UNIVERSITY OF BIRMINGHAM

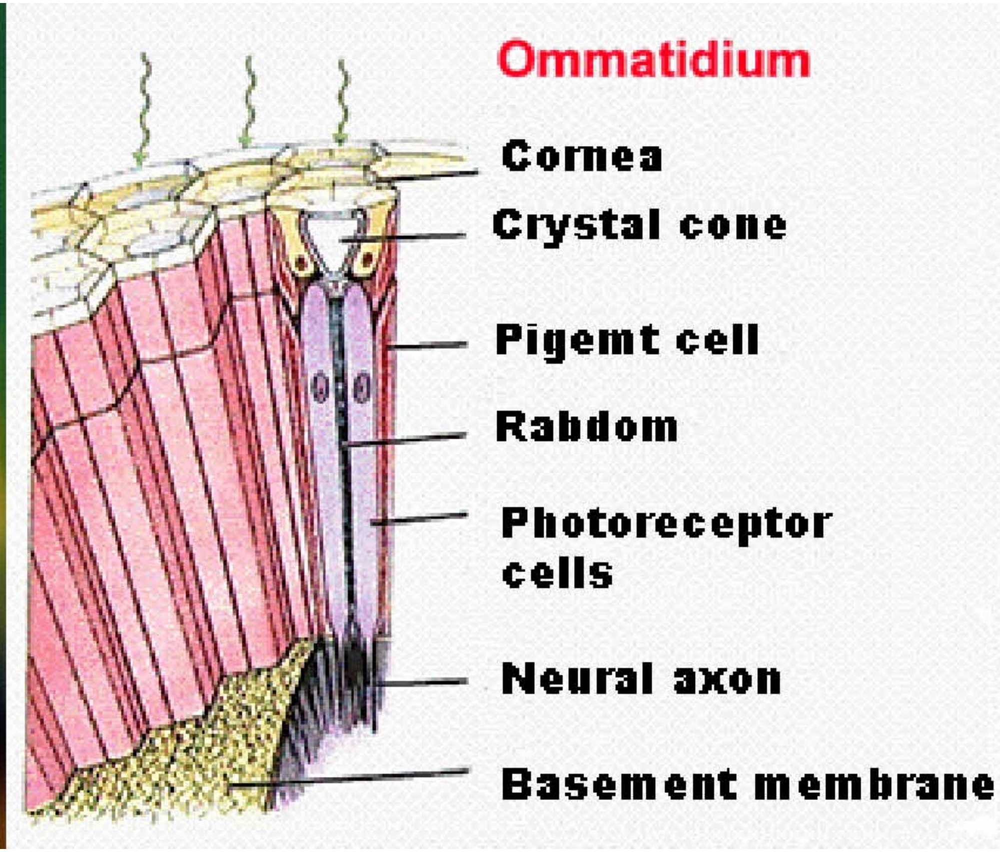
CNCR

Vision

Mind, **B**rain, and **M**odels 2023 May 3, 2023

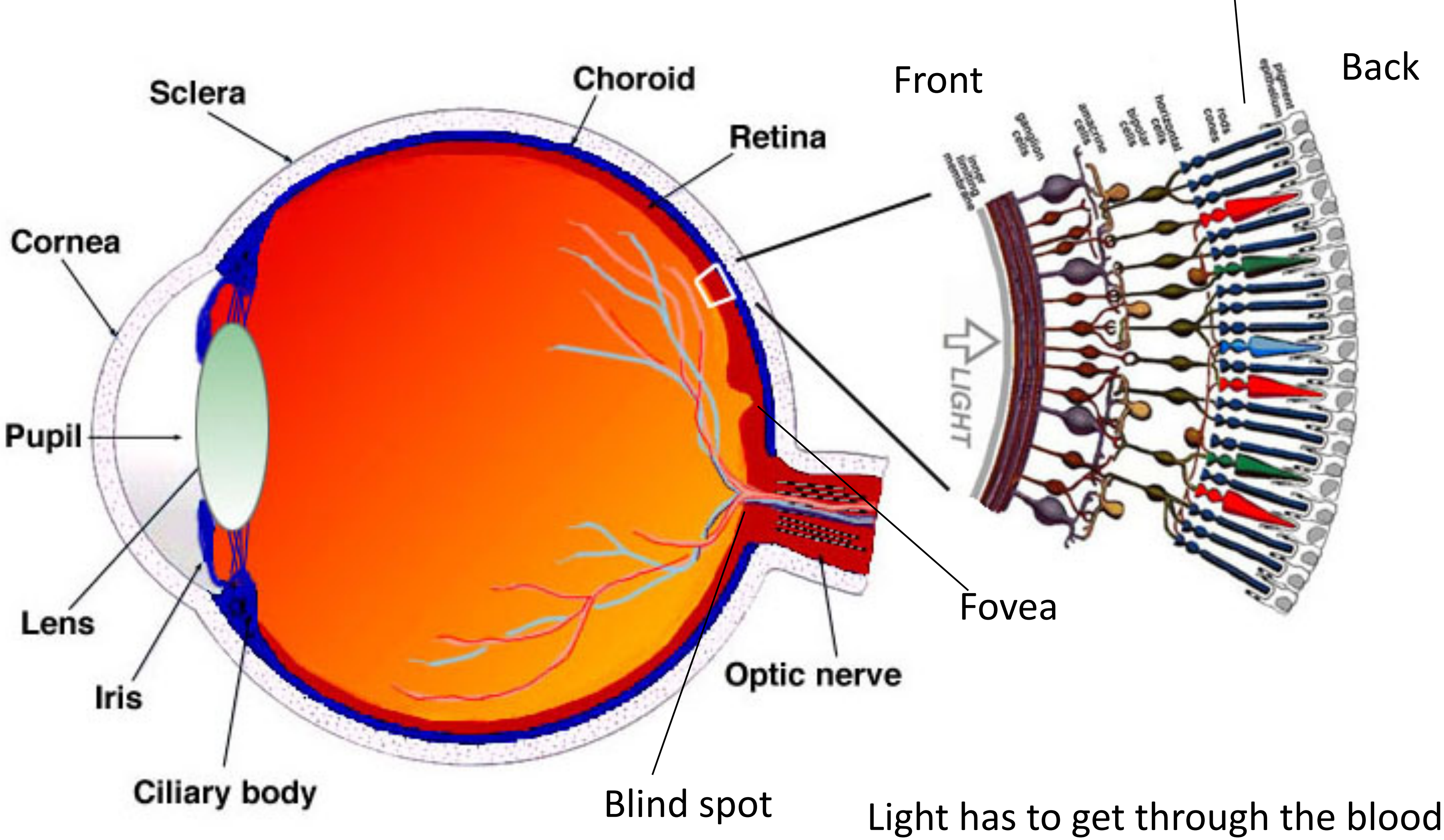
Biological vision

Fly eye



Thomas Shahan (CC BY-NC-ND 2.0)   http://www.biokurs.de/skripten/12/bs12-45.htm

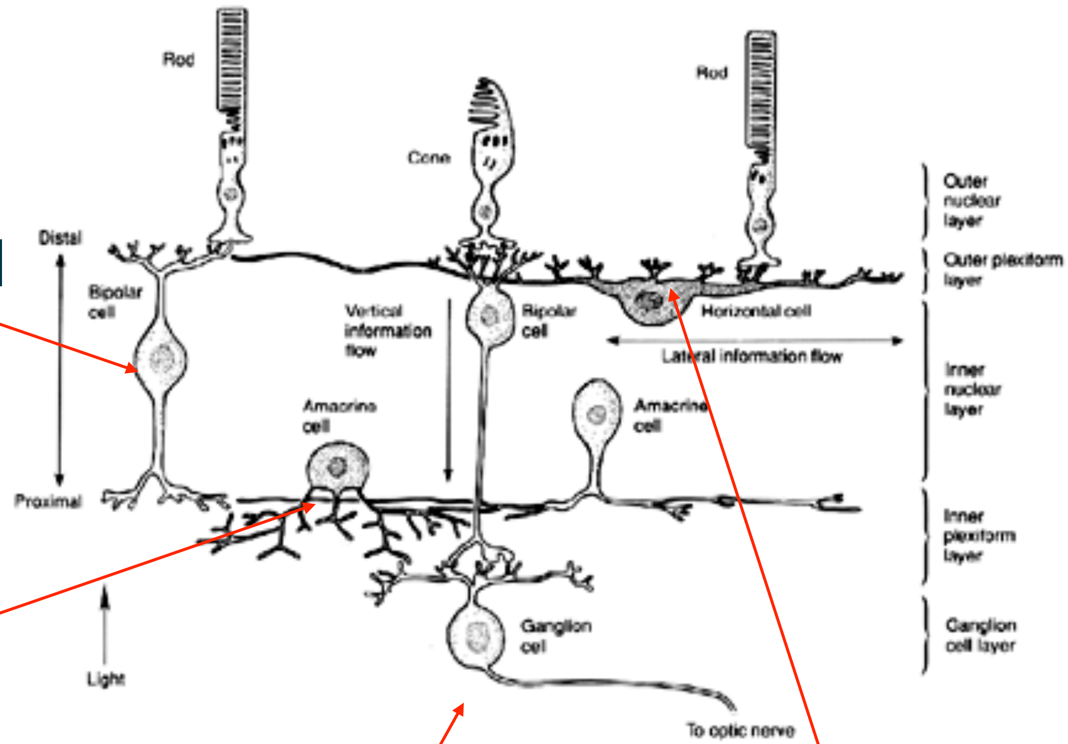
Human eye



Light has to get through the blood vessels and cell bodies to get to the photoreceptor.

Image From http://webvision.med.utah.edu used under Creative Commons Licence

Bipolar cell



Amacrine cell

Ganglion cell

Horizontal cell

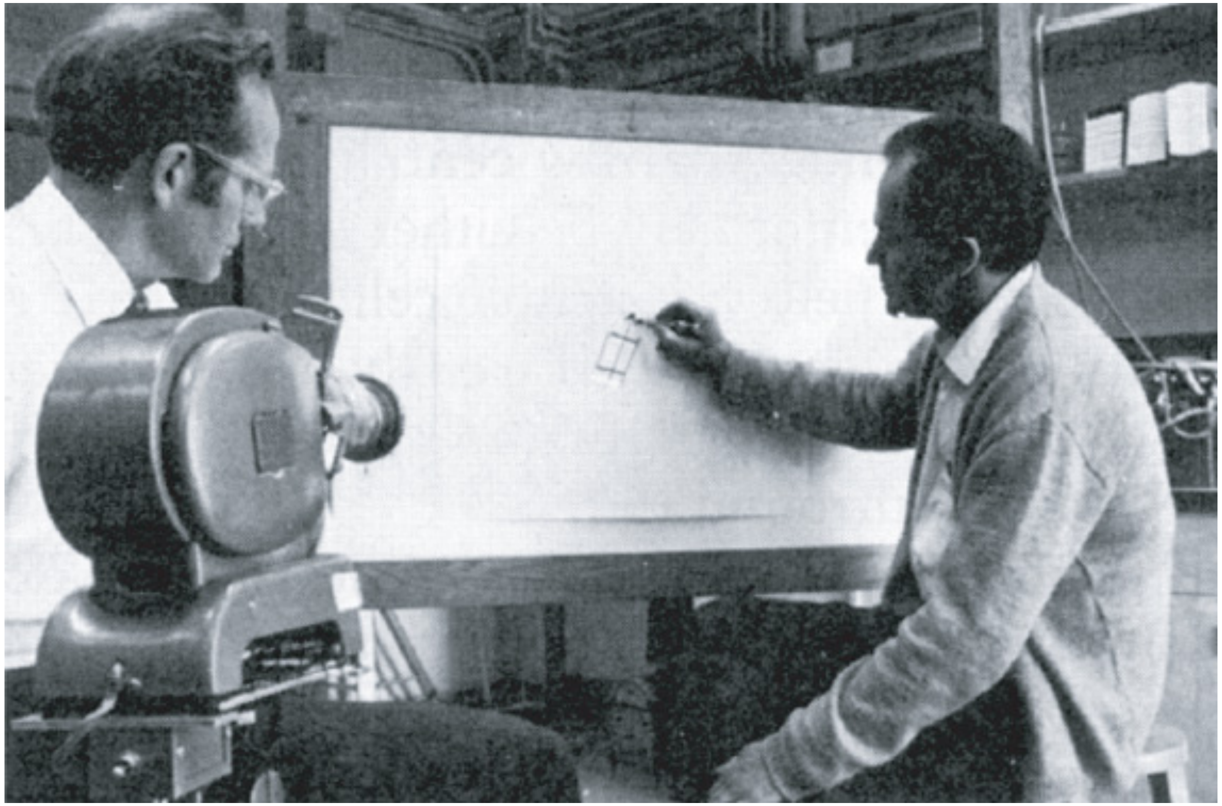
Receptive fields

- Photoreceptors will respond (produce action potentials) when light falls on their receptive field
- No retinal cell ‘sees’ all the image. Each one only gets light from a portion of space
- Brain cells throughout the visual system have receptive fields
- As we investigate deeper into the brain, receptive fields get more structured, specific and respond to more complex properties

From lecture #1

Oriented edges and bars
Area V1 of cortex

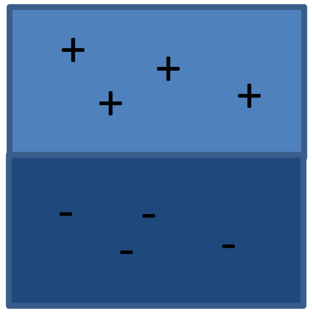
Classical view – Hubel and Wiesel, (1950's) – Won Nobel Prize. (Pub: 1962 Cat, 1968, Monkey)
Simple cells – orientation sensitive, off and on regions
Complex - orientation sensitive, prefers thin lines, but no off region
Hypercomplex – as complex but prefers short lines ('End – stopped')



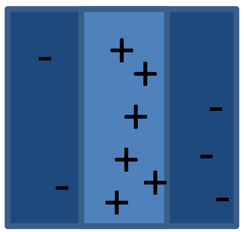
Oriented edges and bars
Cortex Area V1

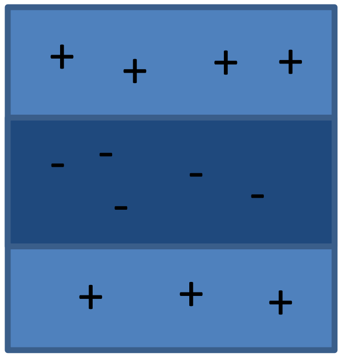
- Simple cell receptive fields
- Structured on / off regions
- Different orientations and sizes

H Edge



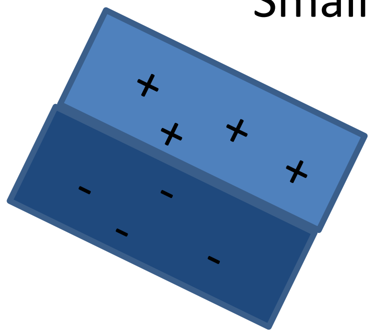
Small Light V bar

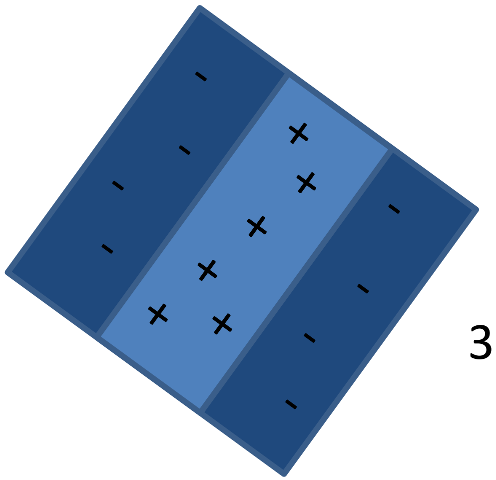




Dark, H bar

Small -15° edge





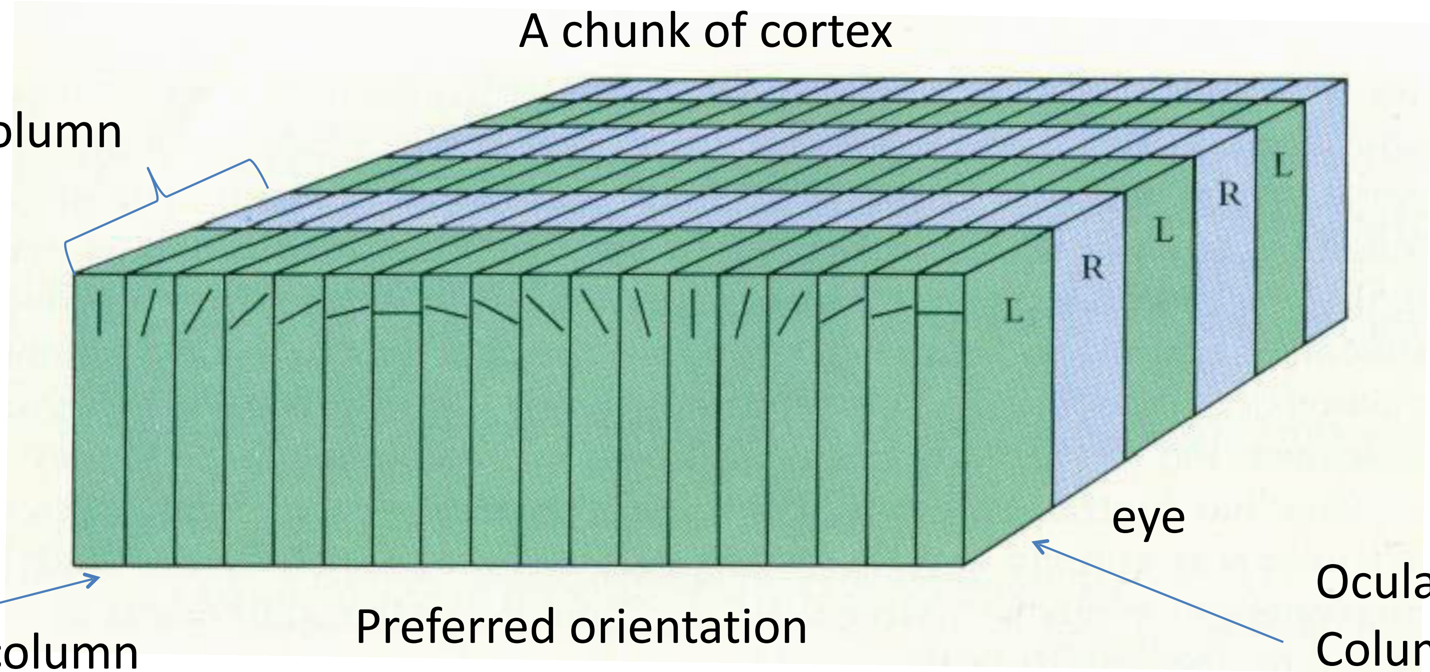
30° light bar

Cortical organisation

Image by Matthew Schmolesky down loaded from http://webvision.med.utah.edu used under Creative Commons Licence

A chunk of cortex

Hyper column



eye

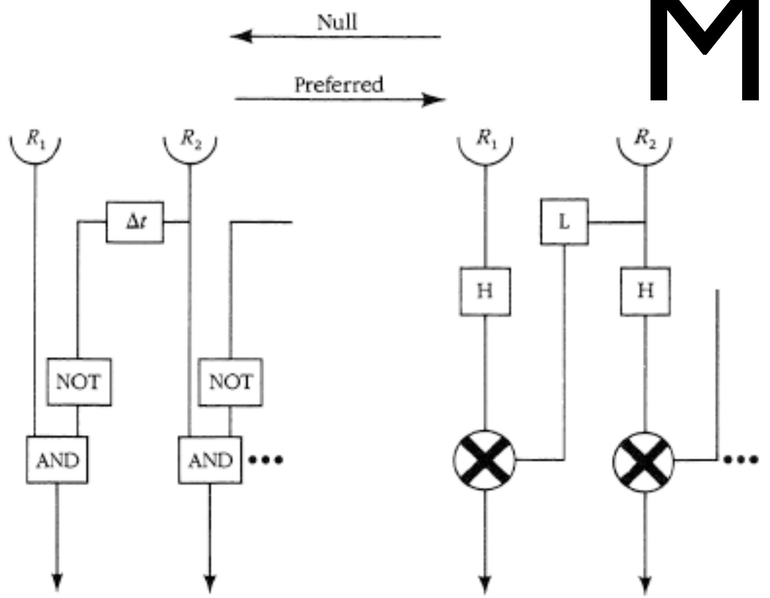
Ocular dominance Column

Orientation column

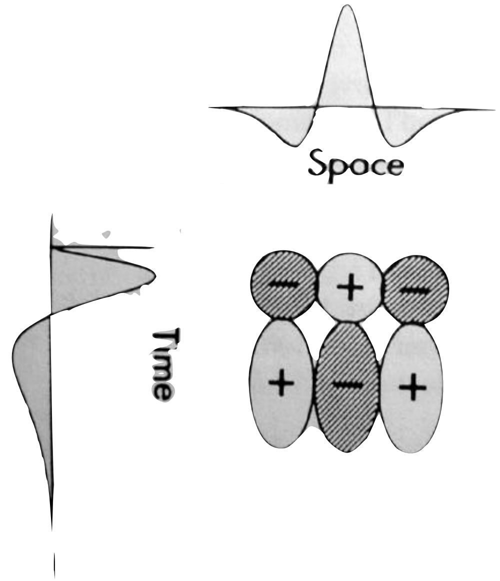
Preferred orientation

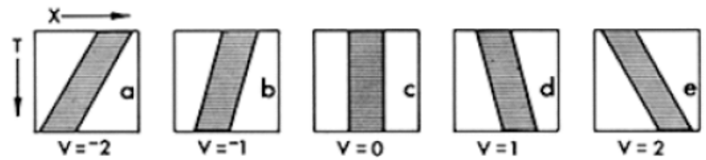
- Hyper columns contain cells with the same receptive field
- Nearby cells have similar preferential features
- Cells are laid out as if ready for further filtering operations

Motion

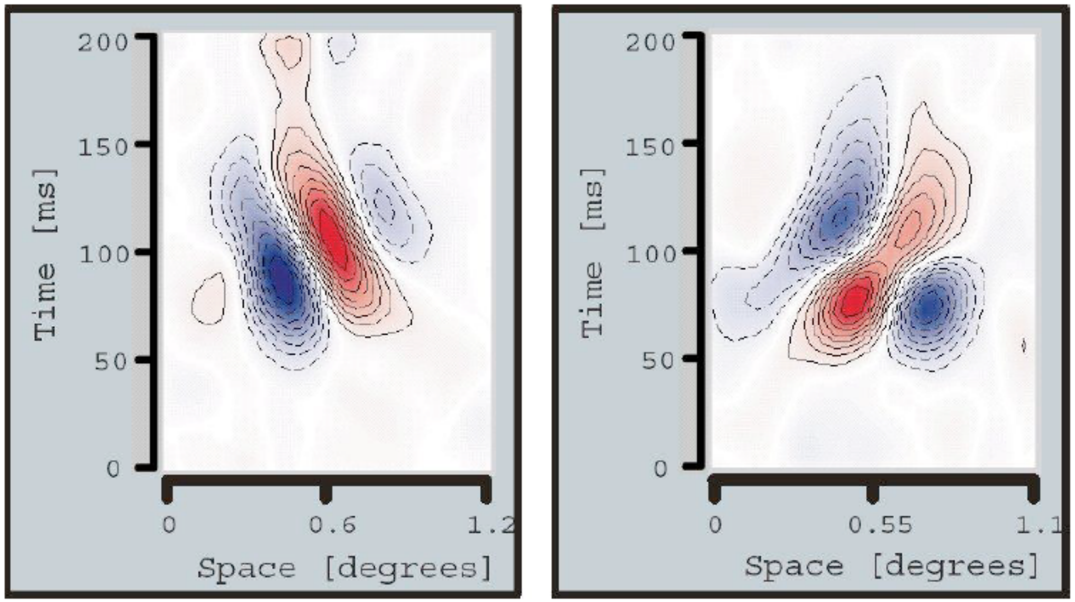


Barlow Levik (1965) Hassenstein Reichardt (1956)

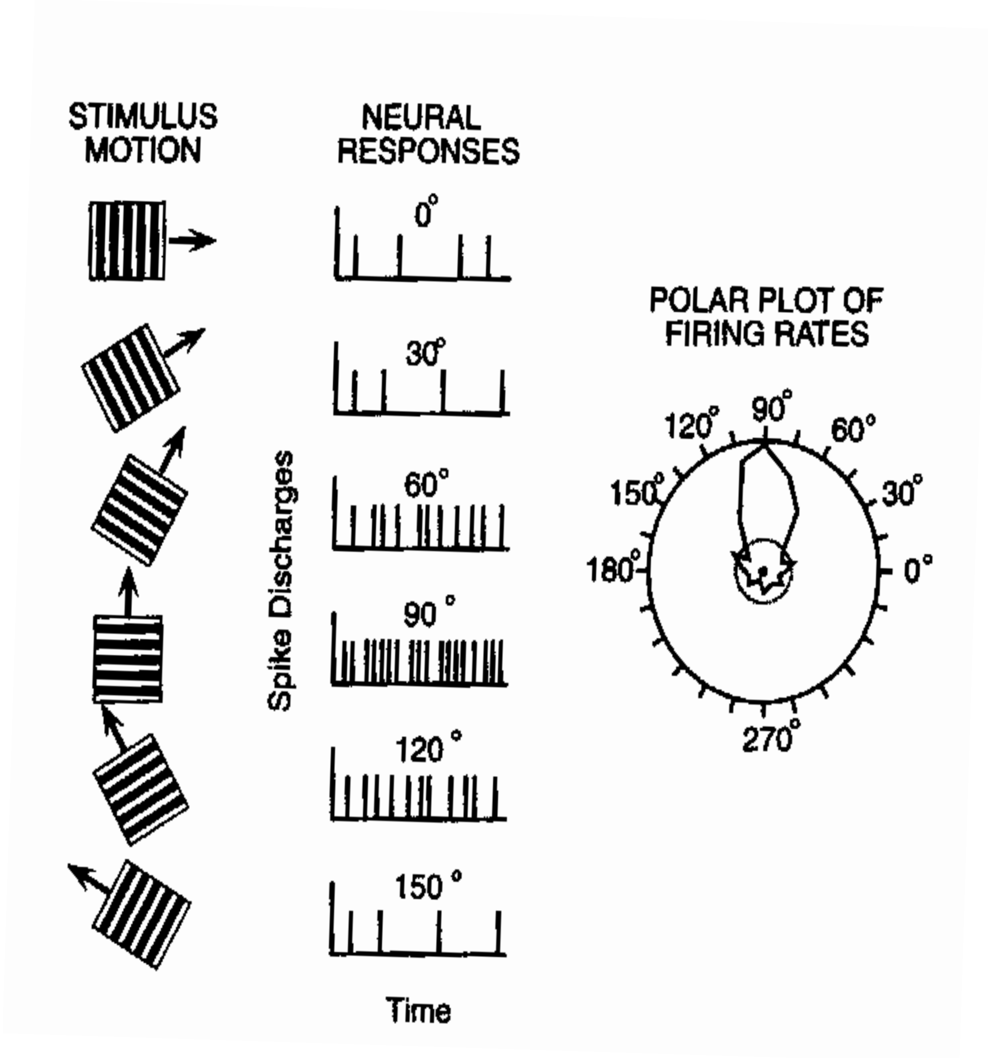




Adelson, p. 269

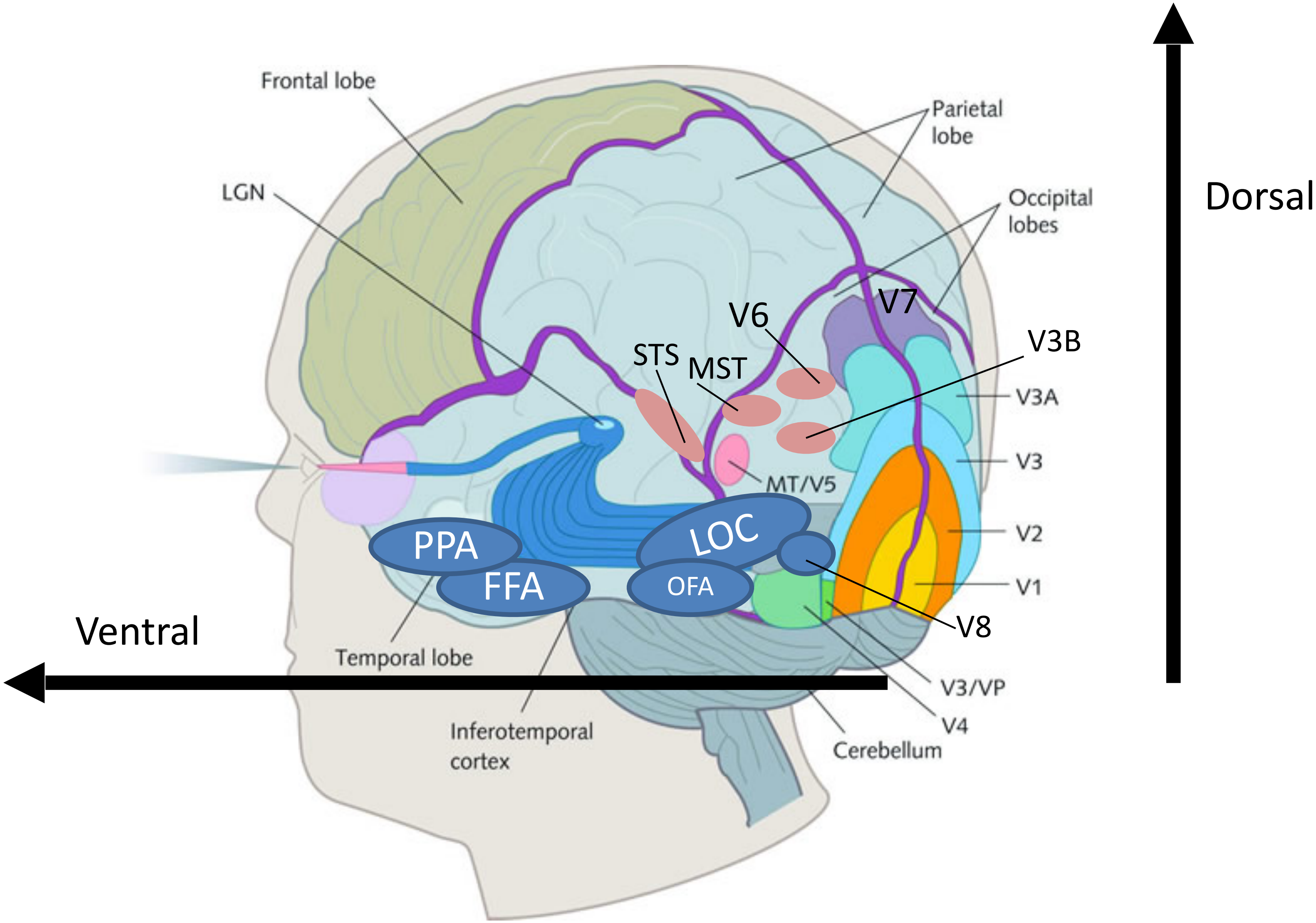


Devalois et al. (2000)



Complex cell in cat cortex

Visual pathways



Computational vision

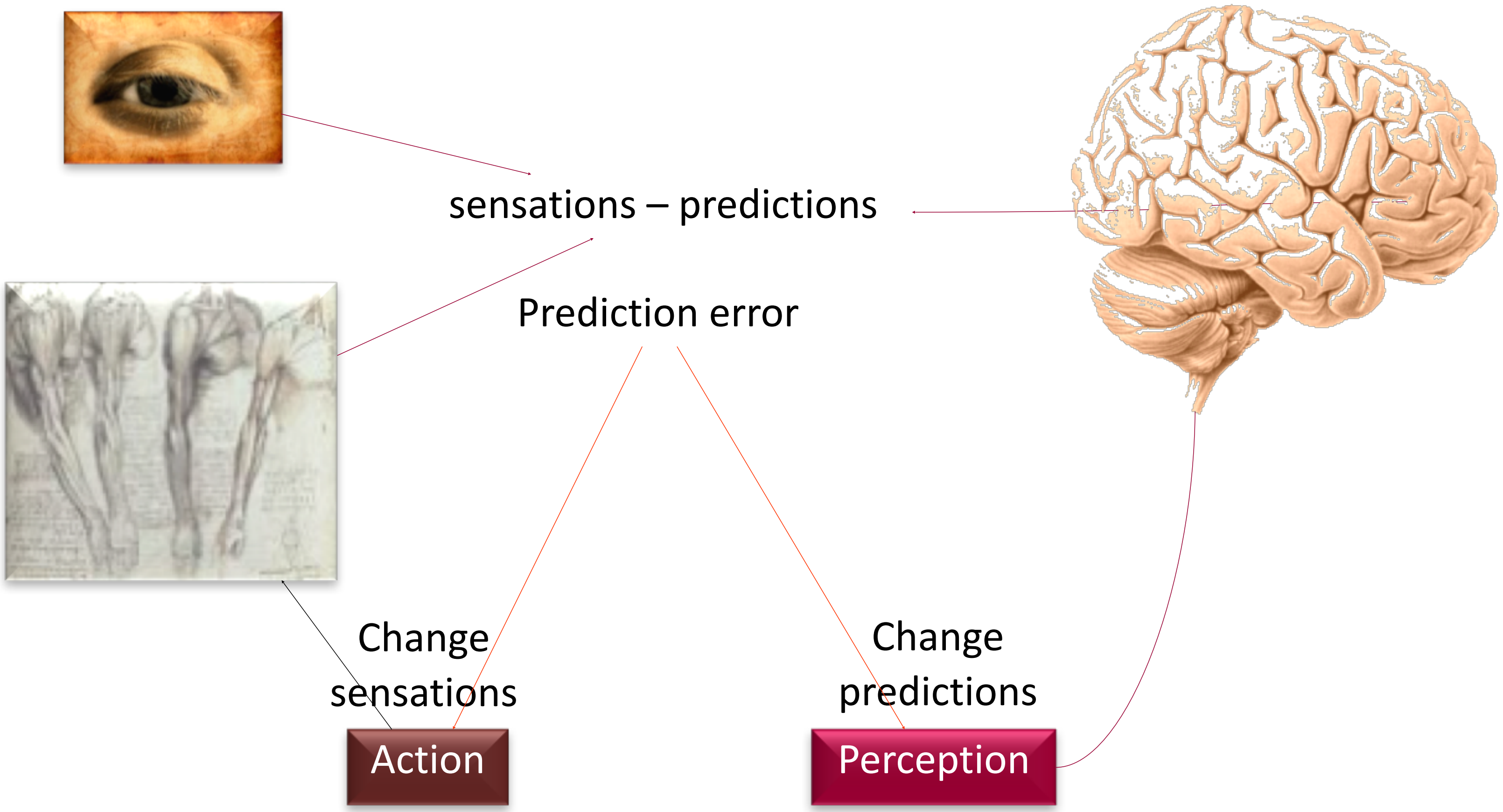
Hypothesis 1

- The visual system can be studied as the sum of separate **modules** that, in first approximation do not interact with each other (cfr Fodor)

- The goal of a module is to make **explicit** some part of the information that is implicitly carried by the image (cfr Friston)

From lecture #2

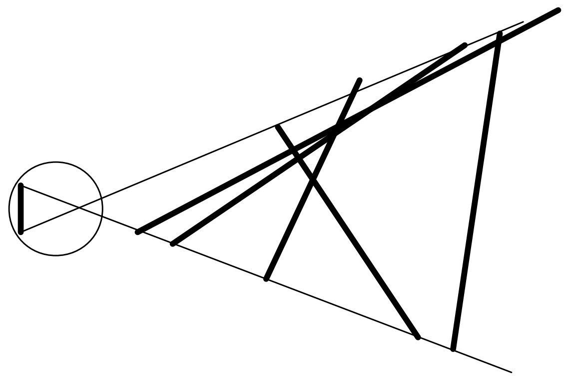
Free Energy Principle



Action and perception minimise surprise

Hypothesis 2

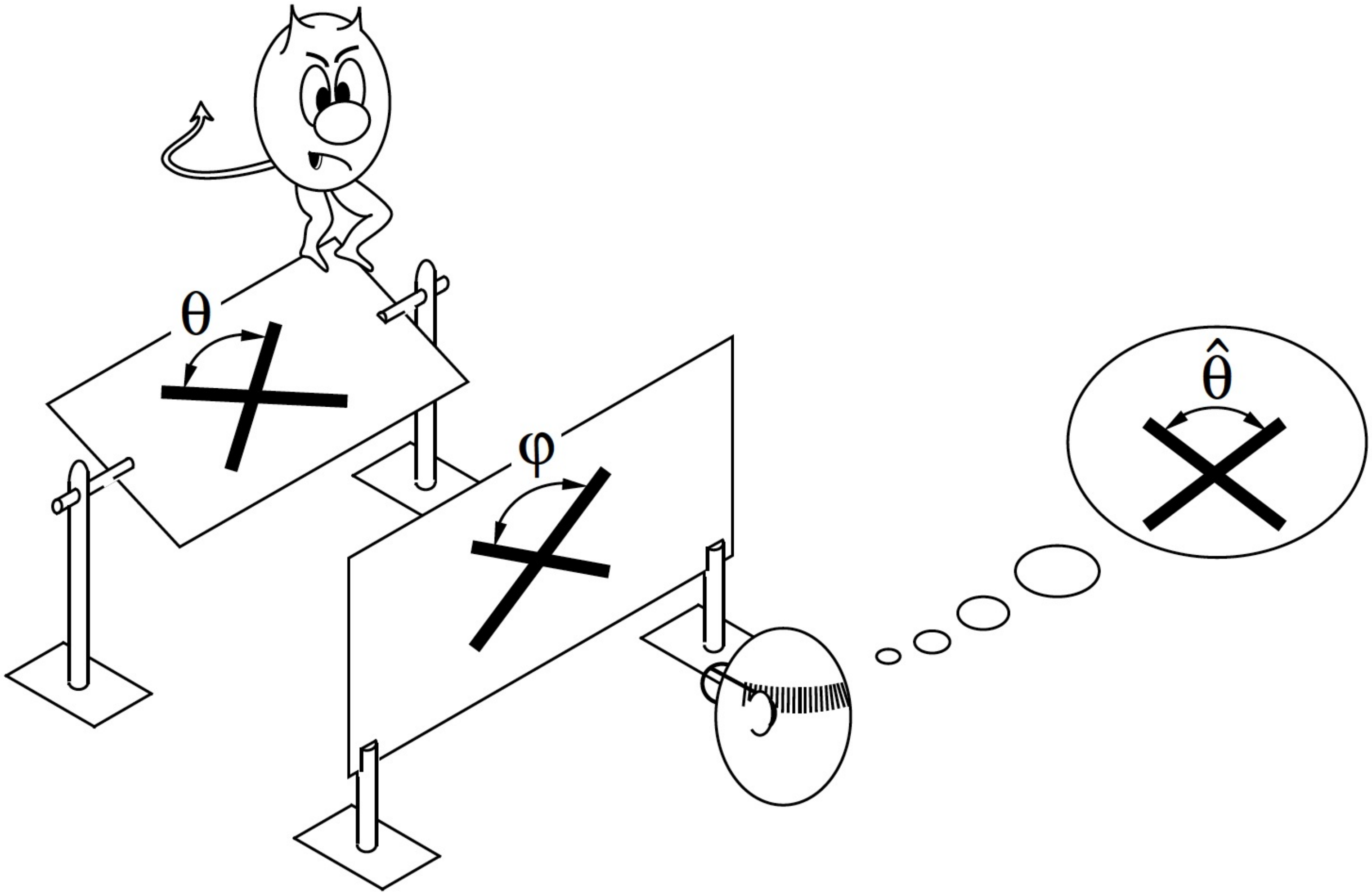
- One-to-many mapping problem
- The visual input is always ambiguous



- It is possible to limit the number of possible interpretations of the visual input by making **assumptions** about the 3D world.

From lecture #2

Bayesian Inference

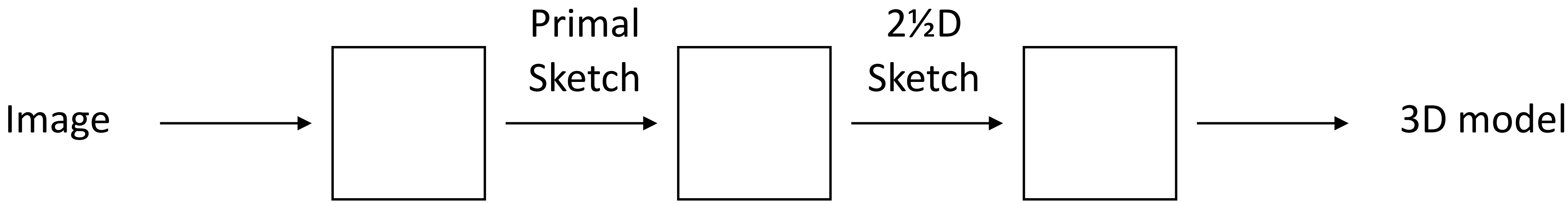


Mamassian Landy Maloney (2002) in Rao Olshausen Lewicki

Hypothesis 3

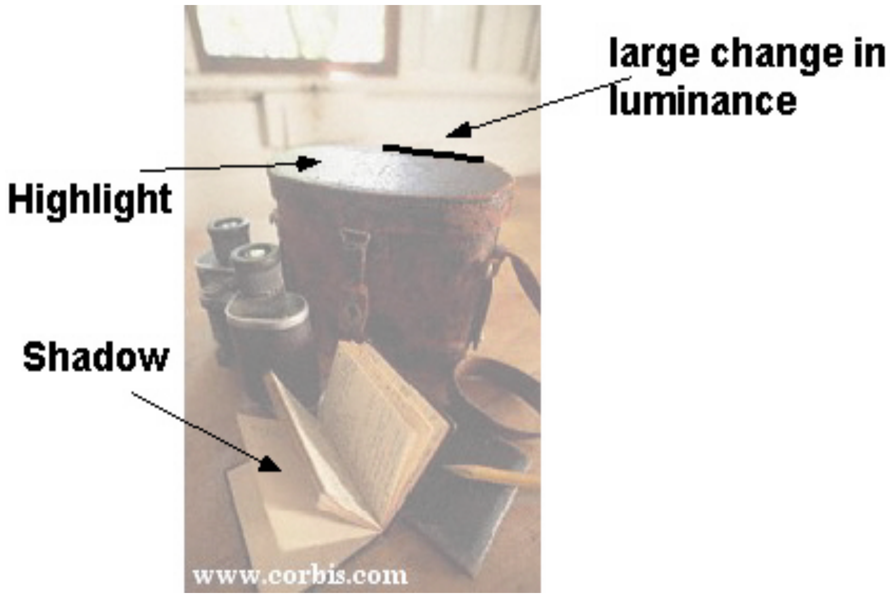
- Modules are organised in a hierarchical structure
- The output of each stage is a **representation** of the input

Marr proposal for deriving shape information from images



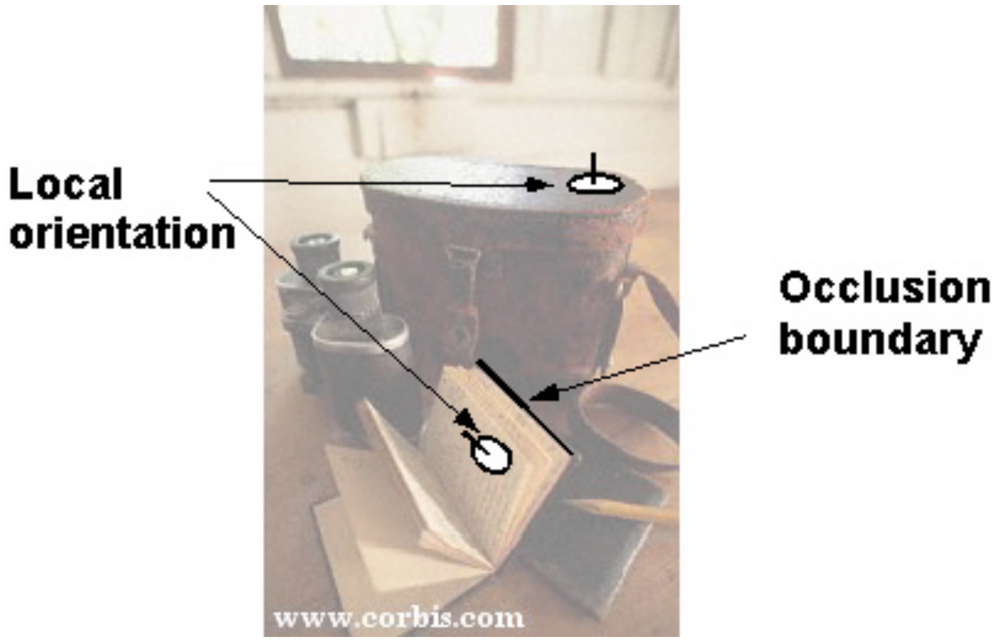
Representation 1: Raw Primal Sketch

"Makes explicit important information about the two-dimensional image, primarily **intensity** changes and their geometrical and spatial **organization**" (Marr, 1982)



Representation 2: 2½-D sketch

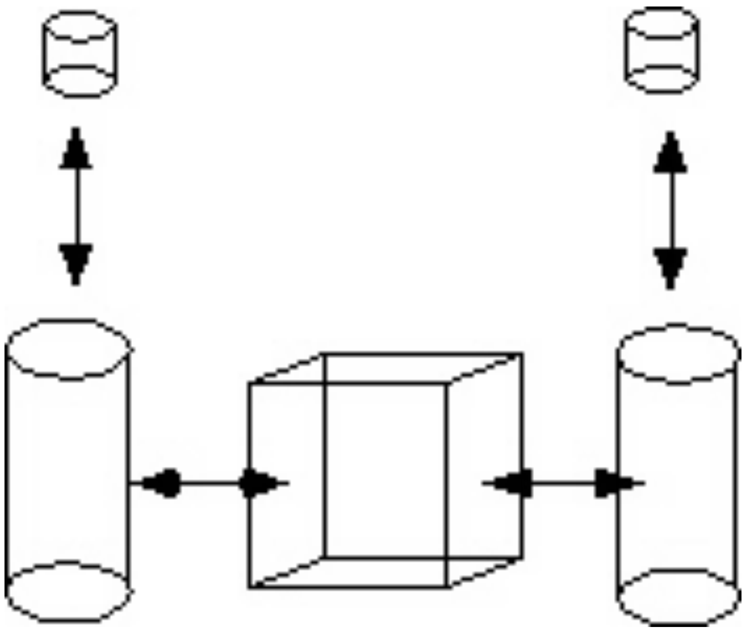
"Makes explicit the **orientation** and rough **depth** of visible **surfaces** and **contours** of discontinuities in these quantities in a **viewer centered** coordinate frame" (Marr, 1982)



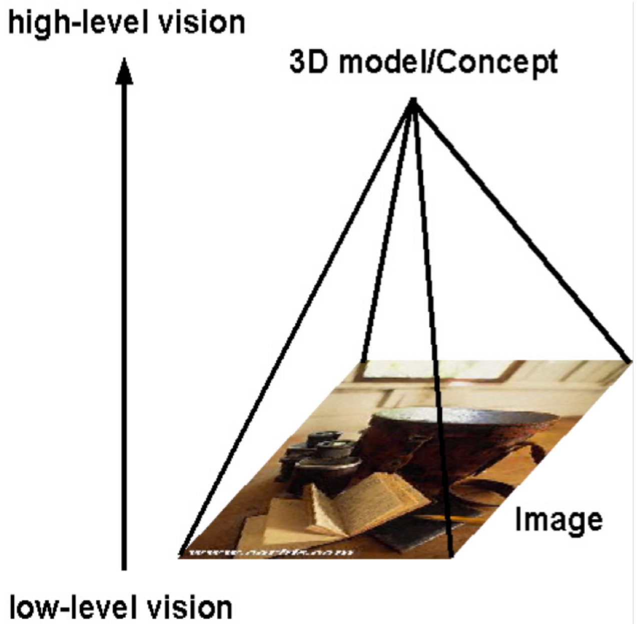
Representation 3: 3D model

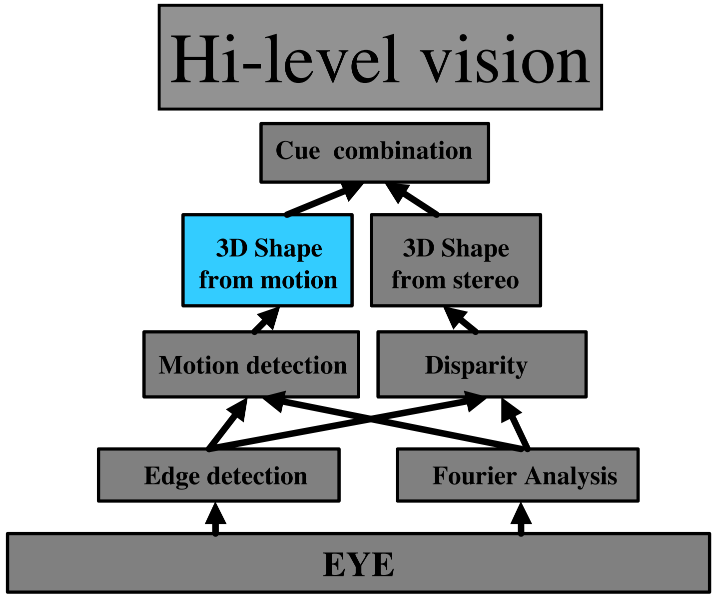
"Describes **shapes** and their spatial **organization** in an **object centered** coordinate frame”

(Marr, 1982)



Pyramidal organisation



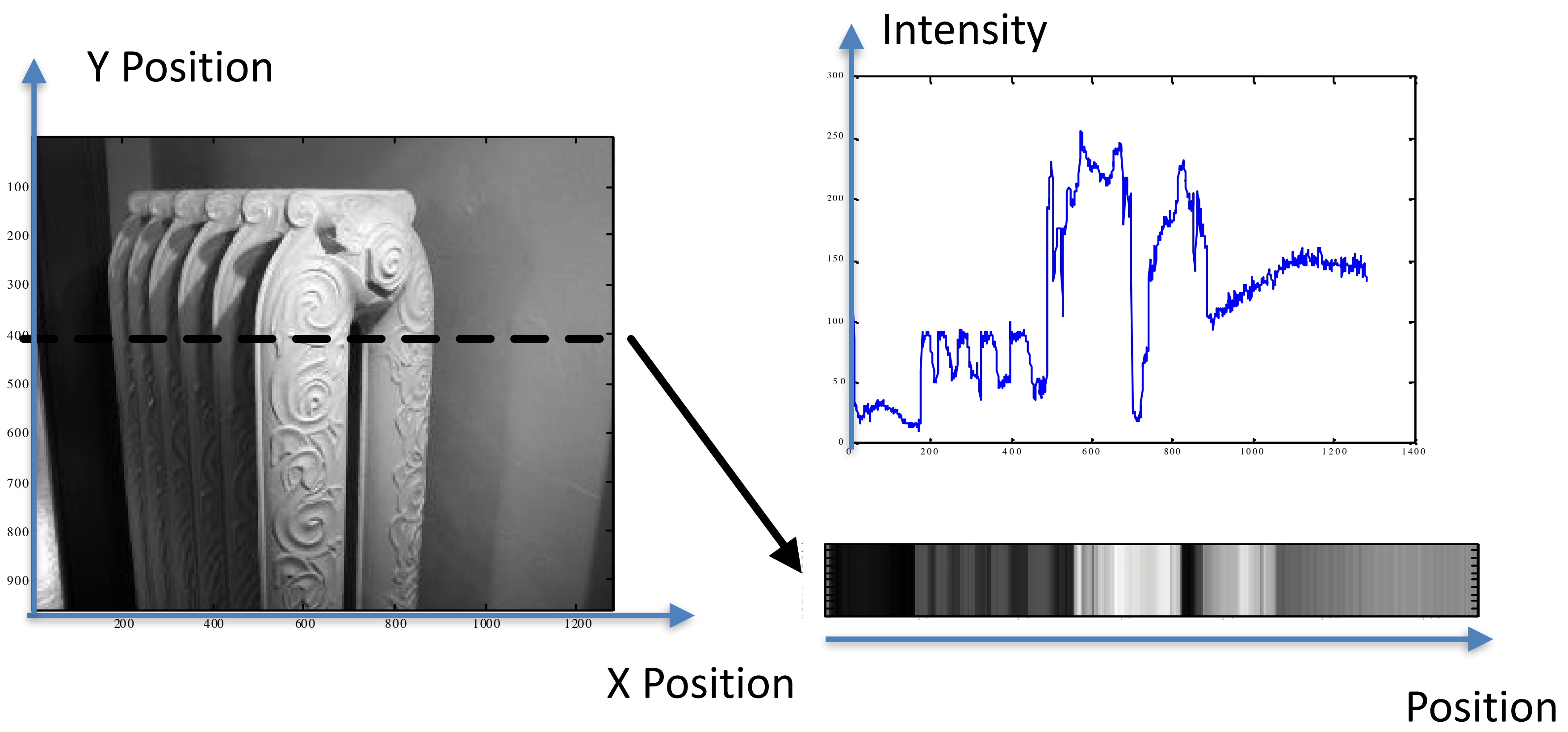


How are the intermediate representations chosen?

1. They correspond to physical changes and they are not the artefact of processing
2. They are "useful" to some action or perception goals
3. Computations can be performed efficiently
4. They have a semantic relevance

Computer vision

Image representation

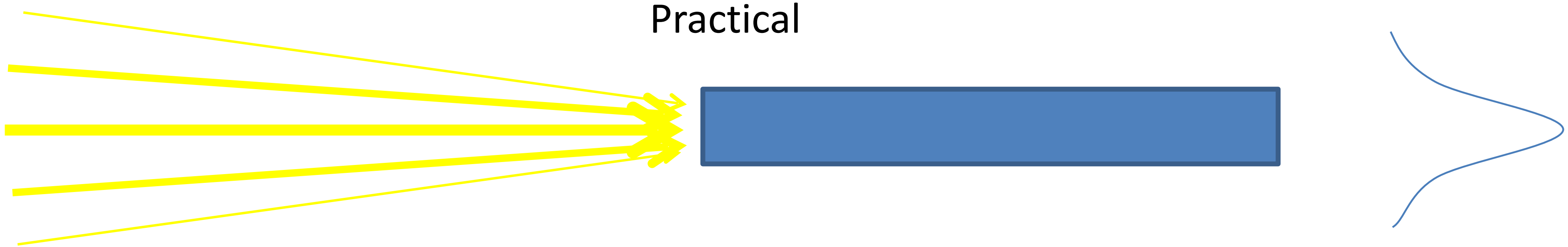


Photoreceptor receptive fields

Ideal

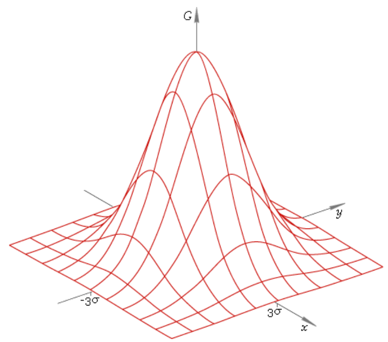
Impulse or delta

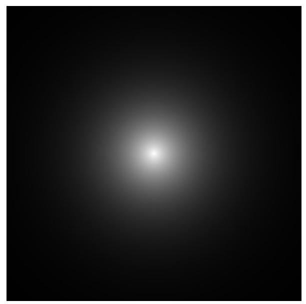
Practical



Gaussian

$$g(x,y)=\frac{e^{-(x^2+y^2)/2\sigma^2}}{2\pi\sigma^2}$$



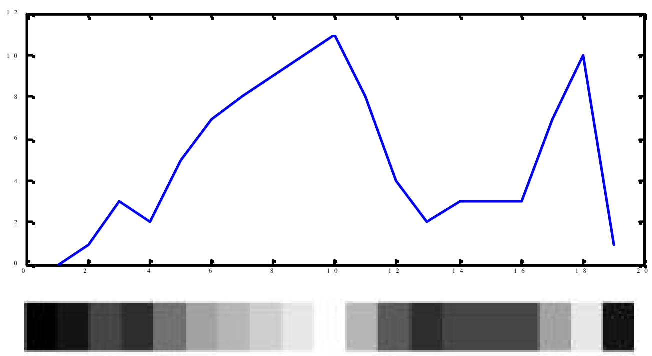


Convolution: intro

Retinal cell activity *J* is obtained by averaging each pixel *i* of an image *I(i)* through a weighted sum where the weights are specified by a function *G(i)*

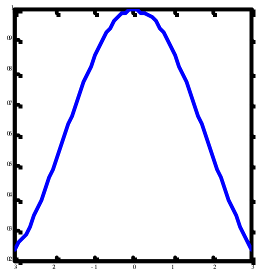
$$J(i) = \sum_{h=-m/2}^{m/2} G(h)\, I(i-h) \qquad\qquad J = G * I$$

$I(i)$



$*$

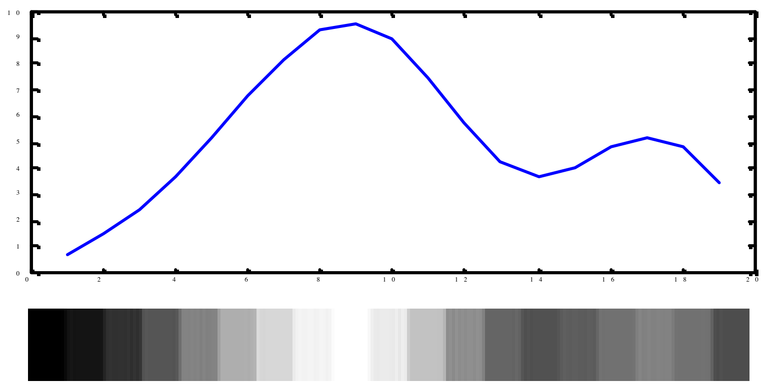
$G(h)$



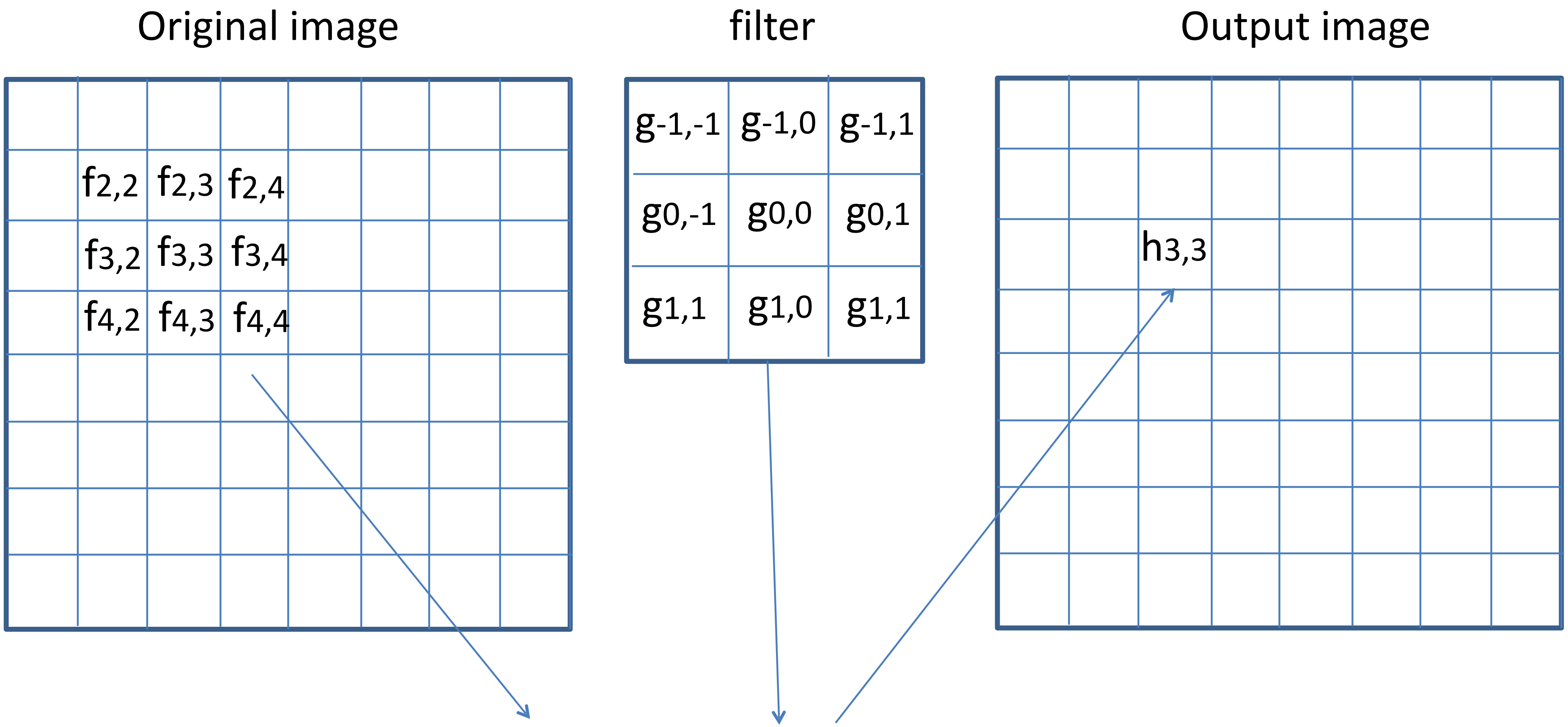
-m/2 0 m/2

$=$

$J(i)$



Convolution: in 2 dimensions



$$h_{3,3}= f_{2,2}g_{-1,-1}+f_{2,3}g_{-1,0}+f_{2,4}g_{-1,1}+f_{3,2}g_{0,-1} \ldots\ldots \quad f_{4,4}g_{1,1}$$

Convolution: formulas

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(m) g(x-m)\, dm$$

$$f(x) * g(x) = \sum_{m=-M}^{M} f(m) g(x-m)$$ Discrete

$$f(x,y) * g(x,y) = \sum_{m=-M}^{M} \sum_{n=-N}^{N} f(m,n) g(x-m, y-n)$$ Discrete 2D

MATLAB: conv2(F,G)

Convolution: Image blurring

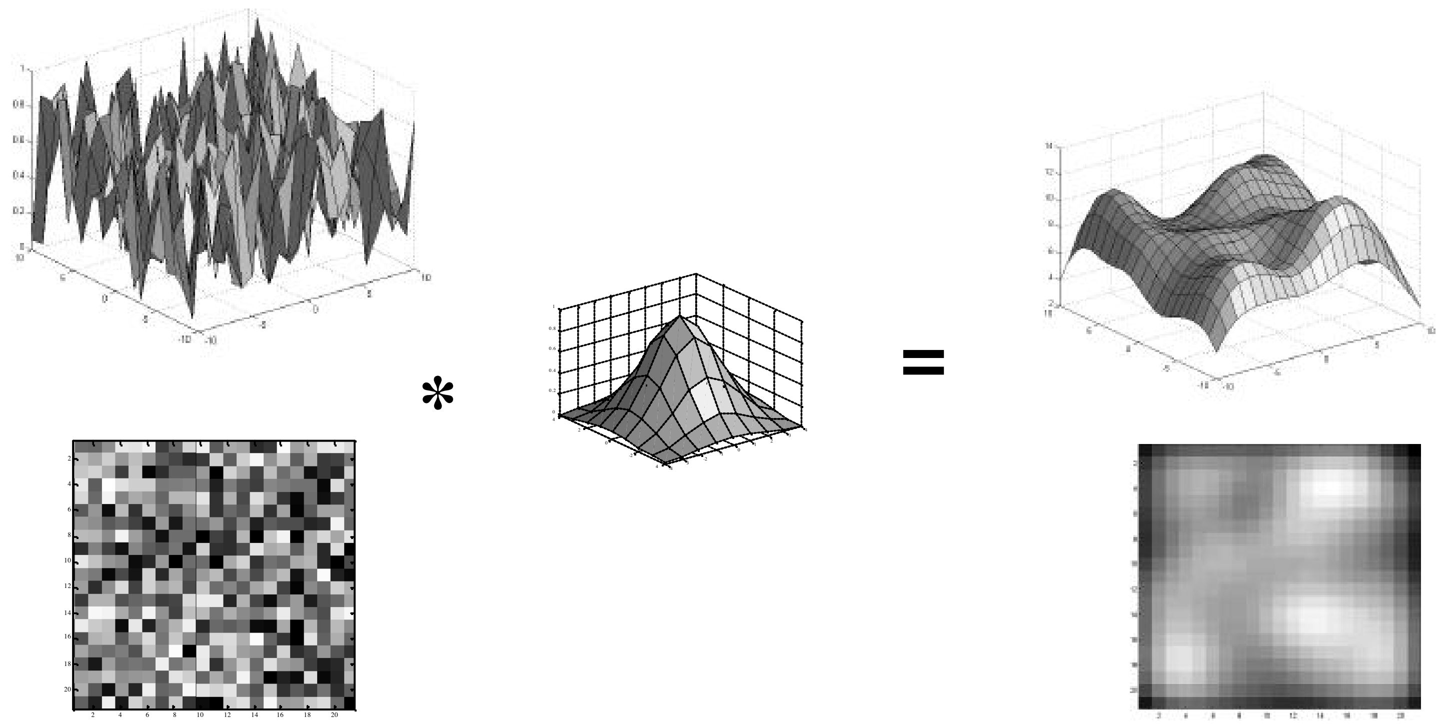
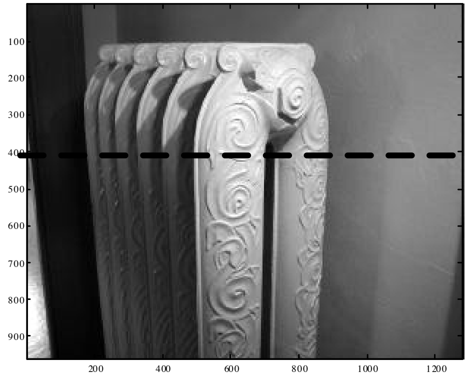
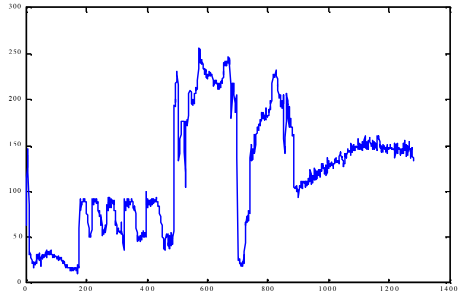
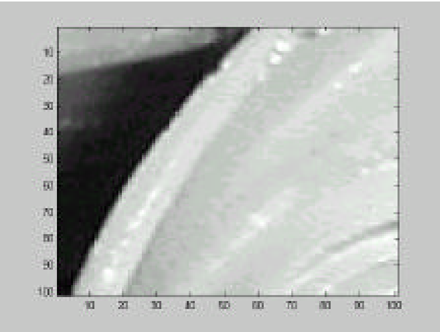
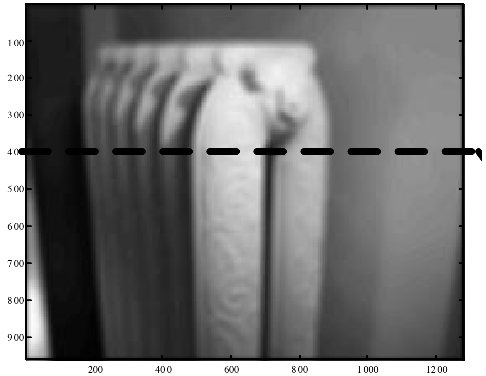


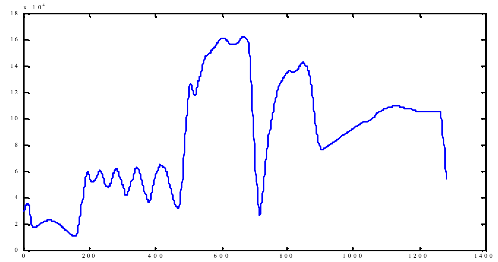
Image blurring

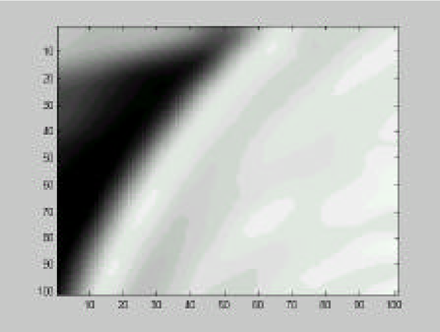












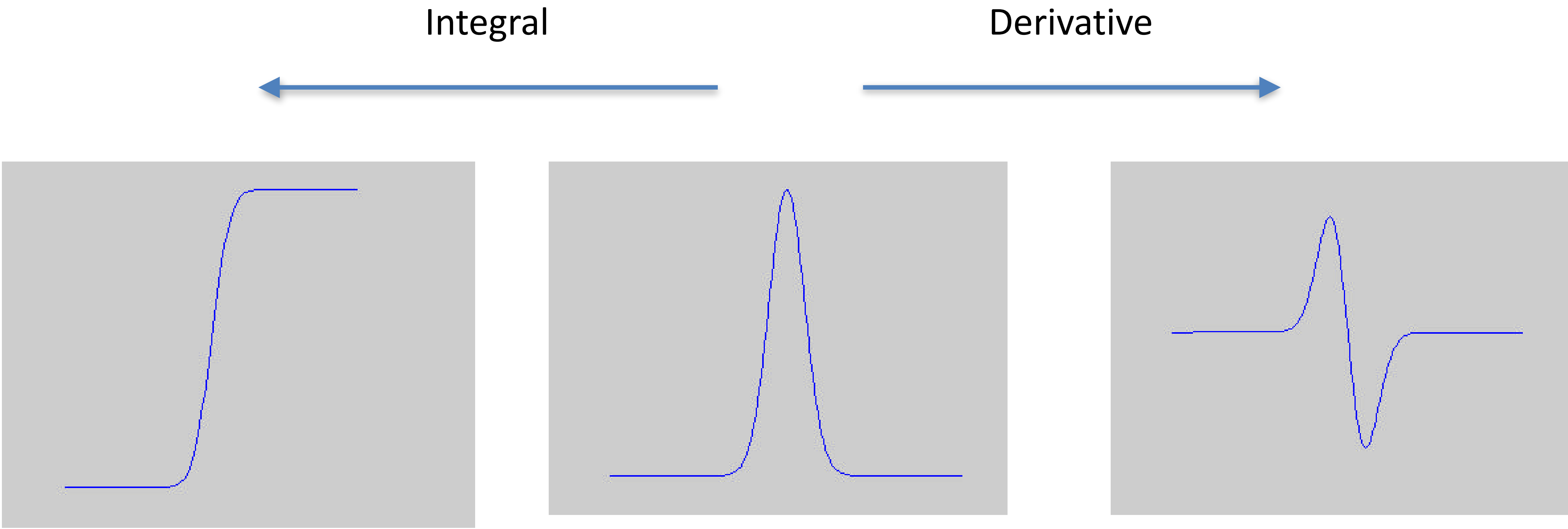
Edges

Changes in the image

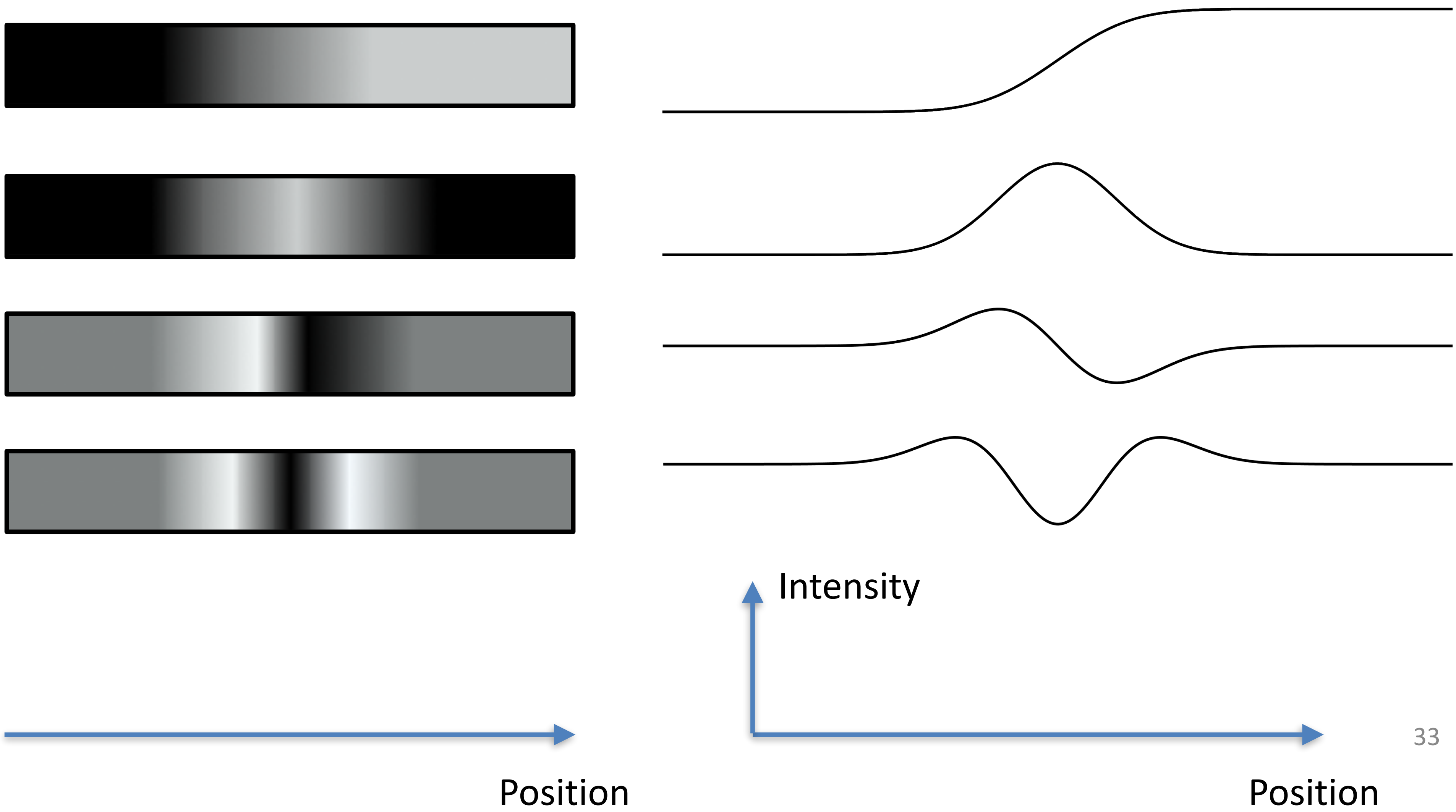
Can reflect:

- Geometry of the 3D world
- Illumination changes
- Reflectance differences

Review



Edges as derivatives



Edges: zero-crossing method

- Taking the Laplacian ∇^2 (second-order derivative) and rectifying can be a way to find edges in all directions, as low values correspond to zero-crossing in the image

Edges: Scale

$$G(x)=e^{-\frac{x2}{2\pi\sigma 2}}$$

Original Image

Fine scale

Medium scale

Large scale

Edges detection: summary

1) Blur the image through the convolution with a gaussian

$$I\_blurred = G * I$$

2) Detect the zero crossing with the Laplacian

$$I\_z = \nabla^2 I\_blurred$$

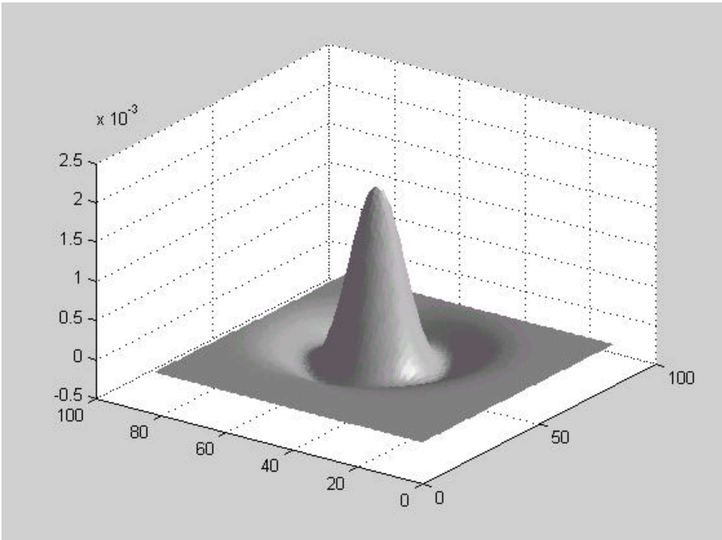
3) Absolute and close to 0

Edges: operation transitivity

$$\nabla^2 (G * I) = (\nabla^2 G) * I$$

The “mexican-hat-shaped” operator, the laplacian of a Gaussian

$$\nabla^2 G$$

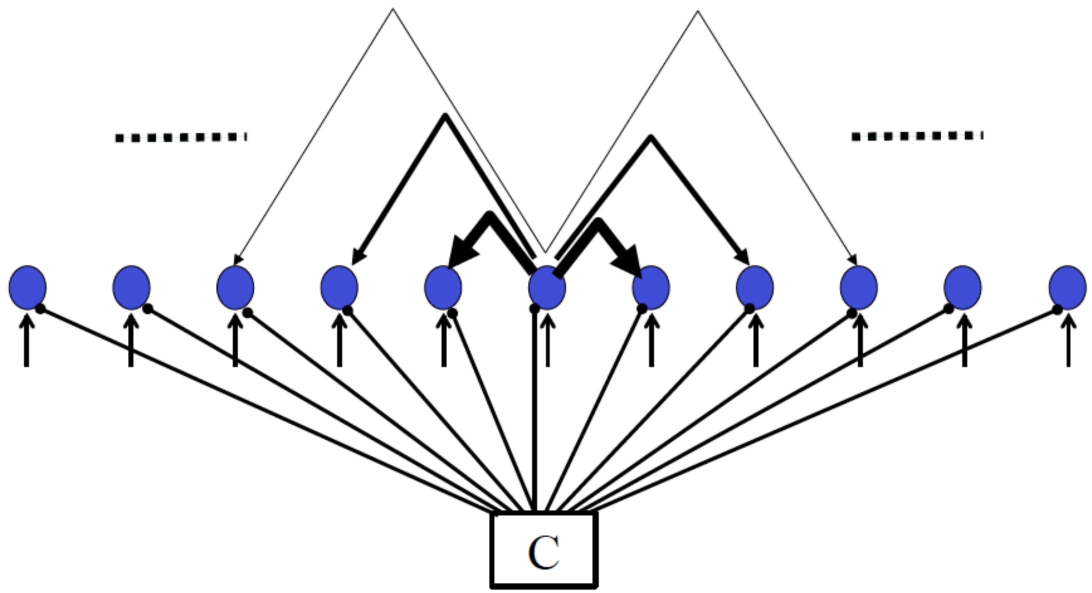


We can convolve the image $\mathbf{I}$

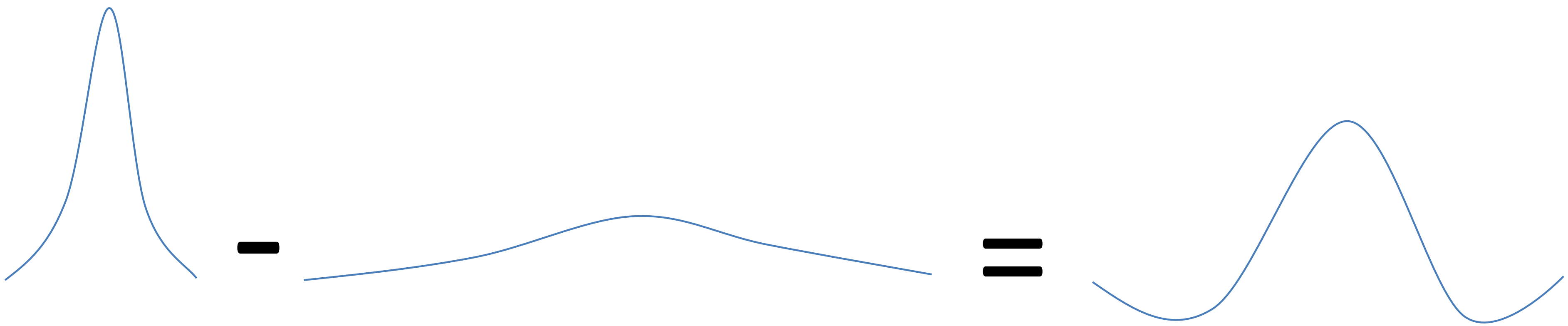
with the operator $\nabla^2 G$

From lecture #10

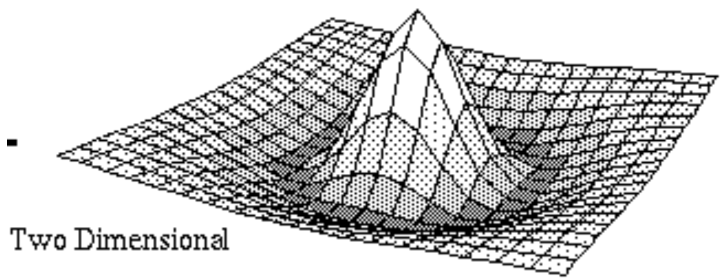
Lateral connections



The Difference of Gaussians



$$g(x,y)=\frac{e^{-(x^2+y^2)/2\sigma_1^2}}{2\pi\sigma_1^2}-\frac{e^{-(x^2+y^2)/2\sigma_2^2}}{2\pi\sigma_2^2}$$



Biological plausibility

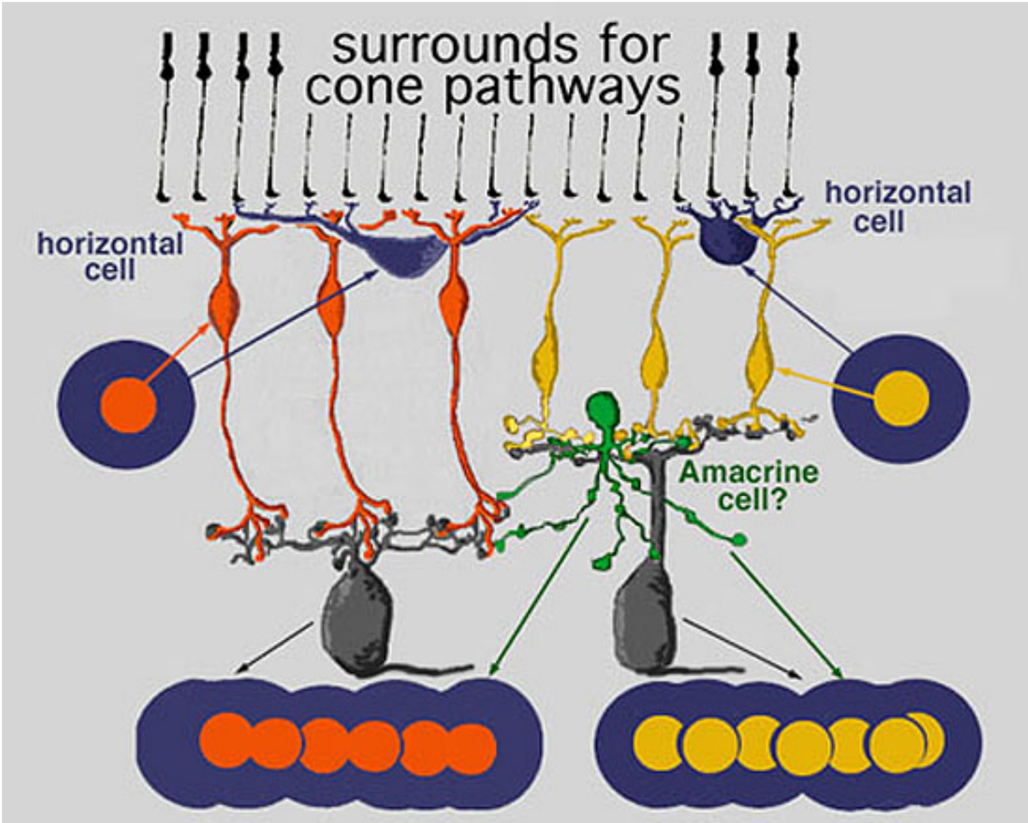
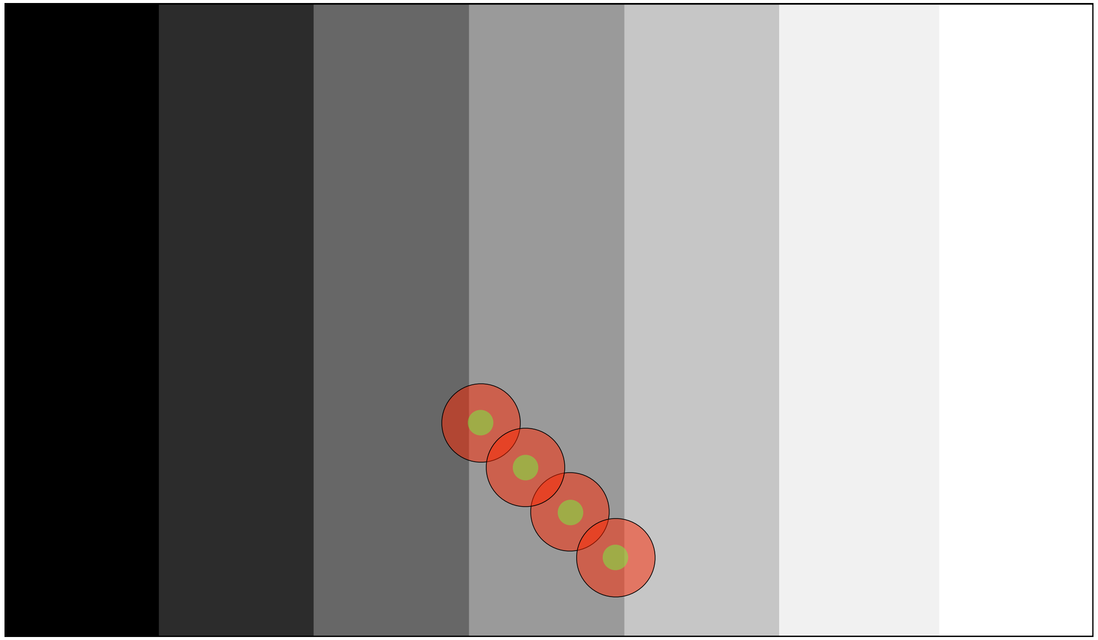
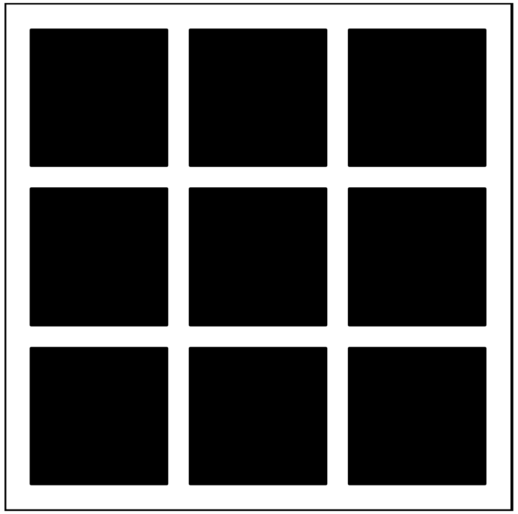


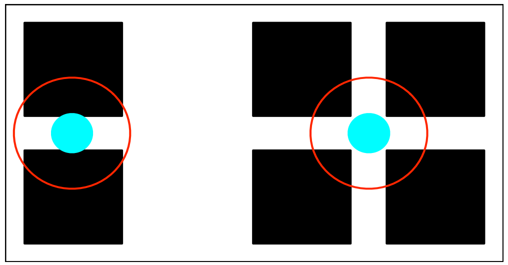
Image From http://webvision.med.utah.edu used under Creative Commons Licence

Perceptual phenomena

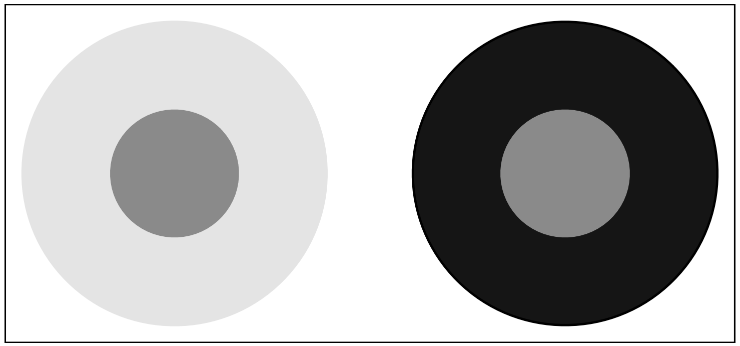


Mach Bands





Hermann grid



Simultaneous contrast

Spatial frequency theory

Fourier theorem applied to images:
Images can be described as the sum of sinusoidal gratings

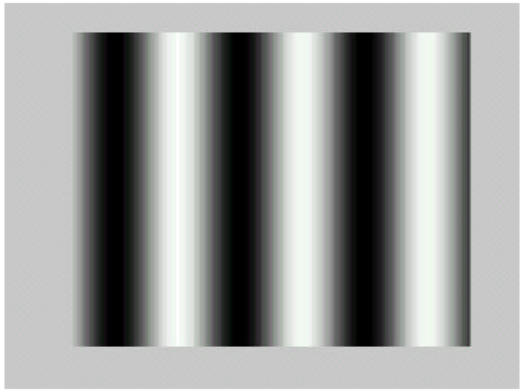
1. **Spatial Frequency.** Light/Dark cycles per degree of visual angle

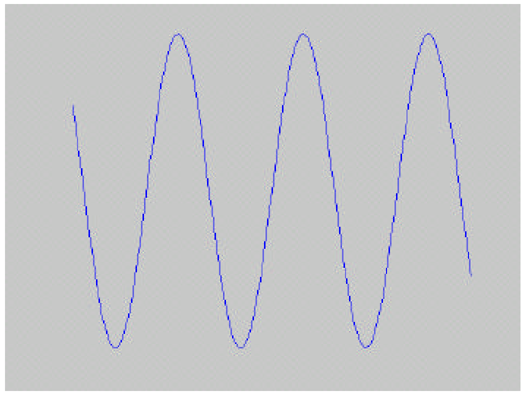
2. **Orientation** of the light and dark bars

3. **Amplitude**. Difference in luminance

between the lightest and darkest parts

4. **Phase**. Position of the sinusoid with respect to a reference point





Spatial frequency theory

The visual system decomposes the image into a sum of sinusoidal basis functions having different frequencies and orientations

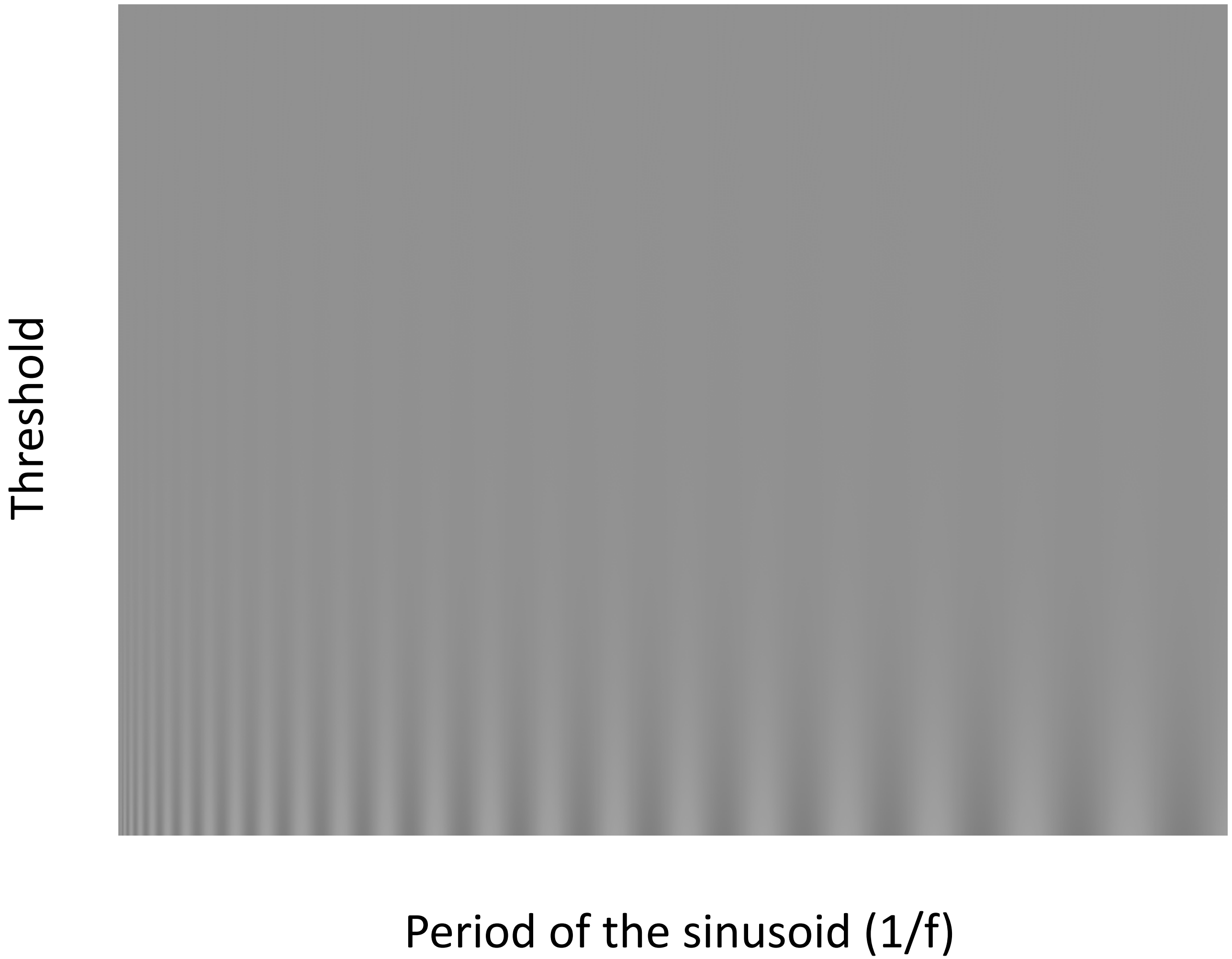
This decomposition may be localised in each position of the visual field (Gabor Patches)

- Psychophysical evidence

1. Contrast Sensitivity Function (Campbell and Robson, 1968)

2. Selective adaptation (Blakemore and Campbell, 1969)

Spatial frequency theory

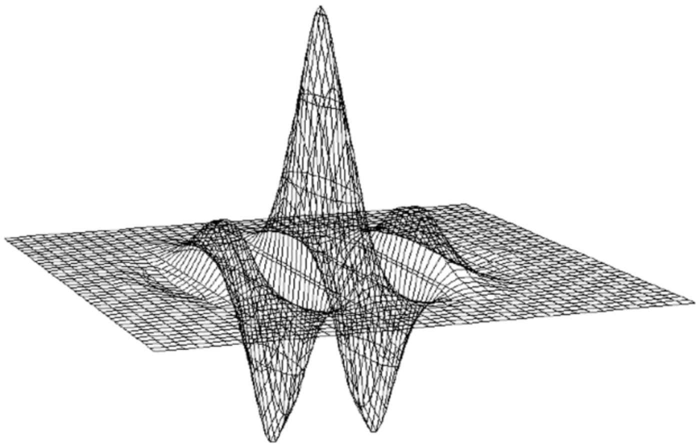


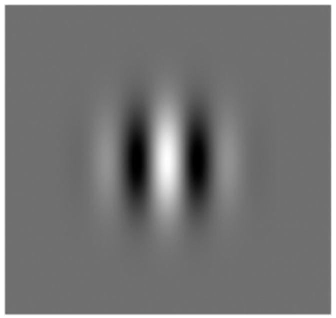
The Gabor Function

$$g(x, y) = \cos(2\pi f x' + \psi) e^{-(x'^2 + \gamma^2 y'^2)/2\sigma^2}$$

$$x' = x\cos(\theta) + y\sin(\theta)$$

$$y' = -x\sin(\theta) + y\cos(\theta)$$





Frequency: operation transitivity

$$F\left(f\left(x,y\right)*g\left(x,y\right)\right)=F\left(f(x,y)\right)F\left(g\left(x,y\right)\right)$$

Fourier transform of convolution =
Fourier transform of image x Fourier transform of filter

$$f\left(x,y\right)*g\left(x,y\right)=F^{-1}\left(F\left(f(x,y)\right)F\left(g\left(x,y\right)\right)\right)$$