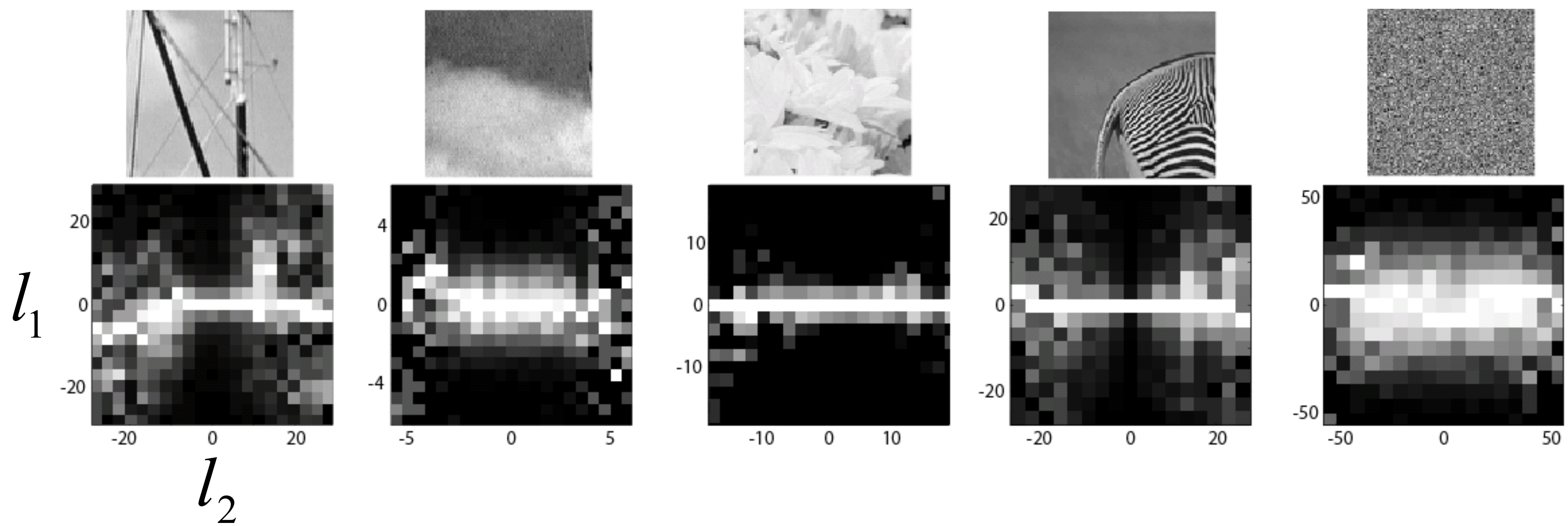


Bottom-up Statistics

Filter pair and different image patches…

$\longrightarrow X_1$

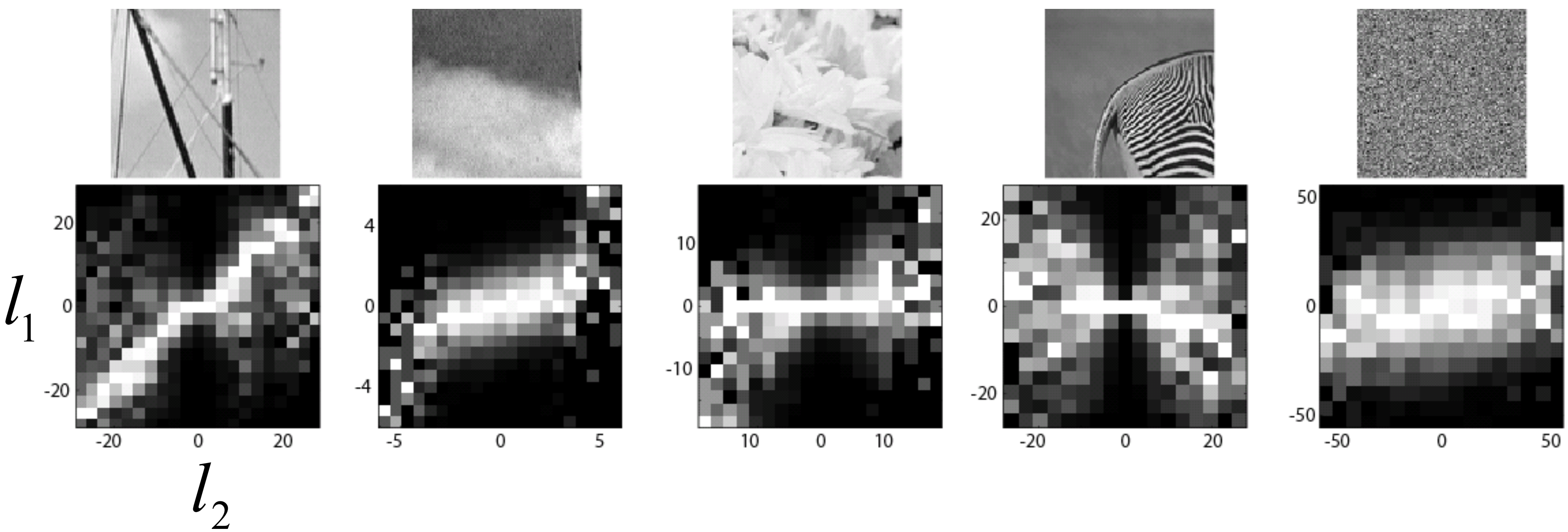
$\longrightarrow X_2$



Bottom-up Statistics

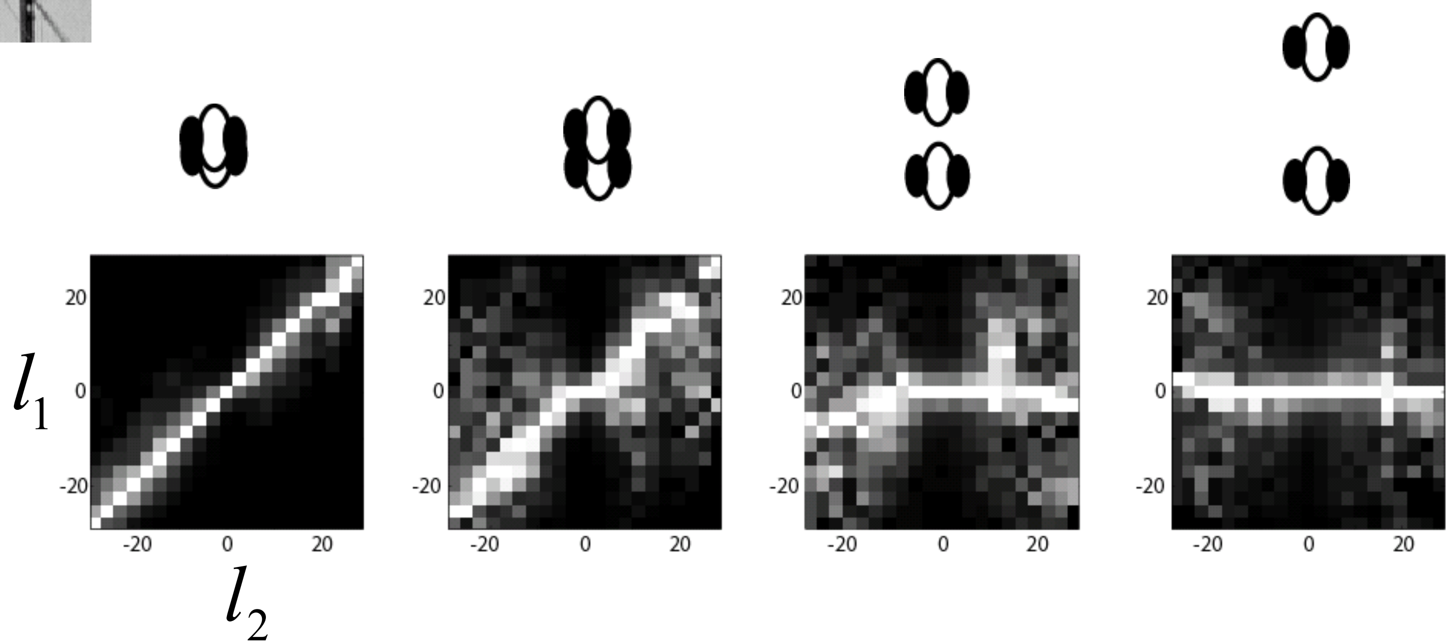
Overlapping filter pair and different image patches…

$$\longrightarrow X_1$$
$$\longrightarrow X_2$$



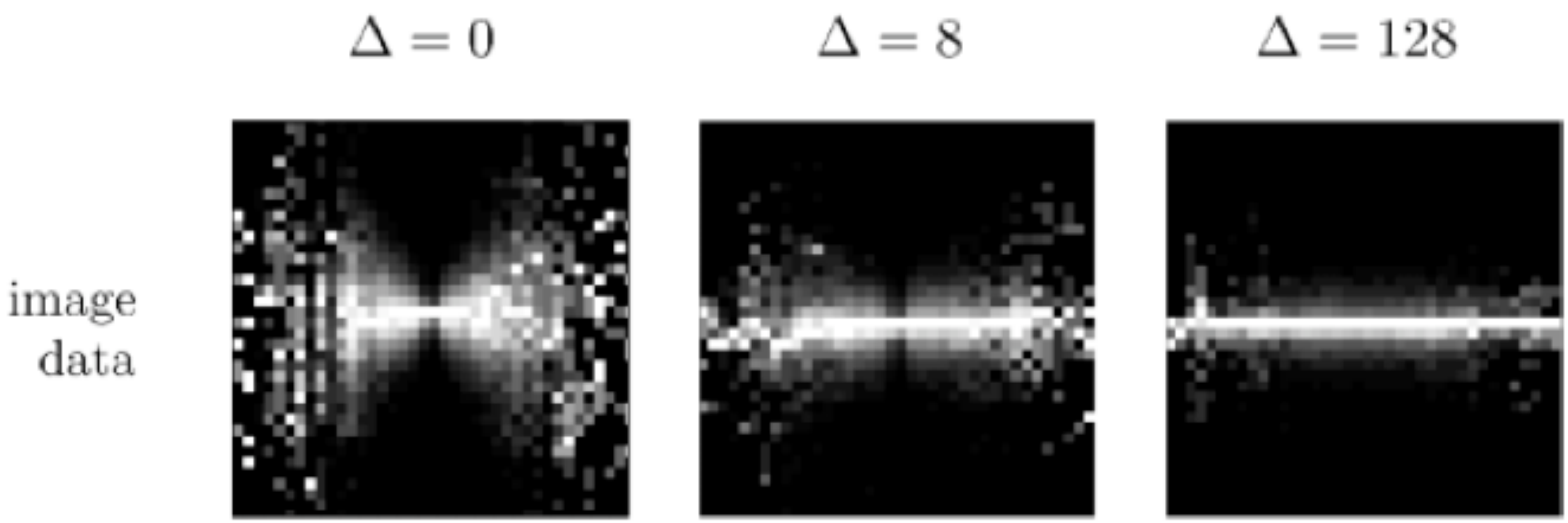
Bottom-up Statistics

Image patch and different filter pairs…



Bottom-up (contour plots)

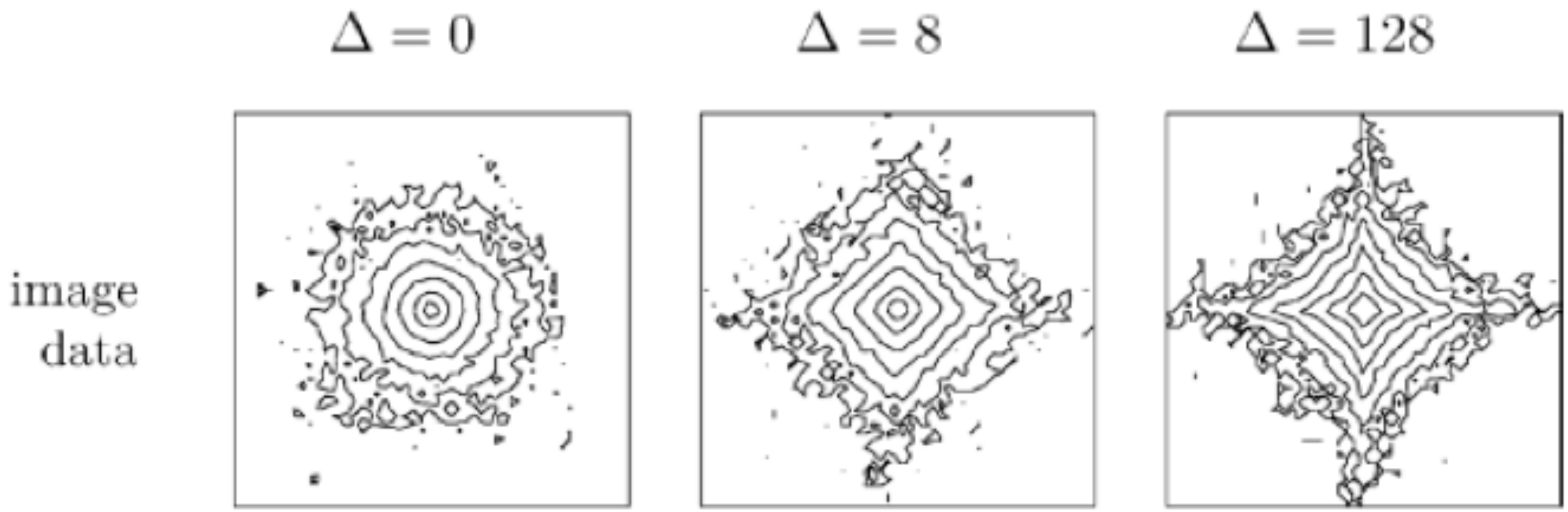
Image patch and different filter pairs (even and odd Phase at different distances) …



Wainwright et al. 2001

Bottom-up (contour plots)

Image patch and different filter pairs (even and odd Phase at different distances) …



Wainwright et al. 2001

Bottom-up neural model

Going beyond linear models

Neural model

assume independence hypothesis

consistent?

Neural data

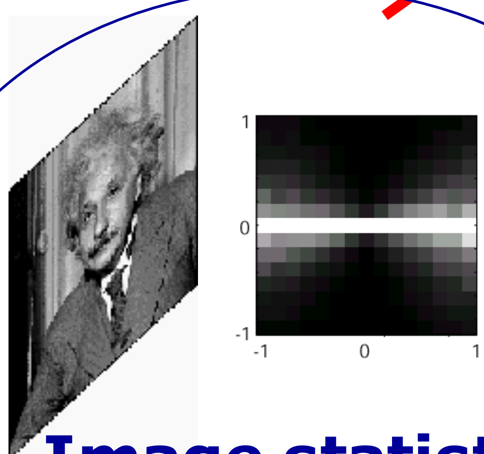
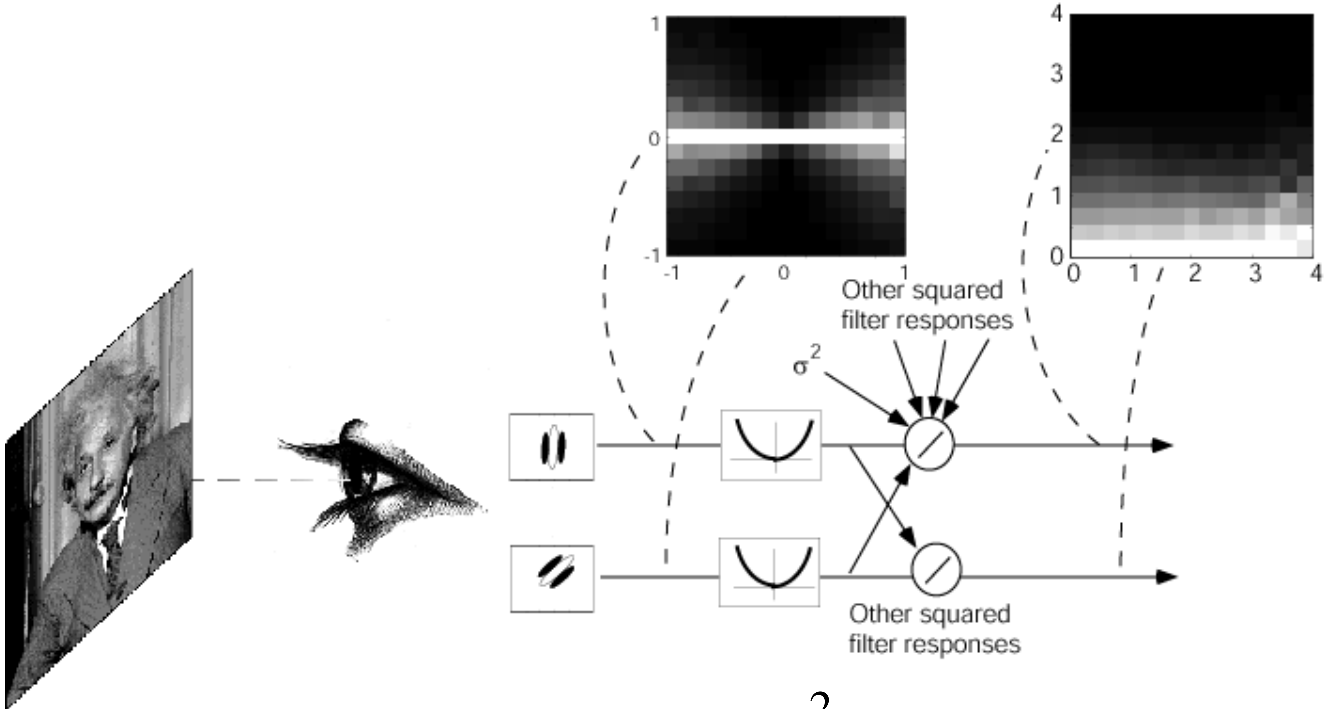


Image statistics

Divisive normalization Model

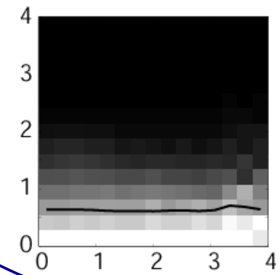


$$R_i = \frac{x_i^2}{\sum_j w_{ji} x_j^2 + \sigma^2}$$

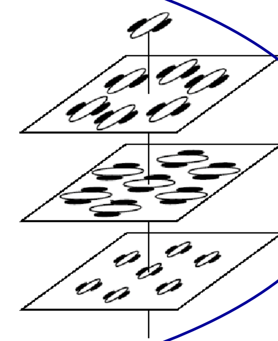
Divisive normalization has been applied in descriptive and mechanistic models (eg, Heeger)

Bottom up

Model



$$R_i = \frac{X_i^2}{\sum_j w_{ji} X_j^2 + \sigma^2}$$



assume
independence
hypothesis

consistent?

Neural data

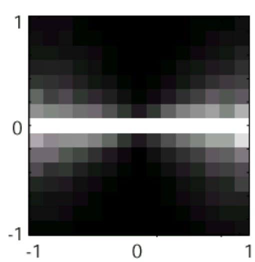


Image statistics

Schwartz and Simoncelli, 2001

Summary

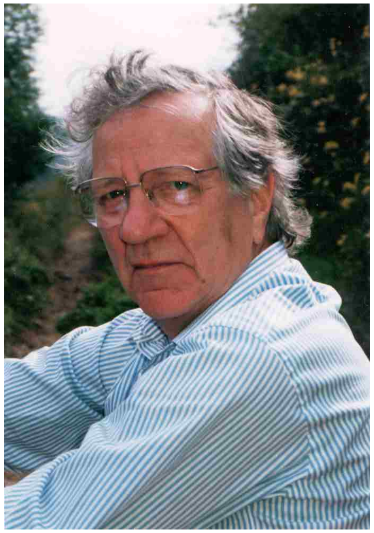
- We've considered bottom-up scene statistics, efficient coding, and relation of linear transforms to visual filters

- This class: nonlinearities

- This class: generative, top-down, perspective

Theoretical Approaches

• Sensory systems aim to form an *efficient code* by reducing the redundancies and statistical dependencies of the input; influenced by information theory in the 1950s



Barlow (also Attneave)

Theoretical Approaches

• Sensory processing as inference of properties of the input (can be formalized via probabilistic *Bayesian inference*)

Helmholtz

Bayes

Scene statistics approaches

Two main approaches for studying scene statistics

1. Bottom-up (e.g., Choose and manipulate projections, to optimize probabilistic and information-theoretic metrics)

2. Top-down, generative (e.g., probabilistic characterizations of the processes by which the signals are generated)

Scene statistics approaches

Two main approaches for studying scene statistics

1. Bottom-up

2. Top-down, generative (This class!)

Generative Model

probabilistic characterizations
of the processes by which the signals
are generated

- If we can capture dependencies in images,
  we should be able to not only reduce them, but
  to model probabilistically...

Image models as generative

Hidden (latent) variables

↓ generative

I Image
(or some observed variable)

Image models as generative

Hidden variables (S)

generative

I image

Hidden variables (S)

recognition

I image

Incorporate prior probabilities for the hidden variables

Bayesian inference and generative

Hidden variables (S)

$\downarrow$ generative

I image

Hidden variables (S)

$\uparrow$ recognition

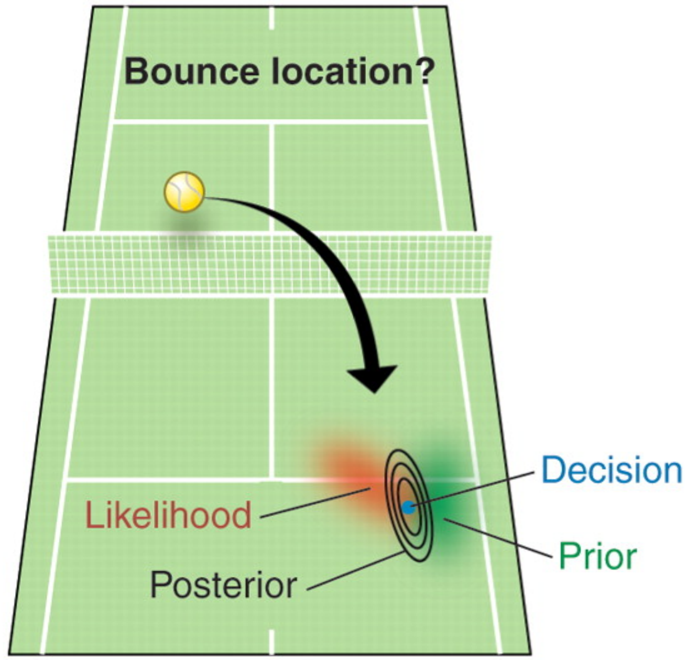
I image

posterior likelihood prior

$$P(S \mid I) = \frac{P(I \mid S)P(S)}{P(I)}$$

$$P(I) = \sum_j P(I \mid S_j)P(S_j)$$

Bayesian inference



Koerding 2007

Generative models

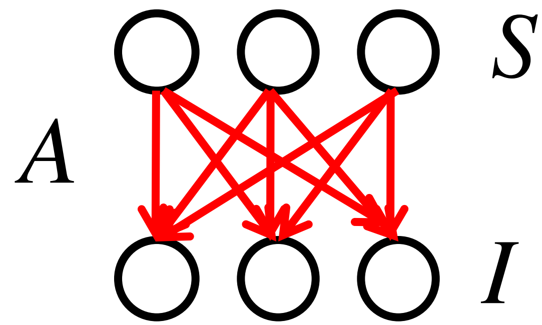
What linear transforms do we already know from efficient coding?

PCA as a generative model

Hidden variables
(here principle components)

generative

I image



$$I = AS$$

$p(S_i)$ Gaussian with variance equal to that of principle component i.

S_i Uncorrelated

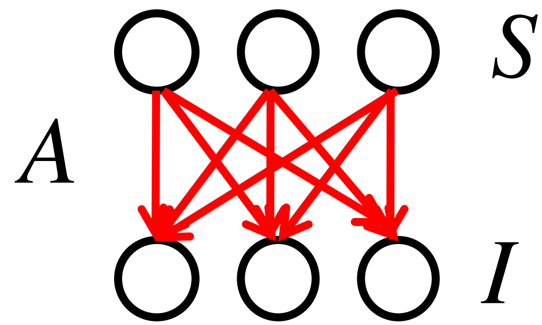
Dimensionality S <= I

ICA as a generative model

Hidden variables

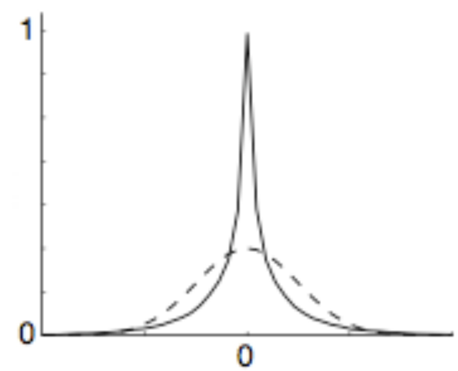
generative

I image



$$I = AS$$

$p(S_i)$ sparse (e.g., super Gaussian)



$$p(S) = \prod_i p(S_i)$$

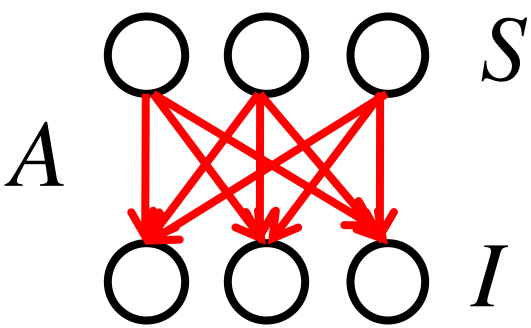
ICA as a generative model

Hidden variables (S)

$\downarrow$ generative

I image

$$I = AS$$

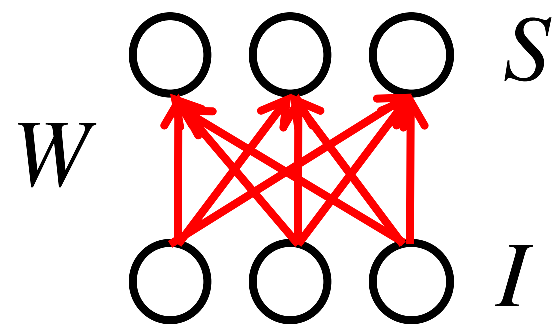


Hidden variables (S)

$\uparrow$ recognition

I image

$$S = WI$$



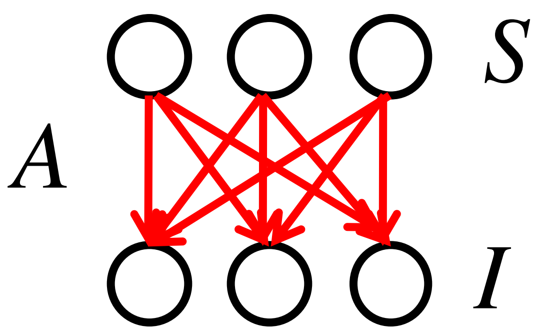
ICA as a generative model

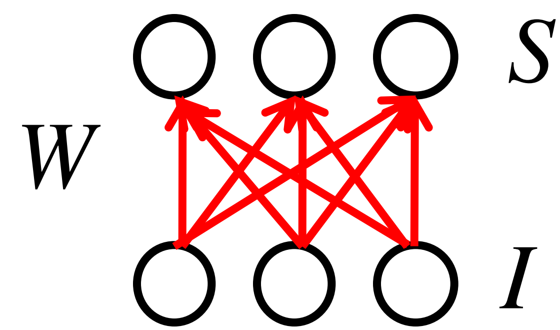
Hidden variables (S) ↓ generative I image

Hidden variables (S) ↑ recognition I image

$$I = AS$$

$$S = WI$$





Given large number of image patches, we'd like to estimate the hidden variables

ICA maximum likelihood

$$S = WI$$

$$p(S) = \prod_i p(S_i)$$ (independence assumption ICA; marginals are sparse)

$$p(I) = \left|\det(W)\right| \prod_i p_i(w_i^T I)$$ (linear transform)

$$P(I_{patches}) = \prod_{t=1}^{num\_patches} p(I_t)$$ (patches independent)

Maximize the (average log) likelihood of the data I over all patches with respect to the weights w…

ICA maximum likelihood

$$p(I) = \left|\det(W)\right| \prod_i p_i(w_i^T I)$$

(linear transform)

$$P(I_{patches}) = \prod_{t=1}^{N} p(I_t)$$

(patches independent)

$$\log P(I_{patches}) = N \log\left|\det(W)\right| + \sum_i \sum_t \log p_i(w_i^T I_t)$$

Related to sparse coding:
Estimated log pdf;
Hyvarinen book