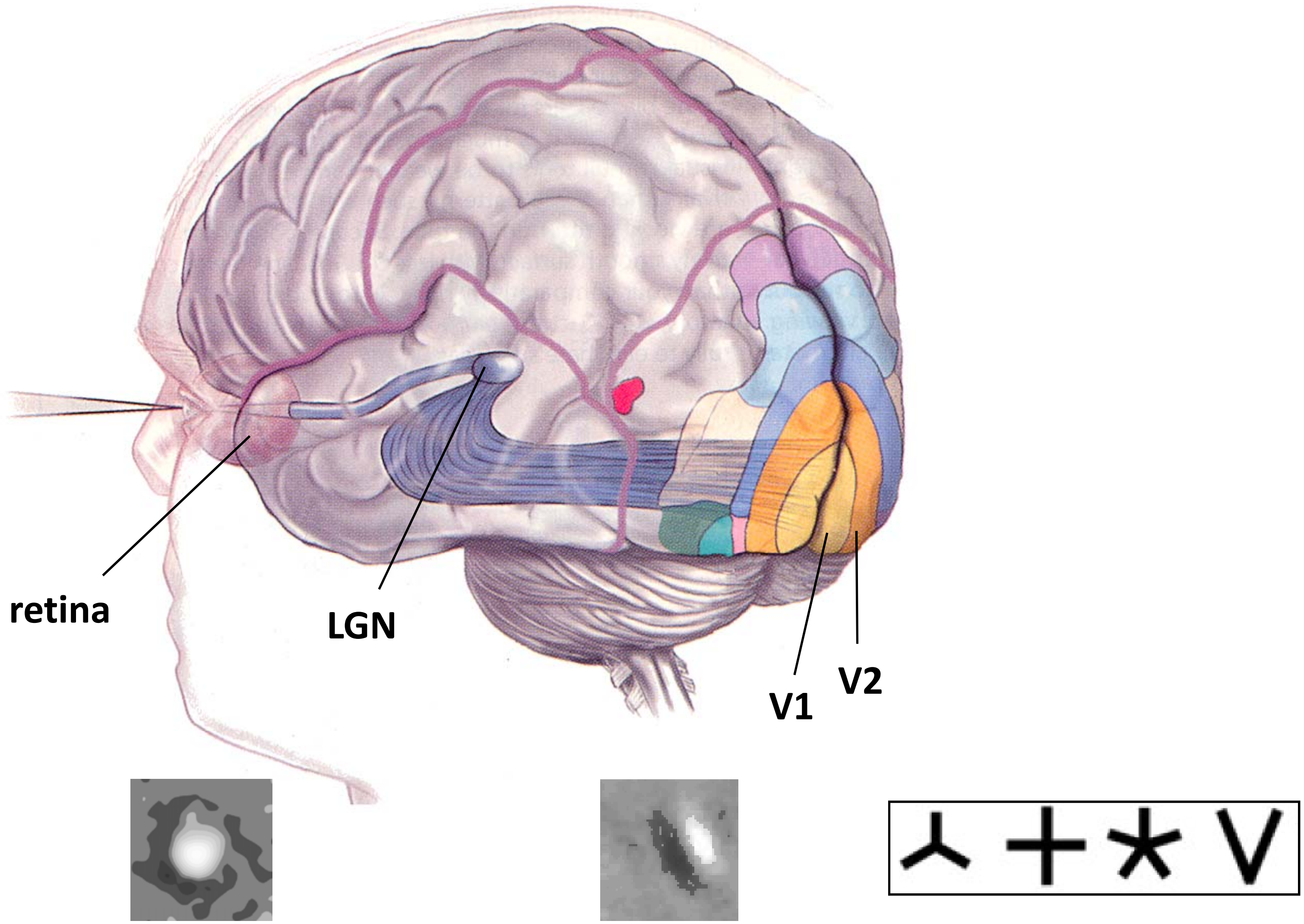
a factor model for learning higher-order features in natural images

yan karklin (*nyu + hhmi*)

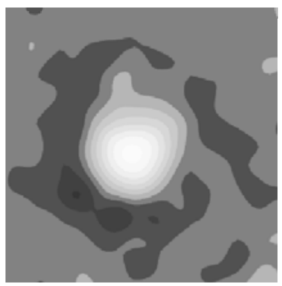
joint work with mike lewicki (*case western reserve u*)

icml workshop on learning feature hierarchies - june 18, 2009


retina
LGN
V1
V2

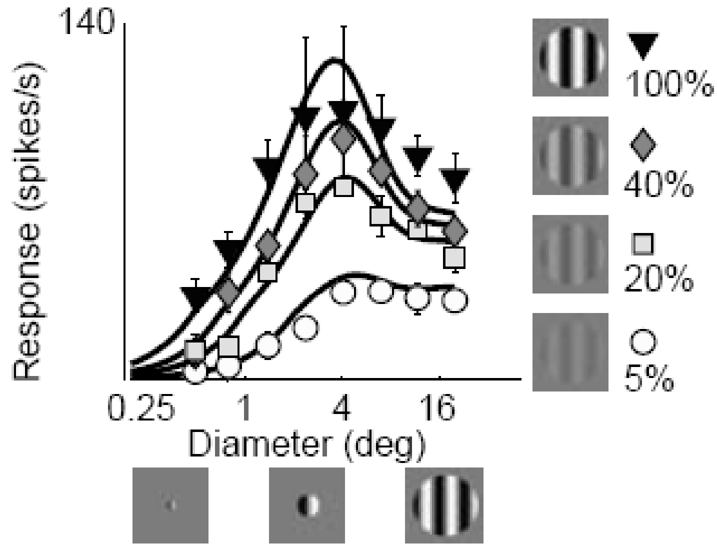
non-linear processing in the early visual system

estimated linear receptive field

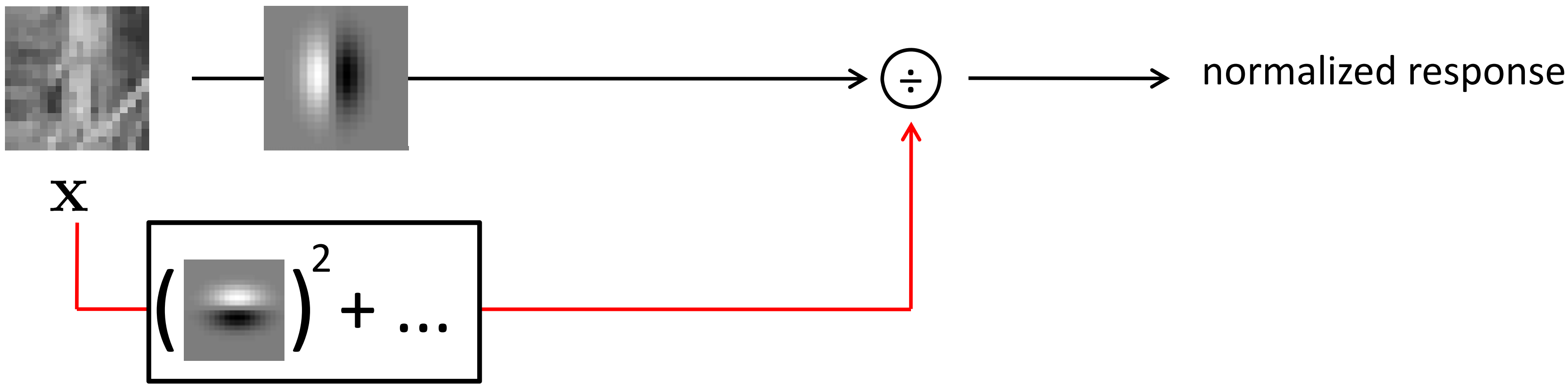


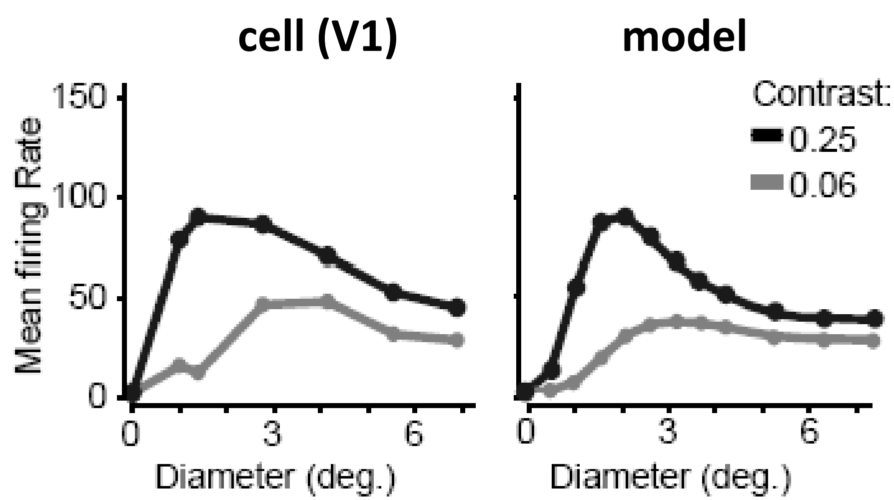
cat LGN (DeAngelis, Ohzawa, Freeman 1995)

contrast/size response function



cat LGN (Carandini 2004)

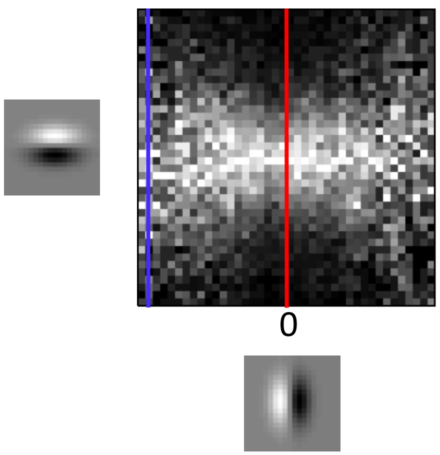


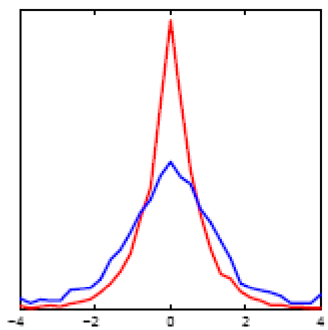


(Schwartz and Simoncelli, 2001)

non-linear statistical dependencies in natural images

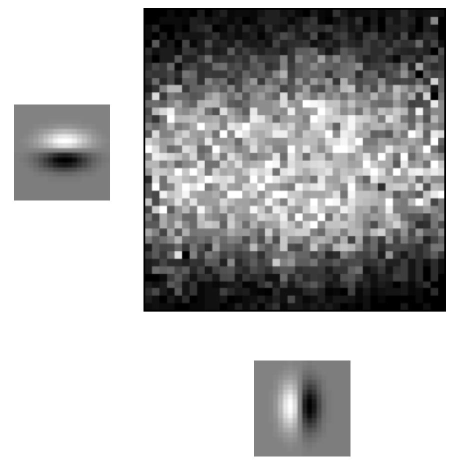
$p(y_2|y_1)$

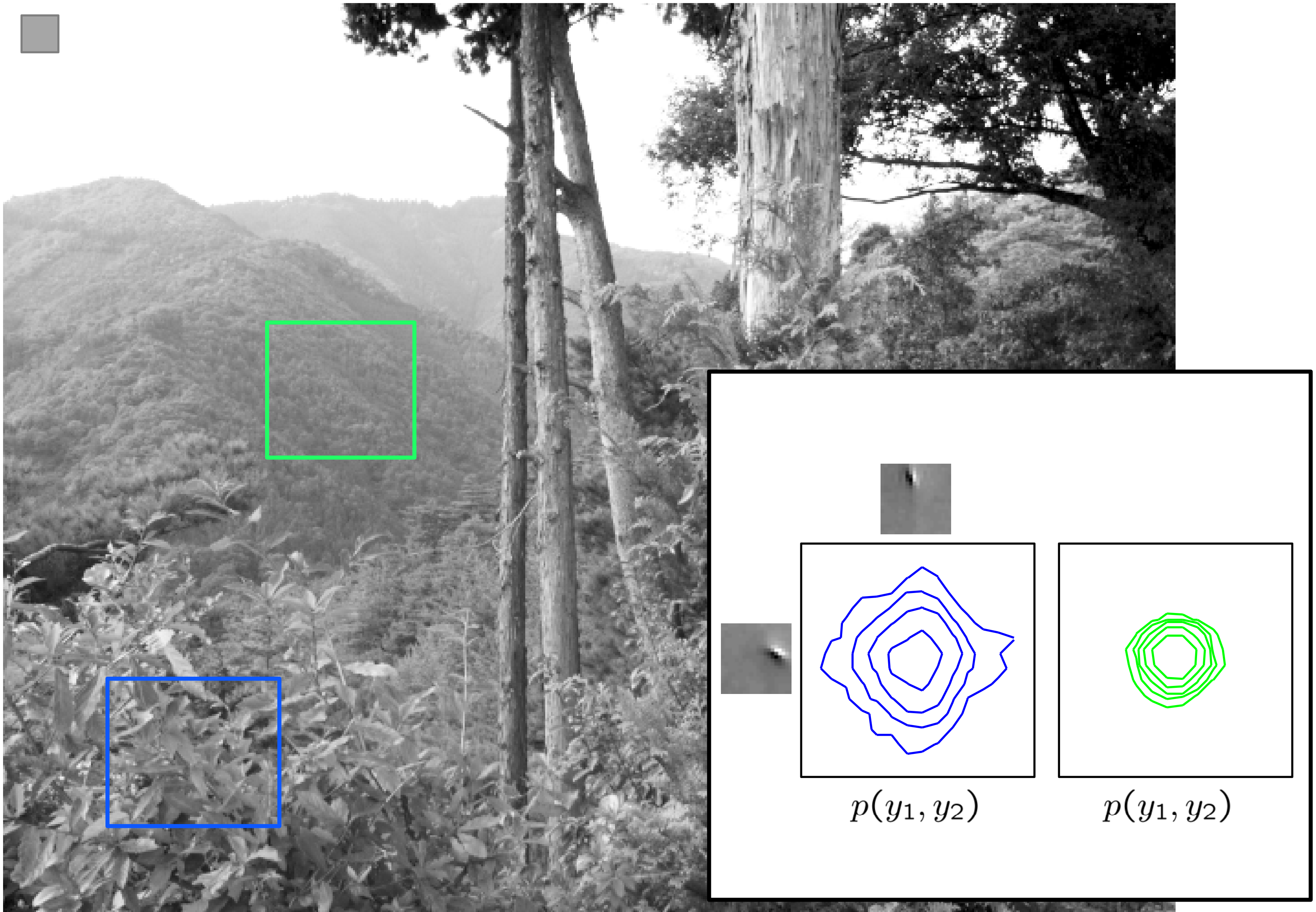


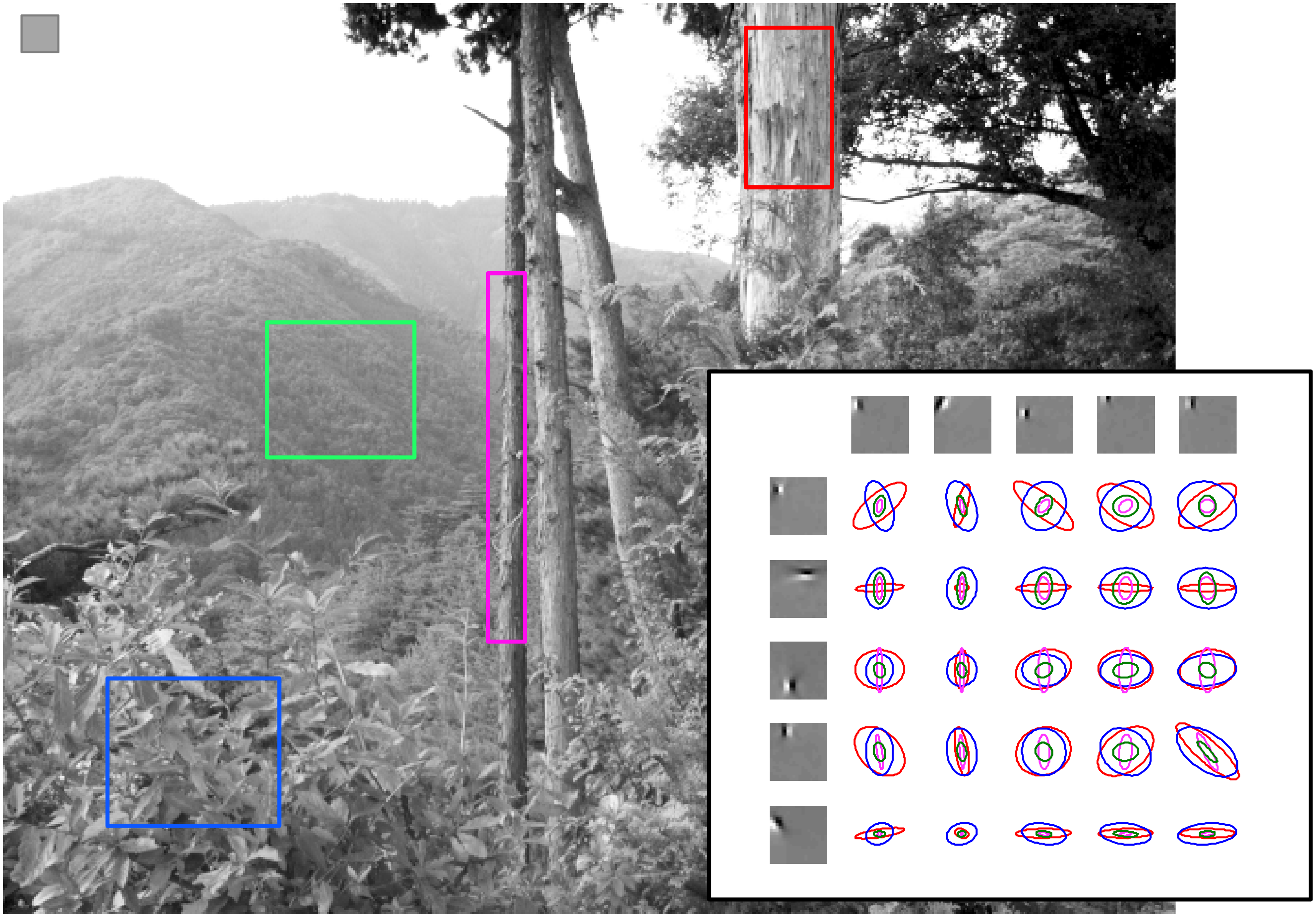


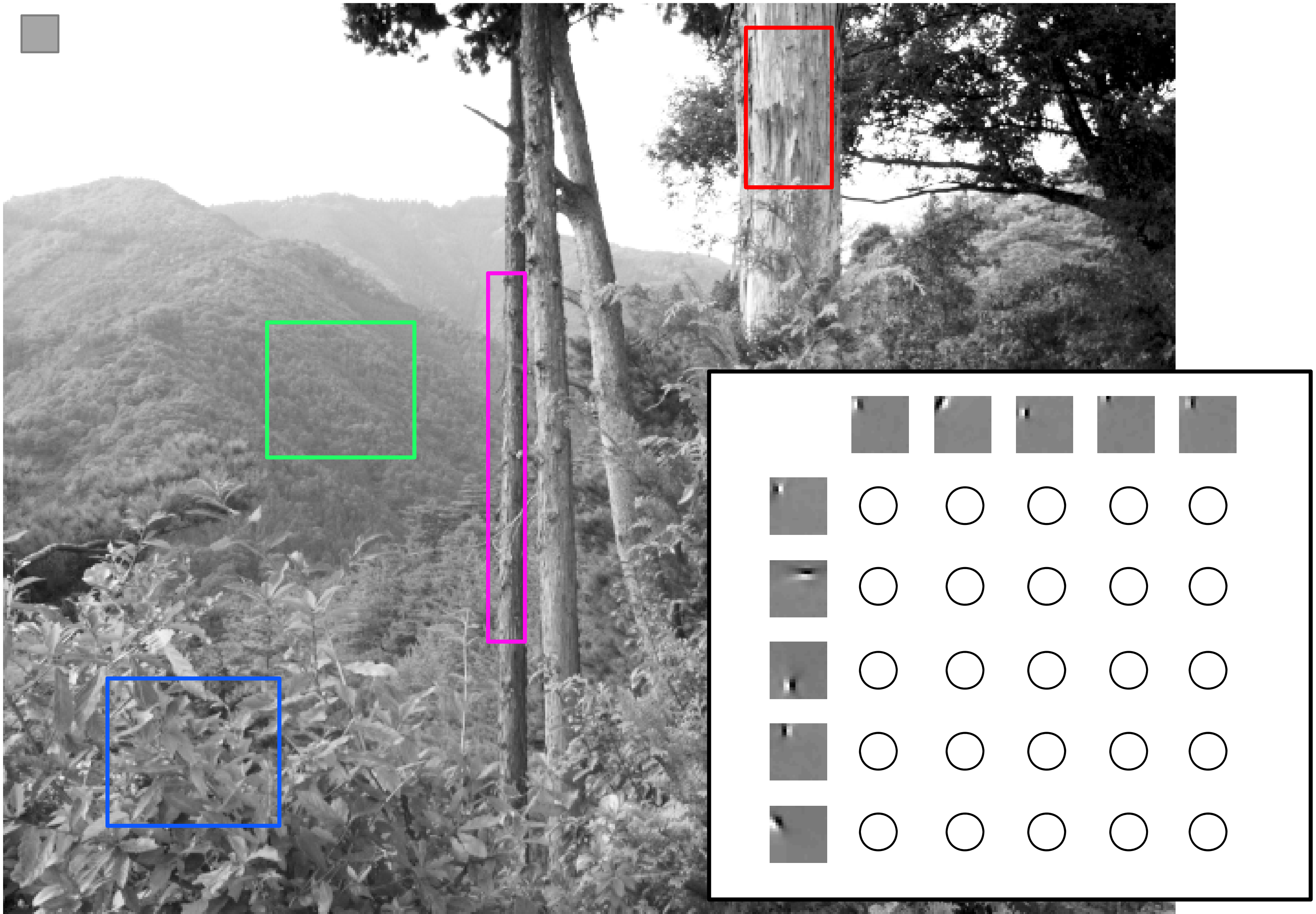
$p(\bar{y}_2|\bar{y}_1)$

$$\bar{y}_i = \frac{y_i}{\beta + \sum w_{ij} y_j^2}$$

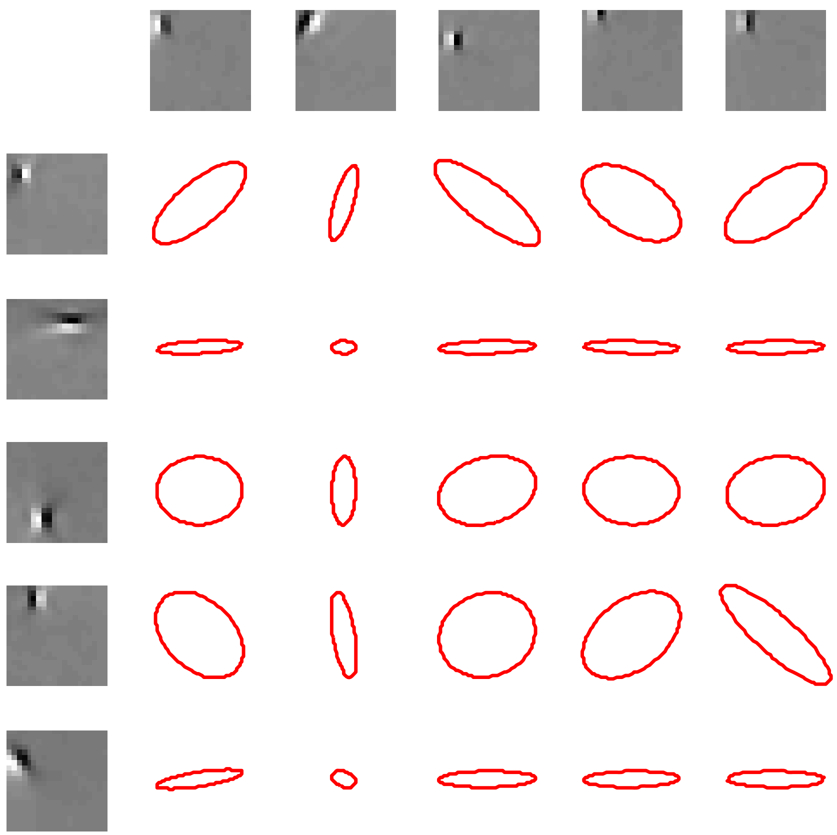



$p(y_1, y_2)$
$p(y_1, y_2)$





a generative model for covariance matrices



multivariate Gaussian model

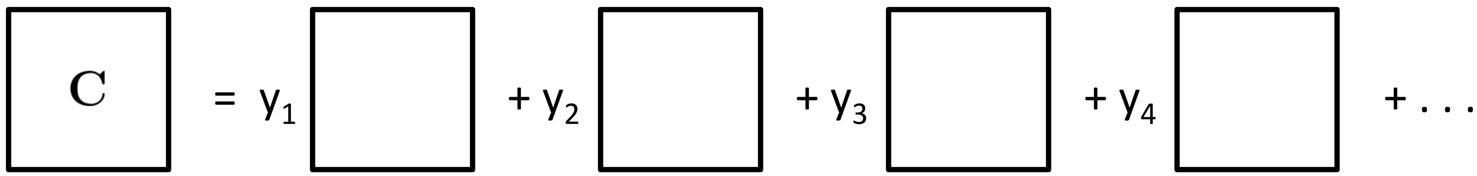
$$\mathbf{x} \sim \mathcal{N}(0, \mathbf{C})$$

latent variables specify the covariance

$$\mathbf{C} = f(\mathbf{y})$$

a distributed code

if only we could find a basis for covariance matrices...



a distributed code

if only we could find a basis for covariance matrices...

$$\mathbf{C} = y_1 \,\square + y_2 \,\square + y_3 \,\square + y_4 \,\square + \dots$$

use the ***matrix-logarithm transform*** $\exp(\mathbf{A}) = \mathbf{I} + \sum \mathbf{A}^k / k!$

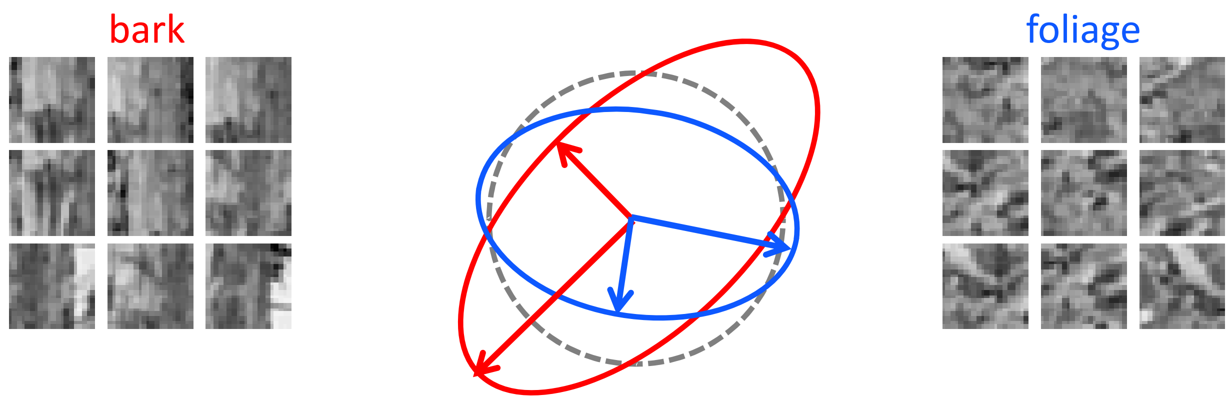
$$\log \mathbf{C} = y_1 \mathbf{A}_1 + y_2 \mathbf{A}_2 + y_3 \mathbf{A}_3 + y_4 \mathbf{A}_4 + \dots$$

$$\mathbf{C} = \exp\left(\sum \mathbf{A}_j y_j\right)$$

$$\mathbf{y} = \mathbf{0} \quad \rightarrow \quad \mathbf{C} = \mathbf{I}$$

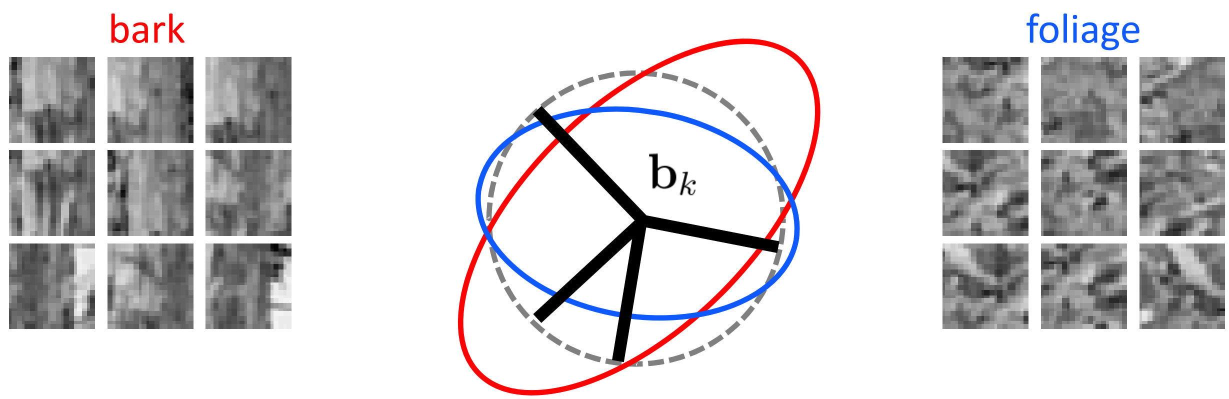
a compact parameterization for cov-components

find ***common directions of change*** in variation/correlation



a compact parameterization for cov-components

find ***common directions of change*** in variation/correlation

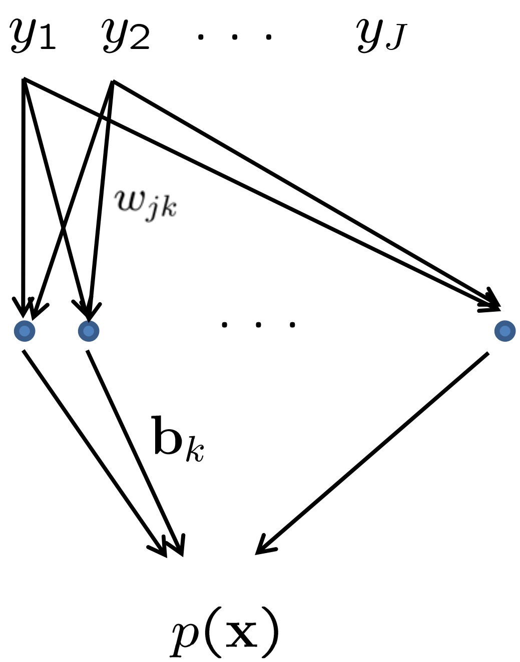


$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

vectors $\mathbf{b}_\mathbf{k}$ allow a more compact description of the basis functions

$\mathbf{b}_\mathbf{k}$'s are unit norm (w_{jk} can absorb scaling)

the full model



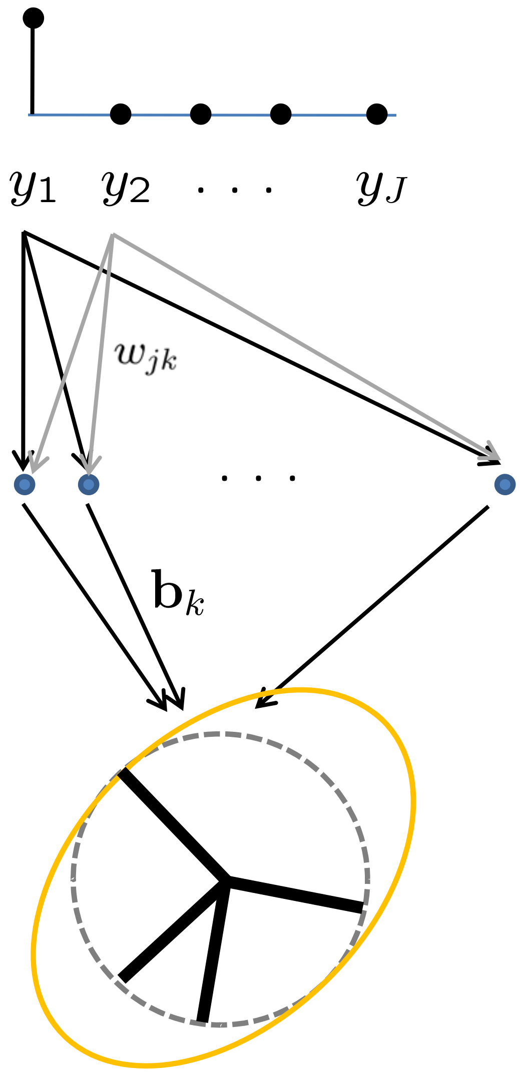
$$p(\mathbf{y}) = \prod p(y_j) \propto e^{-\sum |y_j|}$$

$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

the full model



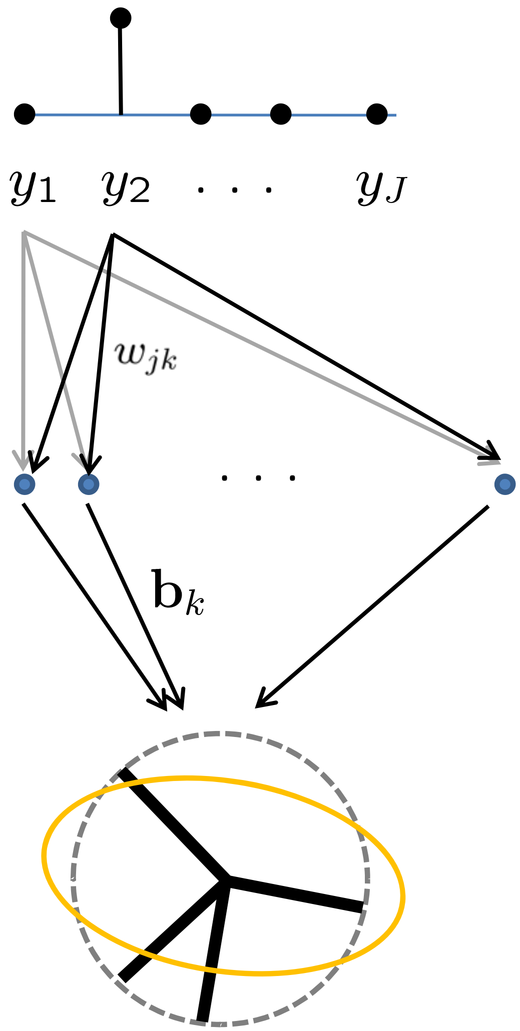
$$p(\mathbf{y}) = \prod p(y_j) \propto e^{-\sum |y_j|}$$

$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

the full model



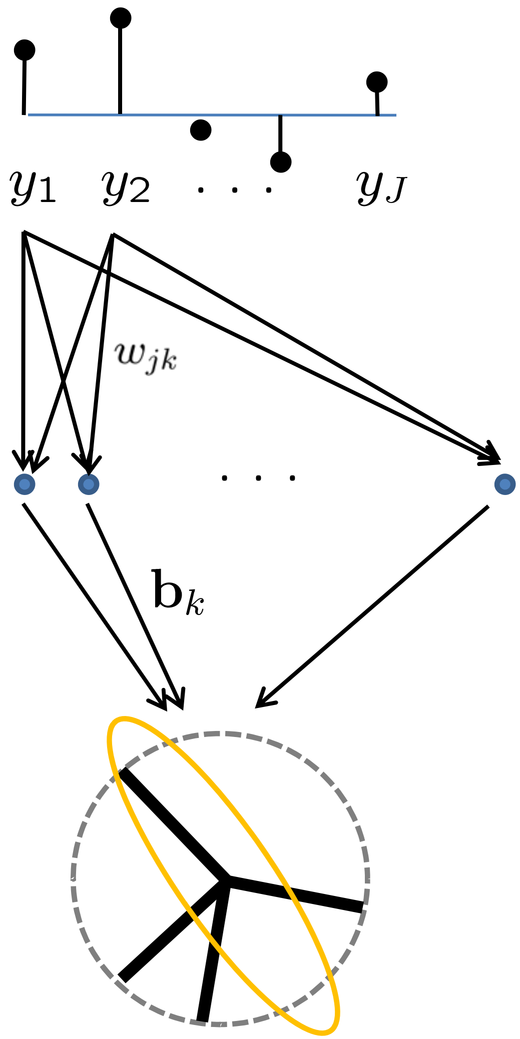
$$p(\mathbf{y}) = \prod p(y_j) \propto e^{-\sum |y_j|}$$

$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

the full model



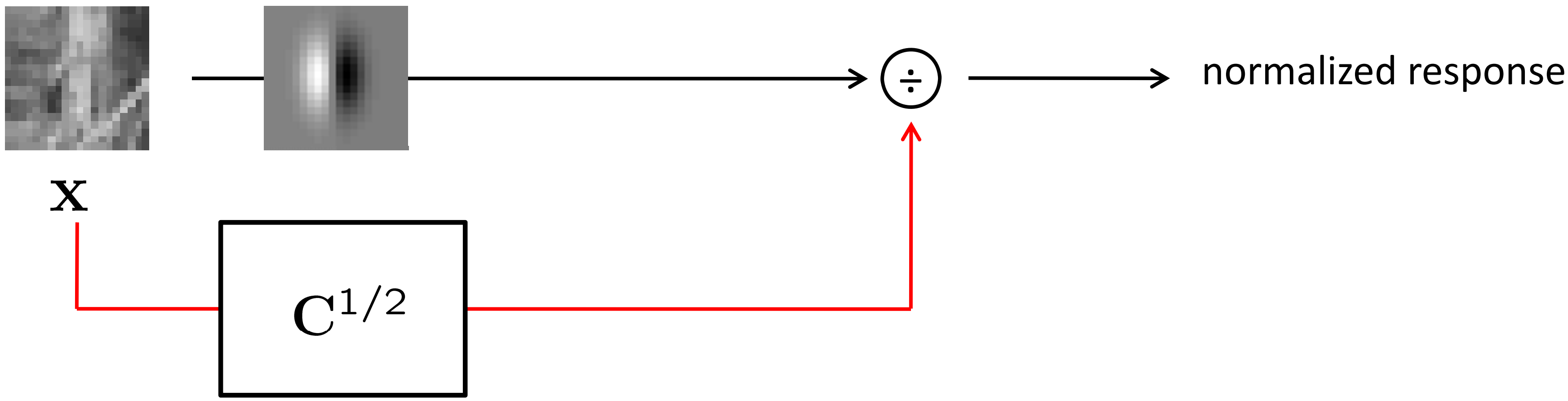
$$p(\mathbf{y}) = \prod p(y_j) \propto e^{-\sum |y_j|}$$

$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

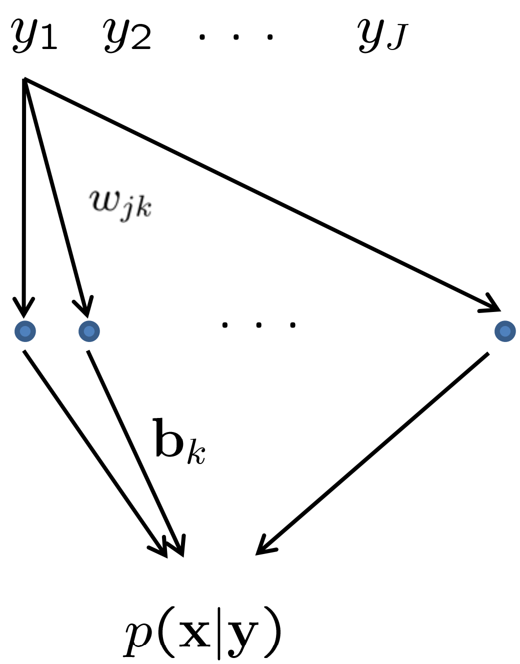
$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

normalization by the inferred covariance matrix



training the model on natural images



$$p(\mathbf{y}) = \prod p(y_j) \propto e^{-\sum |y_j|}$$

$$\mathbf{A}_j = \sum_k w_{jk} \mathbf{b}_k \mathbf{b}_k^T$$

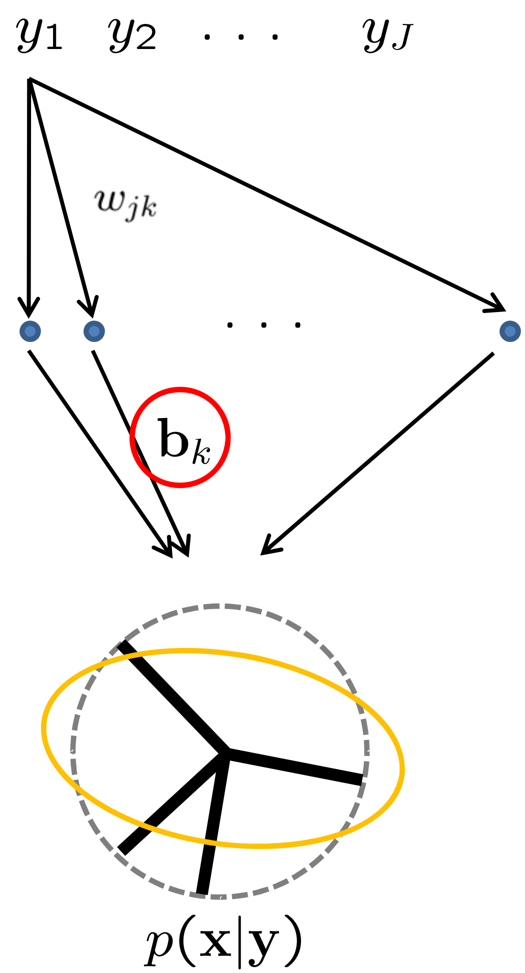
$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

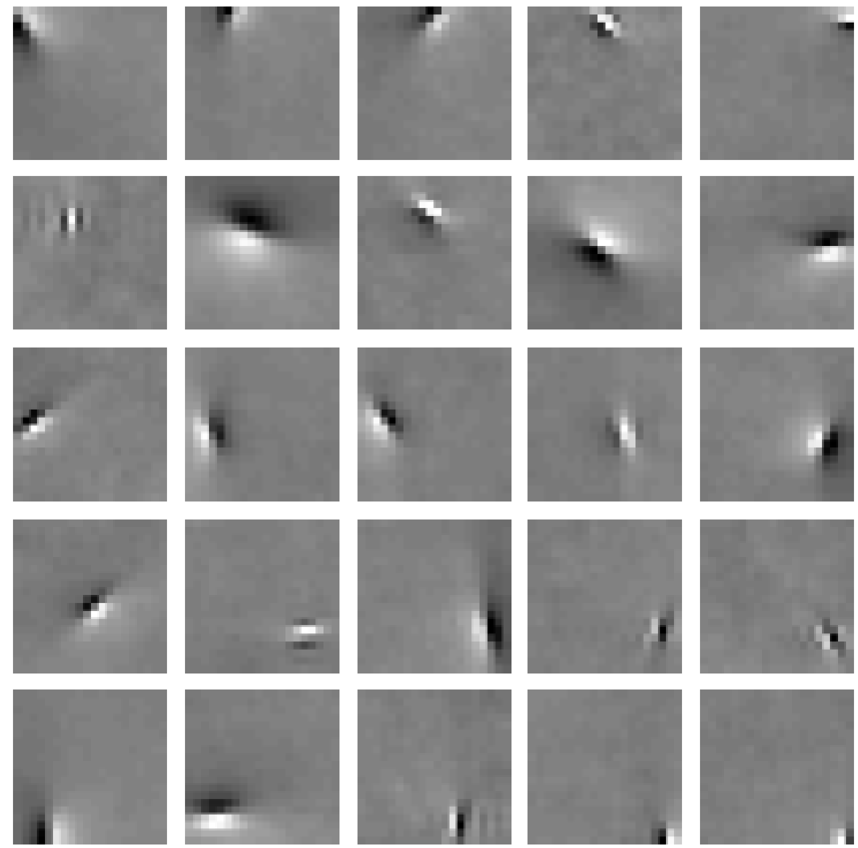
$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

learning details:

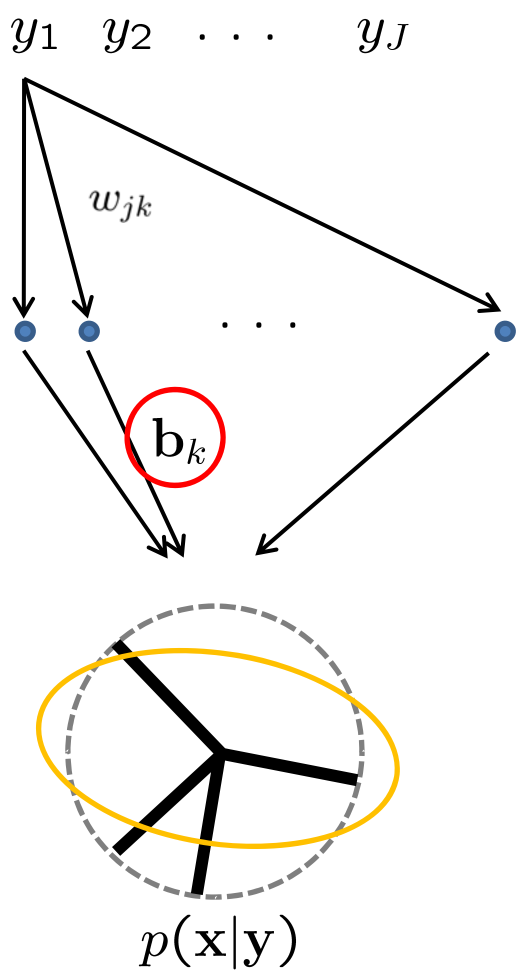
- gradient ascent on data likelihood
- 20x20 image patches from ~100 natural images
- 1000 $\mathbf{b}_k$'s
- 150 y_j's

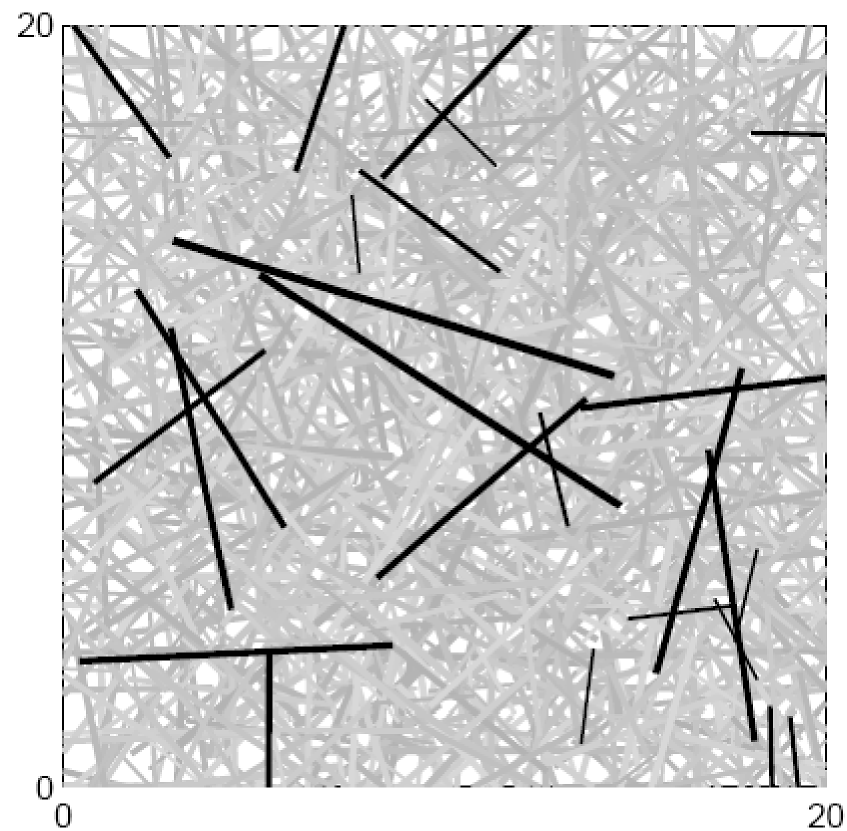
learned vectors





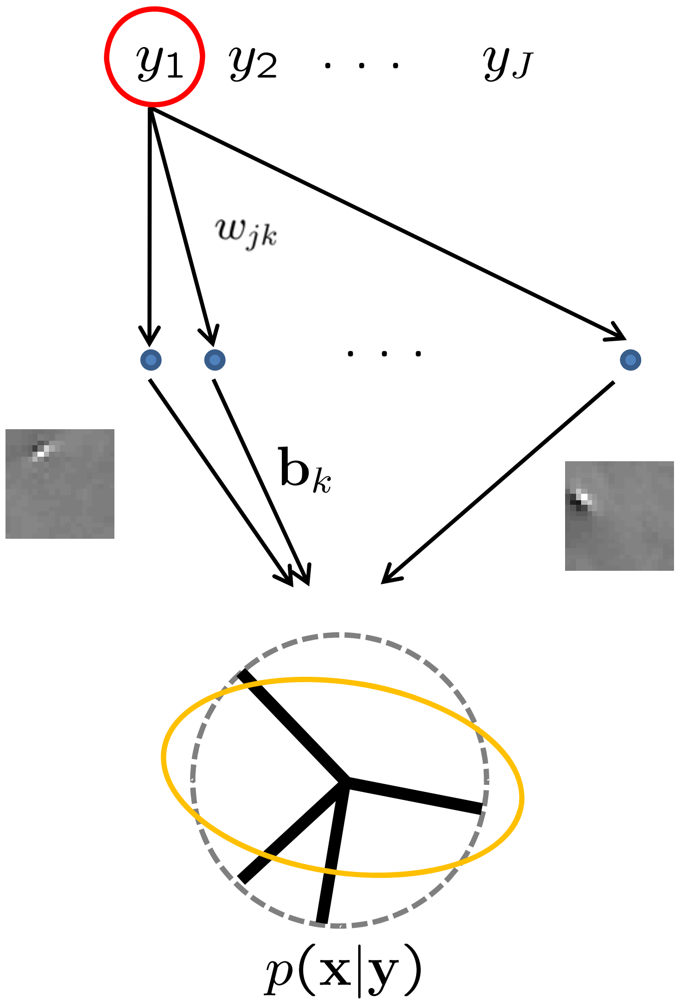
learned vectors

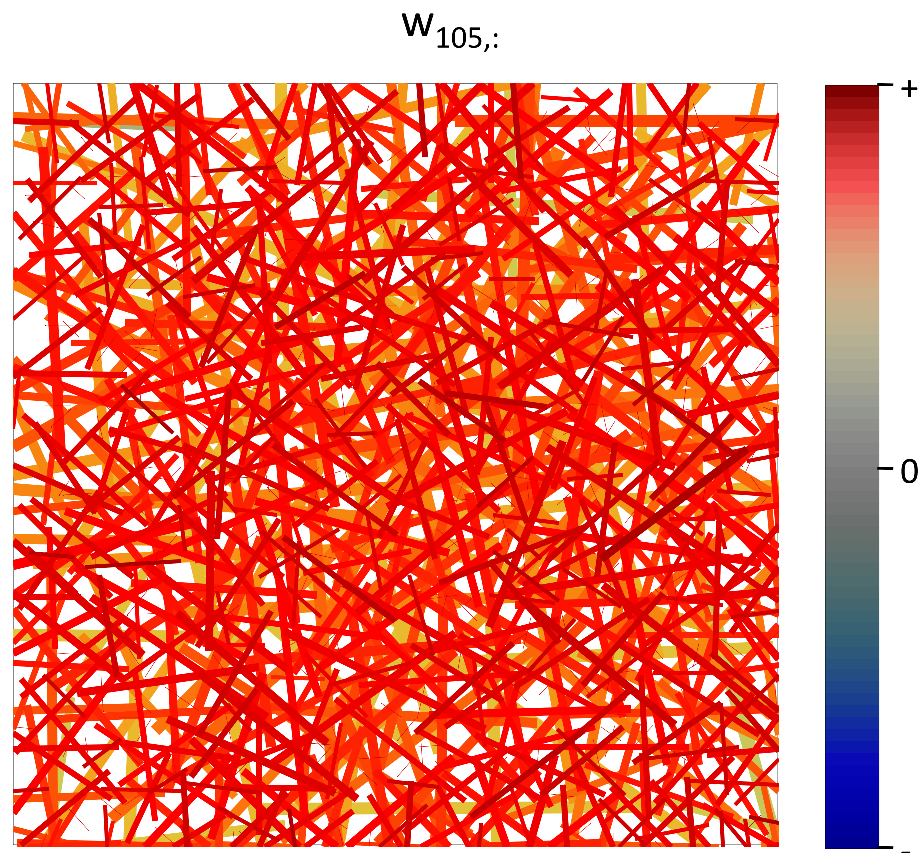




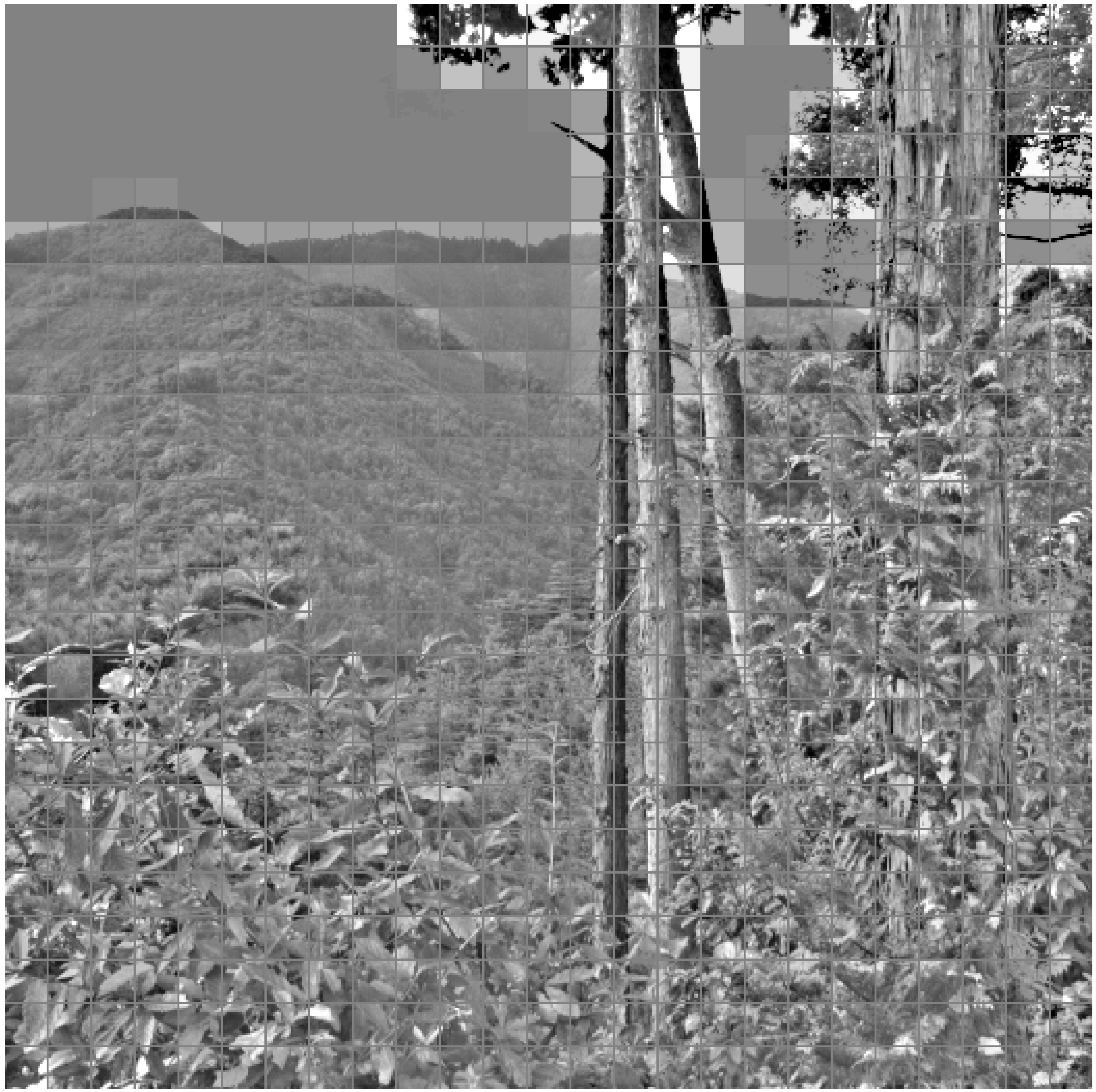
black – 25 example features
grey – all 1000 features in model

one unit encodes global image contrast

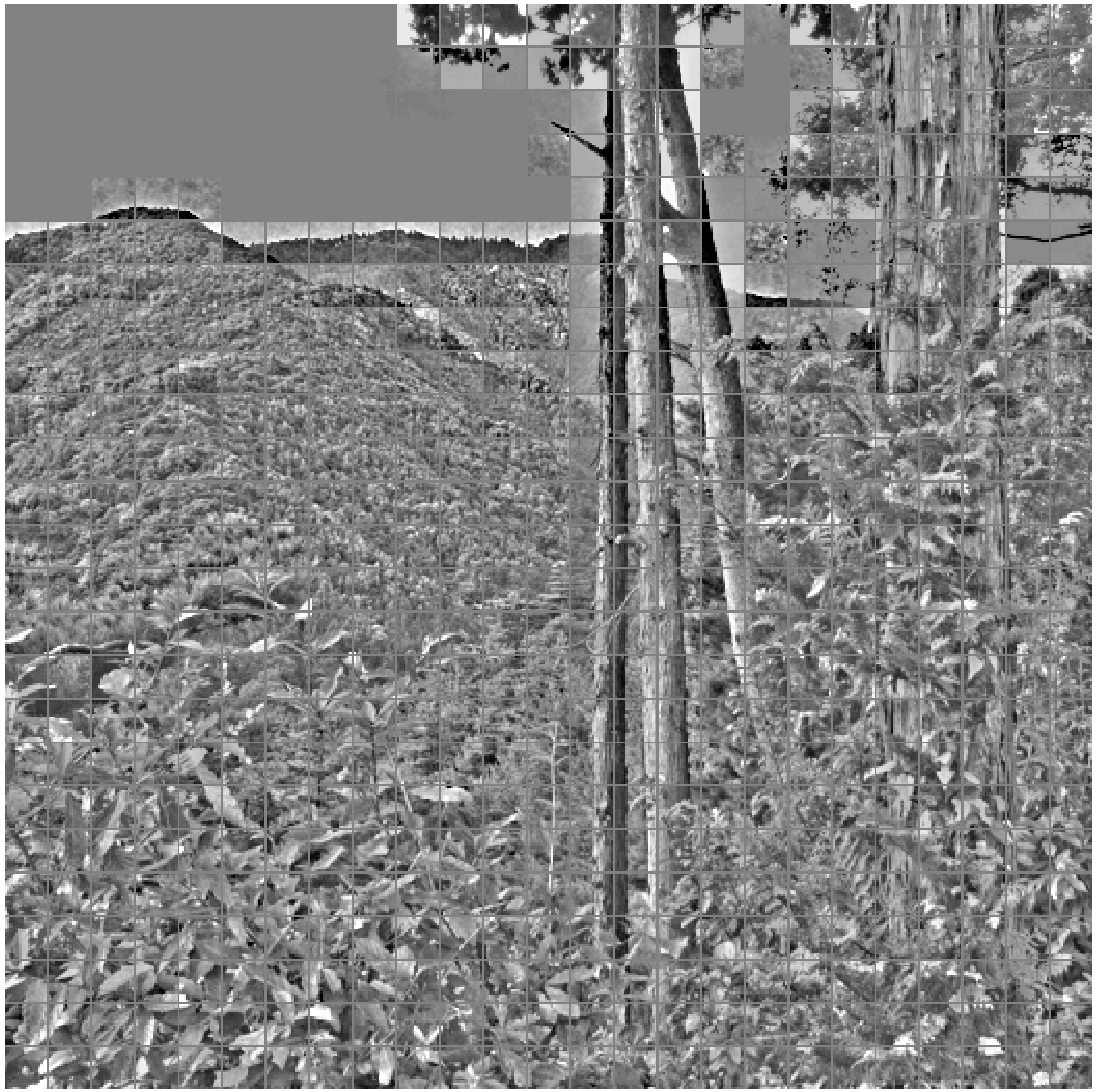


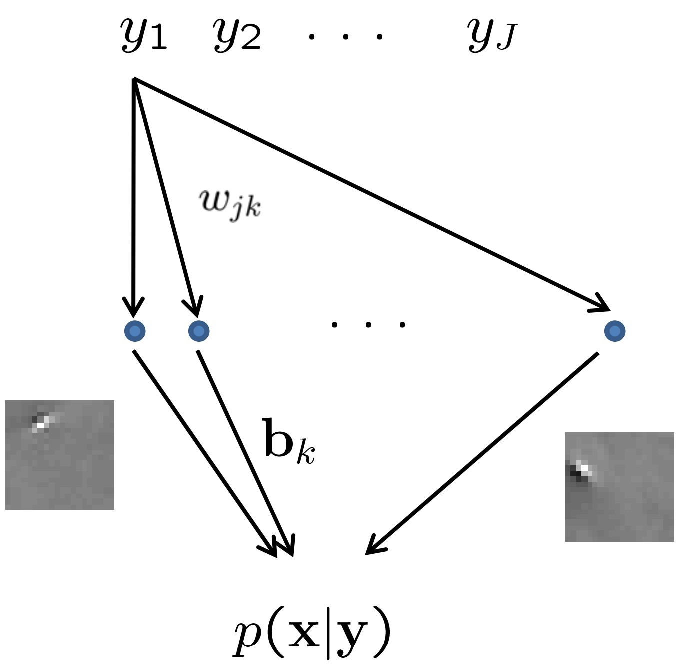


original patches

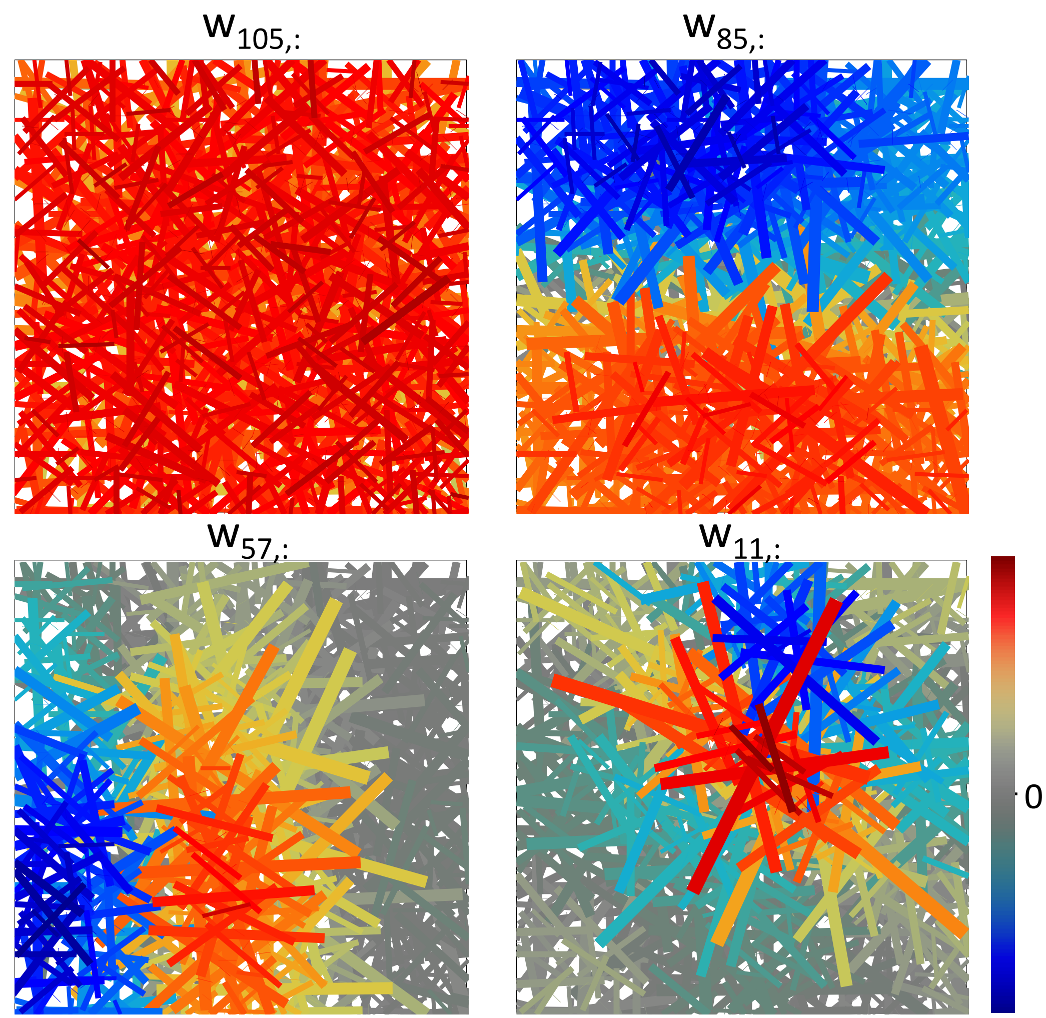


normalization by *inferred contrast*

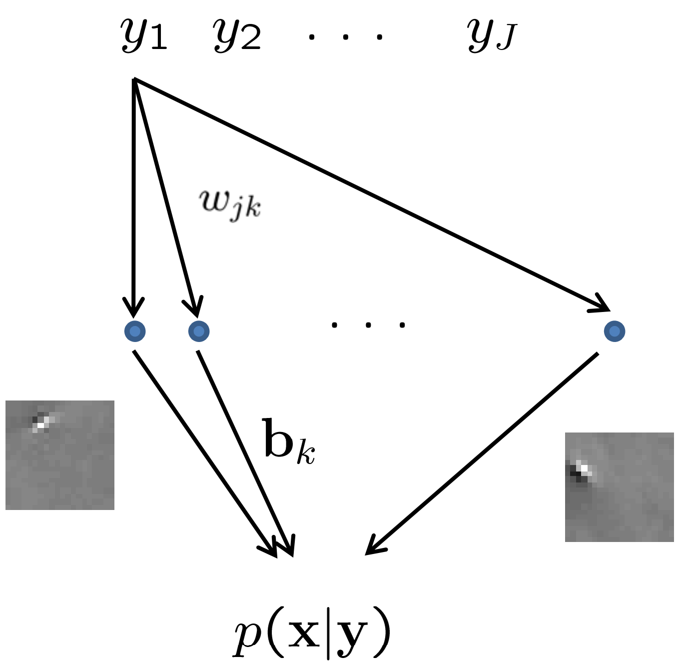


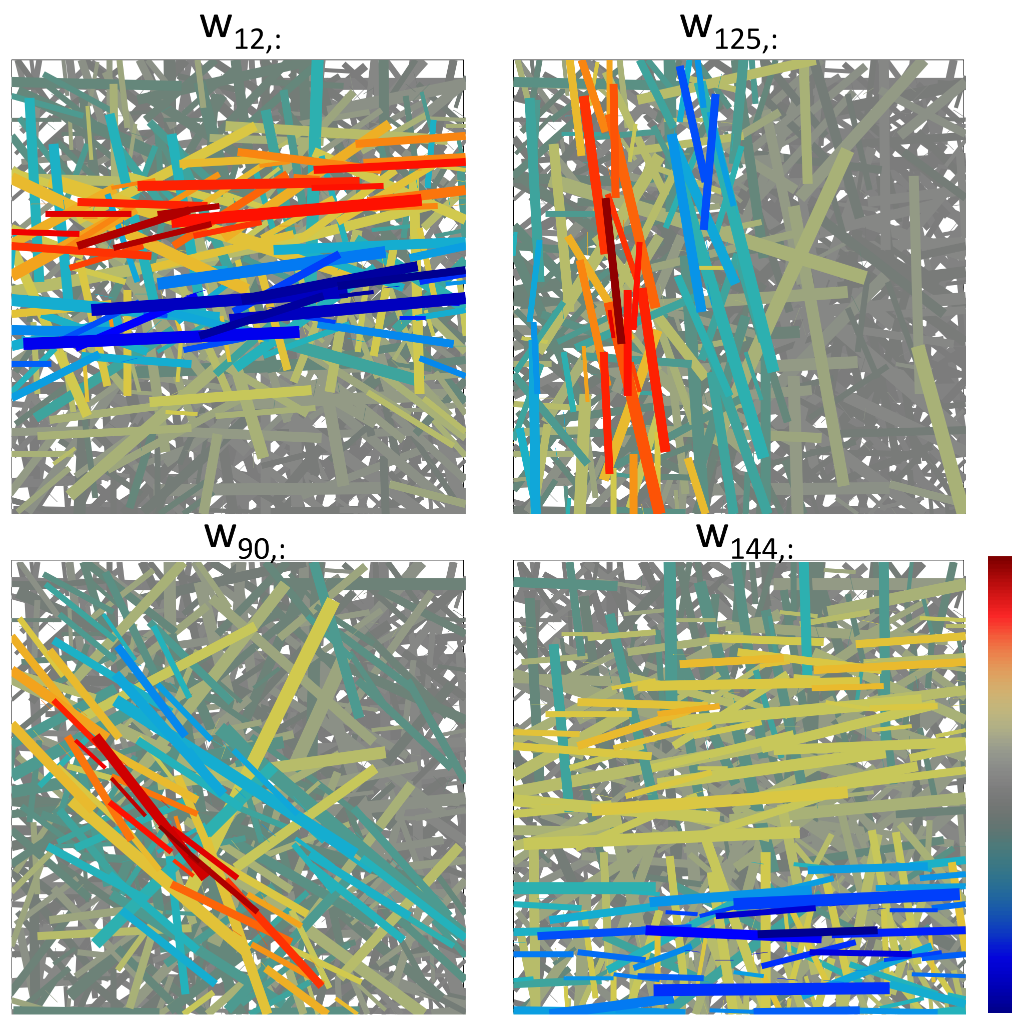


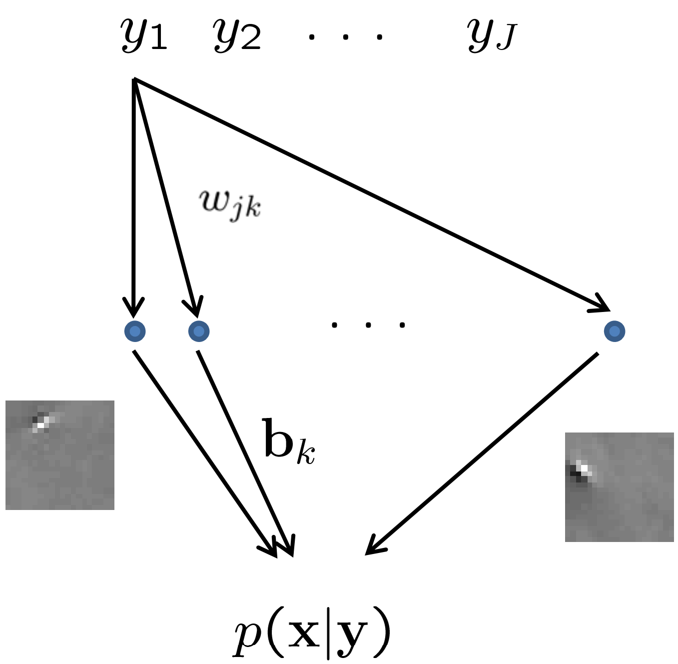
y_1 y_2 $\cdots$ y_J
w_{jk}
$\mathbf{b}_k$
$p(\mathbf{x}|\mathbf{y})$

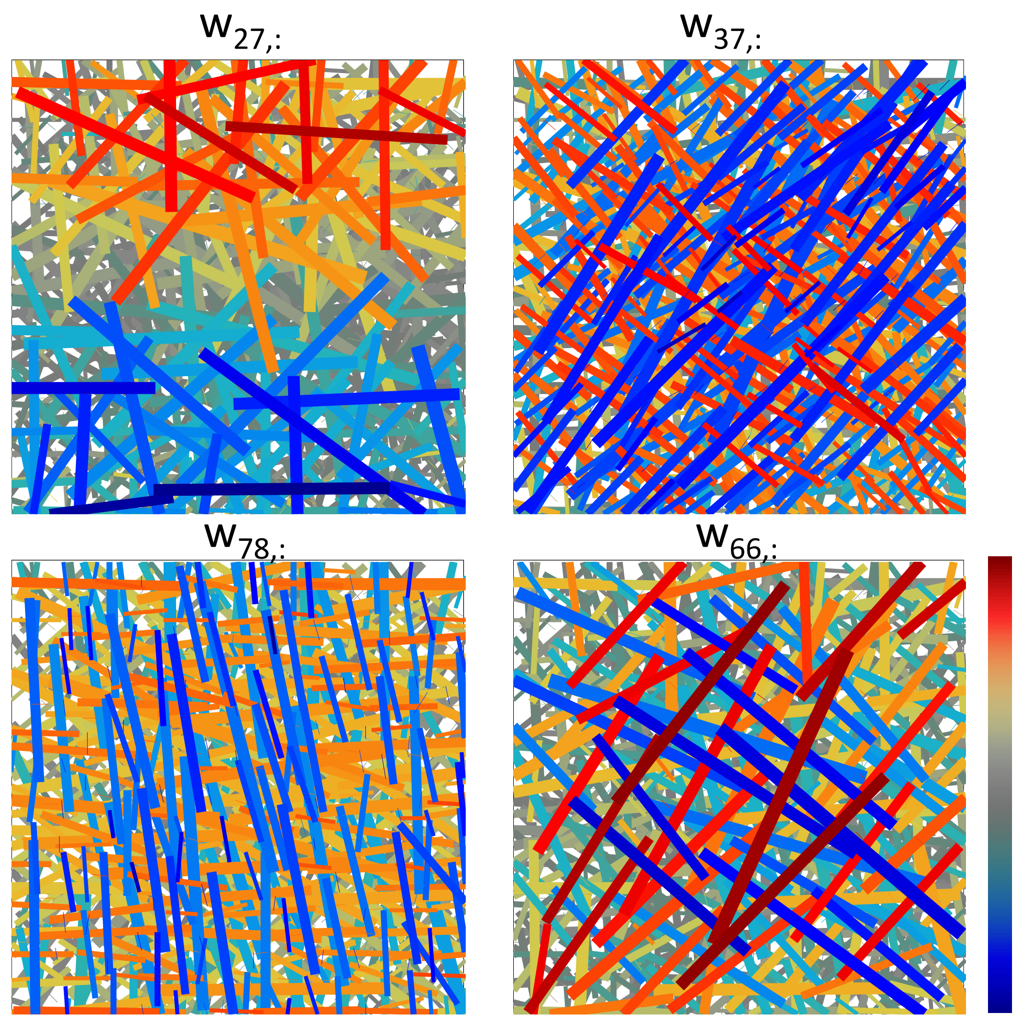


$\mathbf{w}_{105,:}$
$\mathbf{w}_{85,:}$
$\mathbf{w}_{57,:}$
$\mathbf{w}_{11,:}$
0

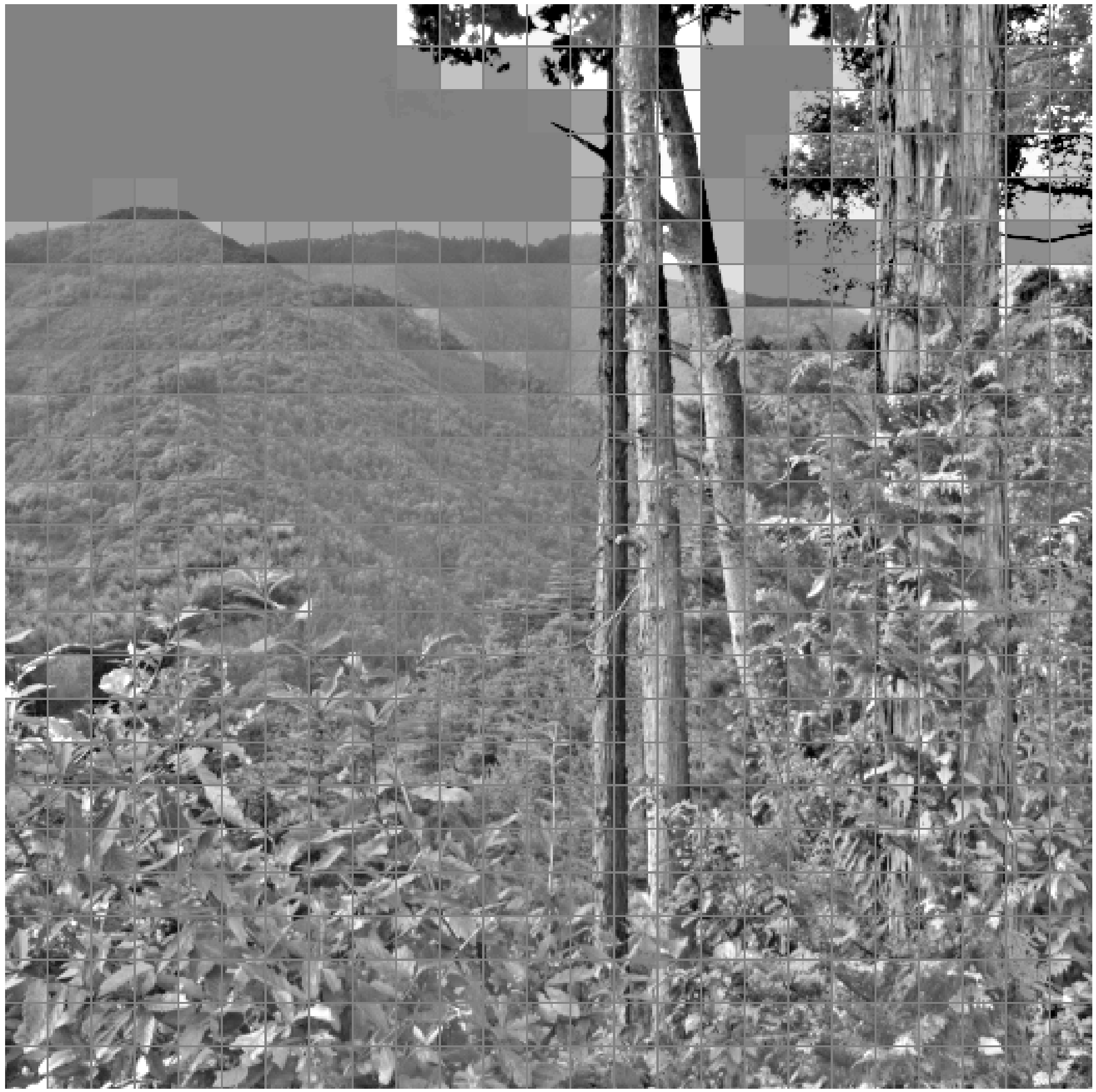

y_1 y_2 $\cdots$ y_J
w_{jk}
$\mathbf{b}_k$
$p(\mathbf{x}|\mathbf{y})$


$\mathbf{w}_{12,:}$
$\mathbf{w}_{125,:}$
$\mathbf{w}_{90,:}$
$\mathbf{w}_{144,:}$

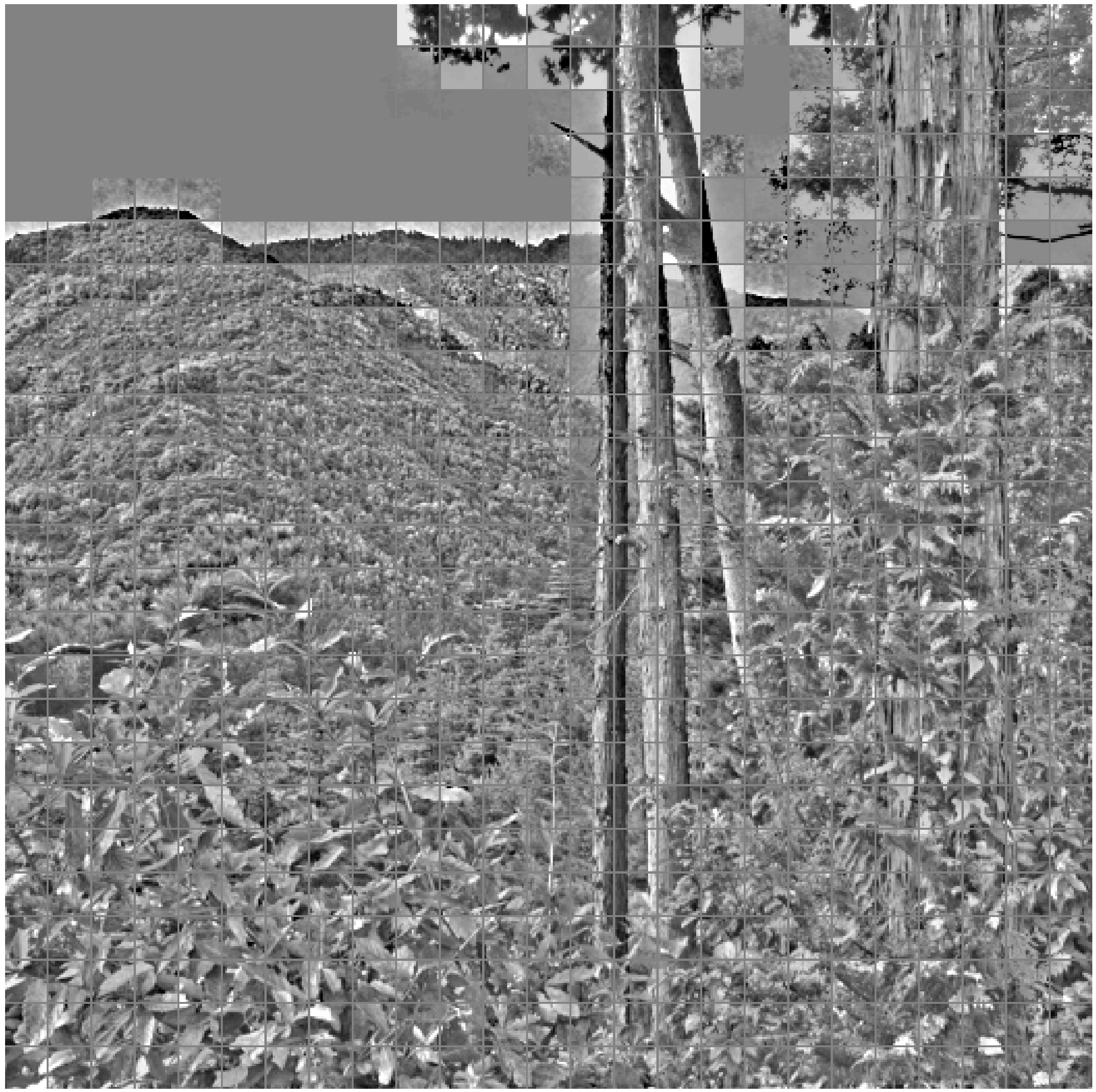

y_1 y_2 $\cdots$ y_J
w_{jk}
$\mathbf{b}_k$
$p(\mathbf{x}|\mathbf{y})$


$w_{27,:}$
$w_{37,:}$
$w_{78,:}$
$w_{66,:}$

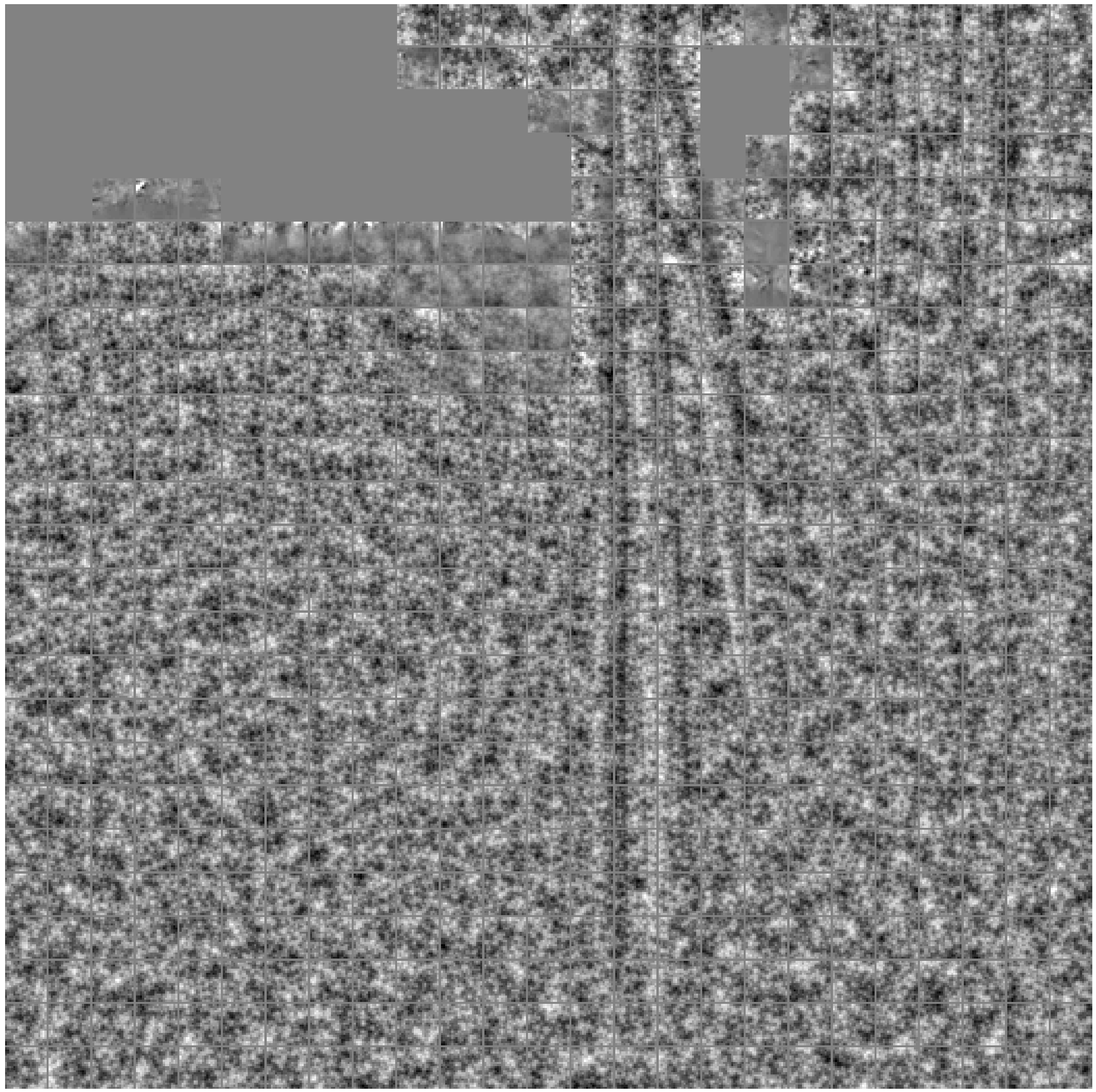
original patches



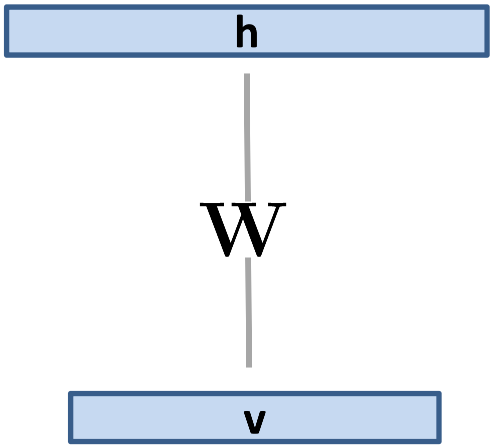
normalization by *inferred contrast*



normalization by *inverse covariance*

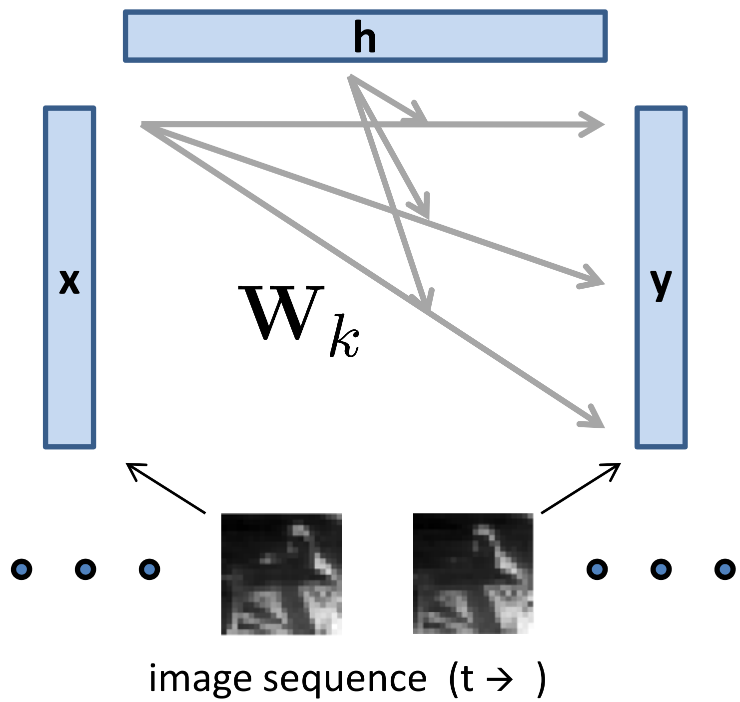


Restricted Boltzmann Machines



$$E(\mathbf{v}, \mathbf{h}) = -\sum W_{ij} v_i h_j$$

gated Restricted Boltzmann Machines *(Memisevic and Hinton, CVPR 07)*



$$E(\mathbf{x}, \mathbf{y}, \mathbf{h}) = -\sum W_{ijk} x_i y_j h_k$$

$$E(\mathbf{x}, \mathbf{y}, \mathbf{h}) = -\mathbf{x}^T \left(\sum h_k \mathbf{W}_k\right) \mathbf{y}$$

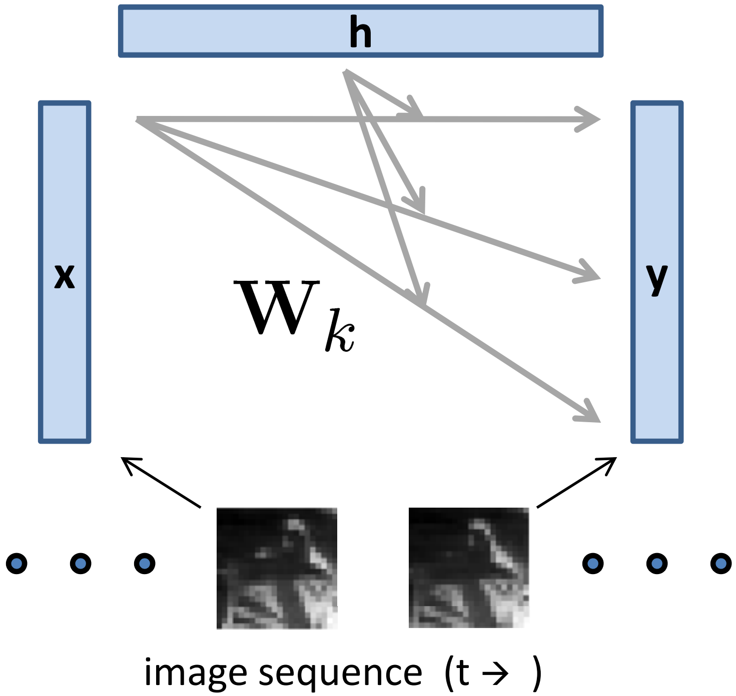
$\mathbf{x}$ - Gaussian (pixels)

$\mathbf{y}$ - Gaussian (pixels)

$\mathbf{h}$ - binary

trained on image sequences

gated Restricted Boltzmann Machines *(Memisevic and Hinton, CVPR 07)*



$$E(\mathbf{x}, \mathbf{y}, \mathbf{h}) = -\sum W_{ijk} x_i y_j h_k$$

$$E(\mathbf{x}, \mathbf{y}, \mathbf{h}) = -[\mathbf{x}; \mathbf{y}]^T \left(\sum h_k \mathbf{W}_k\right) [\mathbf{x}; \mathbf{y}]$$

$\mathbf{x}$ - Gaussian (pixels)

$\mathbf{y}$ - Gaussian (pixels)

$\mathbf{h}$ - binary

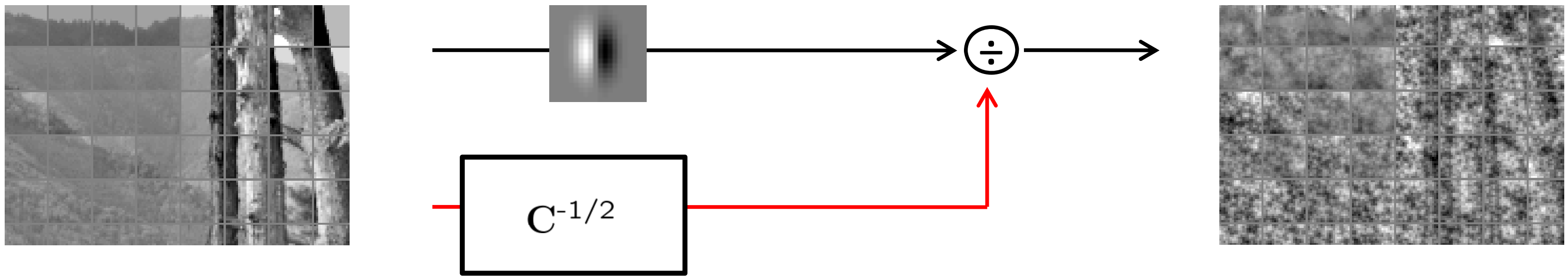
trained on image sequences

log-Covariance factor model

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(0, \mathbf{C})$$

$$\log \mathbf{C} = \sum_j y_j \mathbf{A}_j$$

$$\log p(\mathbf{x}|\mathbf{y}) \propto -\mathbf{x}^T e^{(-\sum y_j \mathbf{A}_j)} \mathbf{x}$$



summary

- modeling non-linear dependencies motivated by statistical patterns
- higher-order models can capture context
- normalization by local “whitening” removes most structure

questions

- joint model for *normalized response* and *normalization (context) signal*?
- how to extend to entire image?
- where are these computations found in the brain?