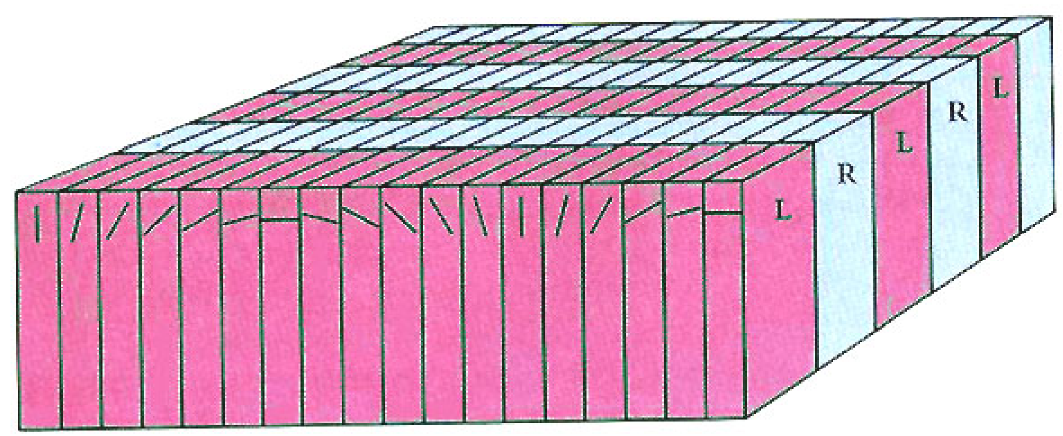
Computational Vision

Primary visual cortex

- Orientation selectivity
- Spatial frequency
- Color opponency
- Normalization

Computing with V1

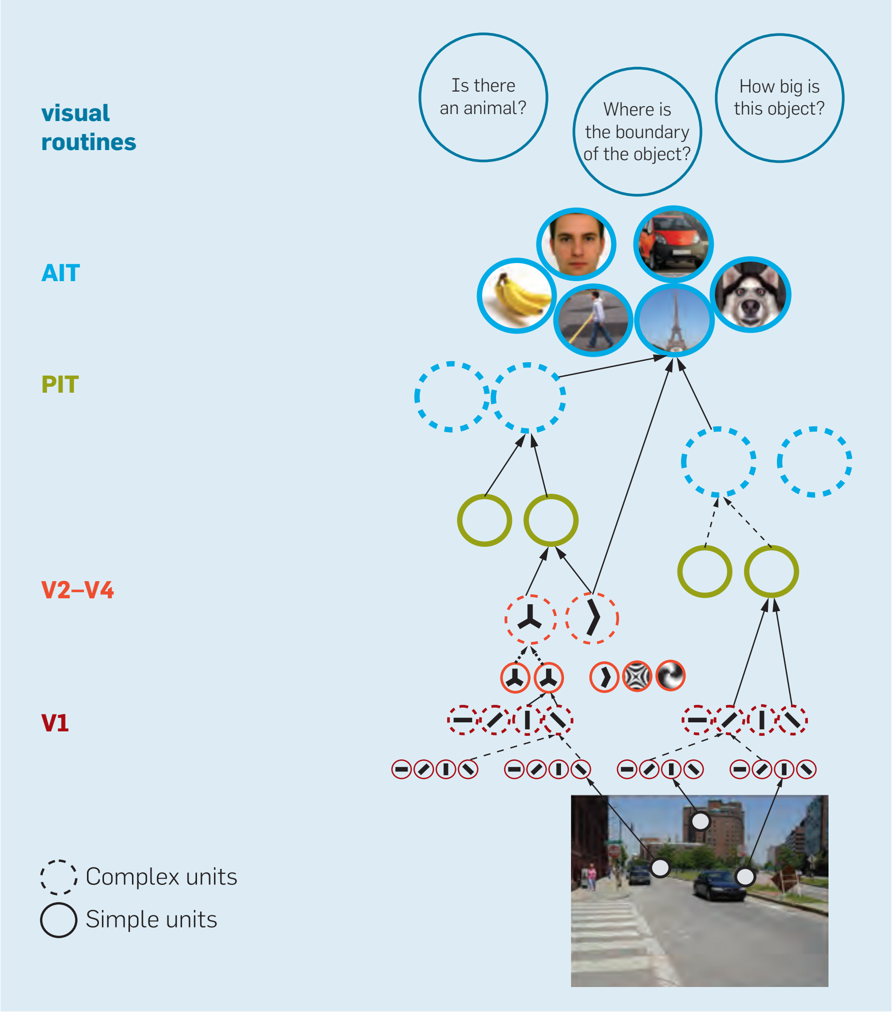
Gabor filters



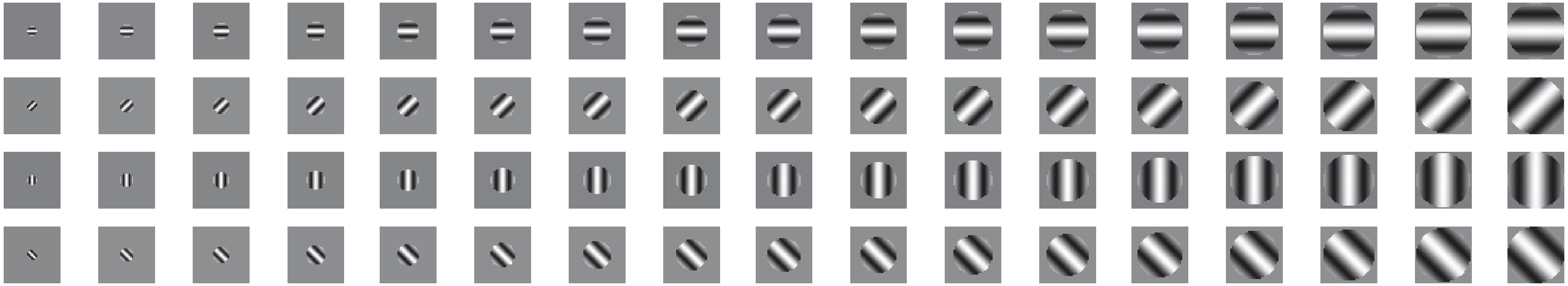
Orientation and ocular dominance columns

Figure 23. The ice-cube model of the cortex. It illustrates how the cortex is divided, at the same time, into two kinds of slabs, one set of ocular dominance (left and right) and one set for orientation. The model should not be taken literally: Neither set is as regular as this, and the orientation slabs especially are far from parallel or straight.

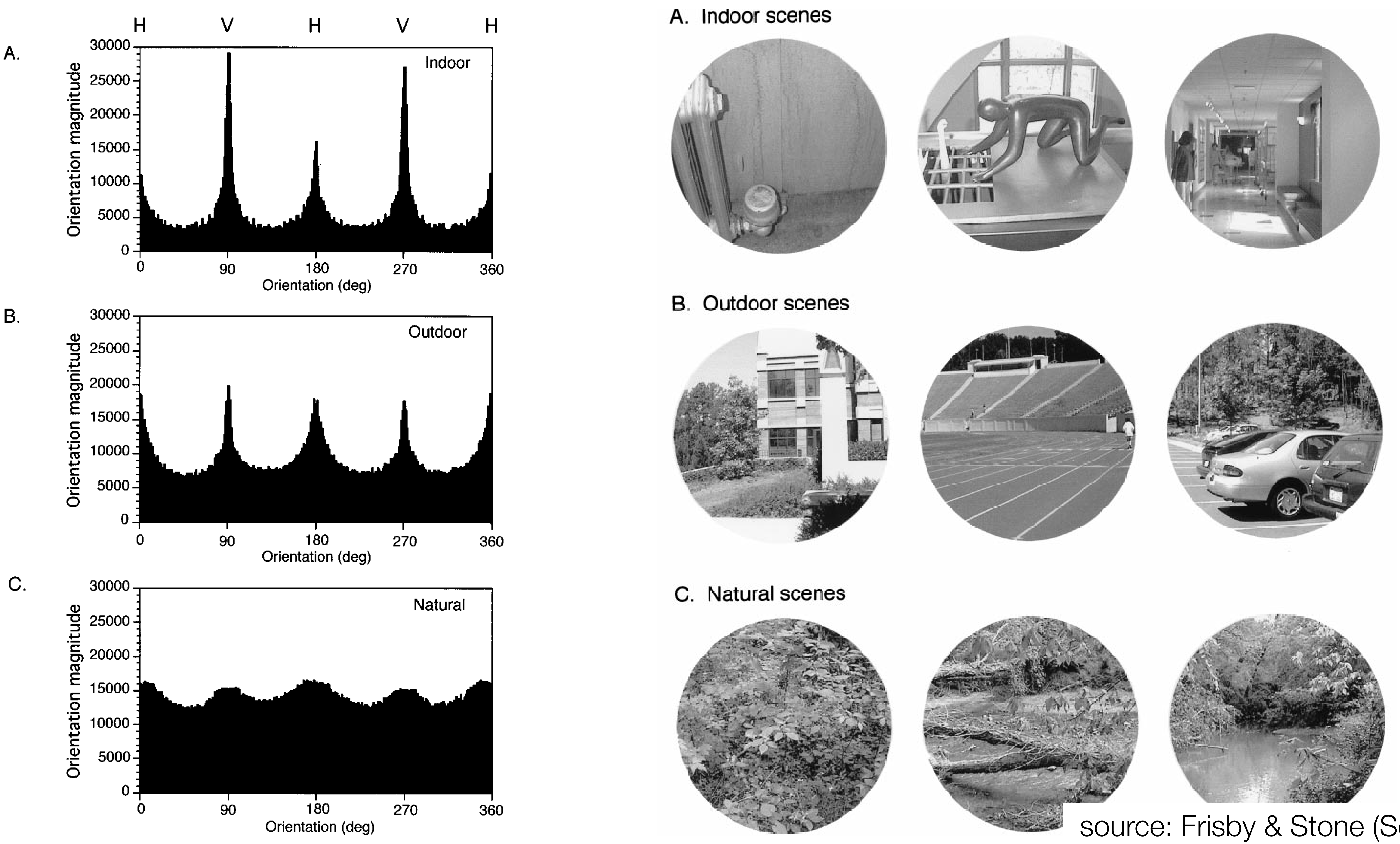
Figure 2. Hierarchical feedforward models of the visual cortex.



Gabor filters at multiple phases (one phase shown), orientations and spatial frequencies/scales (parameters derived from available experimental data)

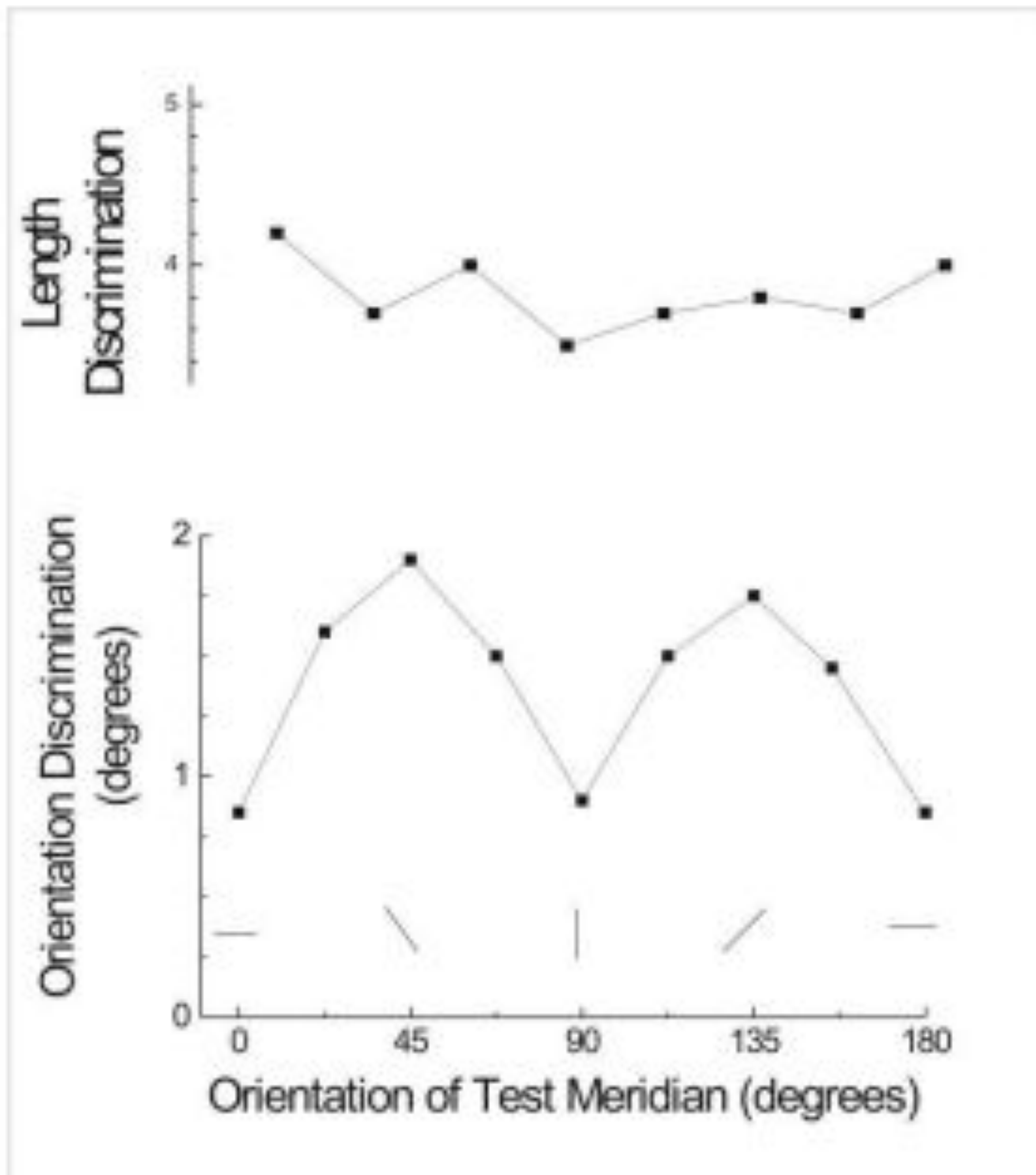


Structure in natural images



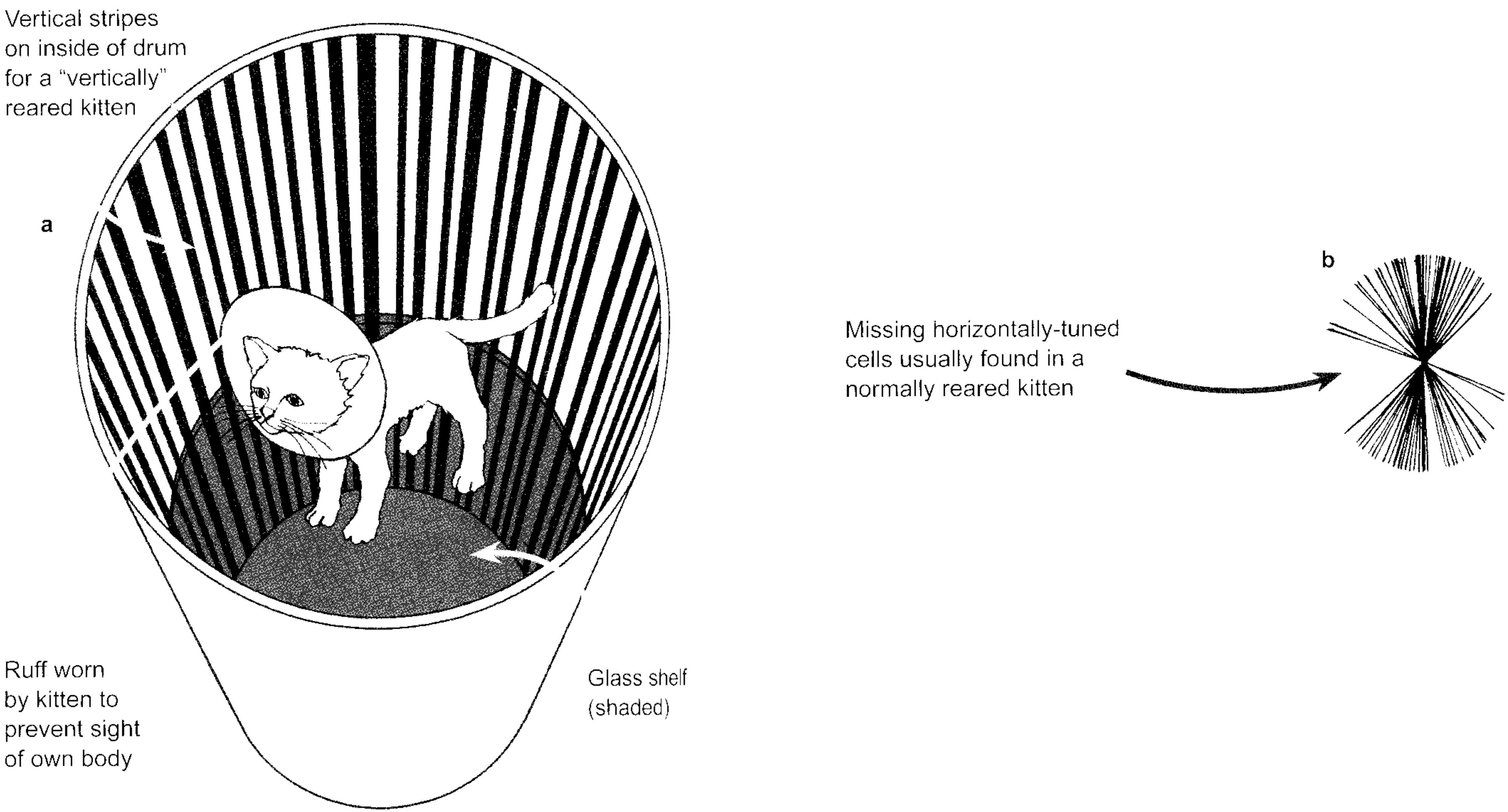
source: Frisby & Stone (Seeing)

The oblique effect



source: wikipedia

Rearing experiments



Blakemore & Cooper '70; source: Frisby & Stone (Seeing)

Natural image statistics cont’d

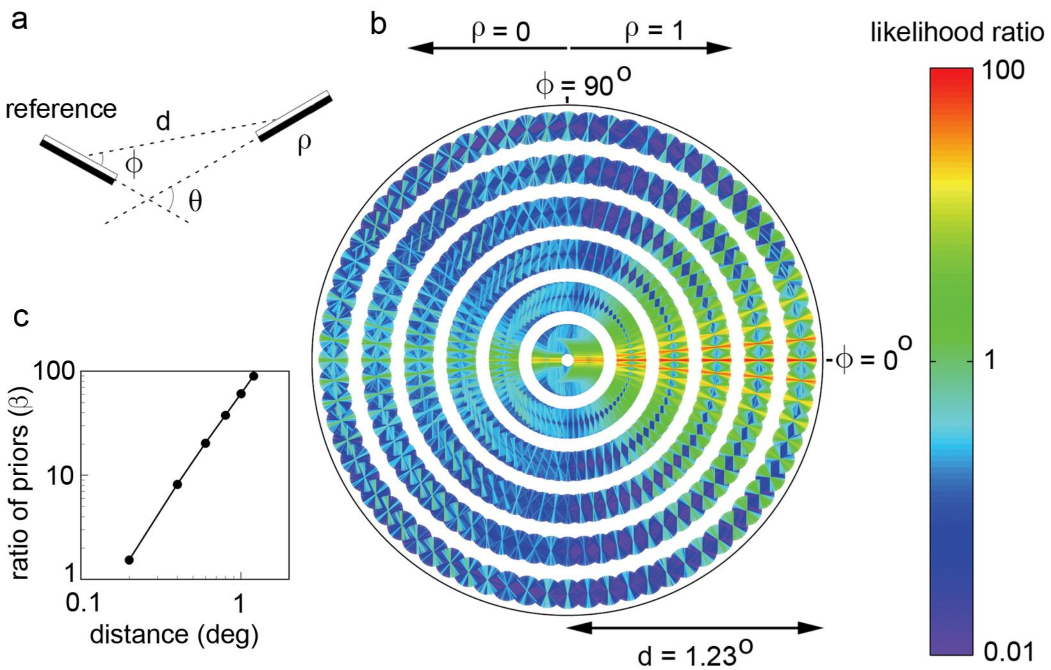
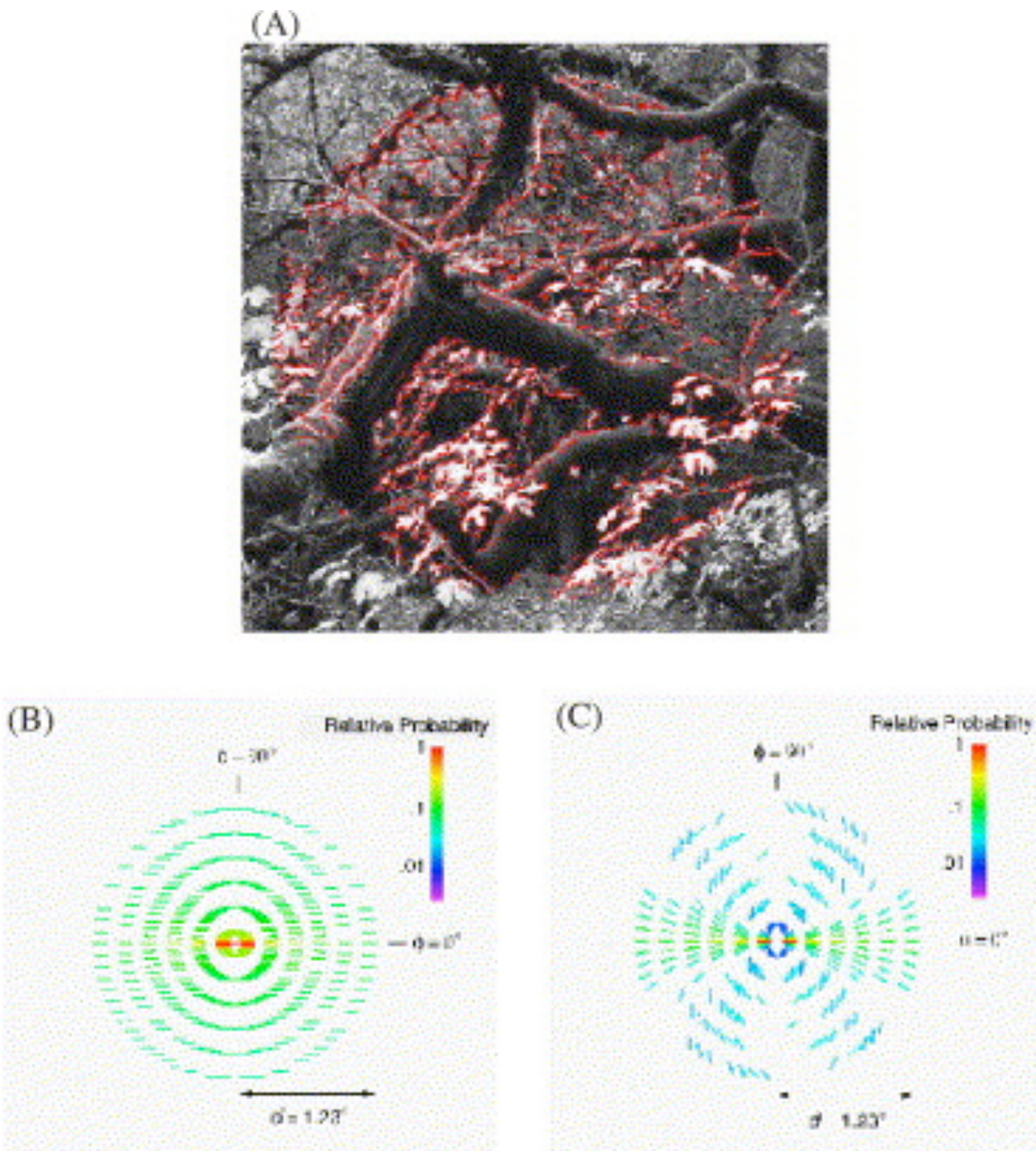
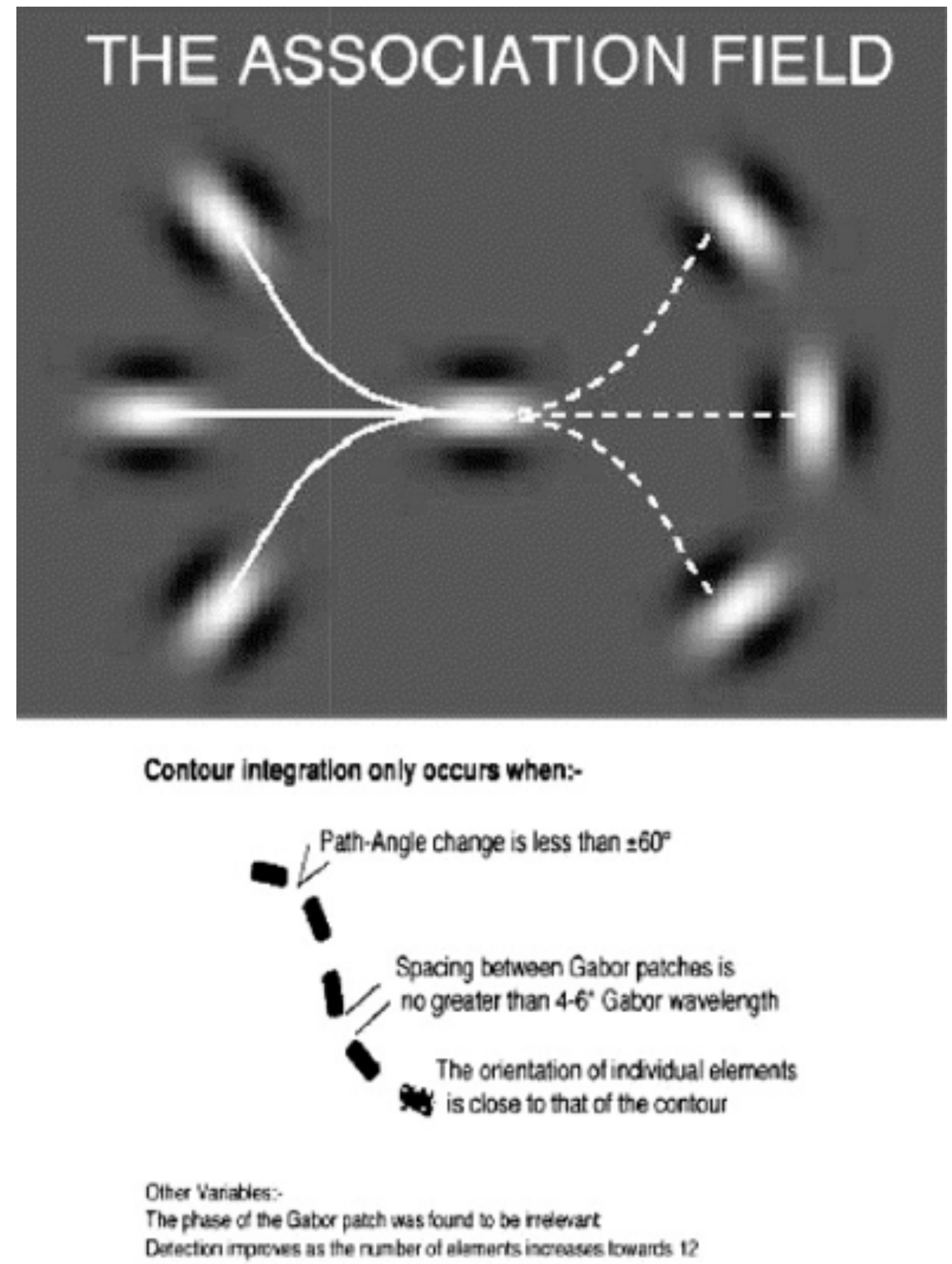


Fig. 2. Co-occurrence statistics of contour elements in natural images. **a**. Definition of parameters describing the geometrical and contrast relationship between a pair of contour elements. **b**. Plot of the likelihood ratio for a given relationship between pairs of contour elements. **c**. Ratio of the prior probabilities that pair of contour elements belong to different versus the same physical source, as a function of distance between the pair of elements. A likelihood ratio greater then 1.0 means (given equal priors) it is more likely that the elements belong to the same physical contour; a ratio less than 1.0 means it is more likely that the elements belong to different physical contours. [For each distance, direction and polarity, the orientation difference bins (line segments) are drawn in rank order starting from the lowest likelihood; thus, the highest likelihoods are the most visible in the plot.]

Geisler 2008

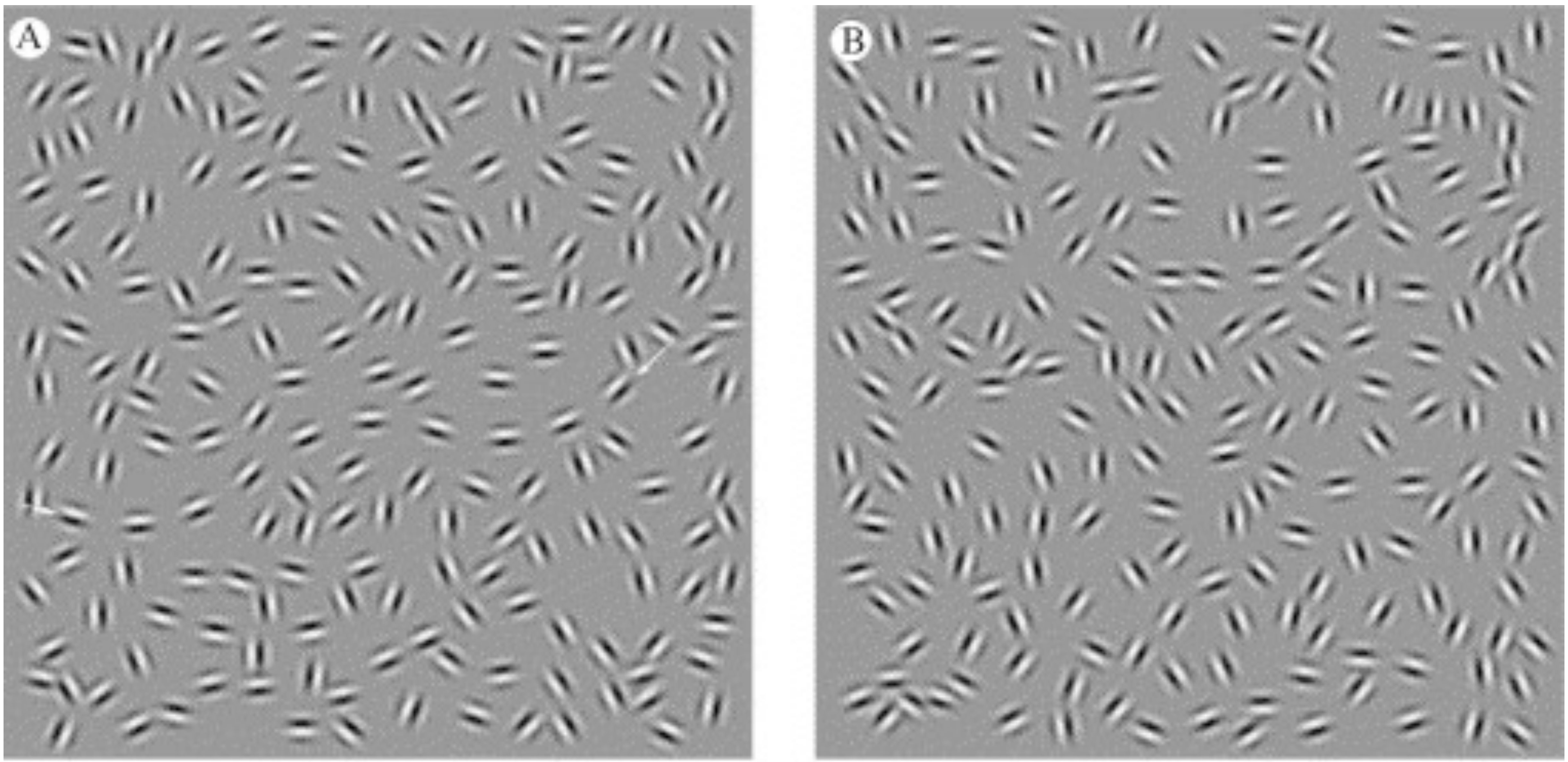
Natural image statistics cont'd





Hess & Field '03

Natural image statistics cont'd

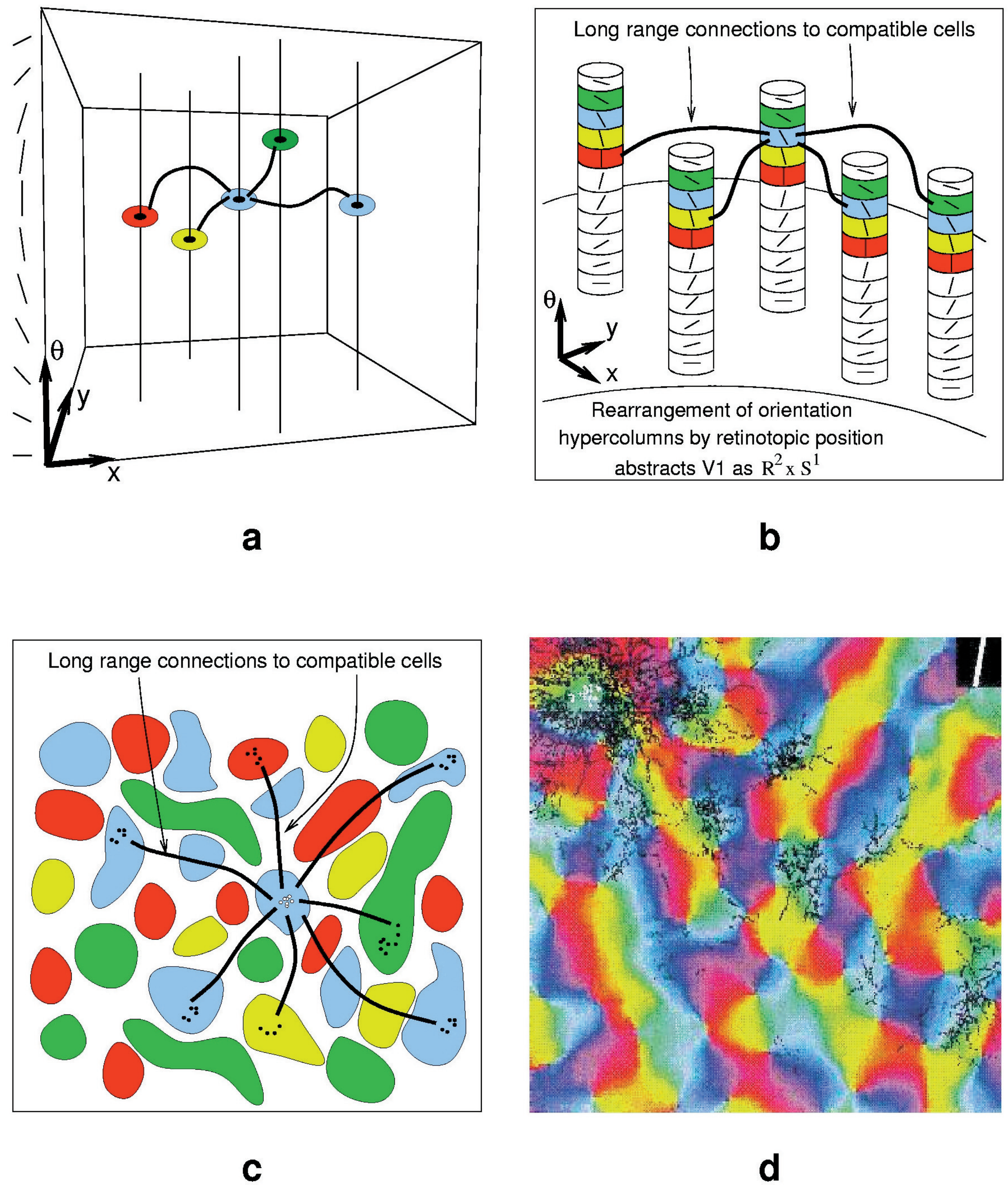


embedded contour

background elements

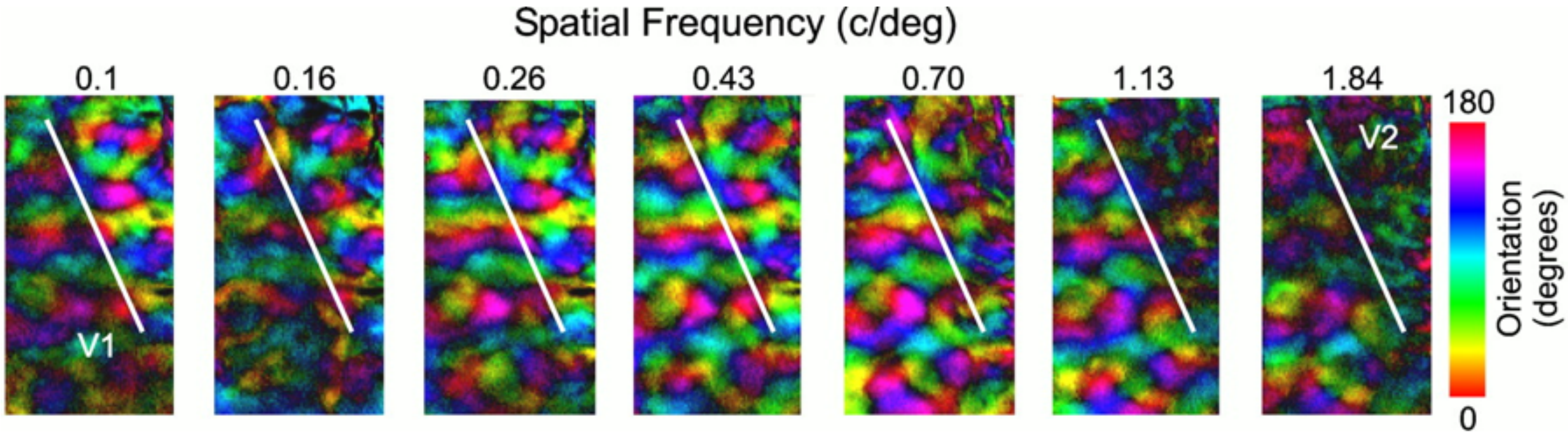
Hess & Field '03

Lateral connections in the primary visual cortex

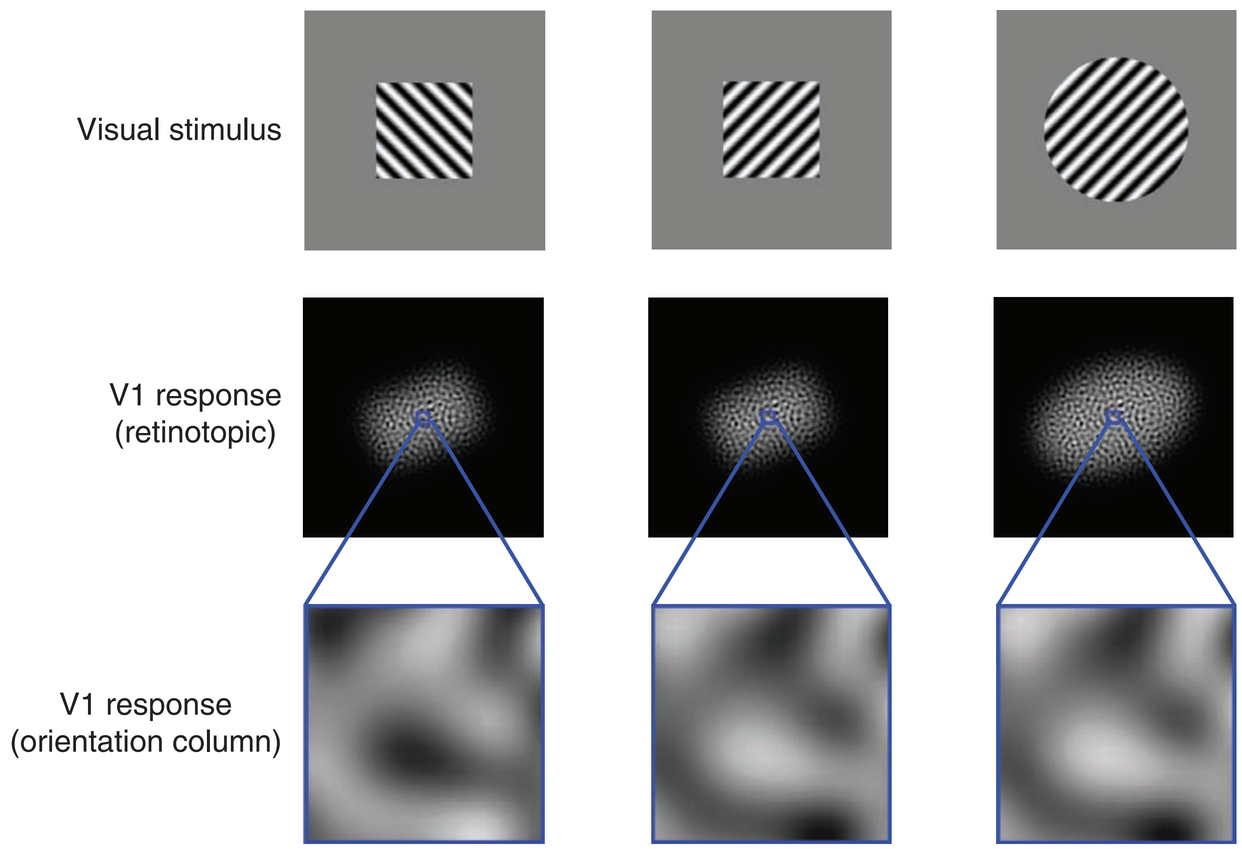


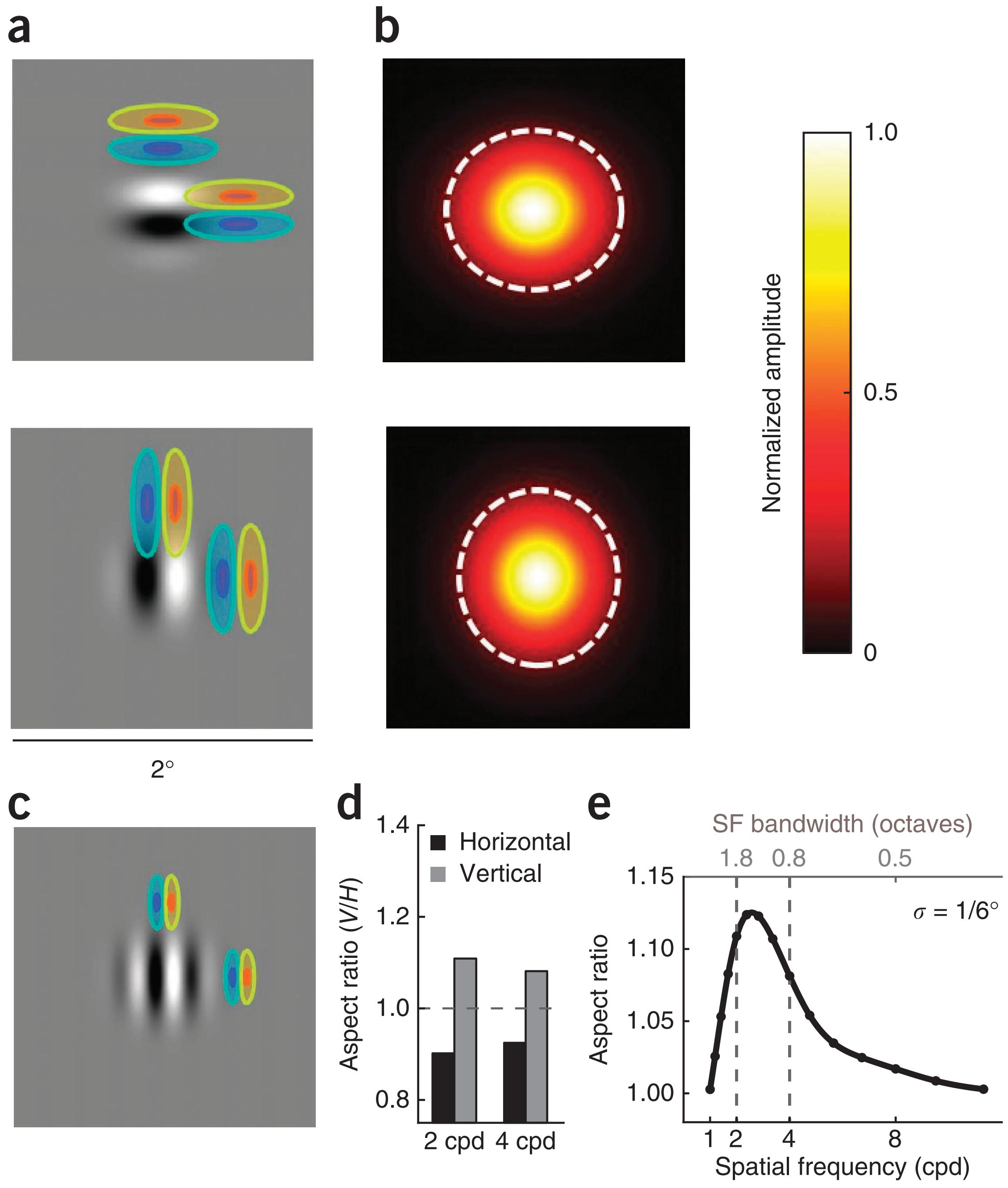
Ben-Shahar & Zucker ’04

Frequency channels



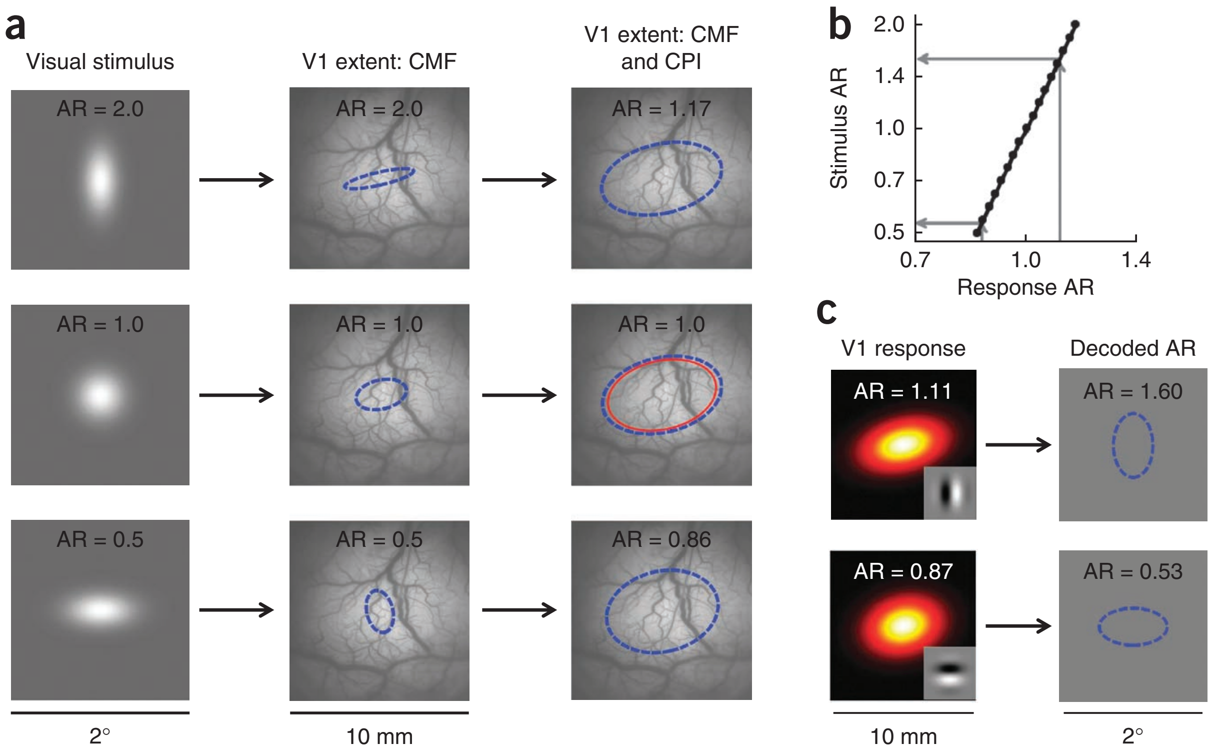
Computing with V1

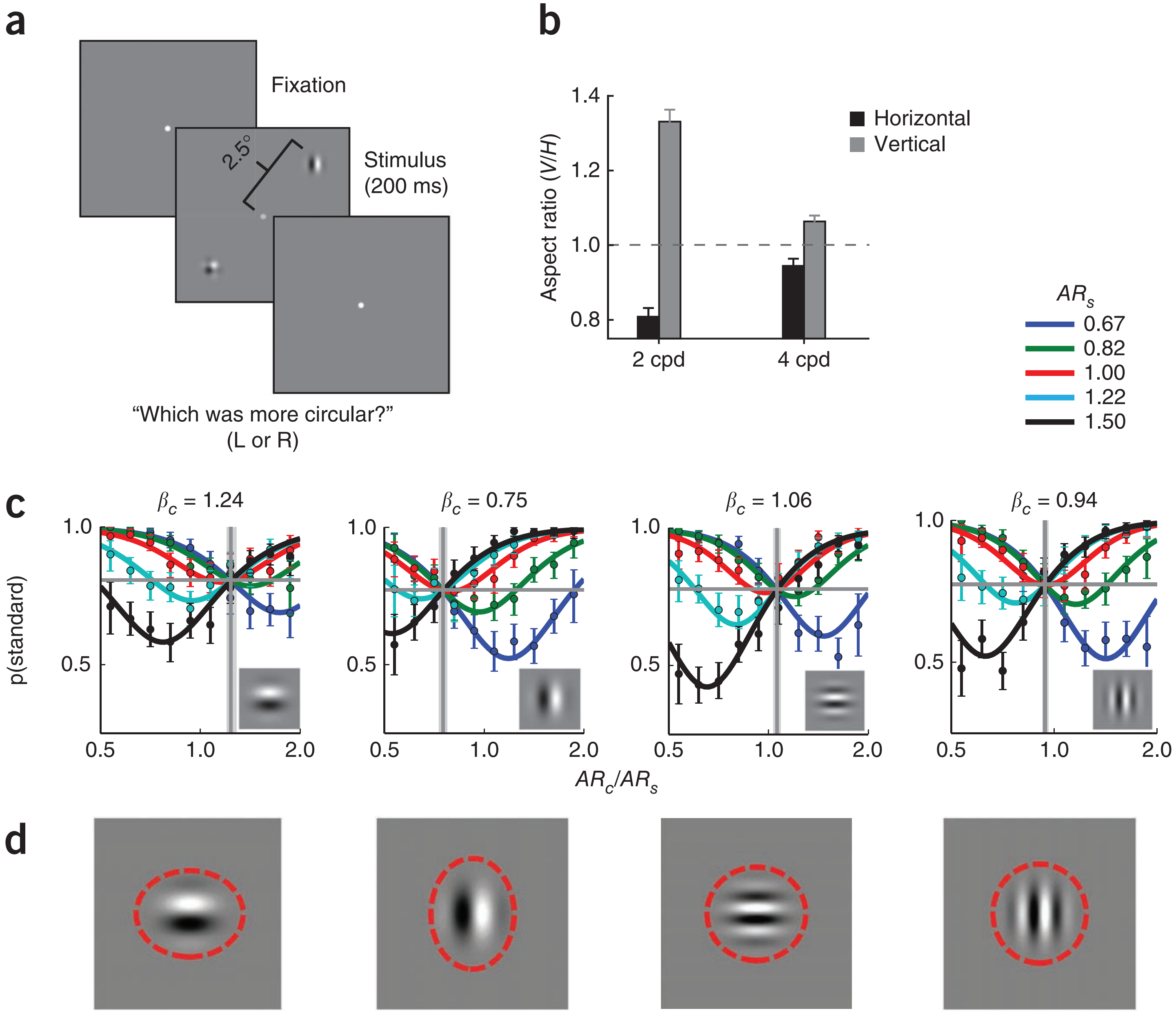




Michel et al 2013

Computing with V1





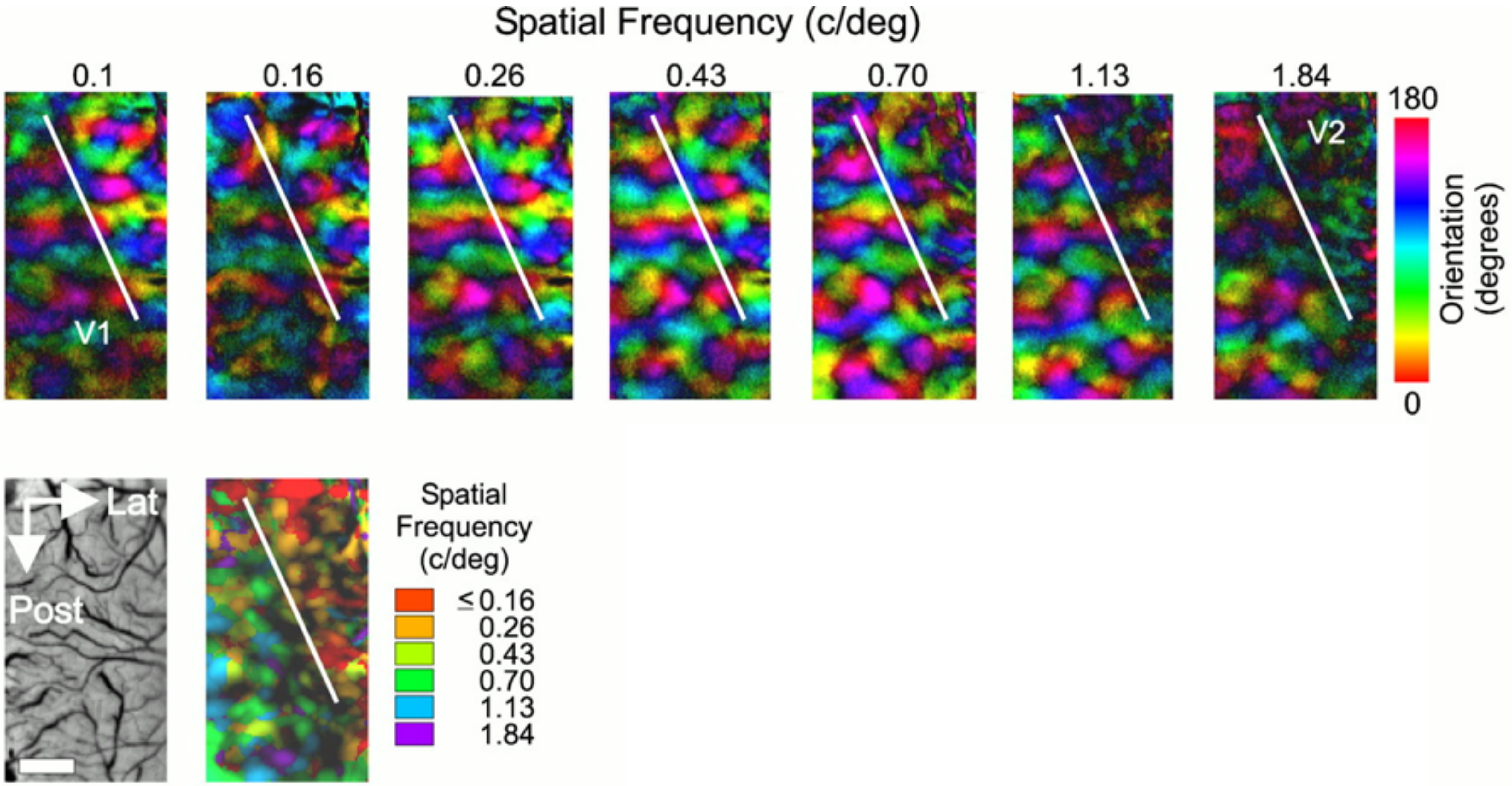
Michel et al 2013

Computational Vision

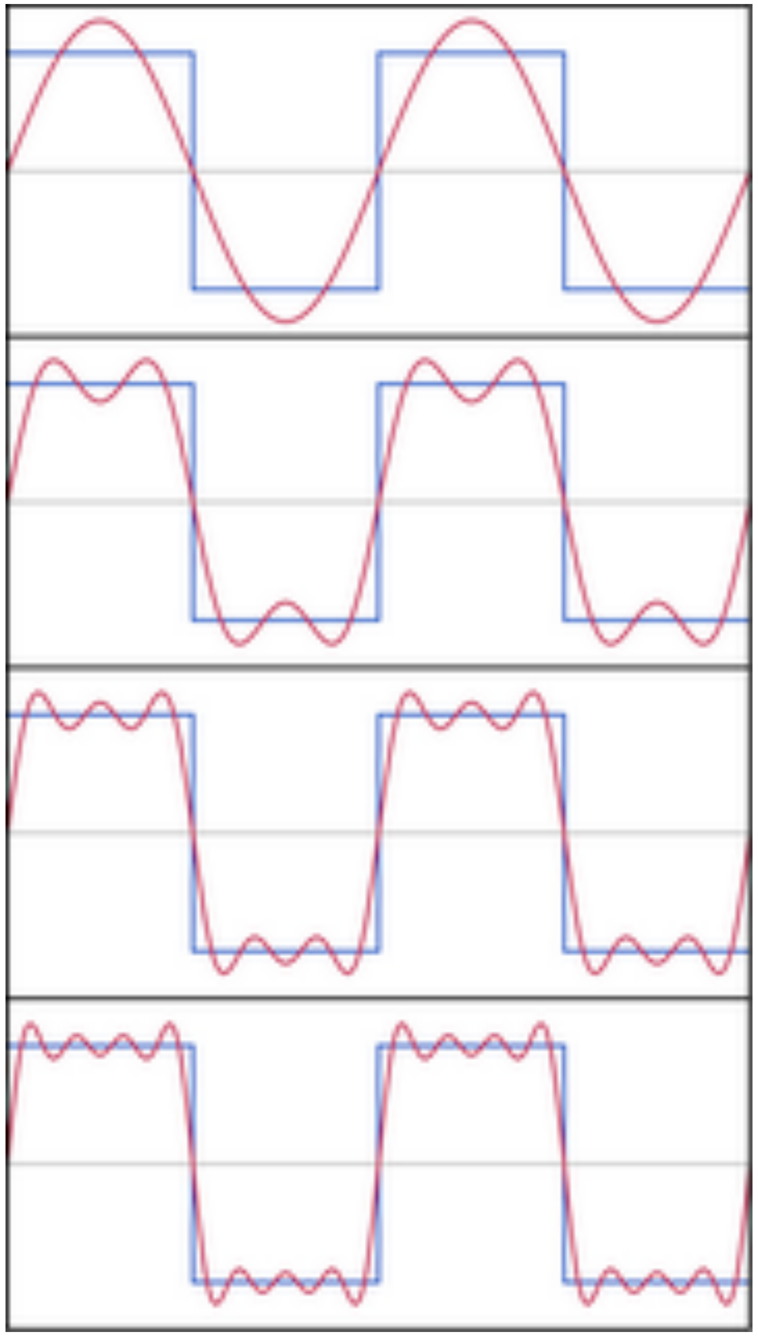
Primary visual cortex

- Orientation selectivity
- Spatial frequency
- Color opponency
- Normalization

Frequency channels

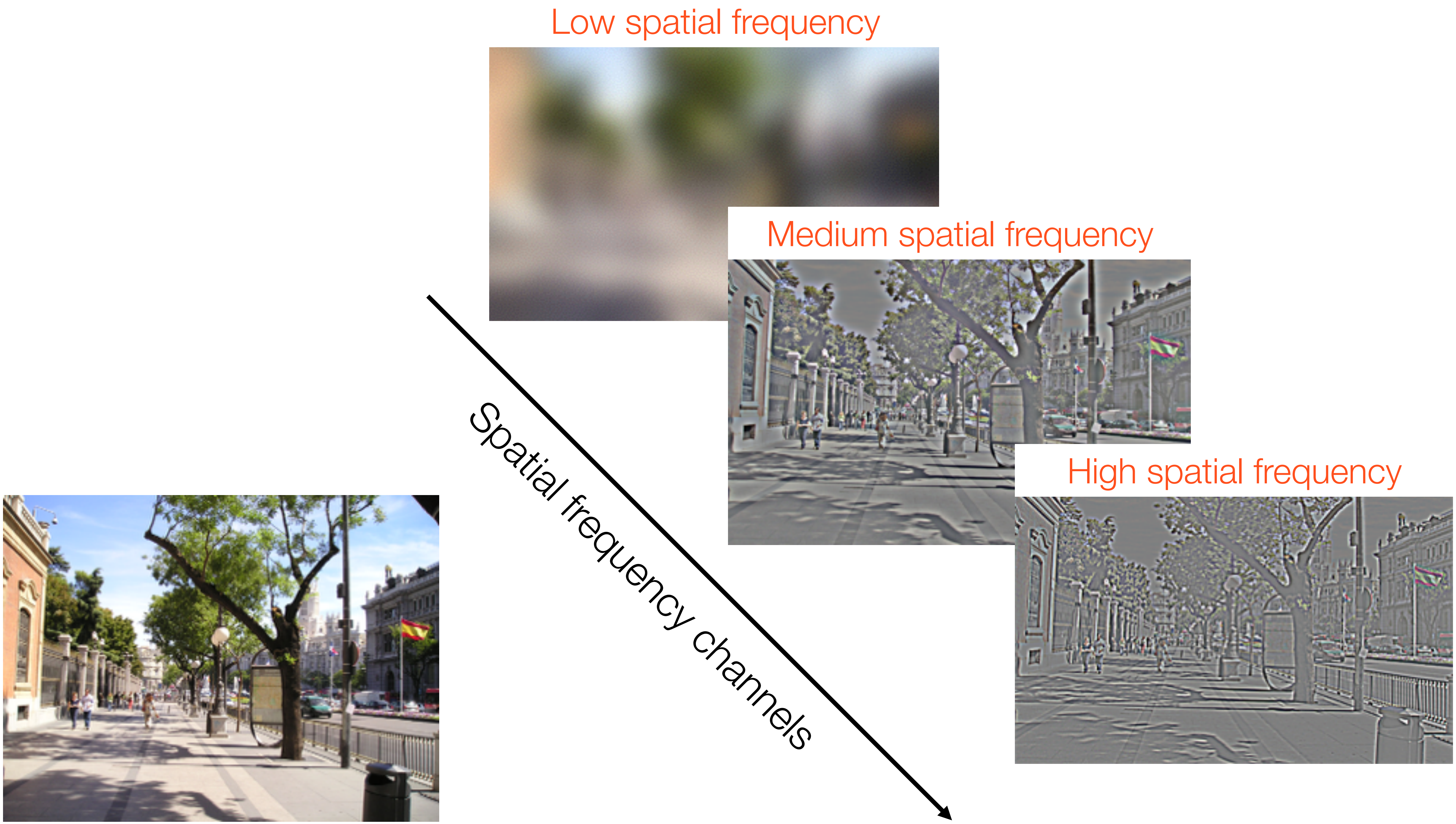


Fourier decompositions



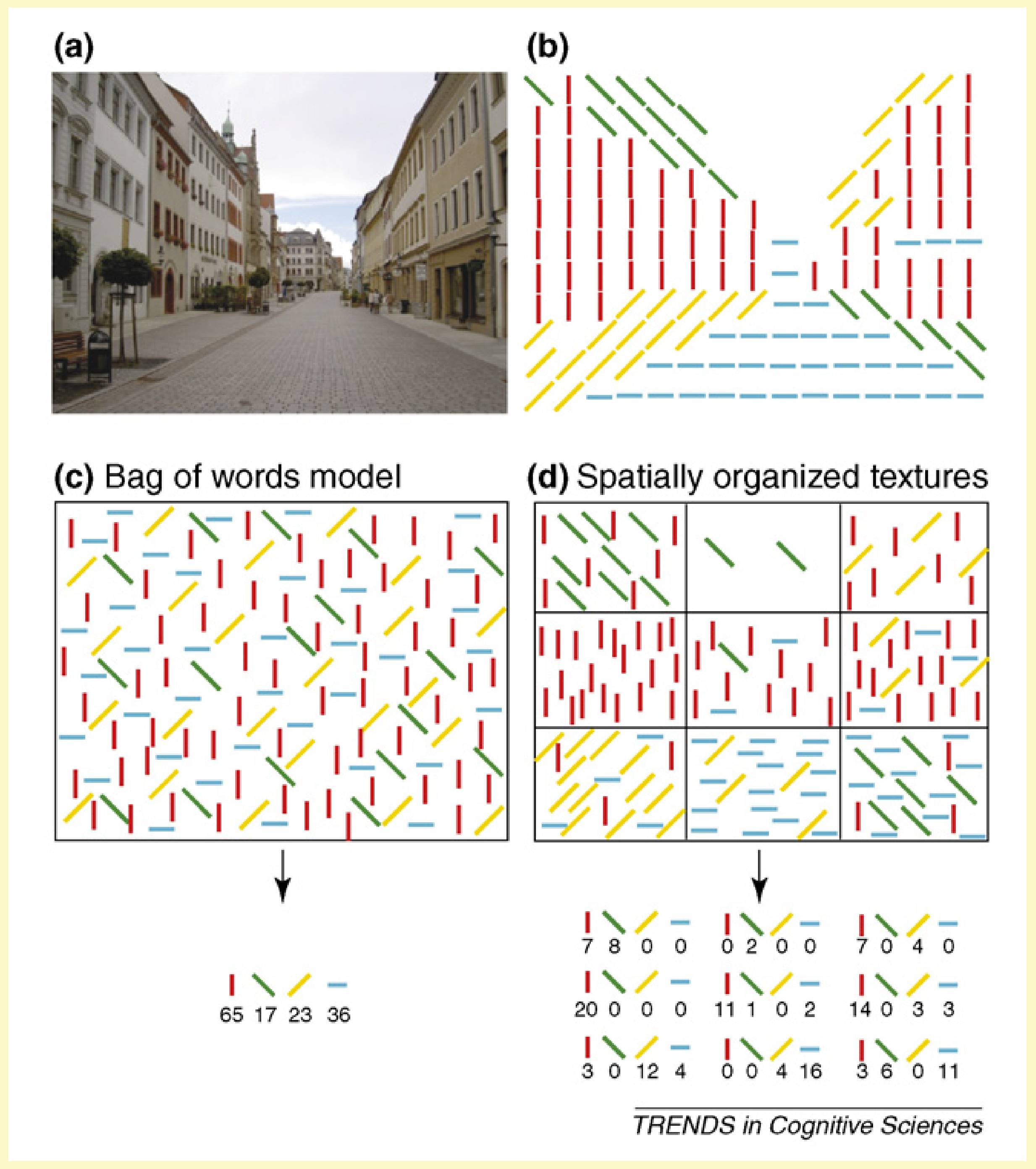
source: http://physicsformom.blogspot.com/2010/04/fourier-analysis-sines-and-integrals.html

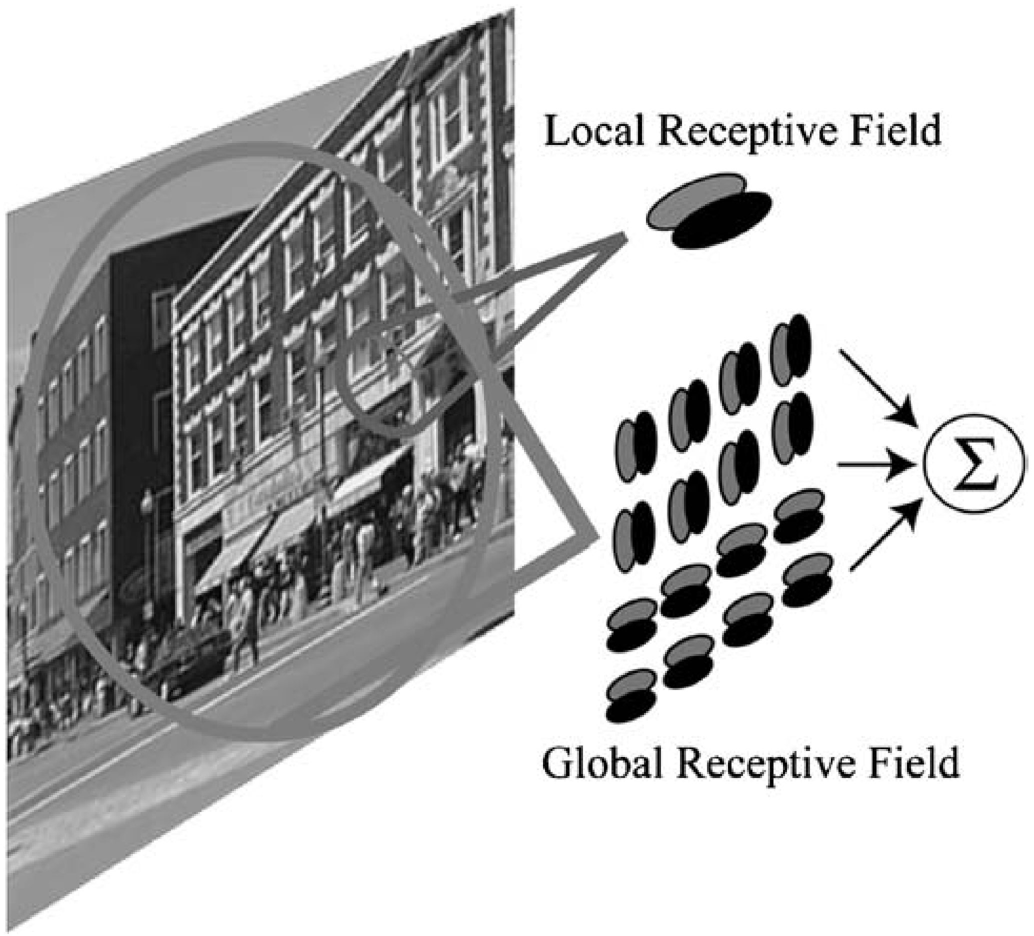
Frequency channels



Computing with V1

Gist descriptor





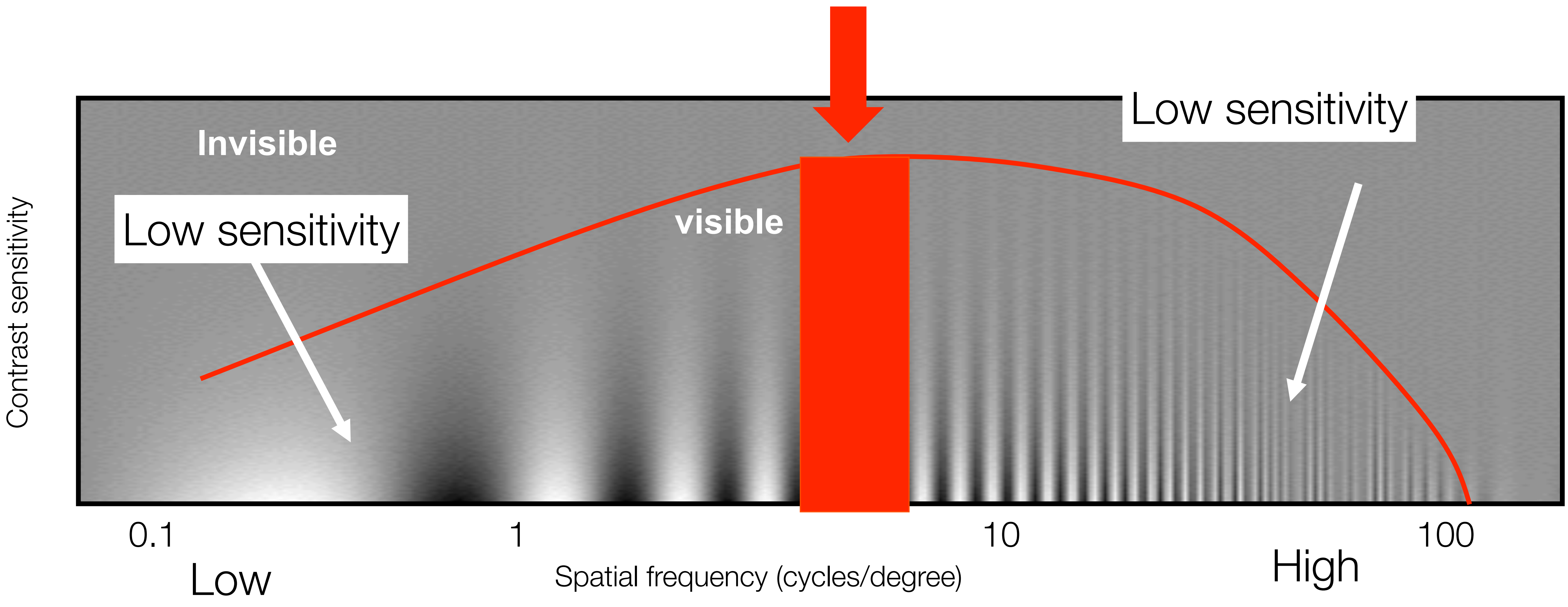
source: Torralba & Oliva 2006 2007

Contrast sensitivity function

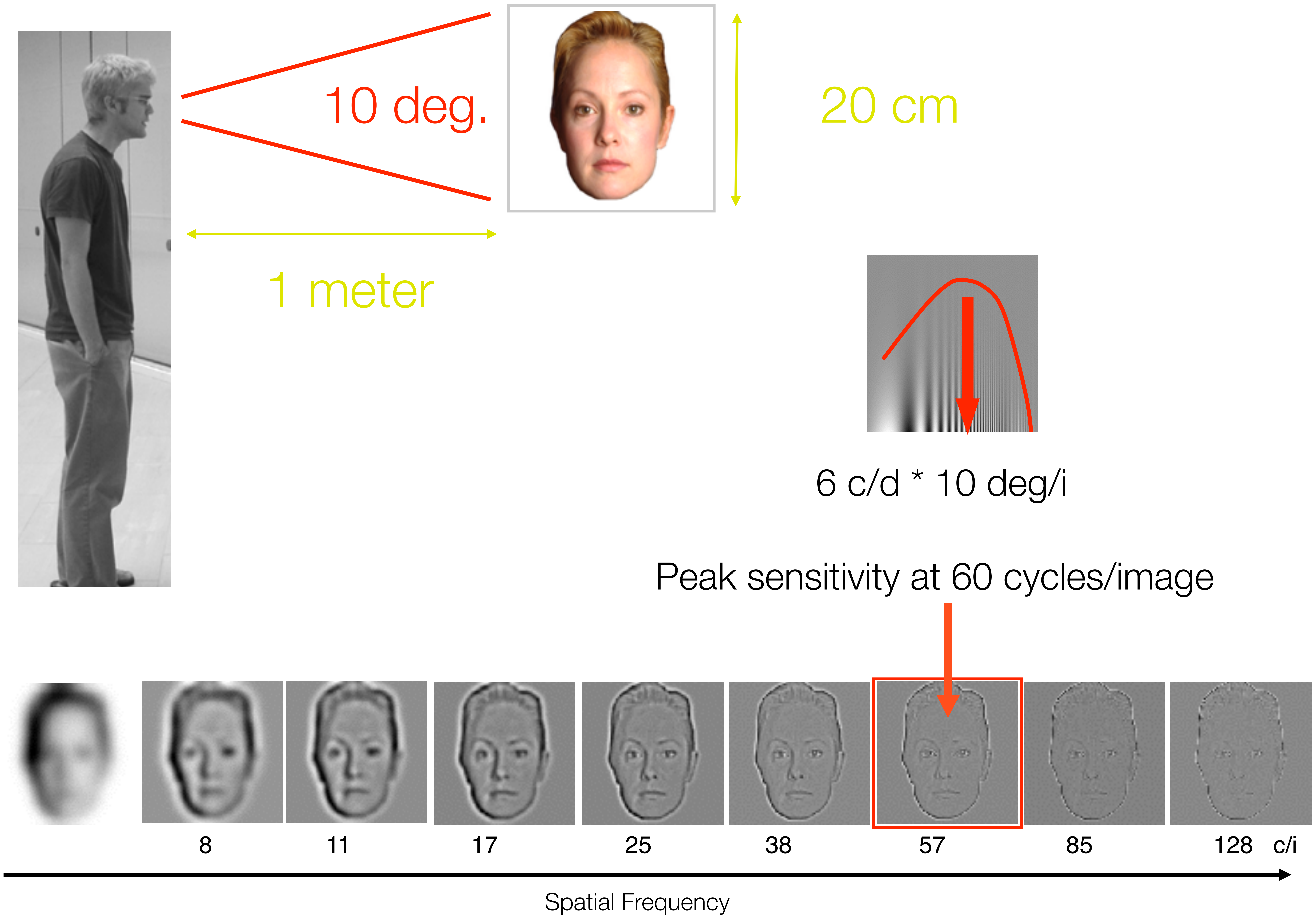
Blackmore & Campbell (1969)

Maximum sensitivity

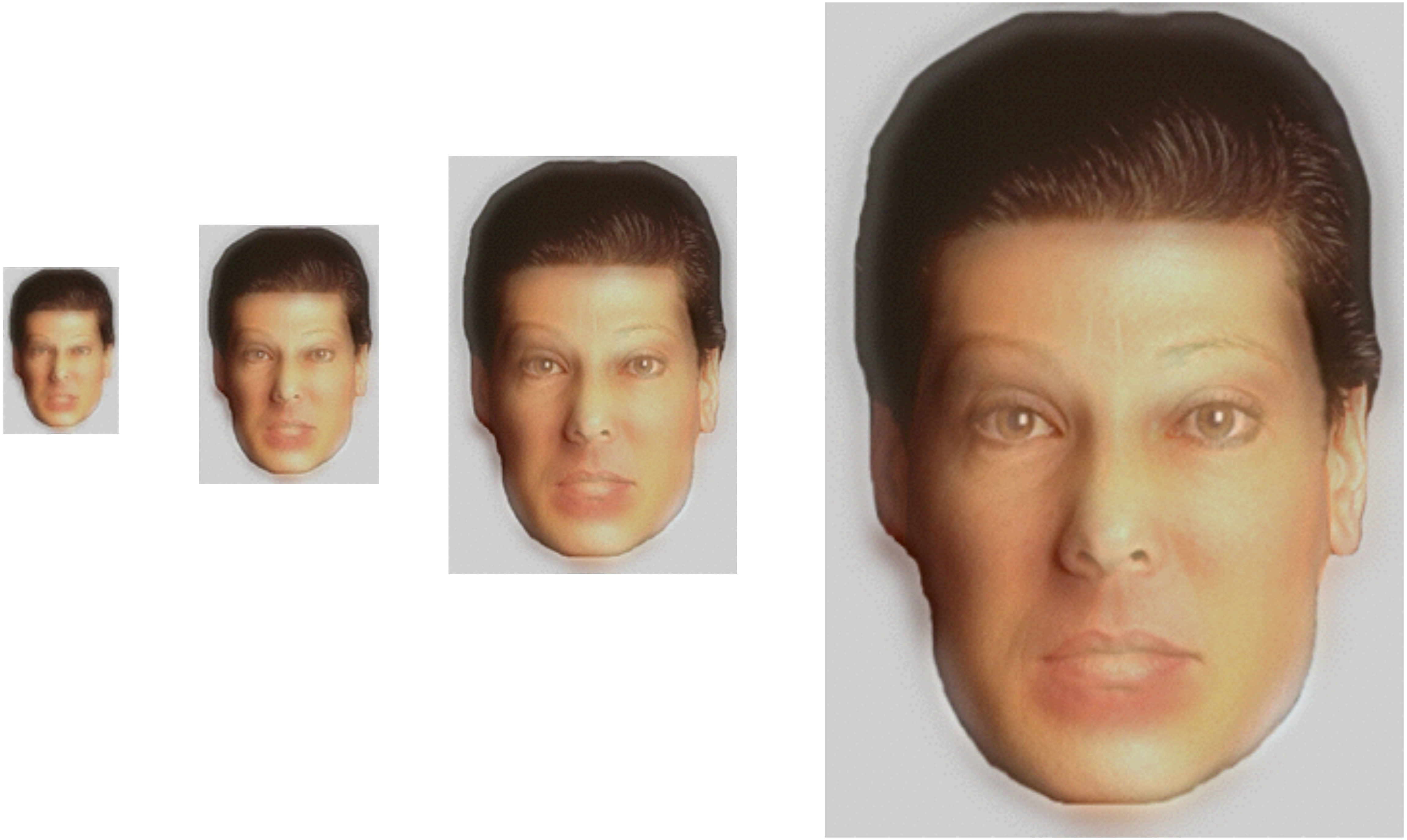
~ 6 cycles / degree of visual angle



source: Aude Oliva



Hybrid Images



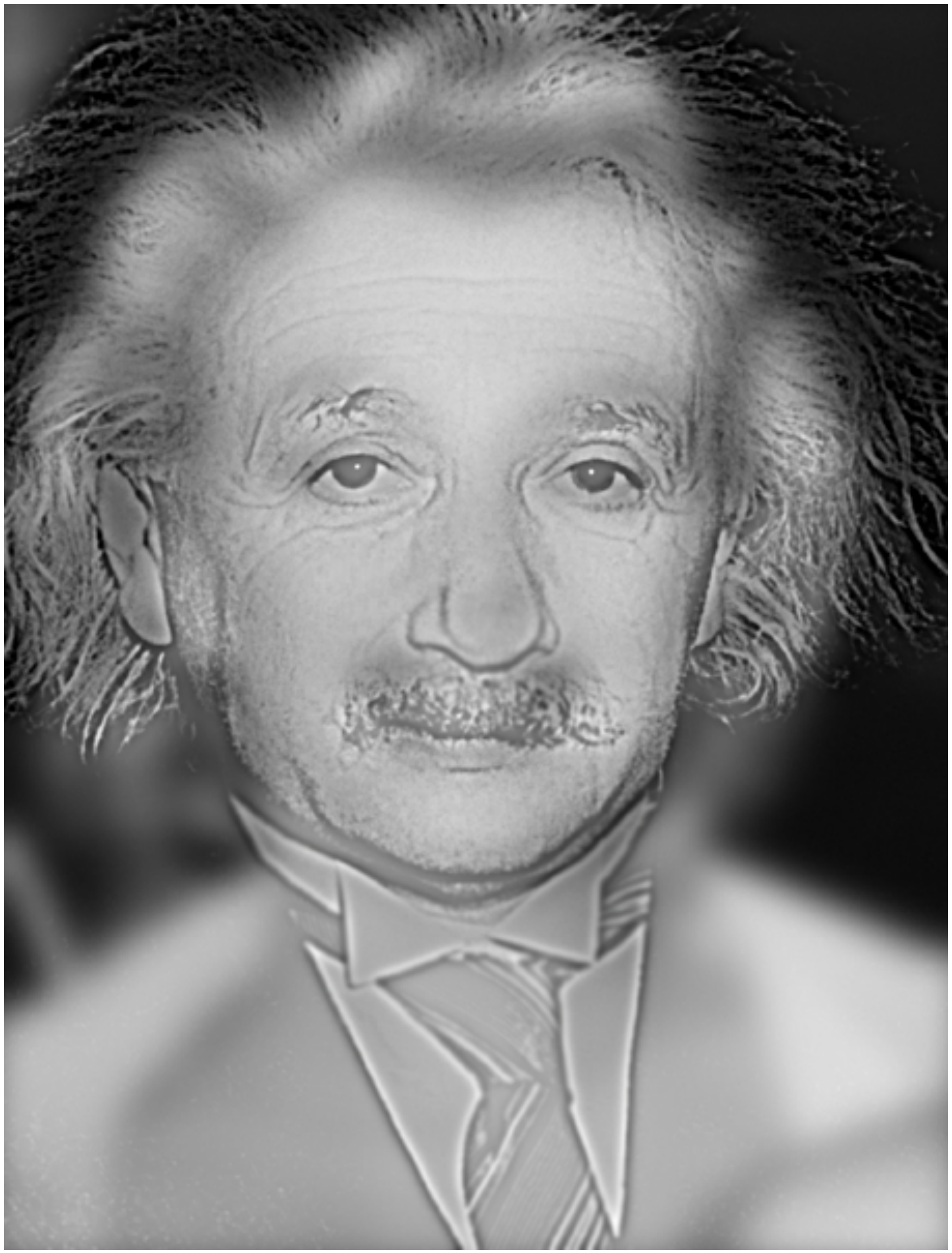
Aude Oliva (MIT)

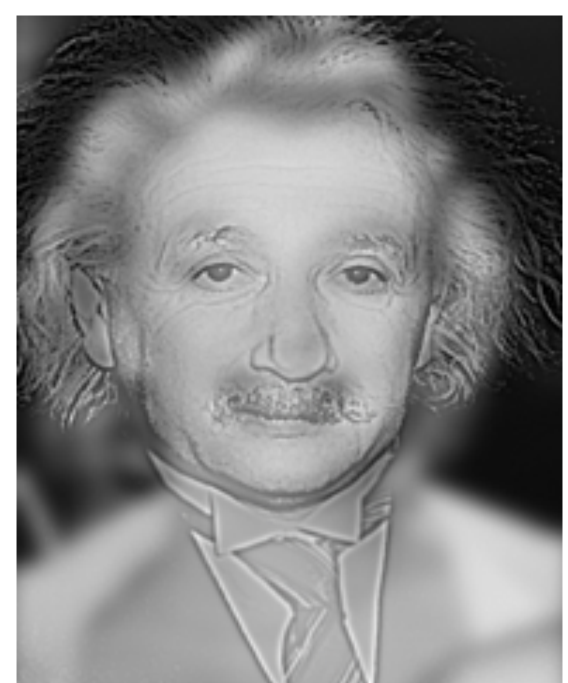
Hybrid Images

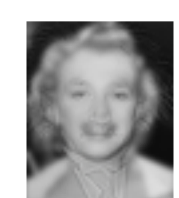
Aude Oliva (MIT)

Hybrid Images

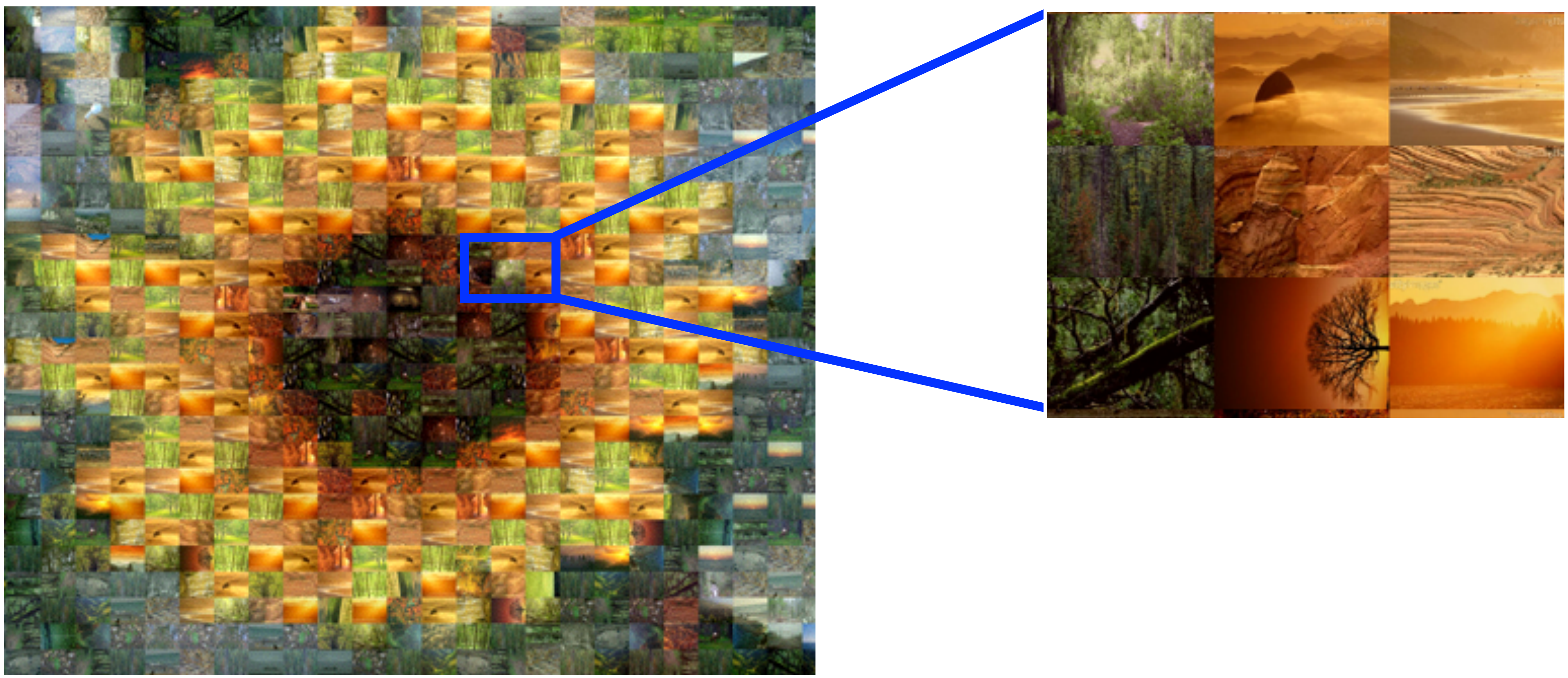
Aude Oliva (MIT)



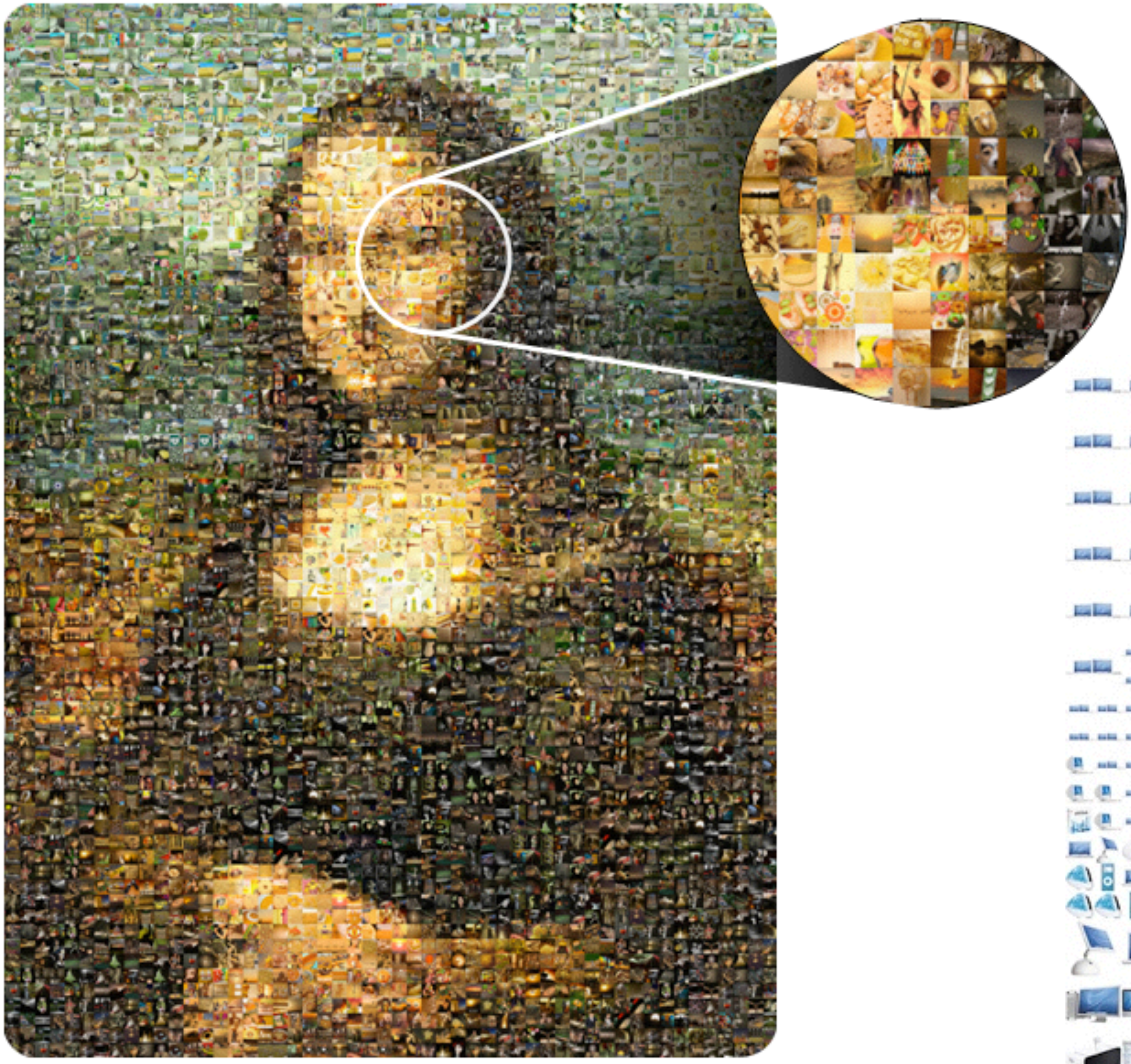


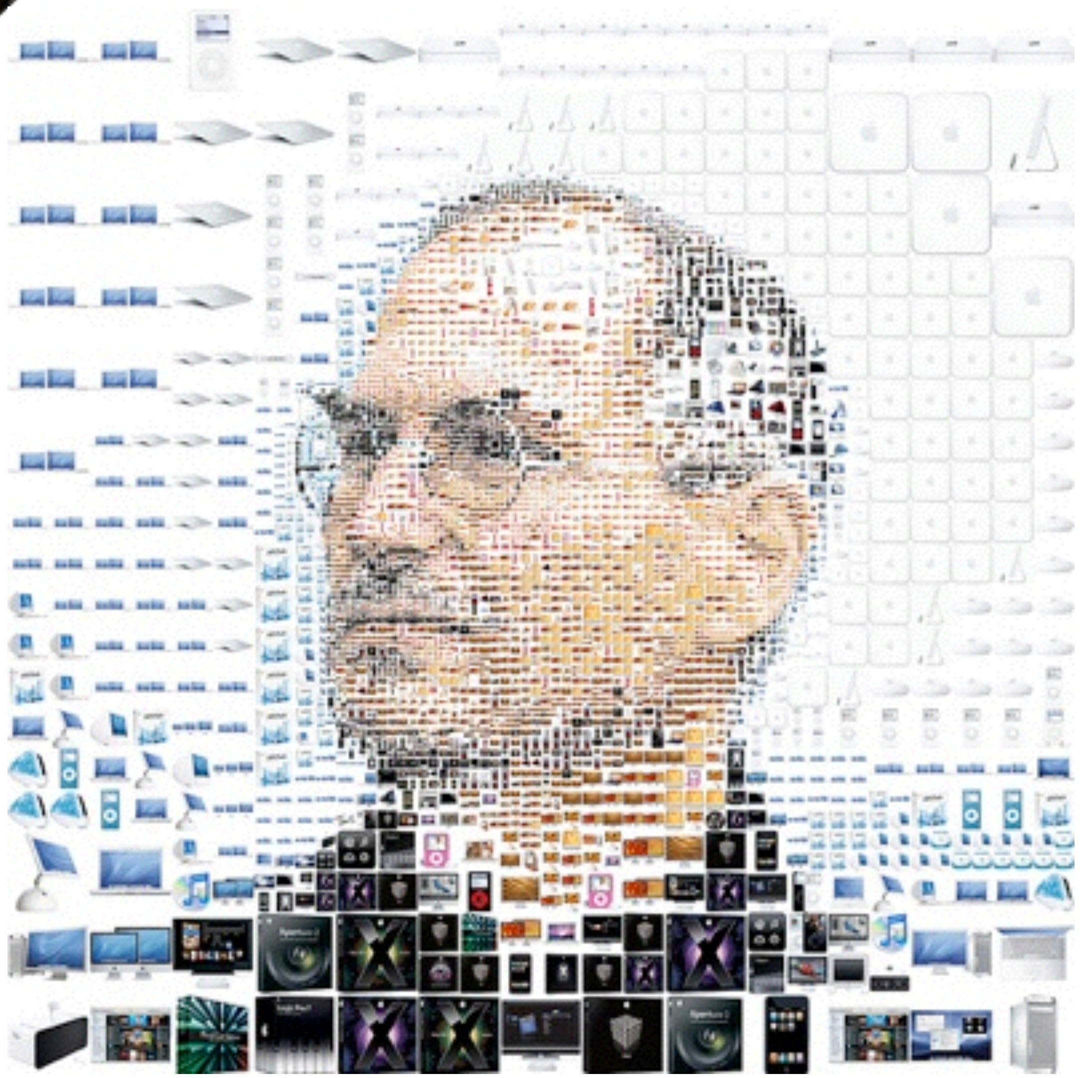


On the perception of spatial frequencies



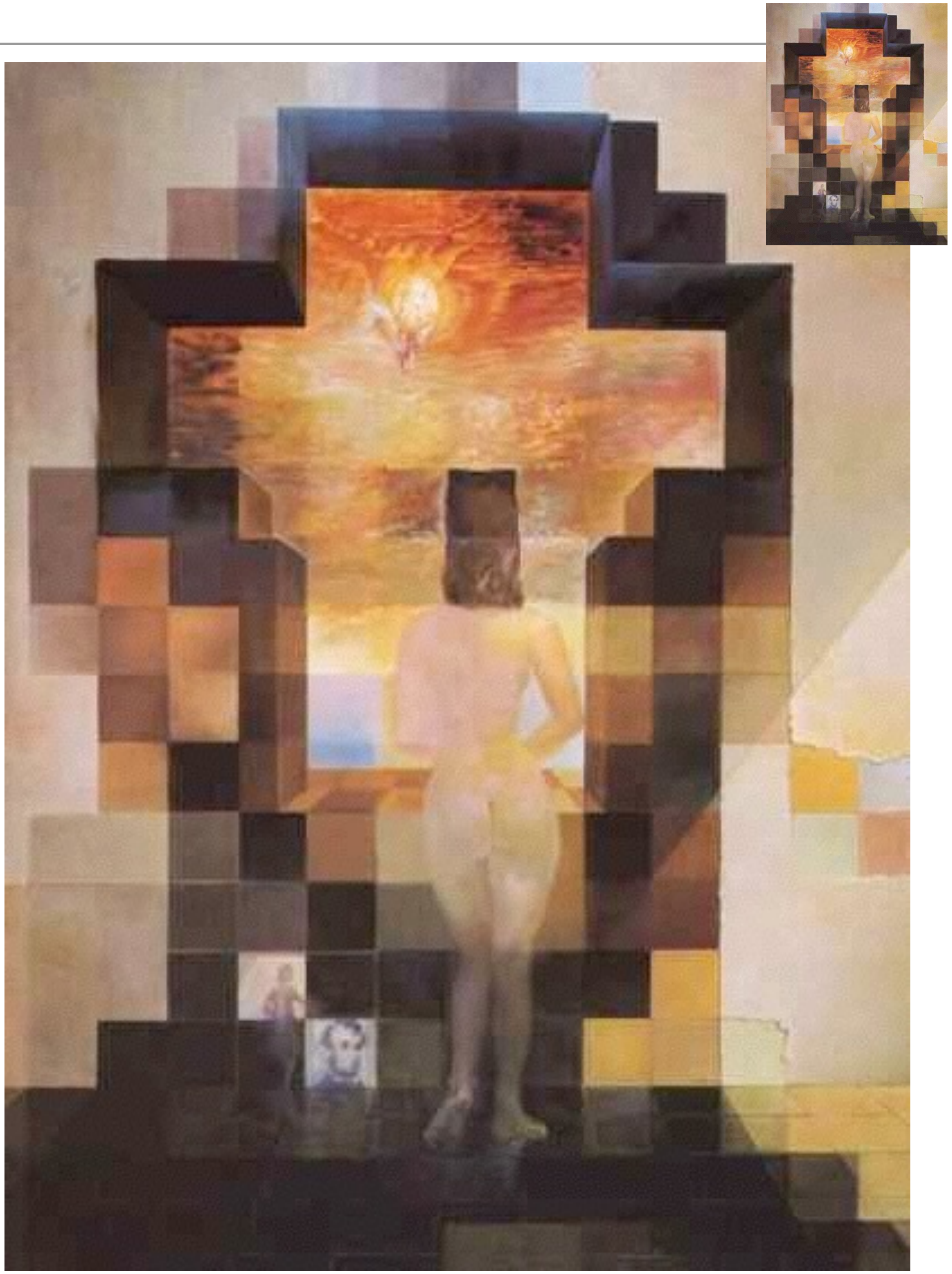
On the perception of spatial frequencies





Dali's *Slave Market with the Disappearing Bust of Voltaire*

On the perception of spatial frequencies



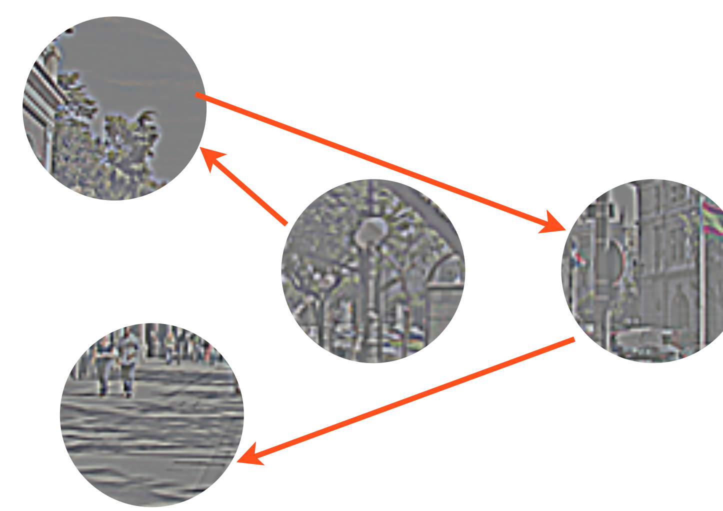
"All Is Vanity", by C. Allan Gilbert 1873-1929.

Frequency channels

Low spatial frequency

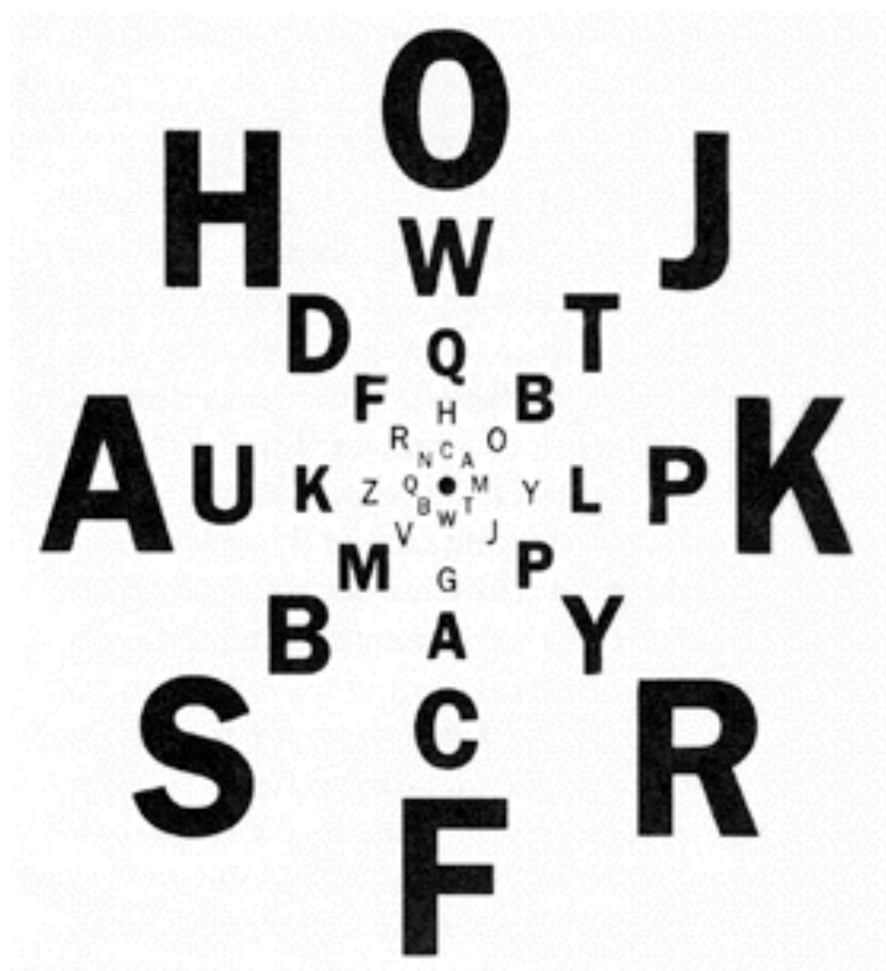
peripheral vision

High spatial frequency



foveal vision

Frequency channels



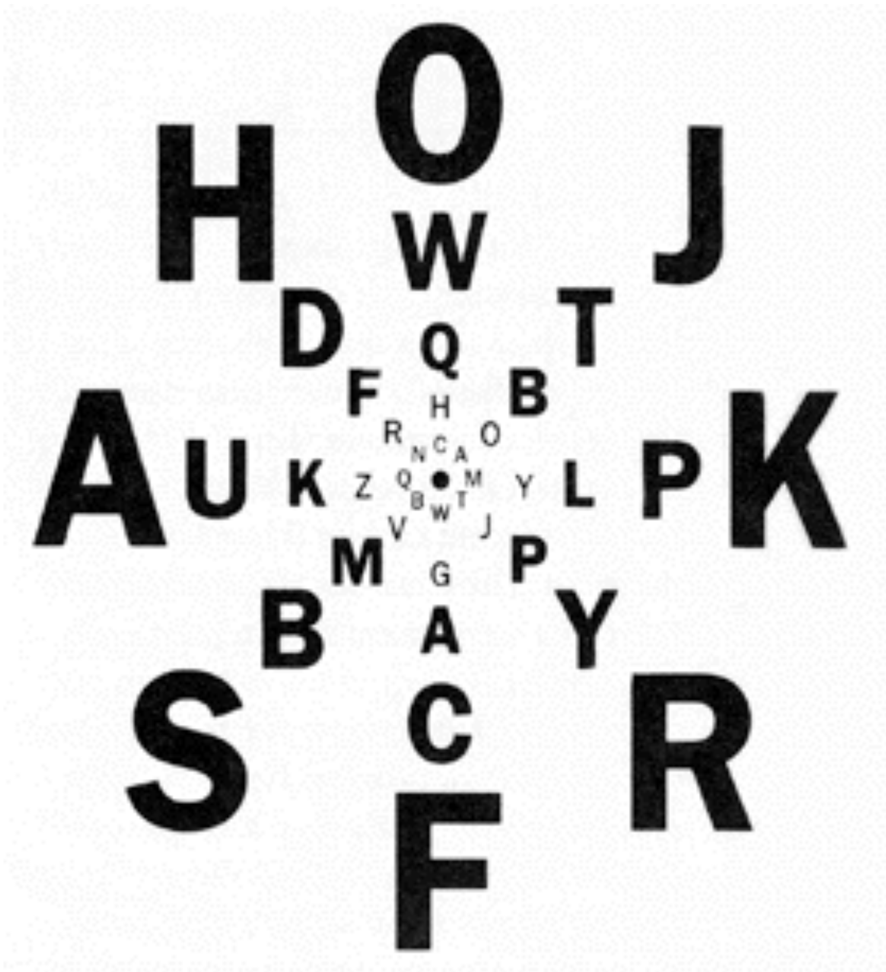
Spatial resolution and acuity

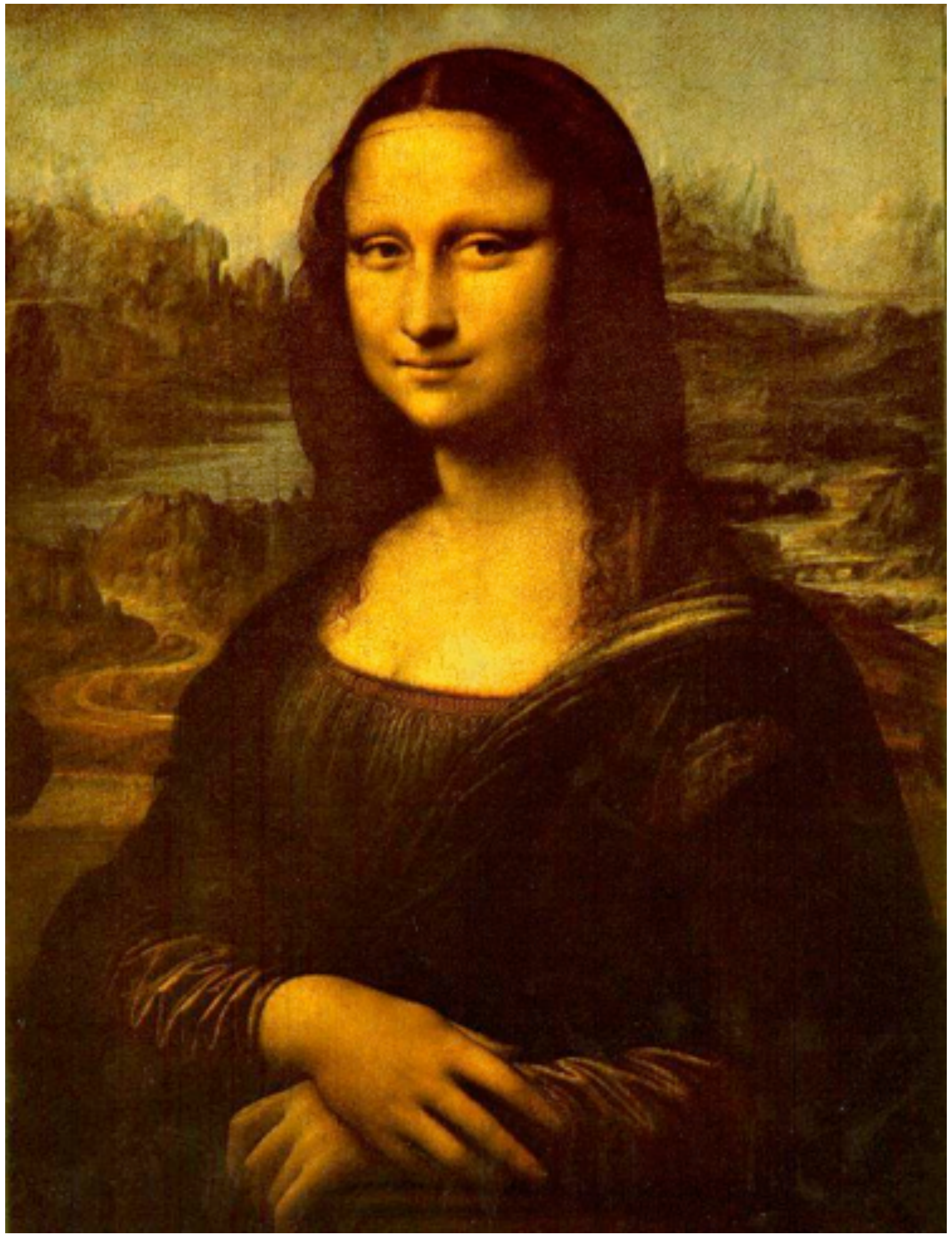
Renoir's Madame Henriot (1876)

Degas' Woman ironing (1869)

Ingres' Mrs Charles Badham (1816)

Spatial resolution and acuity





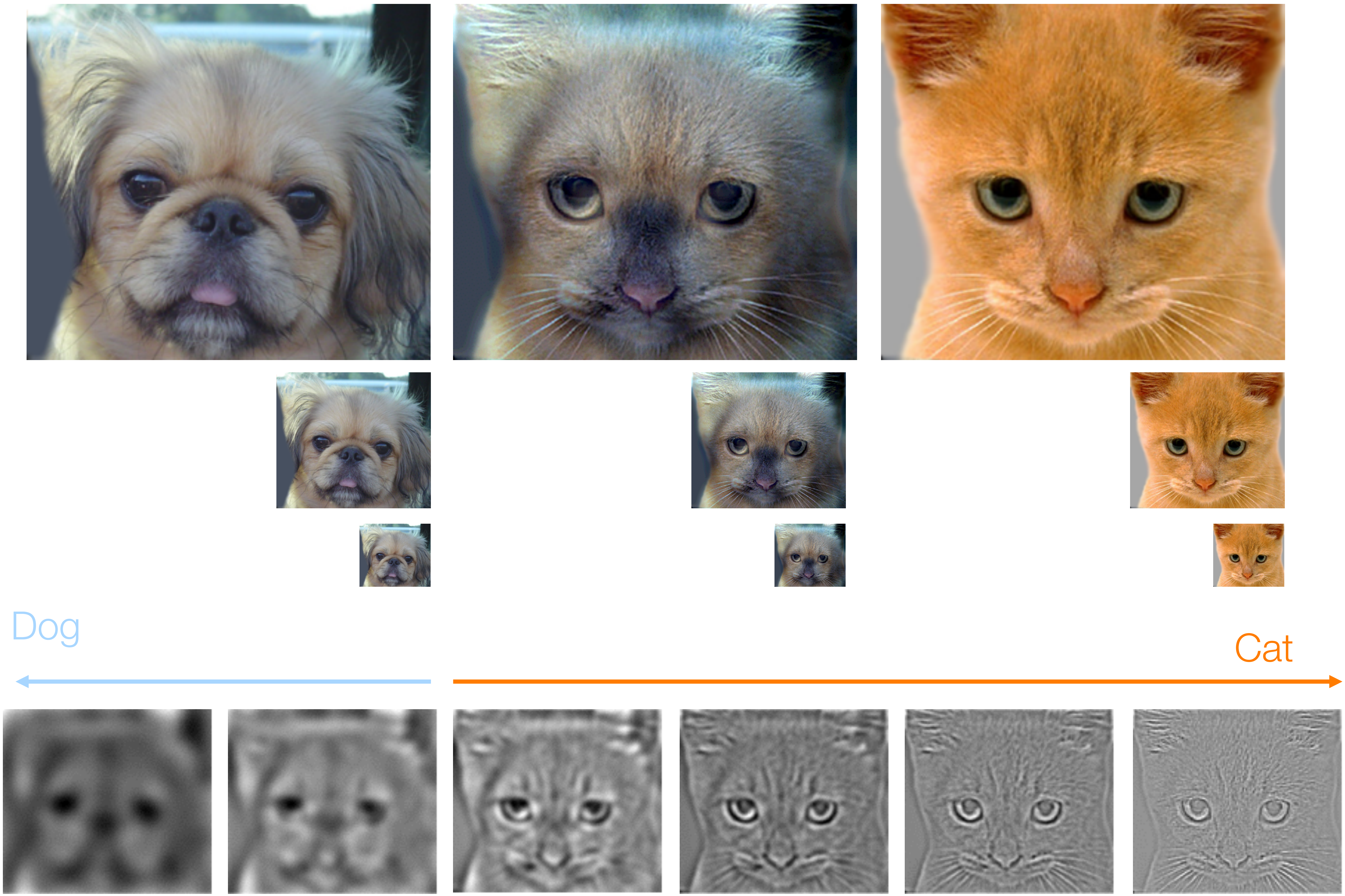
Subtle expression
Leonardo da Vinci's Mona Lisa

Smile | Smile | No smile

Low — Spatial frequency — High

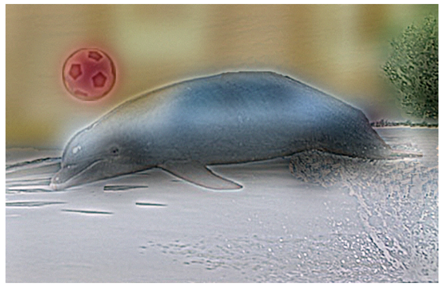
Hybrid Images

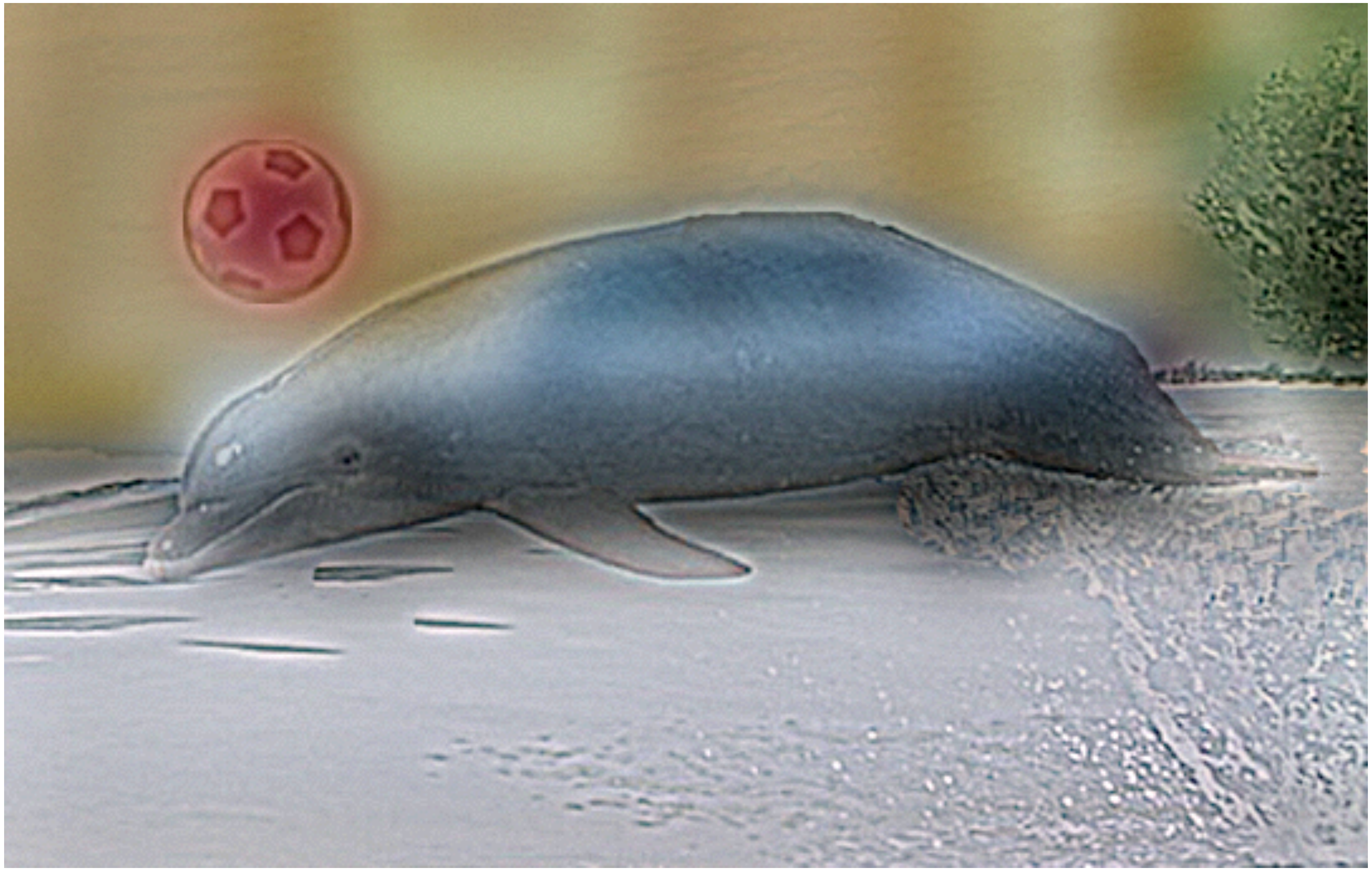
Aude Oliva (MIT)



matlab run_pyr

Hybrid Images

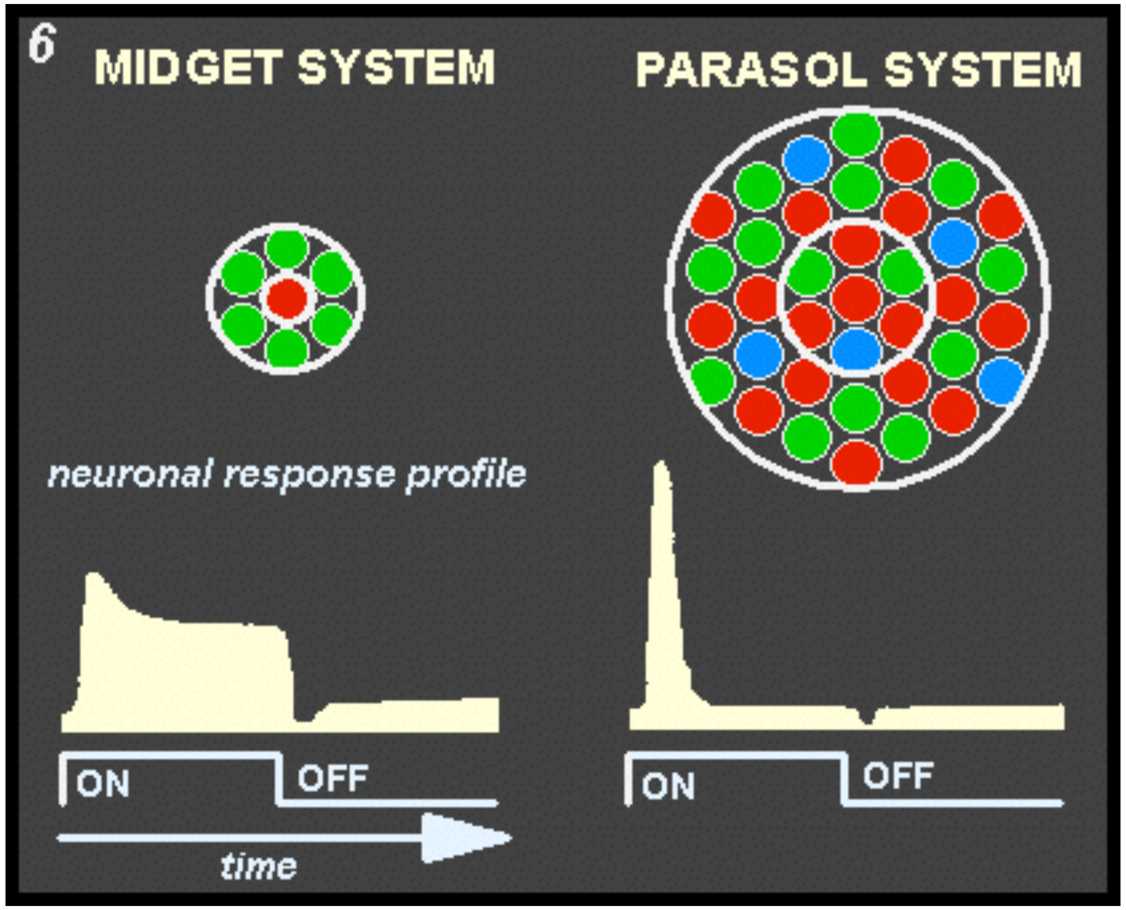




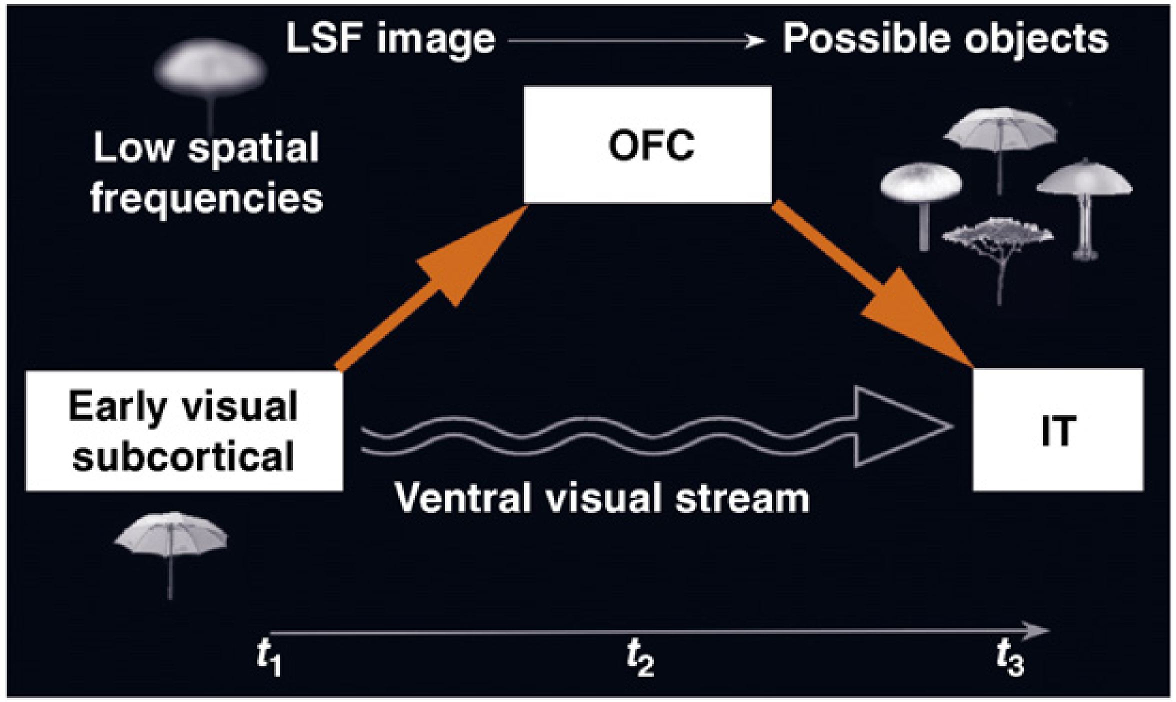
Aude Oliva (MIT)

Theories of object recognition

high SF (slow) low SF (fast)



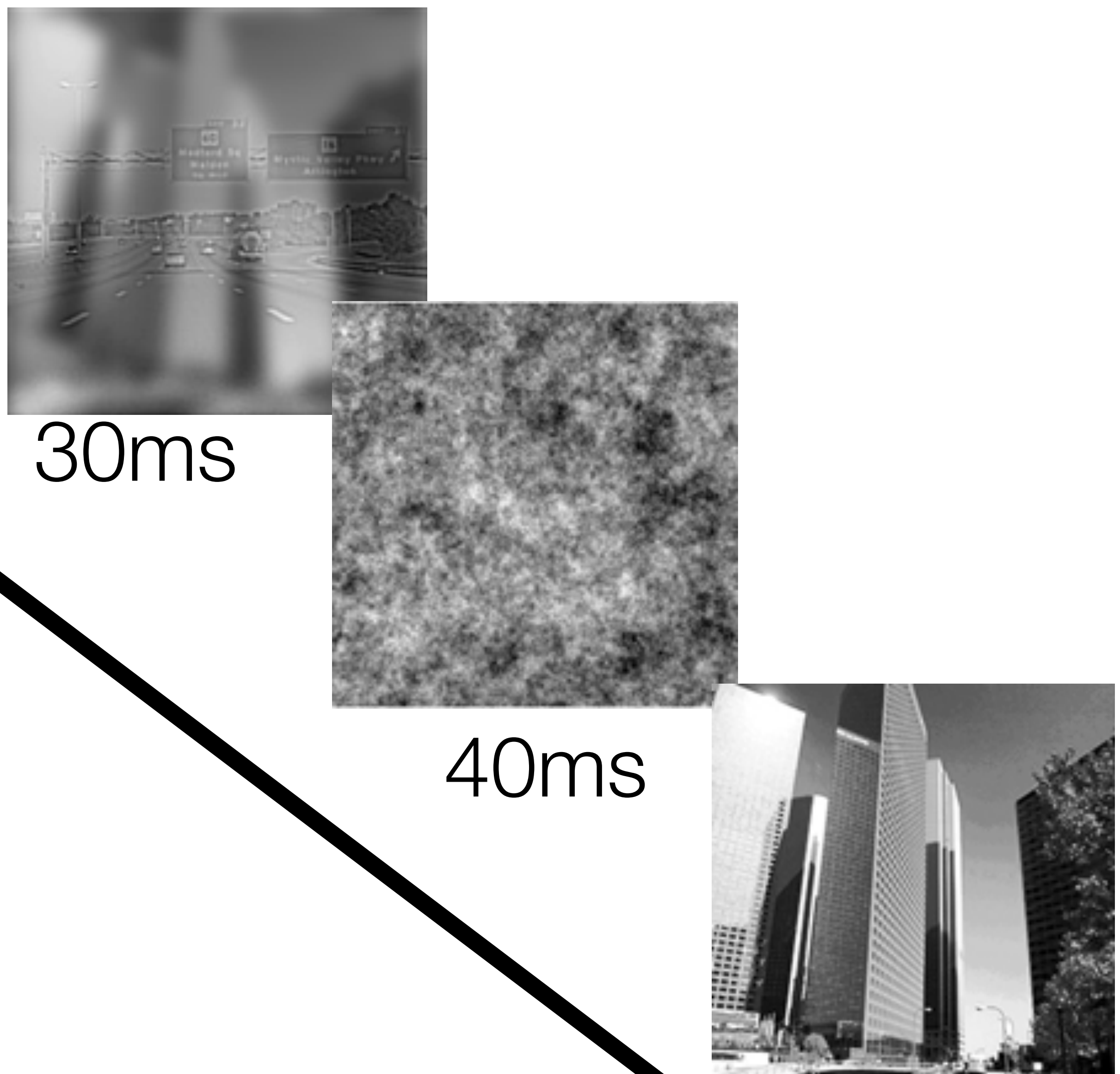
source: Peter Schiller



Bar ’07

Coarse-to-fine processing

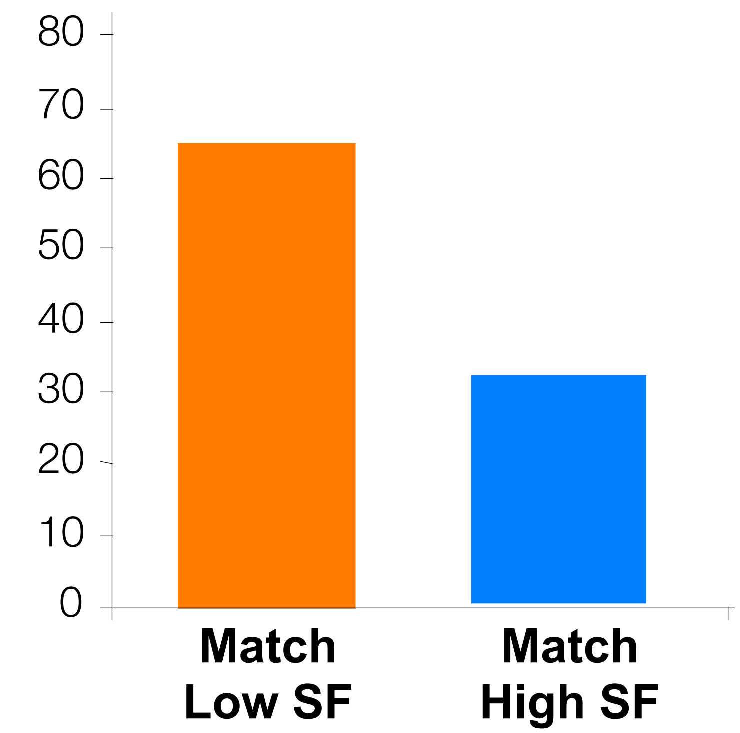
Duration: 30 ms



30ms

40ms

time



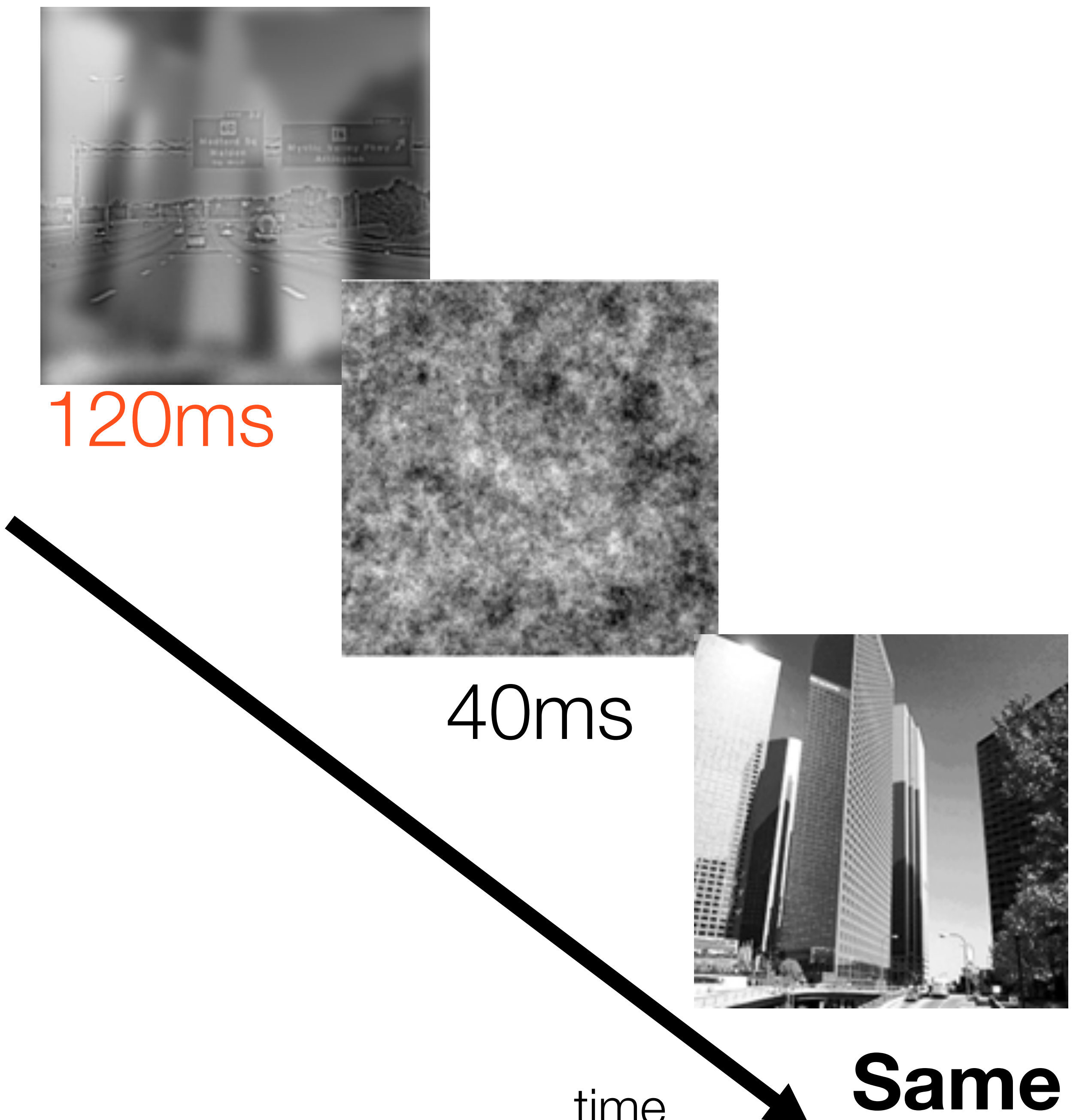
The second image can be:
- New image
- Match to Low SF (city)
- Match to High SF (highway)

Same or different ?

Schyns & Oliva '94

Coarse-to-fine processing

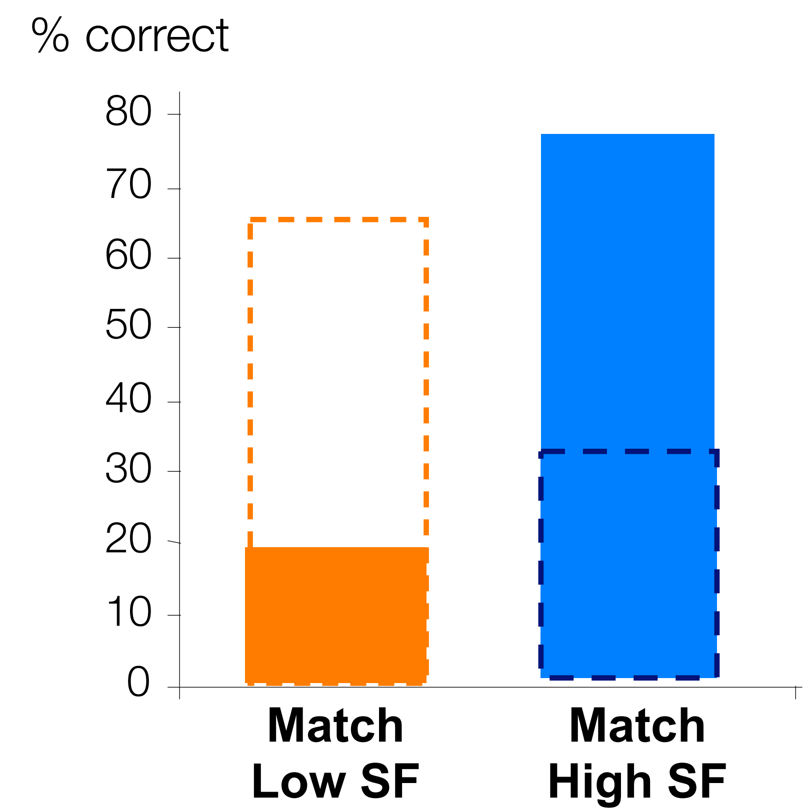
Duration: 120 ms



120ms

40ms

time



The second image can be:
New image
Match to Low SF (city)
Match to High SF (highway)

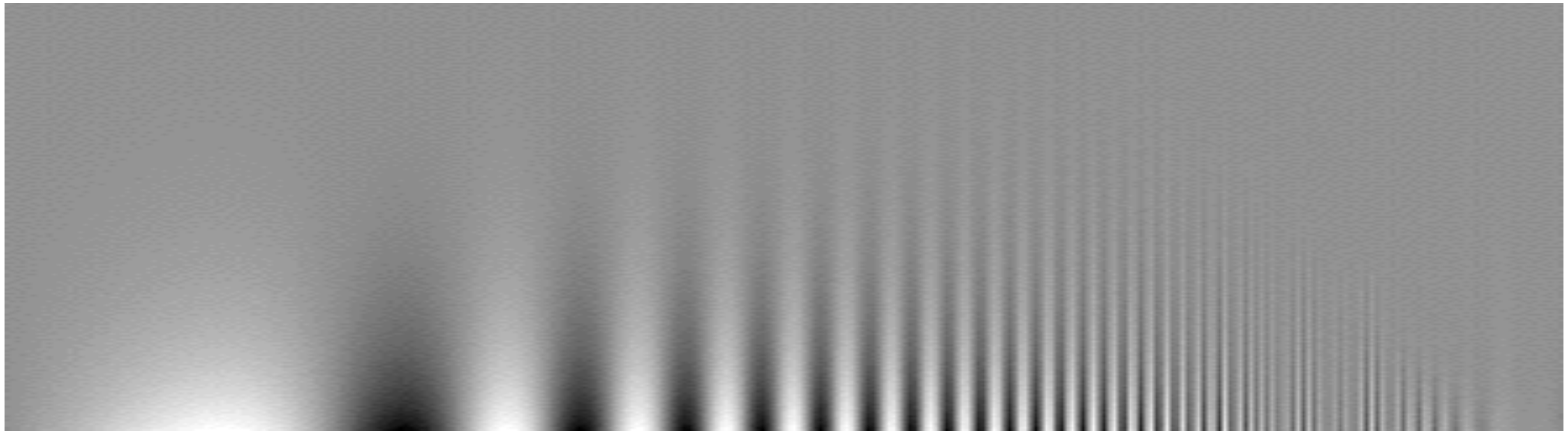
Same or different ?

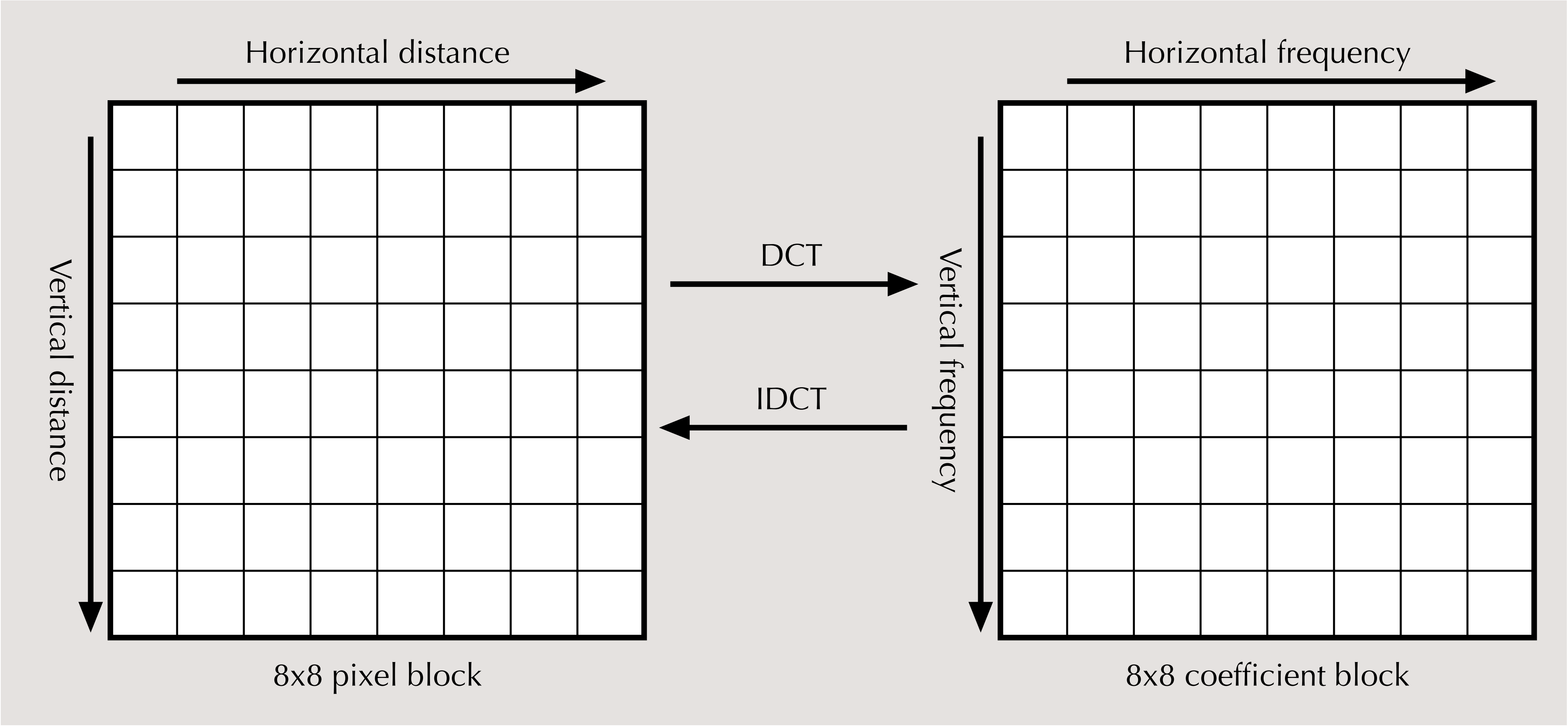
Schyns & Oliva ‘94

Data compression

Based on principles of human vision that we just discussed, how would you use image pyramids/spatial frequency channels for data compression applications?

Contrast sensitivity function



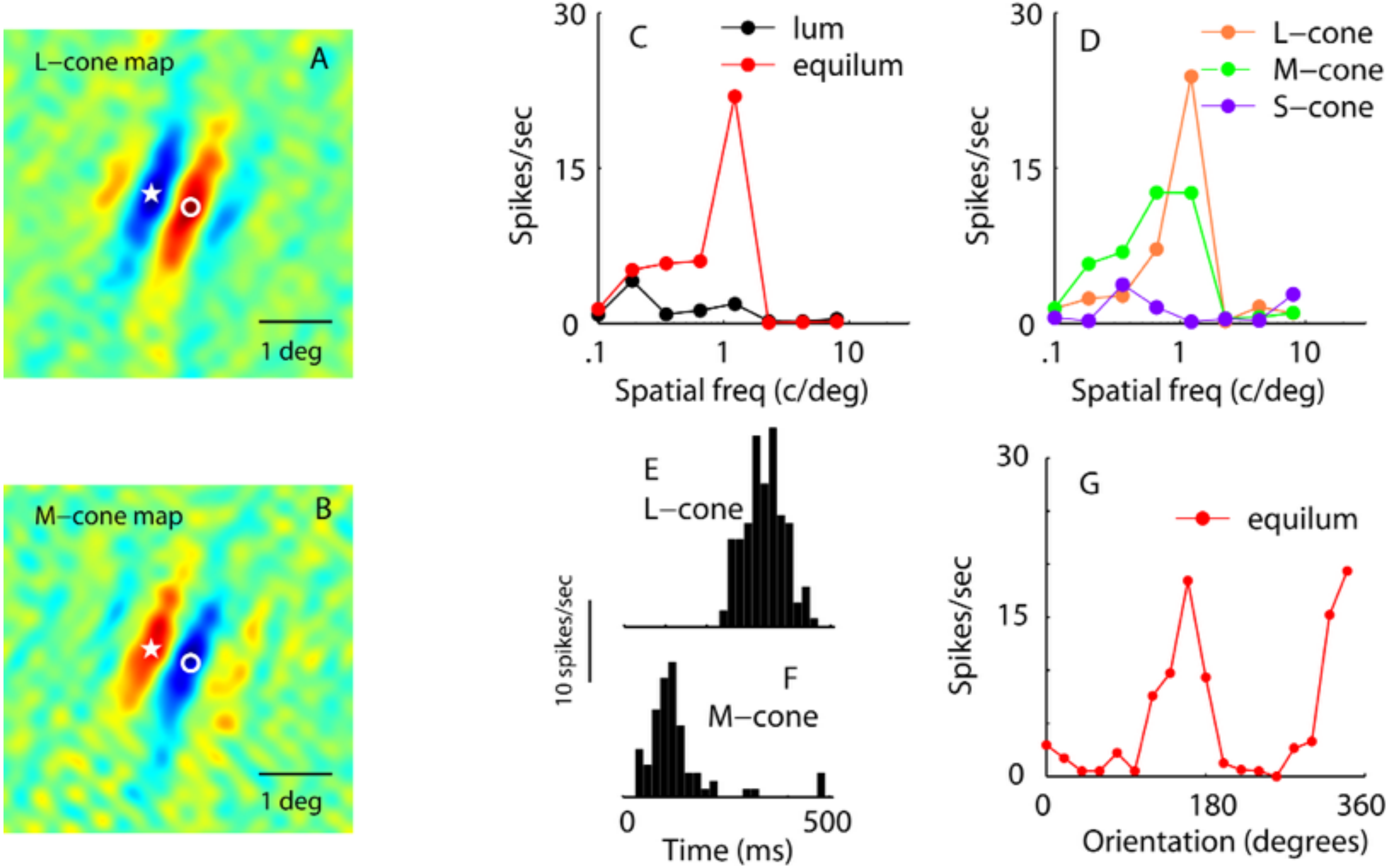


source: http://www.media-matters.net/docs/resources/Digital%20Files/MPEG/MPEG%20Encoding%20Basics.pdf

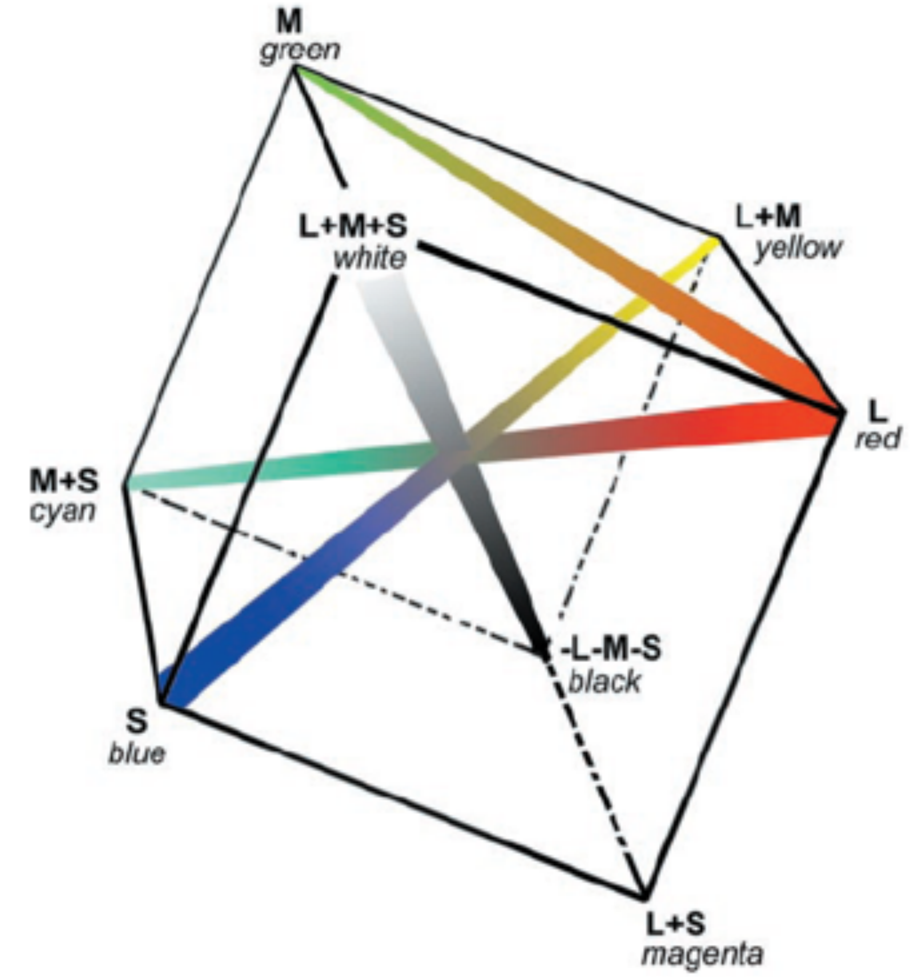
Computational Vision

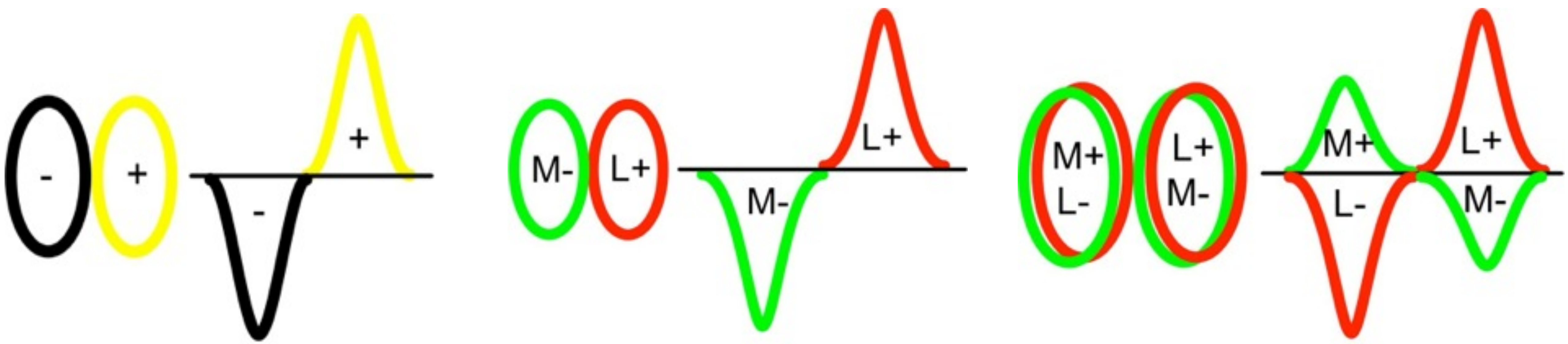
Primary visual cortex

- Orientation selectivity
- Spatial frequency
- Color opponency
- Normalization


L–cone map
A
1 deg
M–cone map
B
1 deg
C
lum
equilum
Spikes/sec
Spatial freq (c/deg)
D
L–cone
M–cone
S–cone
Spikes/sec
Spatial freq (c/deg)
E
L–cone
F
M–cone
10 spikes/sec
Time (ms)
G
equilum
Spikes/sec
Orientation (degrees)

Luminance vs. color selectivity in simple cells





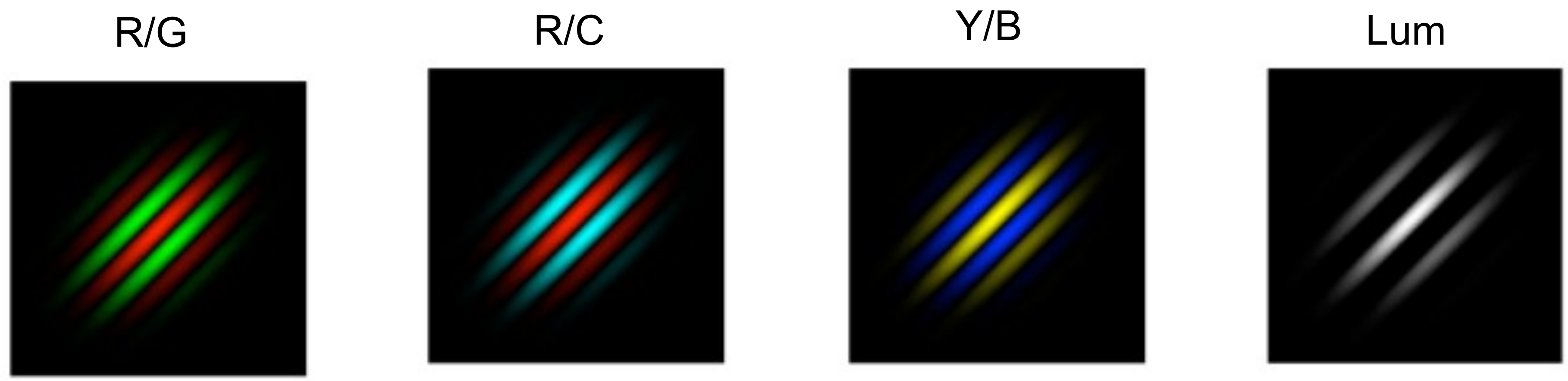
A
Oriented spatial opponent
luminance

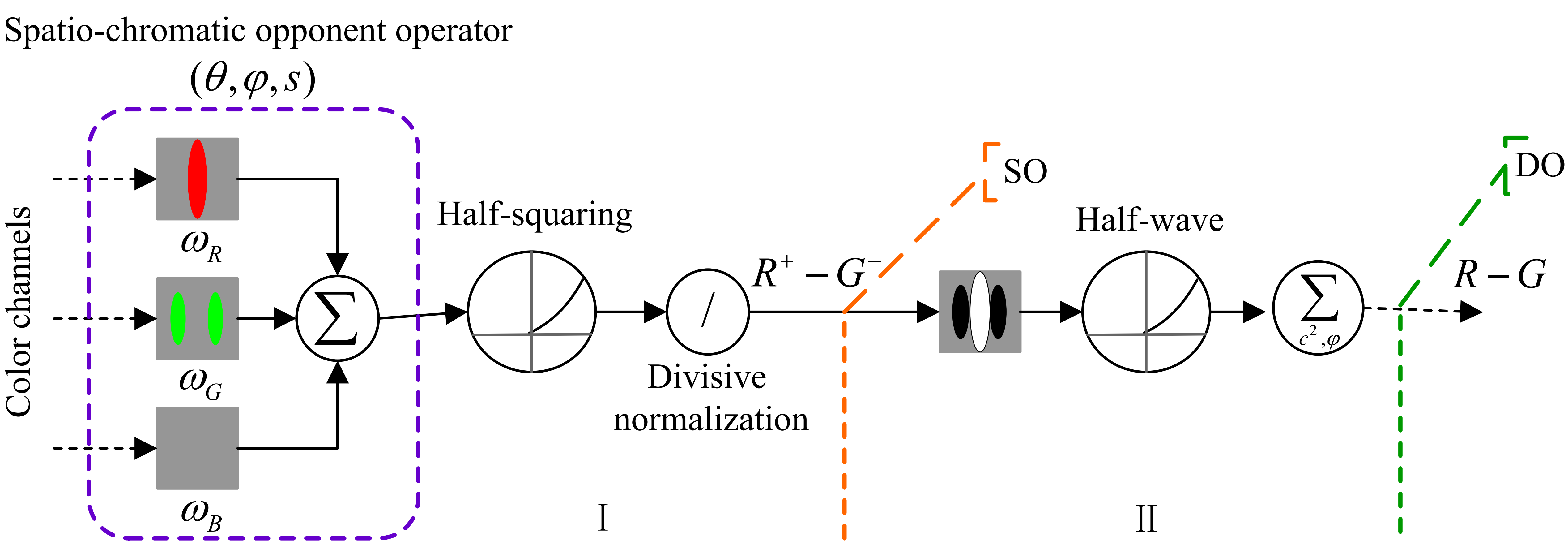
B
Weekly oriented chromatic opponent
Single-Opponency

C
Oriented spatio-chromatic opponent
Double-Opponency

Color processing

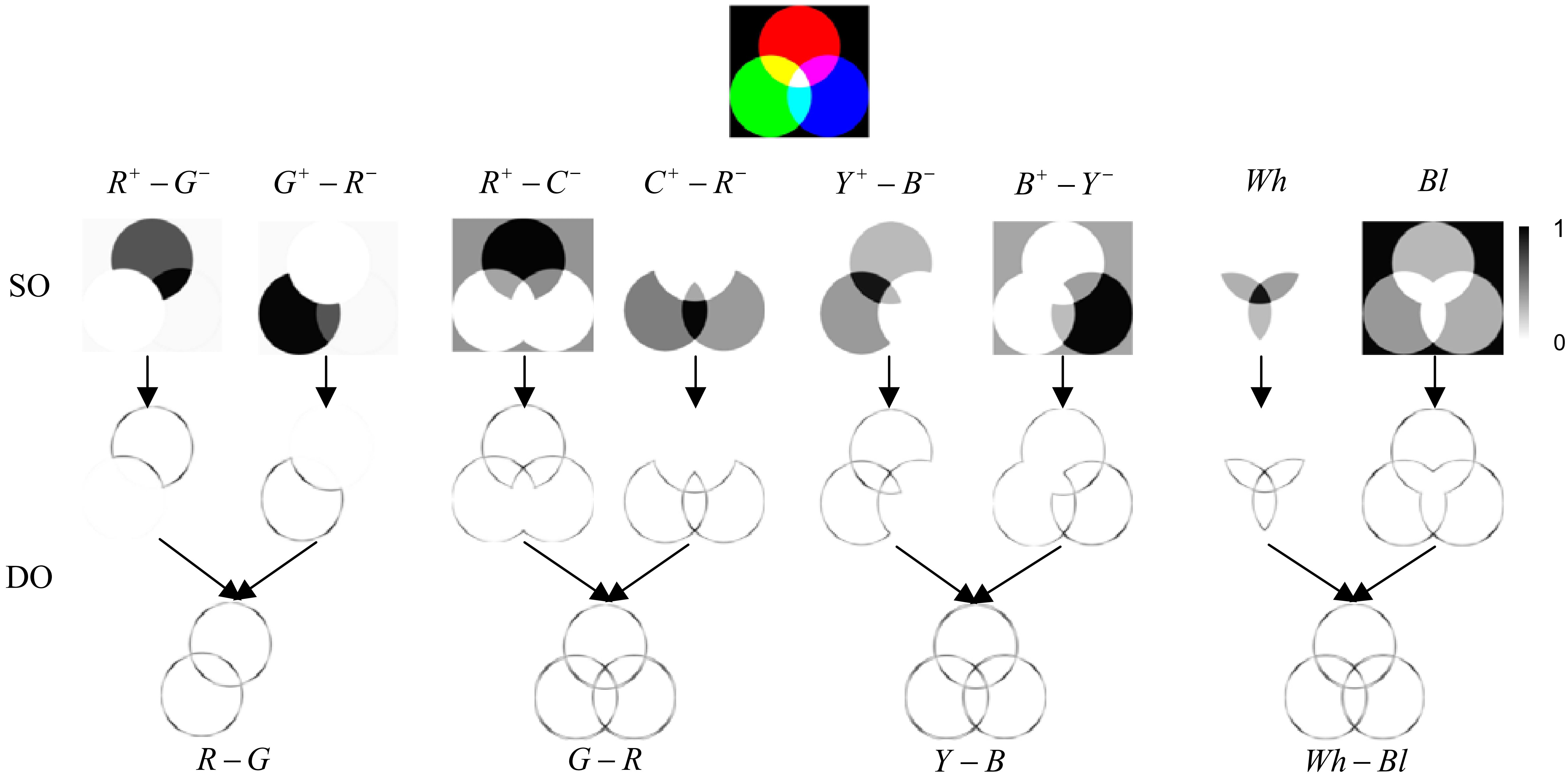
Parameters fitted to psychophysics data on color perception





Zhang Barhomi & Serre ’12

Color processing



Zhang Barhomi & Serre '12

Color processing

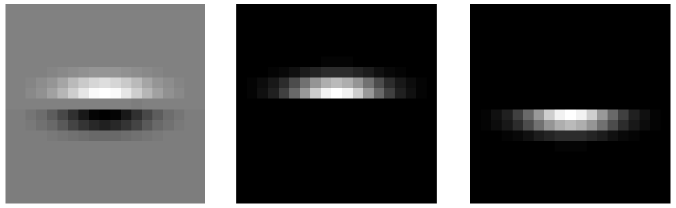
- SO/DO approach improves on all recognition and segmentation datasets tested as compared to existing color representations

- Color datasets

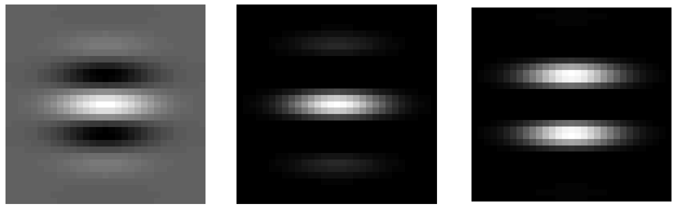
	Soccer team			Flower		
Method	Color	Shape	Both	Color	Shape	Both
Hue/SIFT	69 (67)	43 (43)	73 (73)	58 (40)	65 (65)	77 (79)
Opp/SIFT	69 (65)	43 (43)	74 (72)	57 (39)	65 (65)	74 (79)
SOSIFT/DOSIFT	82	66	83	68	69	79
SOHMAX/DOHMAX	87	76	89	77	73	83

- Pascal challenge

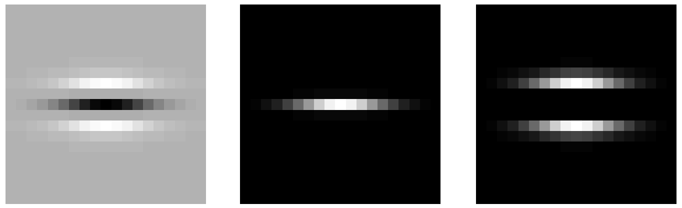
Method	SIFT	HueSIFT	OpponentSIFT	CSIFT	SODOSIFT	SODOHMAX
AP	40 (38.4)	41	43 (42.5)	43 (44.0)	46.5 **(33.3/39.8)**	46.8 **(30.1/36.4)**



A. Gradient used in SIFT



B. Gabor filters used in HMAX



C. Gaussian derivatives used in segmentation

Zhang Barhomi & Serre '12

Computational Vision

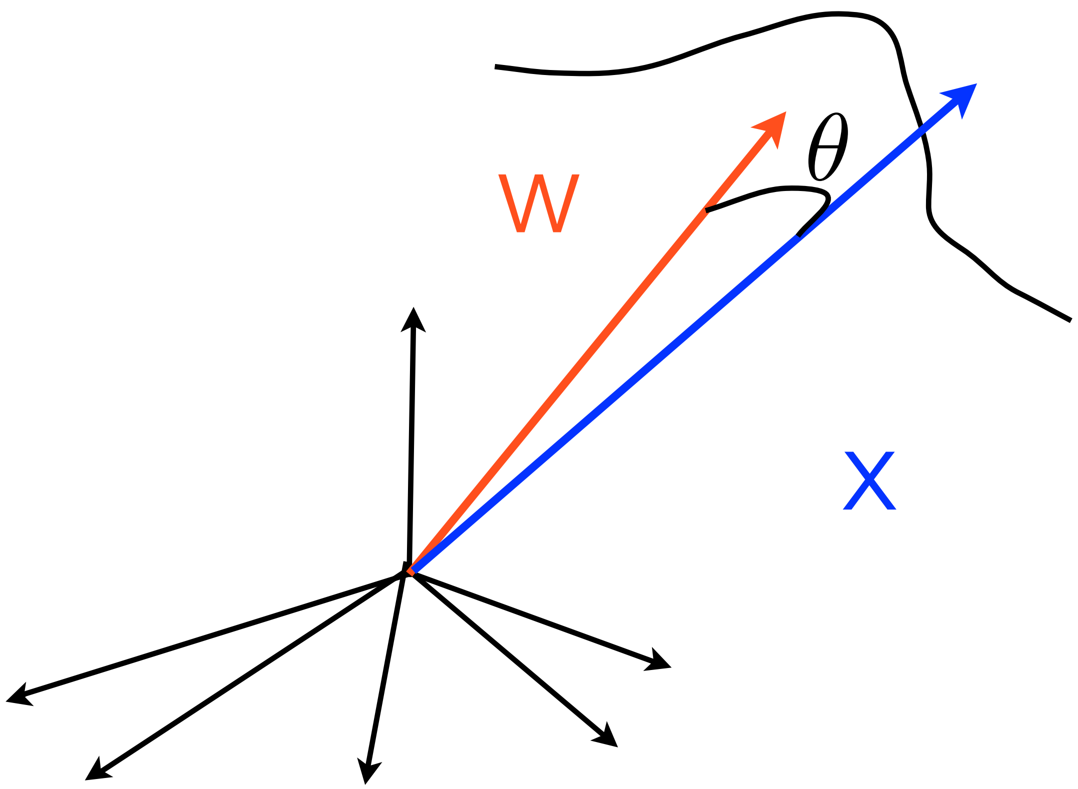
Primary visual cortex

- Orientation selectivity
- Spatial frequency
- Color opponency
- Normalization

On the need for normalization circuits

$$
\begin{aligned}
y(x) &= \sum w_i x_i \\
&= \mathbf{w} \cdot \mathbf{x} \\
&= ||\mathbf{w}||||\mathbf{x}|| \cos(\theta)
\end{aligned}
$$

$$
\begin{aligned}
y(x) &= \sum w_i x_i \\
&= \frac{\mathbf{w} \cdot \mathbf{x}}{||\mathbf{w}||||\mathbf{x}||} \\
&= \cos(\theta)
\end{aligned}
$$

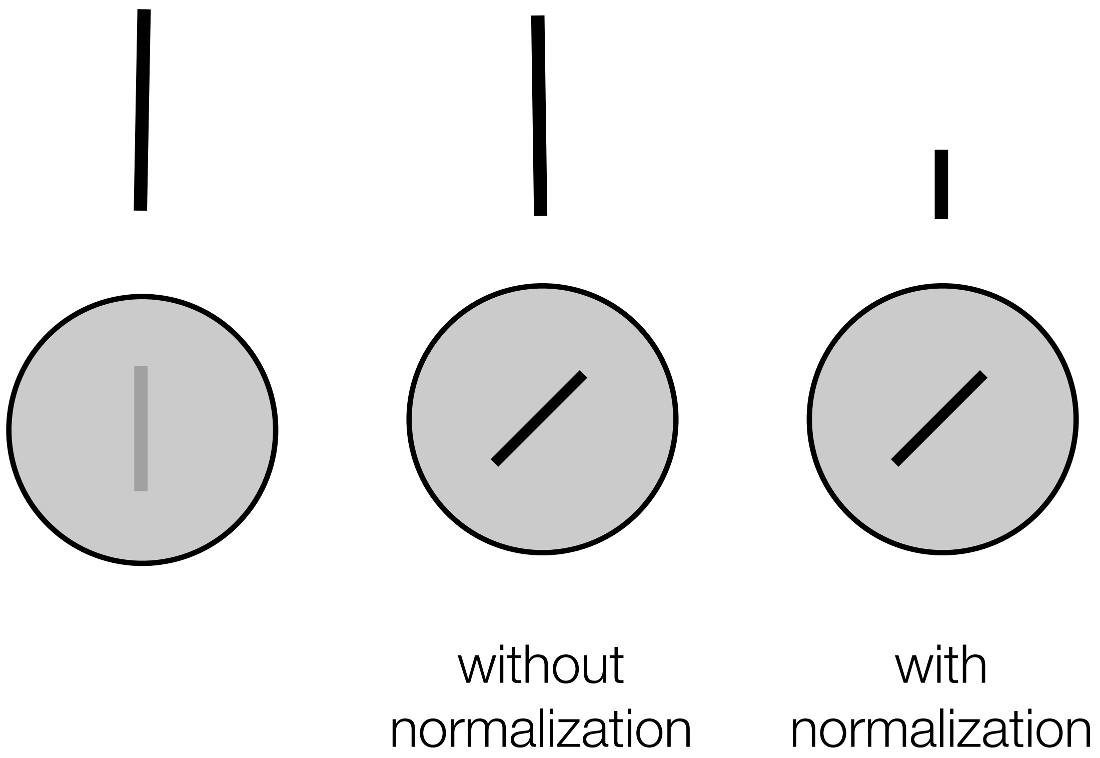


On the need for normalization circuits

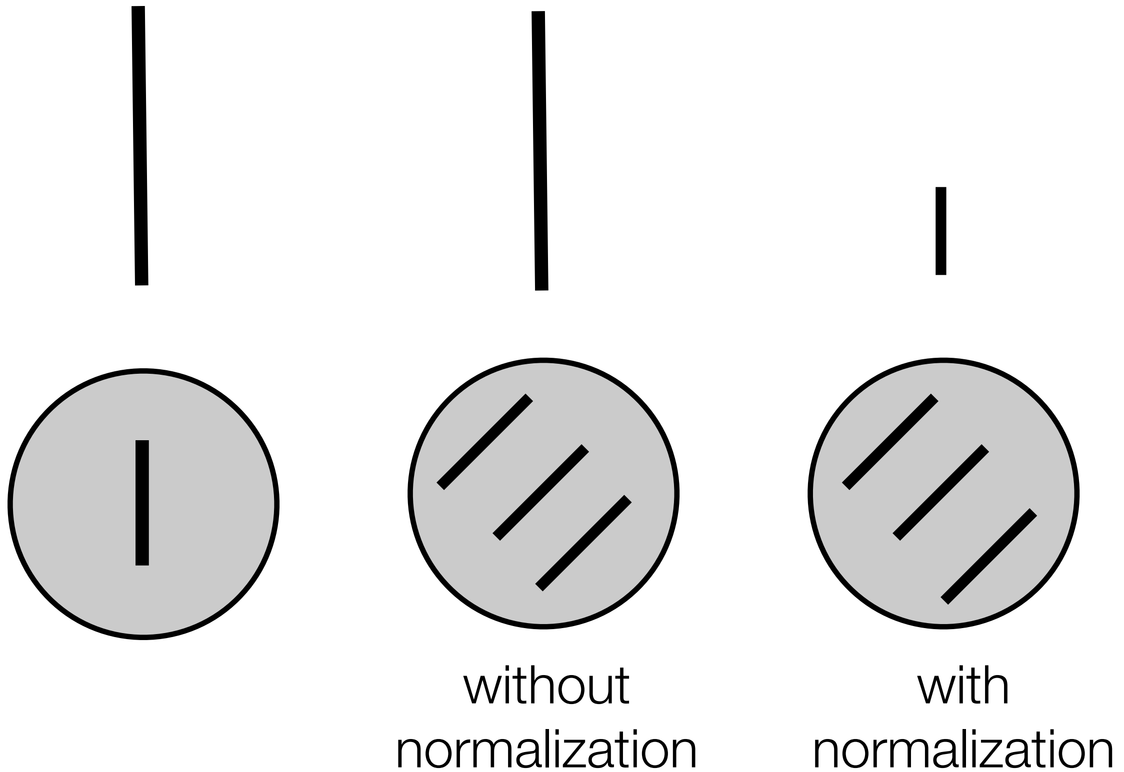
$$
\begin{aligned}
y(x) &= \sum w_i x_i \\
&= \mathbf{w} \cdot \mathbf{x} \\
&= ||\mathbf{w}||\,||\mathbf{x}|| \cos(\theta)
\end{aligned}
$$

$$
\begin{aligned}
y(x) &= \sum w_i x_i \\
&= \frac{\mathbf{w} \cdot \mathbf{x}}{||\mathbf{w}||\,||\mathbf{x}||} \\
&= \cos(\theta)
\end{aligned}
$$

contrast problem



superposition problem

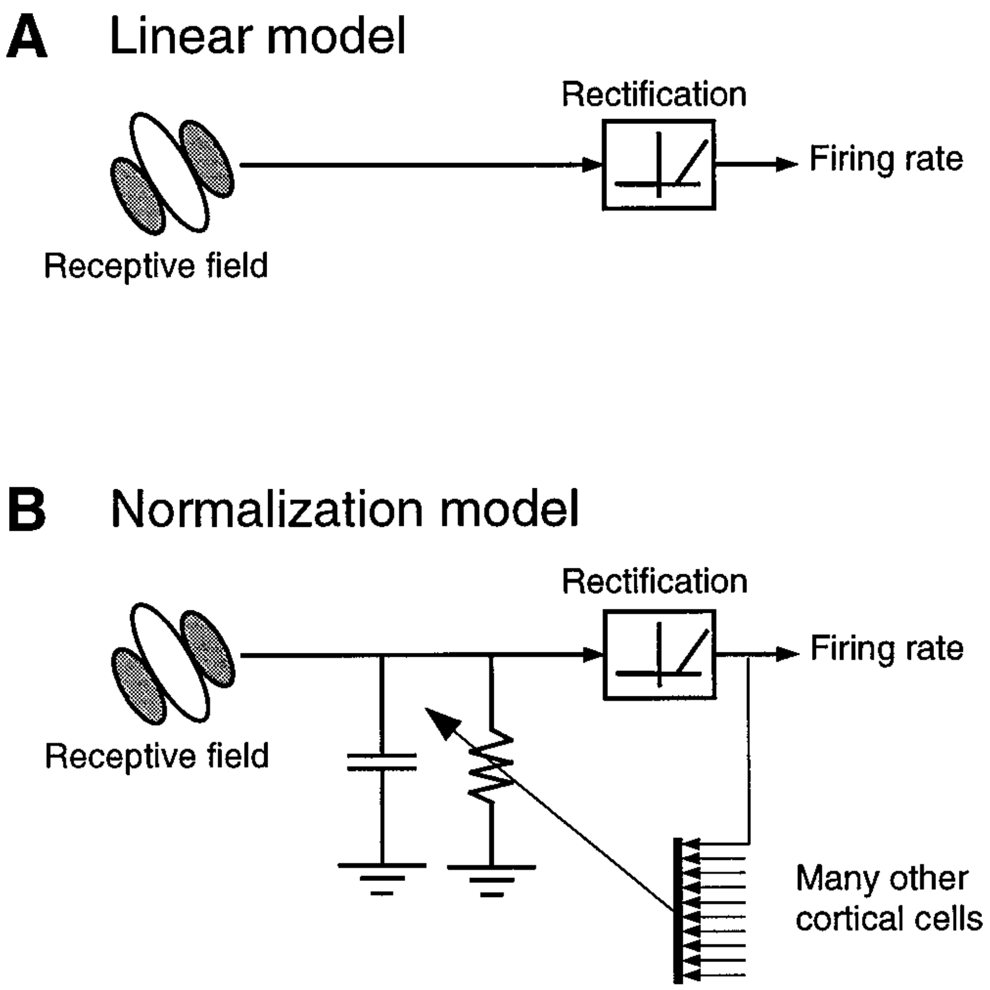


Gain-control circuits

$$y(\mathbf{x}) = \frac{\mathbf{w}^* \cdot \mathbf{x}}{||\mathbf{x}|| + c}$$

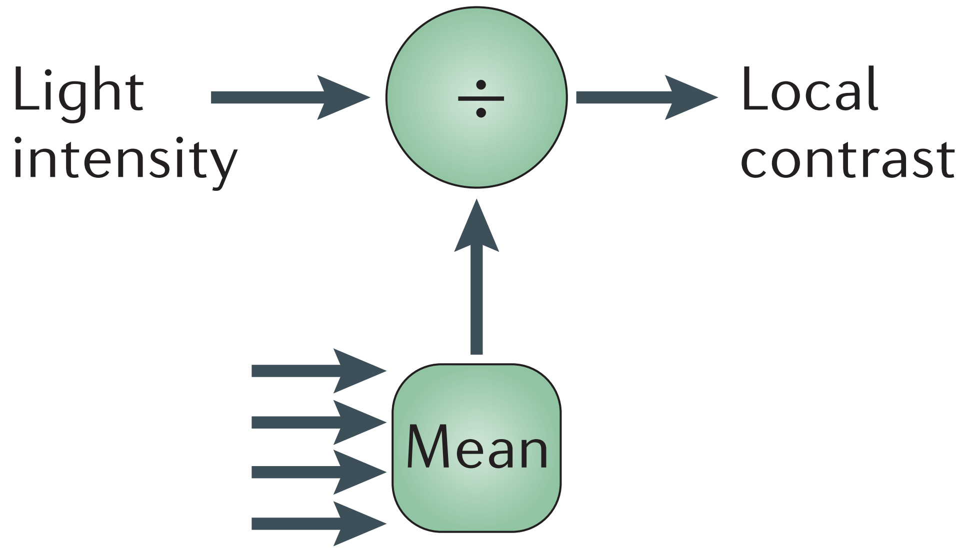
$$y(\alpha\mathbf{x}) = \frac{\mathbf{w}^* \cdot \alpha\mathbf{x}}{||\alpha\mathbf{x}|| + c}$$

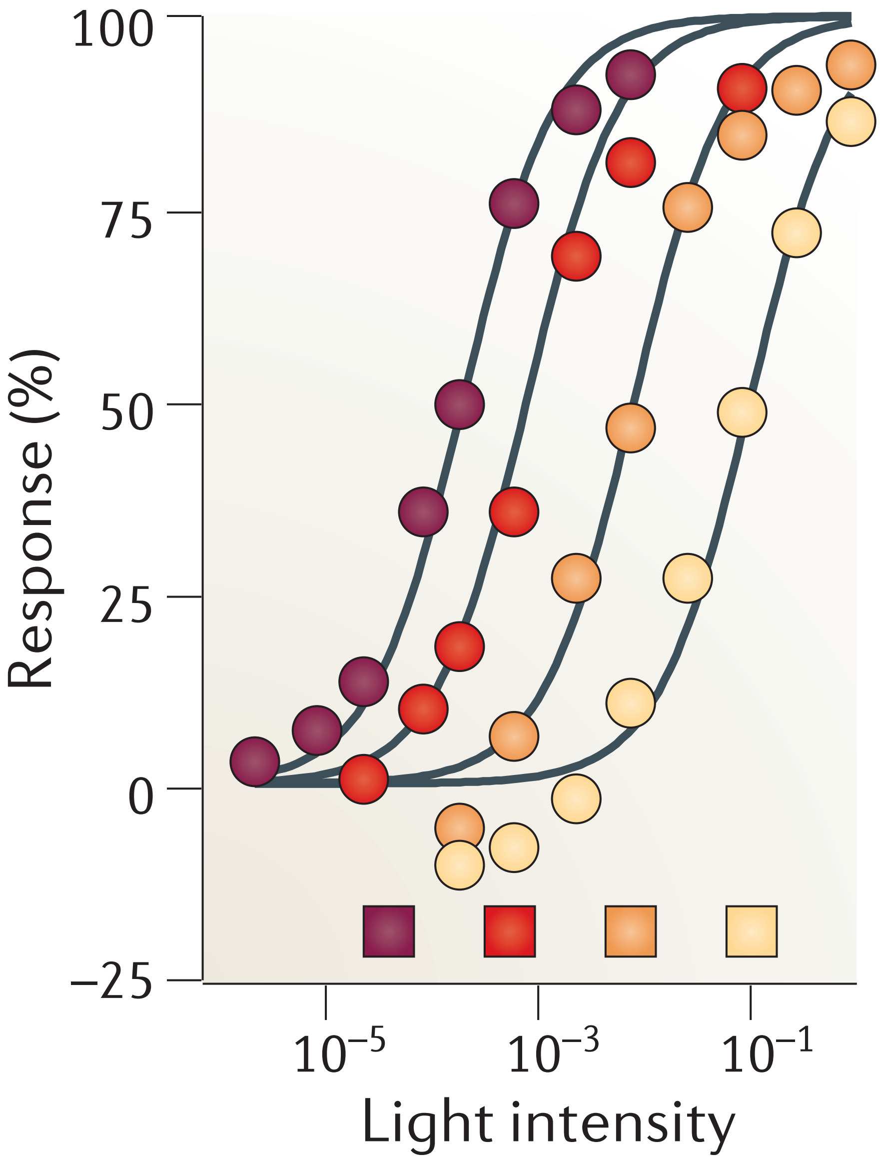
$$= \frac{\alpha\mathbf{w}^* \cdot \mathbf{x}}{\alpha(||\mathbf{x}|| + c/\alpha)}$$



Carandini Heeger & Movshon '96

Gain-control circuits in the (turtle) retina





Carandini & Heeger 2011

Gain-control circuits in V1

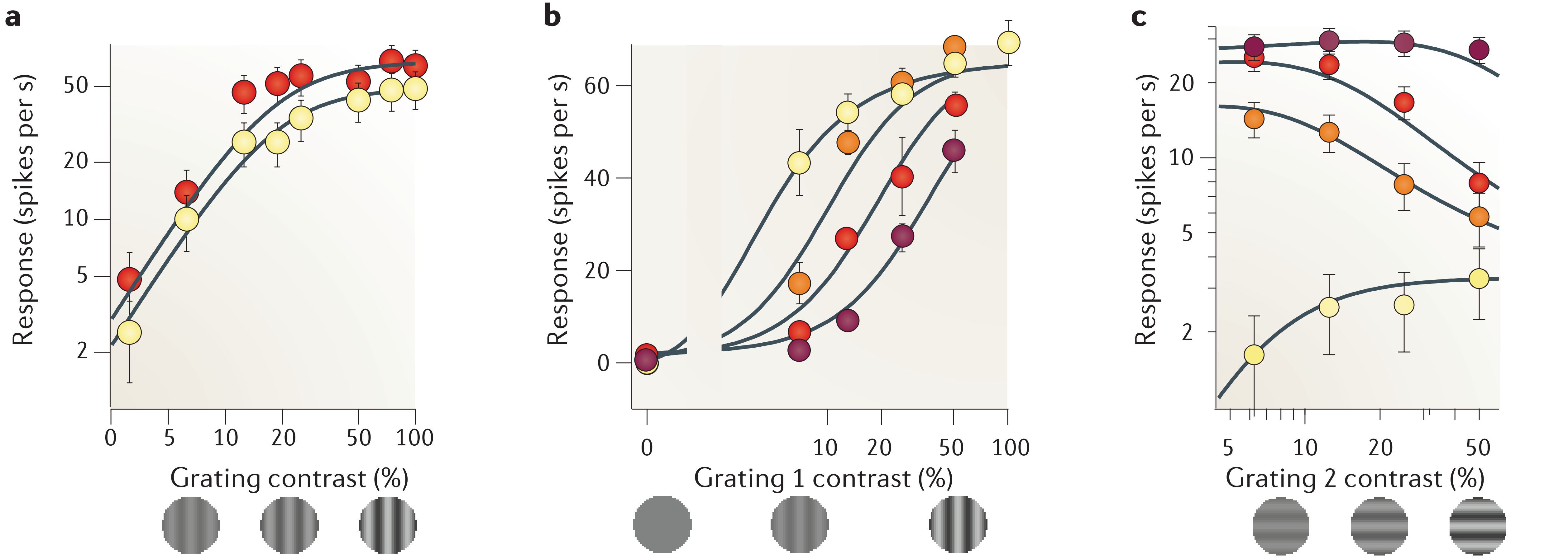


Figure 3 | **Normalization in the primary visual cortex. a** | Contrast saturation. Responses as a function of grating contrast for gratings having optimal orientation (shown in red) and suboptimal orientation (shown in yellow). **b** | Cross-orientation suppression. Responses to the sum of a test grating and an orthogonal mask grating (colours indicate mask contrast, from 0% (shown in yellow) to 50% (shown in dark red)). **c** | Transition from drive to suppression. Grating 1 had optimal orientation and grating 2 had suboptimal orientation. Grating 2 could provide some drive to the neuron when presented alone (shown in yellow) but became suppressive when grating 1 had moderate contrasts (shown in red). **d** | Surround suppression. A grating contained in a central disk was surrounded by a grating in an annulus. The annulus elicited minimal responses when presented alone, but suppressed responses to the central disk. **e** | Effects of normalization on population responses. Each dot indicates the response of a population of neurons selective for a given orientation, and each panel indicates the population responses to a stimulus. Stimuli are gratings of increasing contrast, presented alone (top) or together with an orthogonal grating (bottom). Data in part **a** from REF. 43; data in part **b** from REF. 56; data in part **c** from REF. 43; data in part **d** from REF. 142; data in part **e** from REF. 48.

Carandini & Heeger 2011

Gain-control circuits in V1

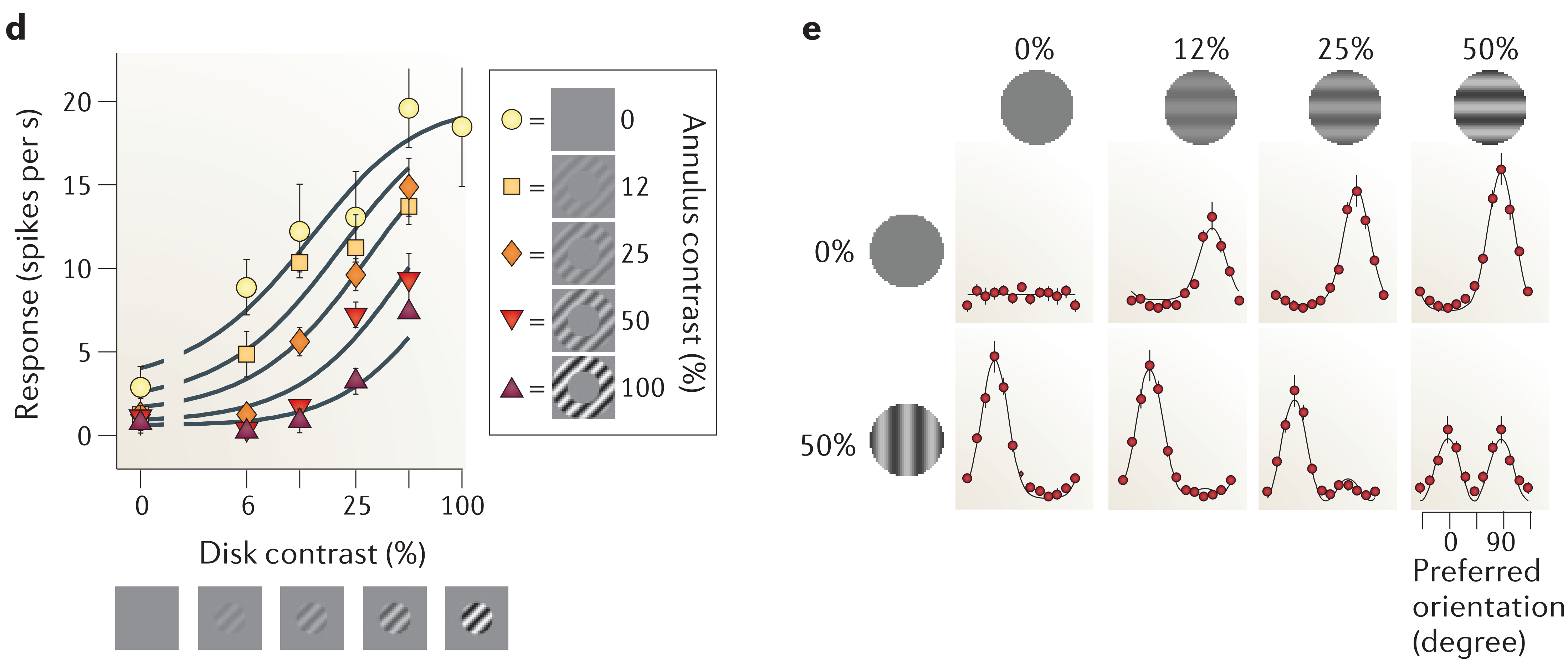


Figure 3 | **Normalization in the primary visual cortex. a** | Contrast saturation. Responses as a function of grating contrast for gratings having optimal orientation (shown in red) and suboptimal orientation (shown in yellow). **b** | Cross-orientation suppression. Responses to the sum of a test grating and an orthogonal mask grating (colours indicate mask contrast, from 0% (shown in yellow) to 50% (shown in dark red)). **c** | Transition from drive to suppression. Grating 1 had optimal orientation and grating 2 had suboptimal orientation. Grating 2 could provide some drive to the neuron when presented alone (shown in yellow) but became suppressive when grating 1 had moderate contrasts (shown in red). **d** | Surround suppression. A grating contained in a central disk was surrounded by a grating in an annulus. The annulus elicited minimal responses when presented alone, but suppressed responses to the central disk. **e** | Effects of normalization on population responses. Each dot indicates the response of a population of neurons selective for a given orientation, and each panel indicates the population responses to a stimulus. Stimuli are gratings of increasing contrast, presented alone (top) or together with an orthogonal grating (bottom). Data in part **a** from REF. 43; data in part **b** from REF. 56; data in part **c** from REF. 43; data in part **d** from REF. 142; data in part **e** from REF. 48.

Carandini & Heeger 2011

Gain control circuits in computer vision!

What is the Best Multi-Stage Architecture for Object Recognition?

Kevin Jarrett, Koray Kavukcuoglu, Marc'Aurelio Ranzato and Yann LeCun
The Courant Institute of Mathematical Sciences
New York University, 715 Broadway, New York, NY 10003, USA
koray@cs.nyu.edu

- Normalization appears to be the most important component of a good computer vision system (over learning algorithms, architecture, etc)