

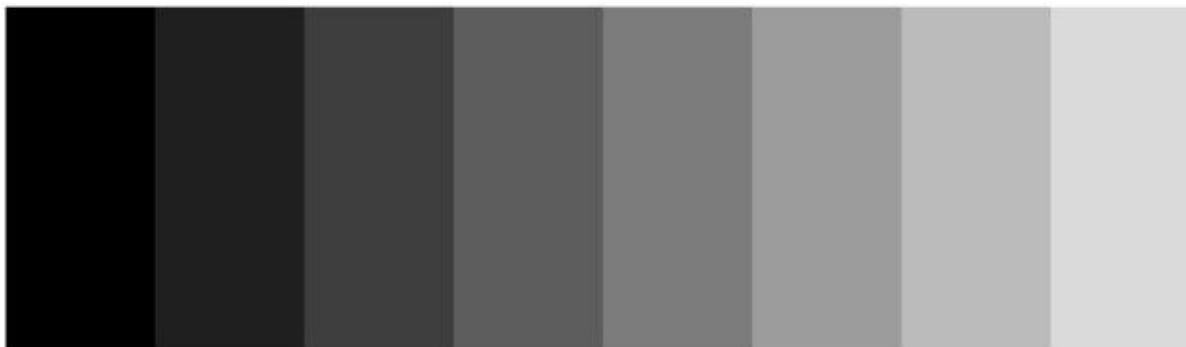
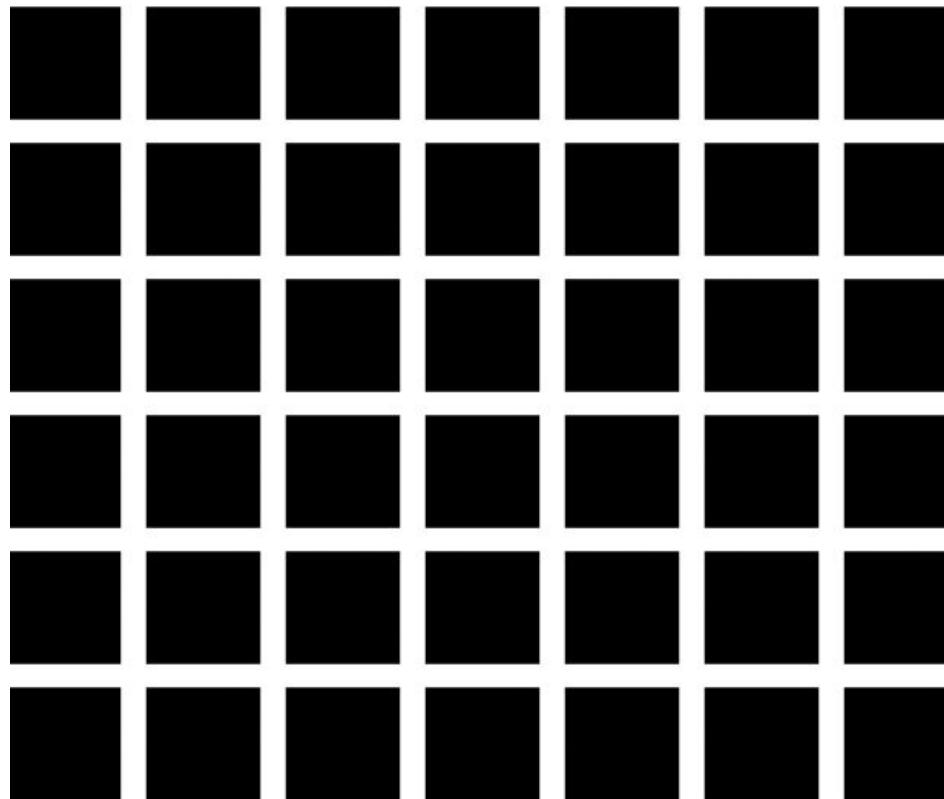
# Circuitos elementales en la percepción

campos receptivos; selectividad a la orientación: células simples y células complejas. Selectividad al movimiento.

# Motivación de la clase

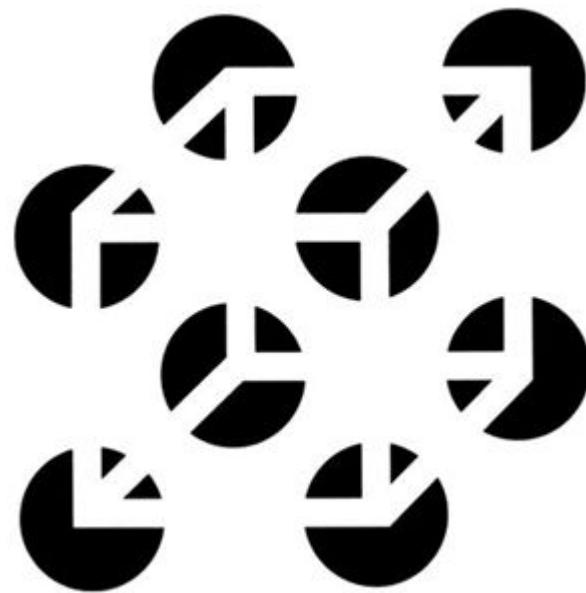
- ¿Cómo se integran las respuestas neuronales individuales en circuitos que implementan cálculos útiles?
- ¿Qué operaciones computacionales extraen información significativa del patrón espacio-temporal de luz en la retina?

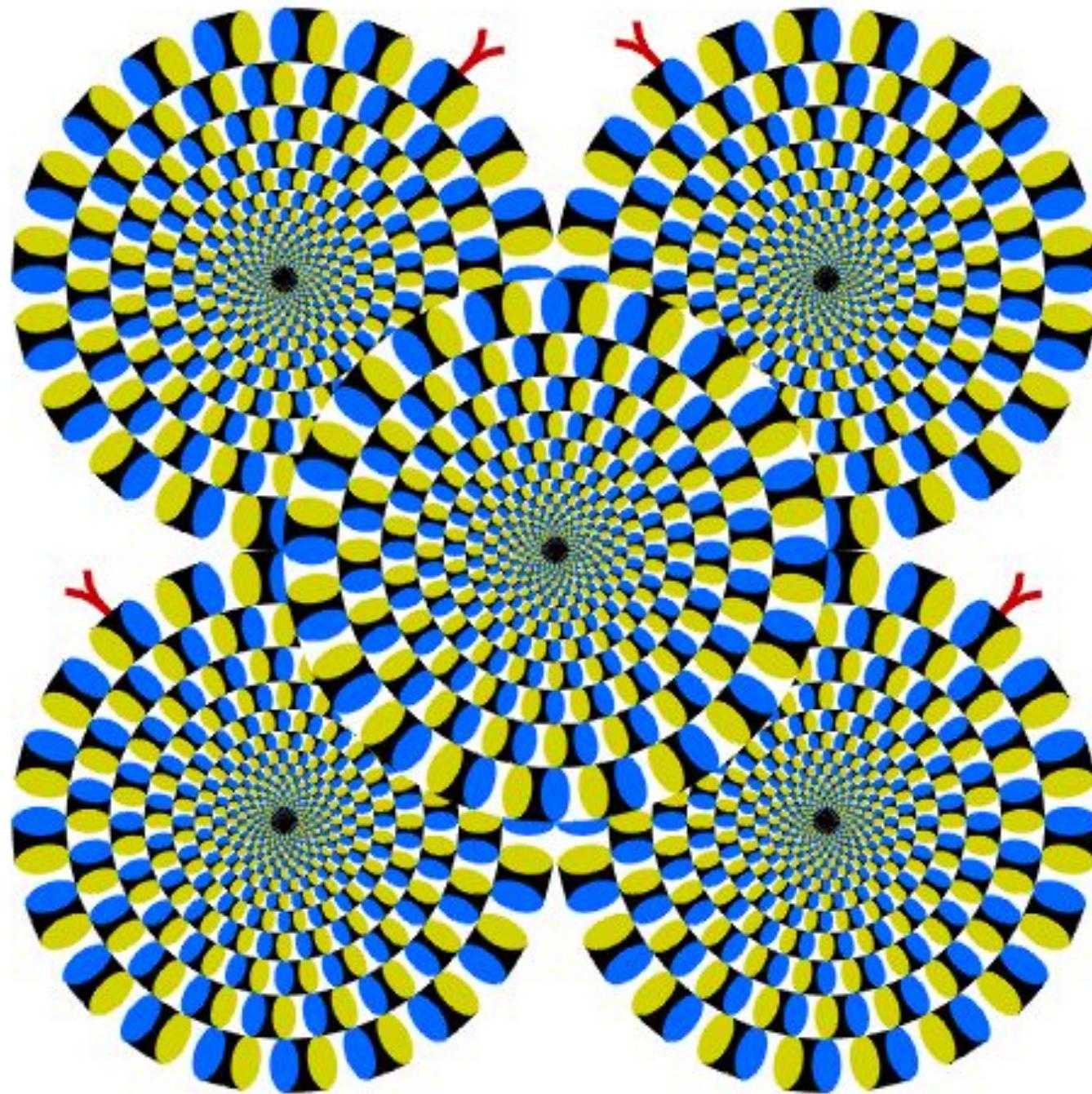
# Grilla de Hermann y bandas de Mach



# Kurt Koffa

- ¿Por qué vemos las cosas como las vemos?
  - Porque el mundo es como es
  - Porque nosotros somos como somos



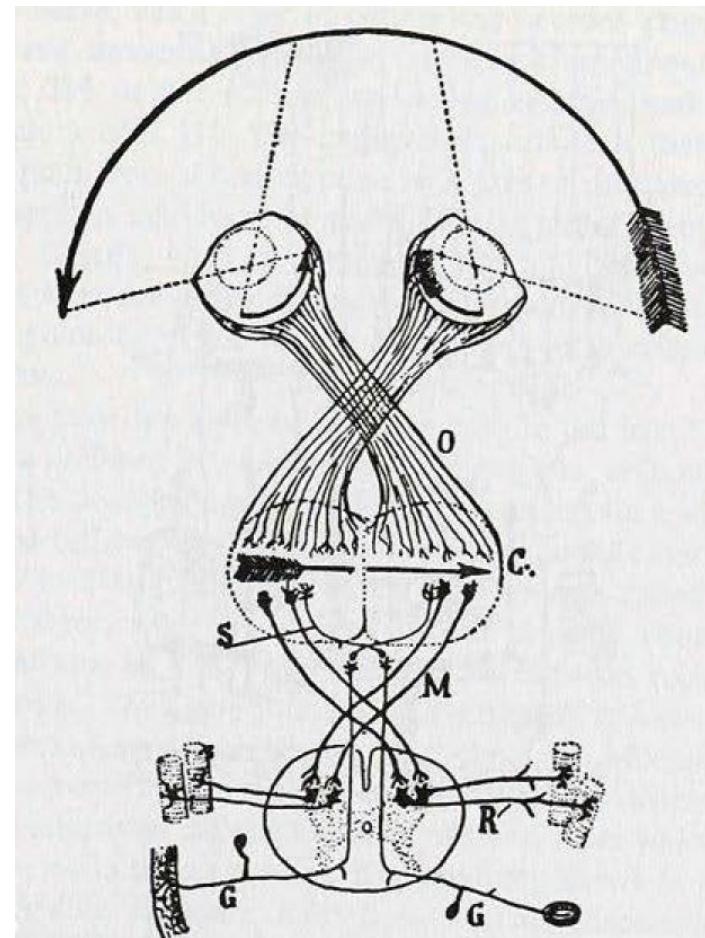


# What the frog's eye tells the frog's brain

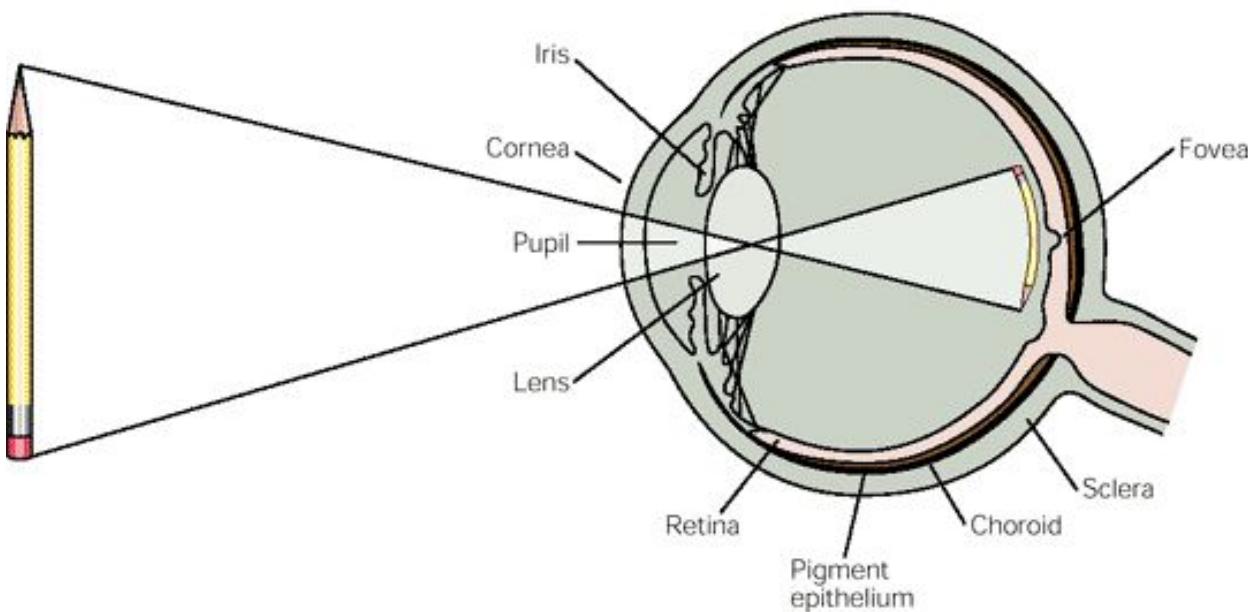
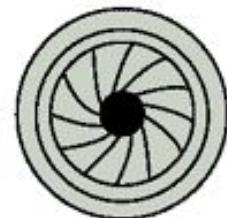
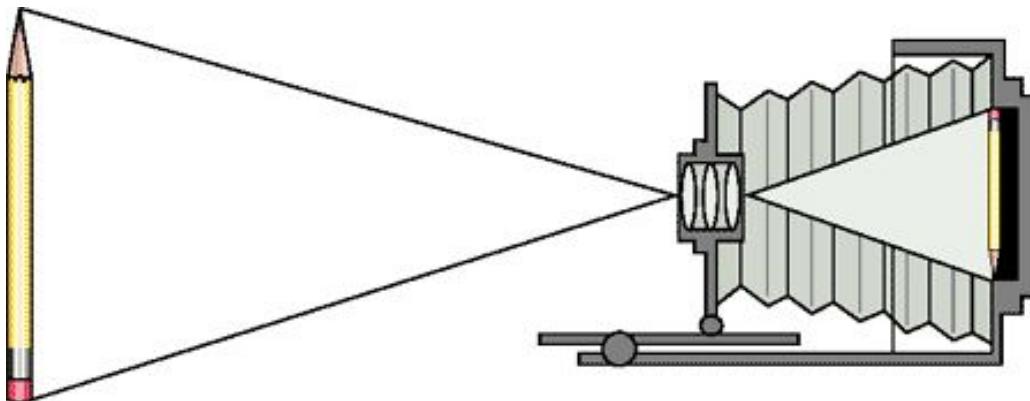
“Hay cuatro tipos de fibras, cada tipo relacionado con un tipo diferente de patrón. Cada tipo se distribuye uniformemente sobre toda la retina de la rana. Por lo tanto, hay cuatro canales distintos distribuidos en paralelo...”

1. Bordes locales y contraste
2. Bordes curvos de objetos oscuros
3. Movimiento de los bordes
4. Oscurecimiento debido al movimiento u oscurecimiento general rápido

Lettvin, Maturana, McCulloch, & Pitts (1959). What the frog's eye tells the frog's brain. In Warrens. McCulloch (Ed.), Embodiments of Mind, pp. 230-255. Cambridge: MIT Press.



# Se parece el ojo a una cámara fotográfica?



# Los niveles de Marr

- Nivel computacional
- Nivel algorítmico
- Nivel de implementación

*"...trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers. It cannot be done" - David Marr (1982/2010, p. 27)*



# VISION

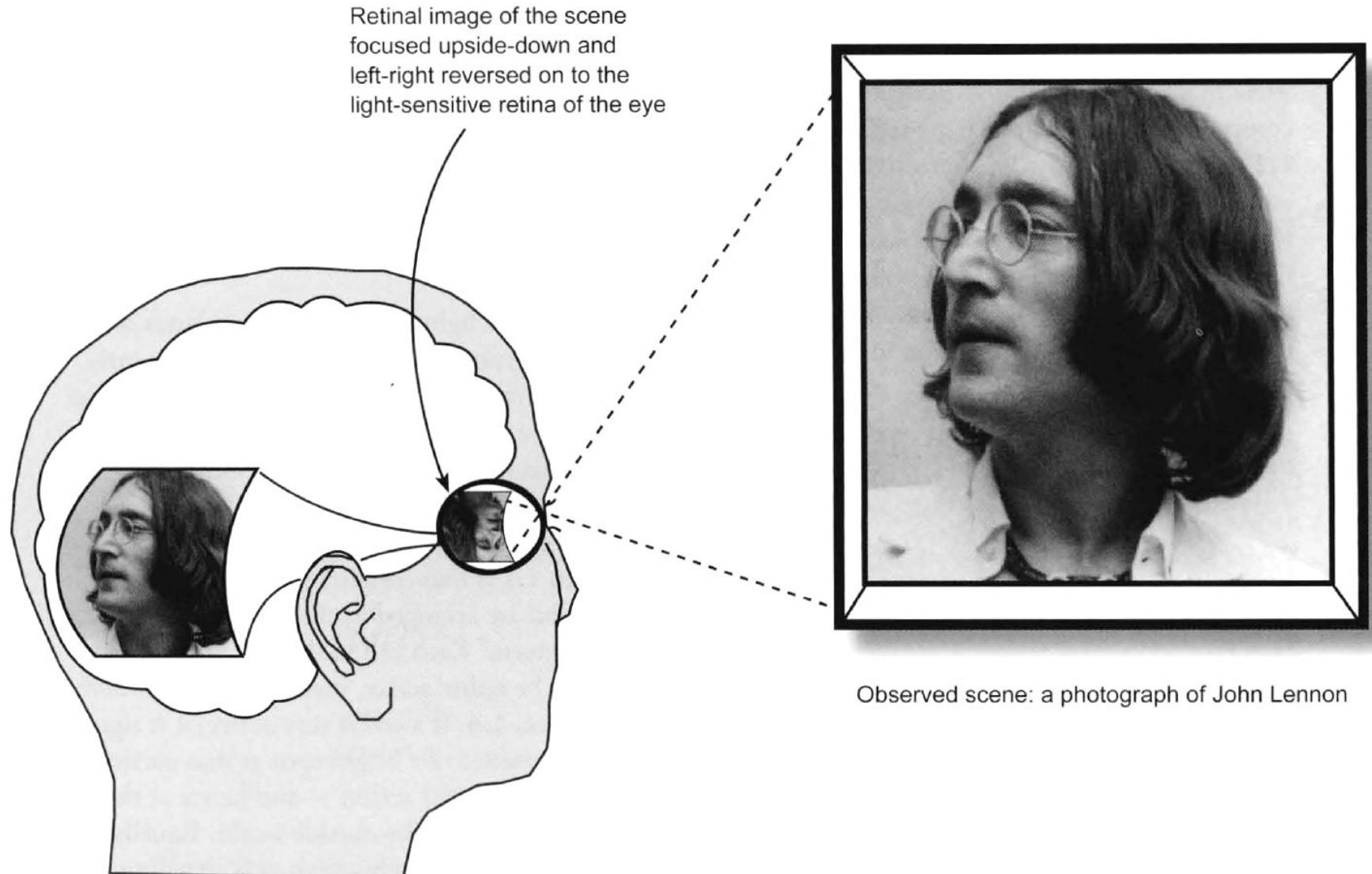


David Marr

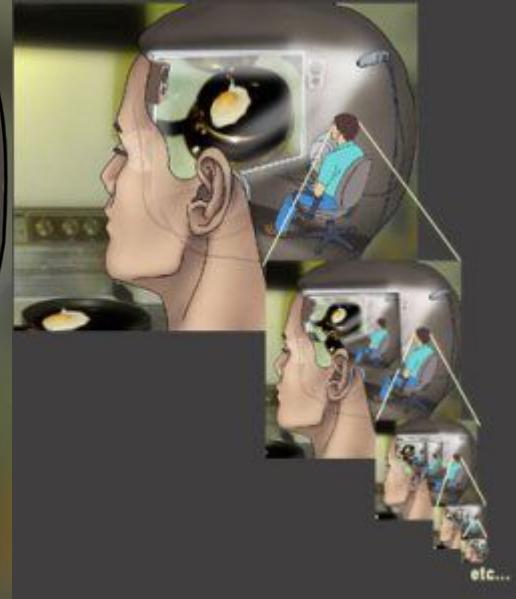
FOREWORD BY  
Shimon Ullman  
AFTERWORD BY  
Tomaso Poggio



# The "inner screen" theory



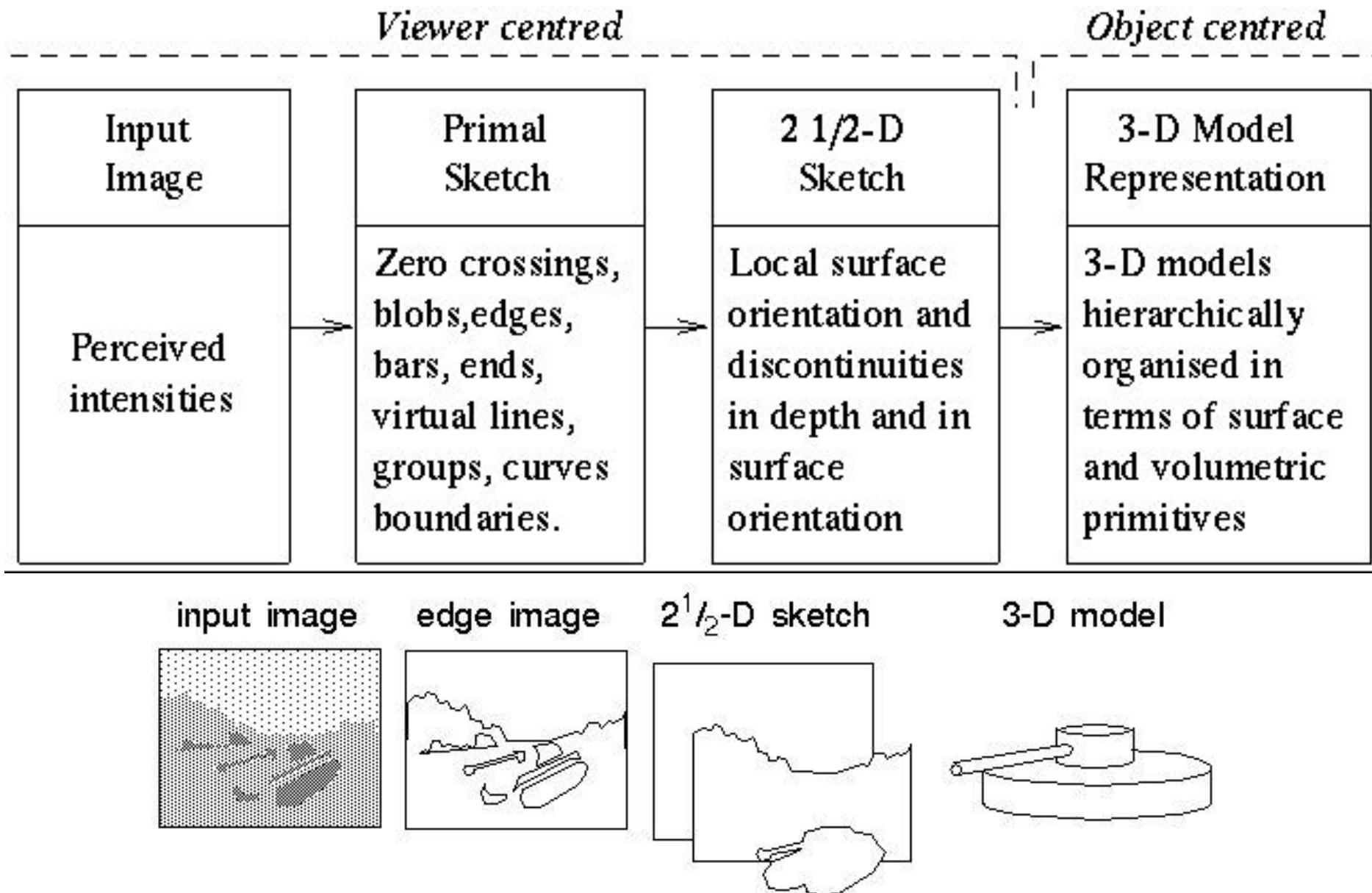
# El problema del Homúnculo o de la regresión infinita:



# El problema:



# Modelo de Marr



# Horace Barlow



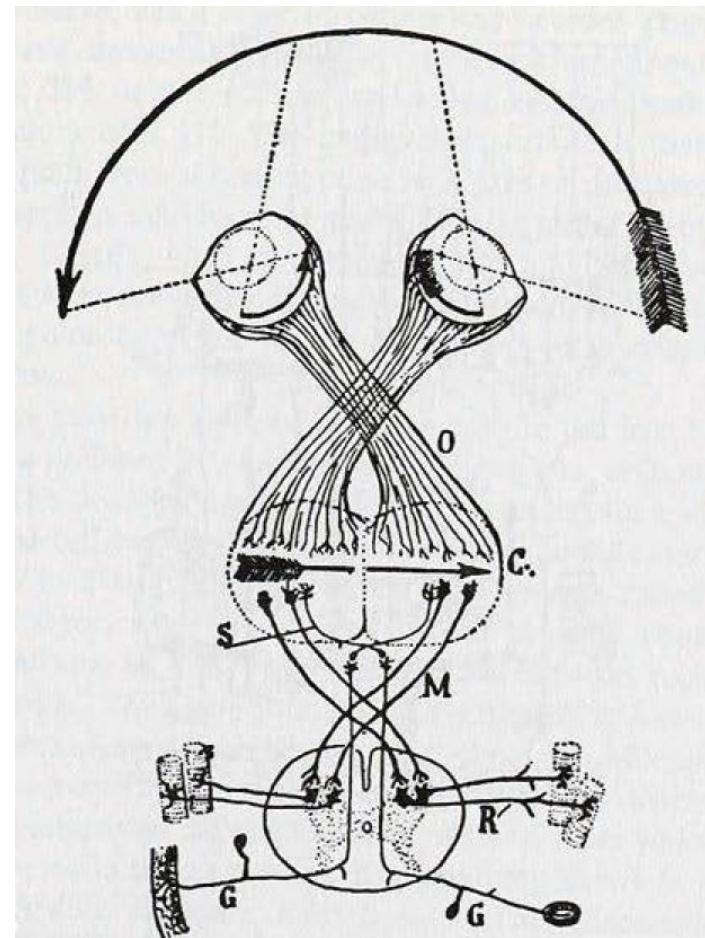
- In 1953 Barlow discovered that the frog brain has neurons which fire in response to specific visual stimuli.
- This was a precursor to the work of Hubel and Wiesel on visual receptive fields in the visual cortex.

# What the frog's eye tells the frog's brain

“Hay cuatro tipos de fibras, cada tipo relacionado con un tipo diferente de patrón. Cada tipo se distribuye uniformemente sobre toda la retina de la rana. Por lo tanto, hay cuatro canales distintos distribuidos en paralelo...”

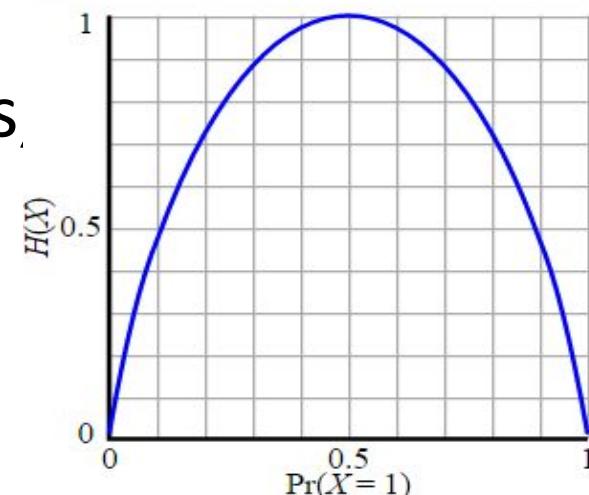
1. Bordes locales y contraste
2. Bordes curvos de objetos oscuros
3. Movimiento de los bordes
4. Oscurecimiento debido al movimiento u oscurecimiento general rápido

Lettvin, Maturana, McCulloch, & Pitts (1959). What the frog's eye tells the frog's brain. In Warrens. McCulloch (Ed.), Embodiments of Mind, pp. 230-255. Cambridge: MIT Press.



# Principios de Barlow

1. Los relevos sensoriales son para detectar, en los mensajes entrantes, ciertas "contraseñas" que tienen un significado particular para el animal.
2. Son filtros, o centros de recodificación, cuyas "características de paso" pueden controlarse de acuerdo con los requisitos de otras partes del sistema nervioso.
3. Recodifican los mensajes sensoriales extrayendo señales de alta entropía relativa de la entrada sensorial altamente redundante.



# Seeing

The Computational Approach  
to Biological Vision

second edition

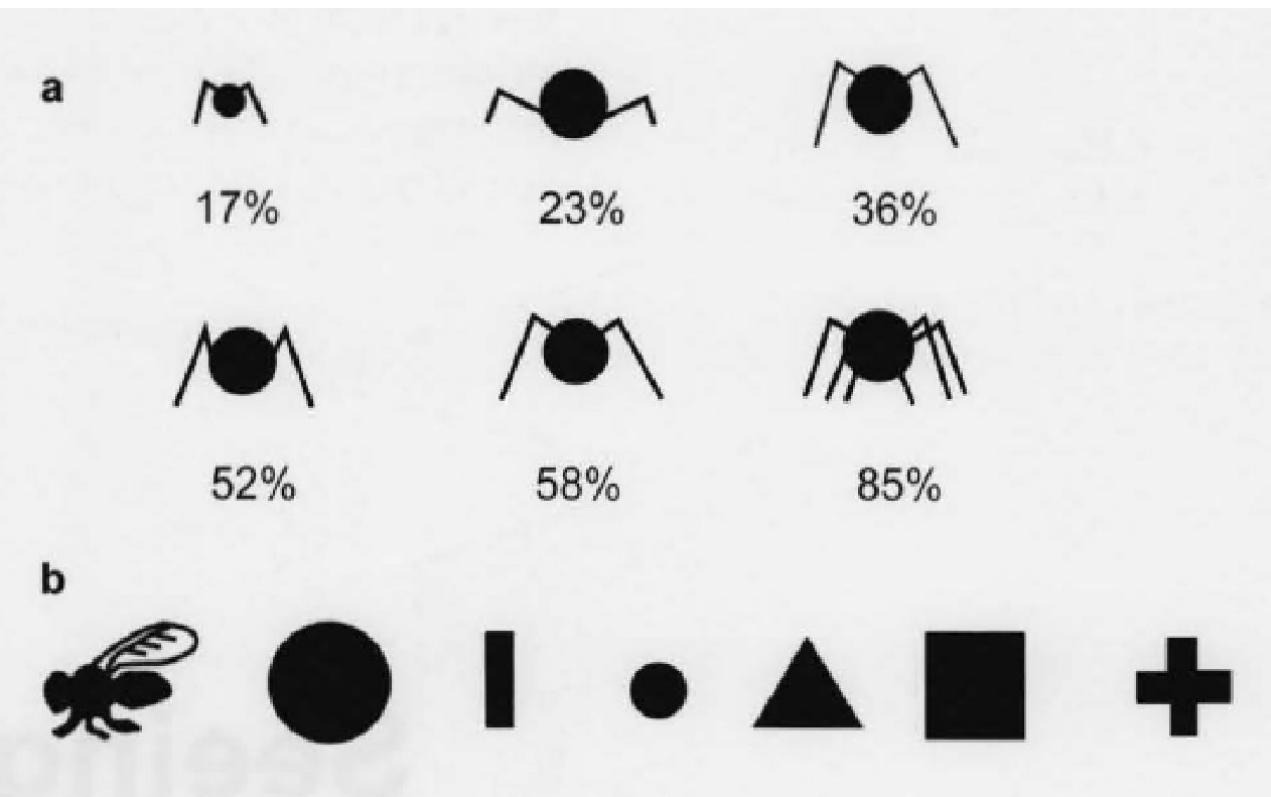


John P. Frisby and James V. Stone

# Araña saltadora



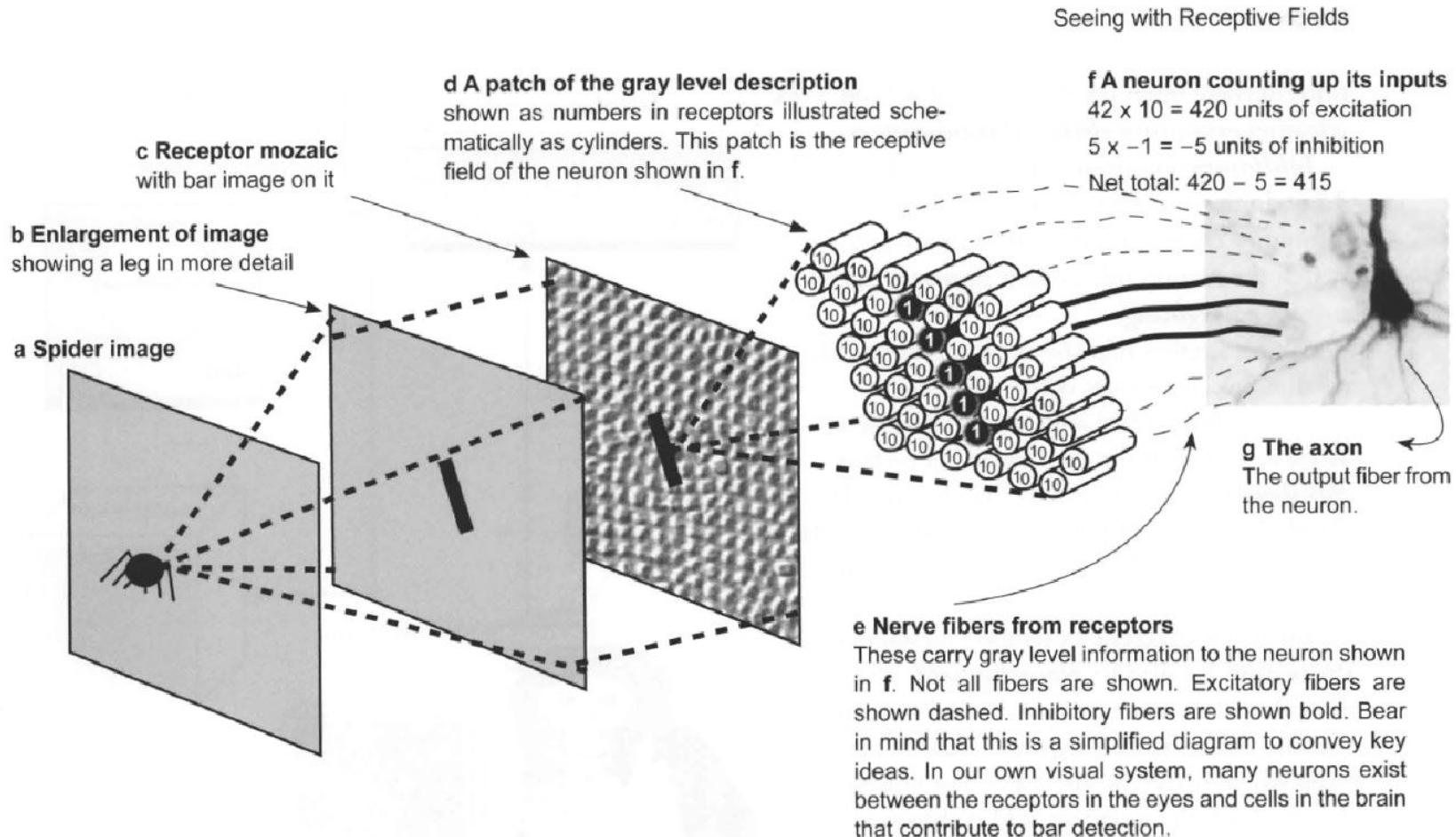
Michael Land (1969) observó que las arañas saltadoras atacan cualquier objeto pequeño y oscuro que se mueva, **a menos** que el objeto tenga **dos líneas oblicuas** a los lados de un punto central, que parecen “**piernas**” y “**cuerpo**”. En ese caso, lo reconocen como **otra araña saltadora** y el macho responde con una **conducta de cortejo** (Land y Nilsson, 2001).



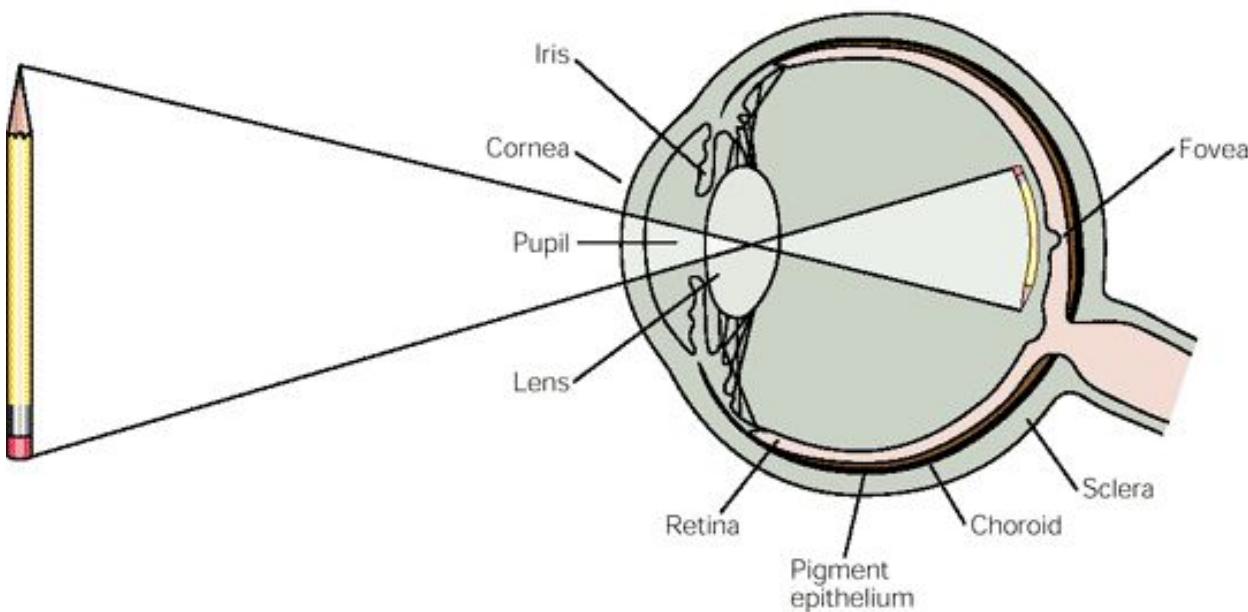
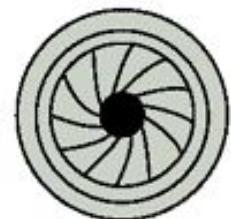
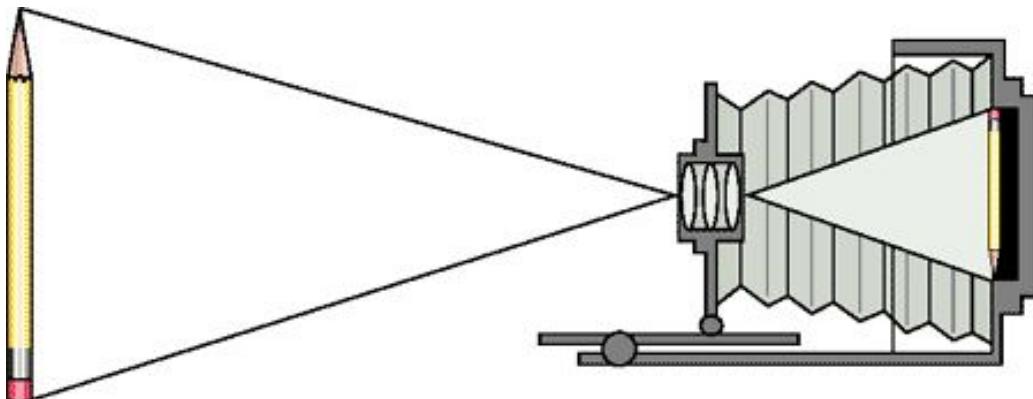
# “bar” detector



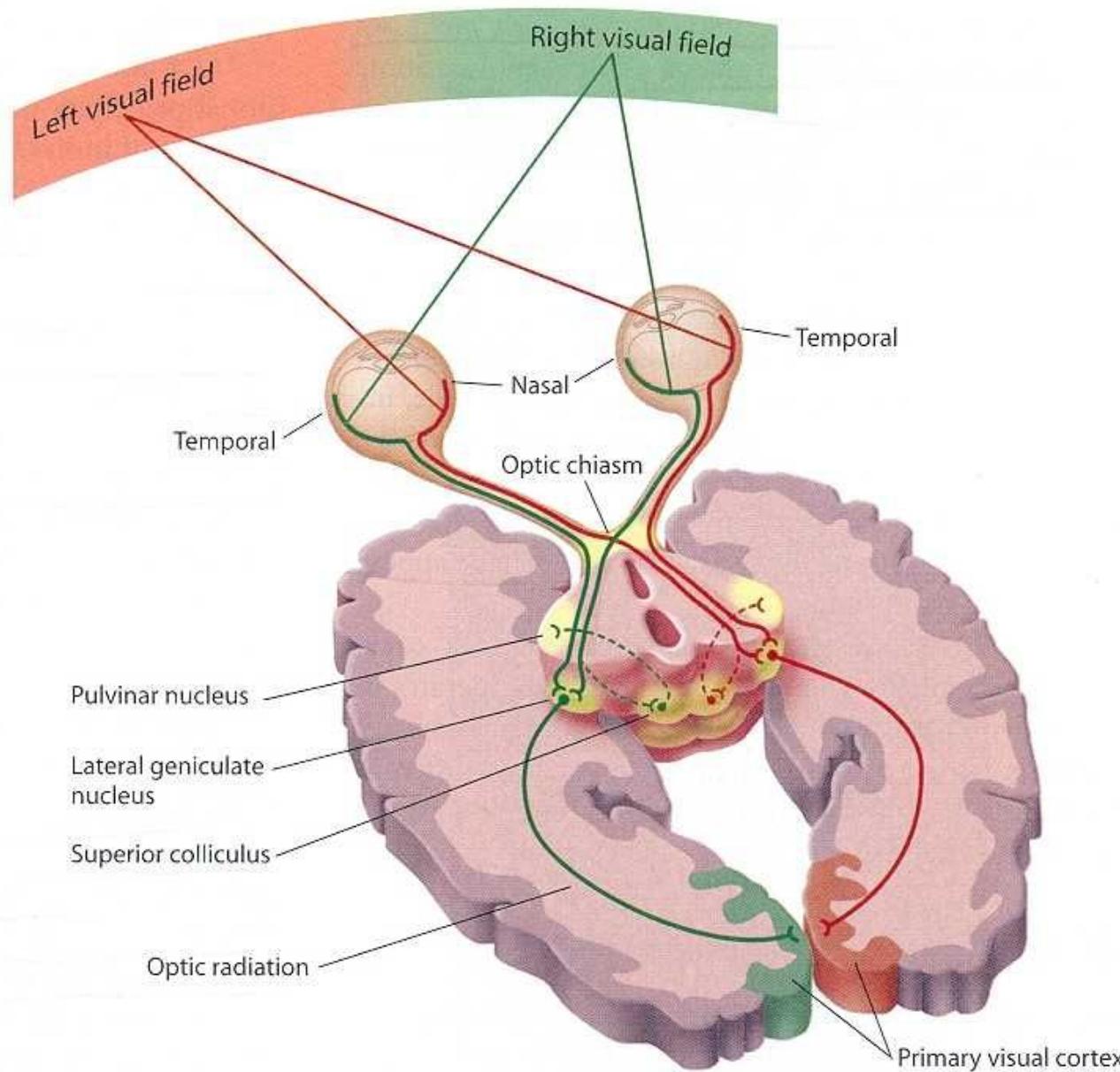
# Template matching



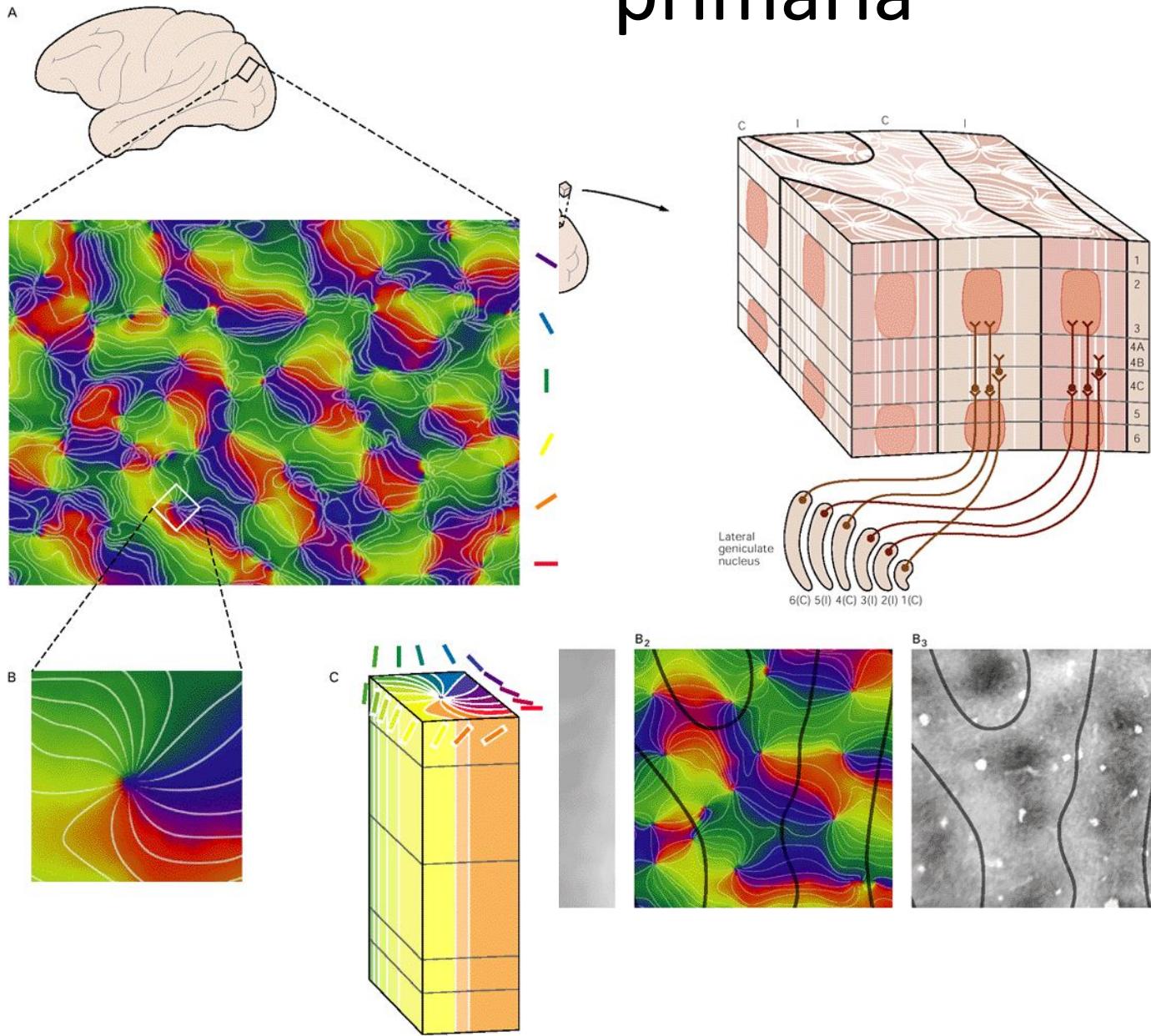
# Se parece el ojo a una cámara fotográfica?



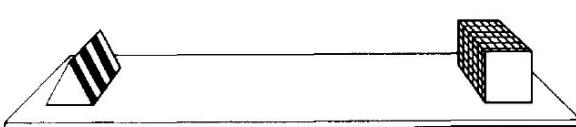
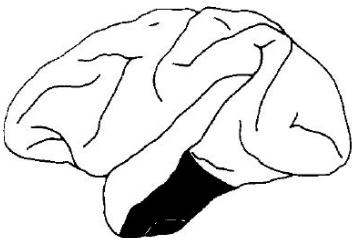
# Vías visuales



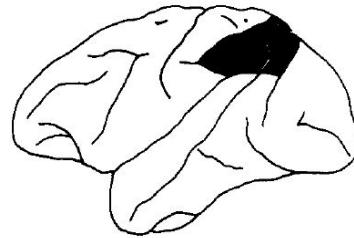
# Organización de la corteza visual primaria



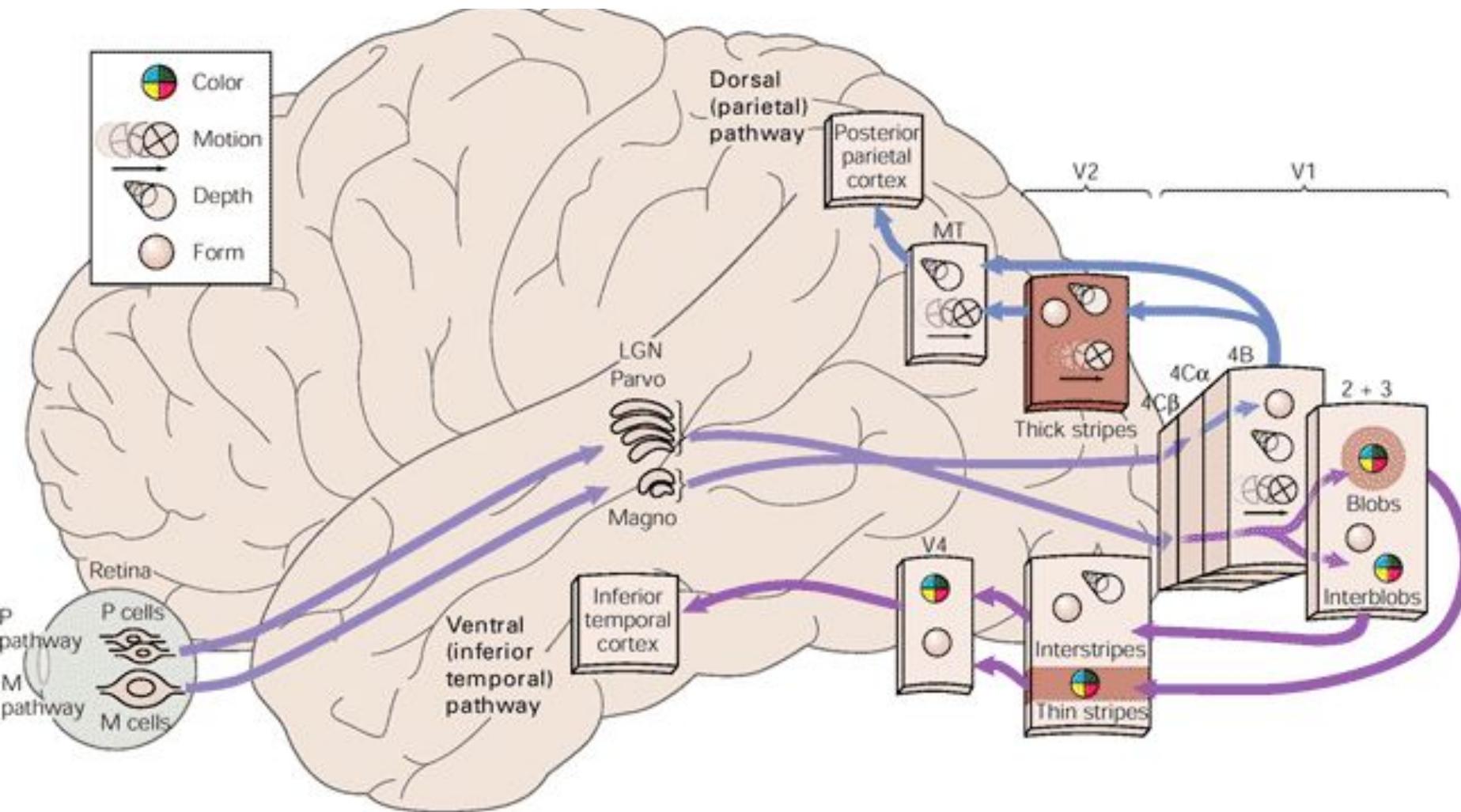
A

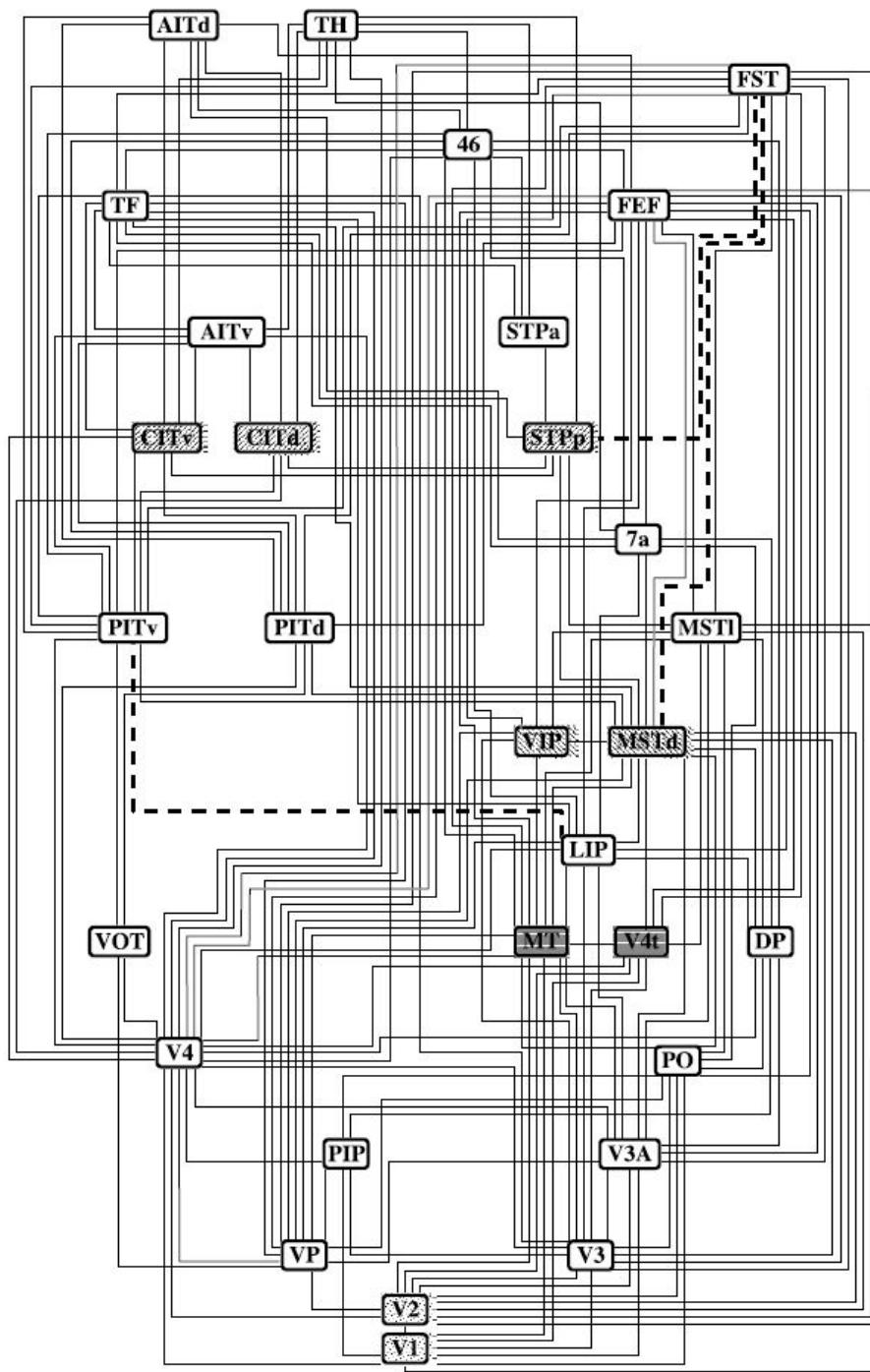


B

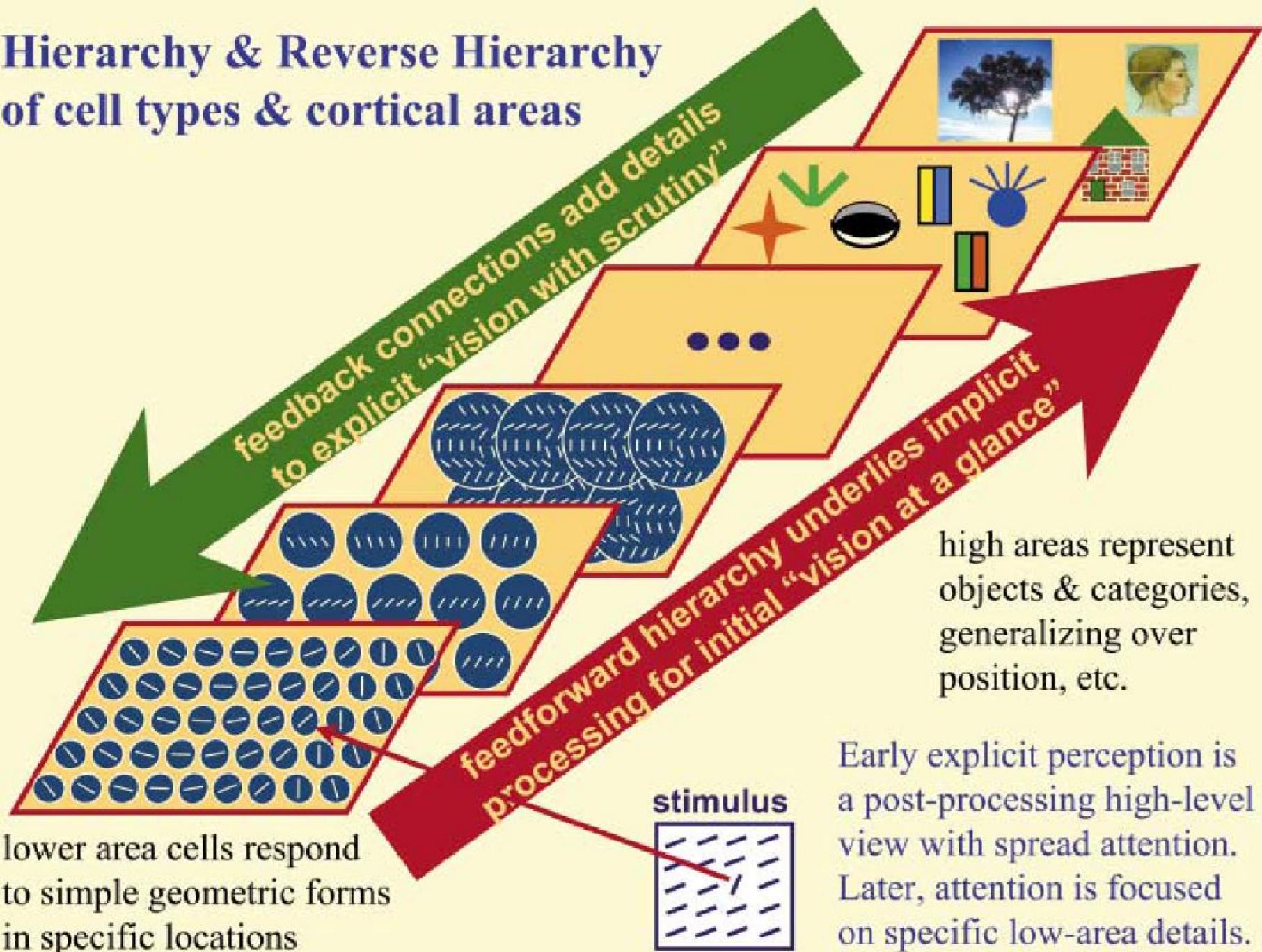


**Mortimer Mishkin,  
Leslie G. Ungerleider**





# Hierarchy & Reverse Hierarchy of cell types & cortical areas



# Campo receptivo de una neurona

El campo receptivo se define como la porción del campo visual en la cual la actividad de disparo de la neurona es modulada mediante el estímulo visual.

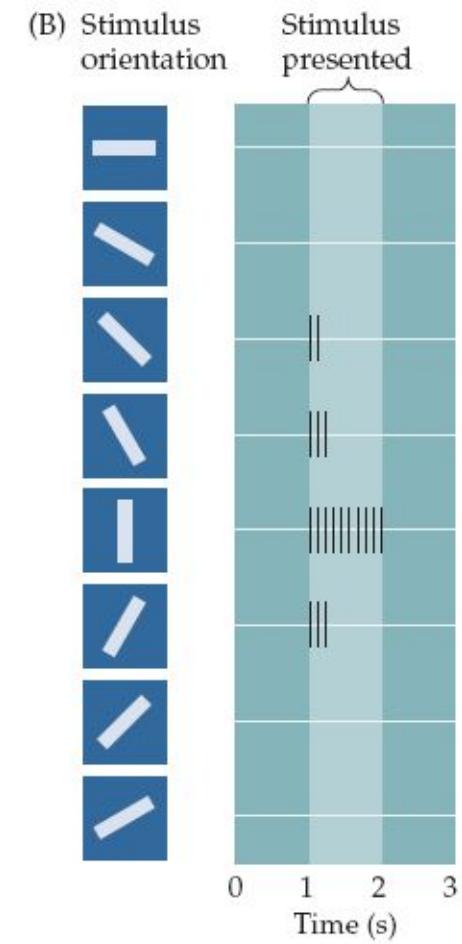
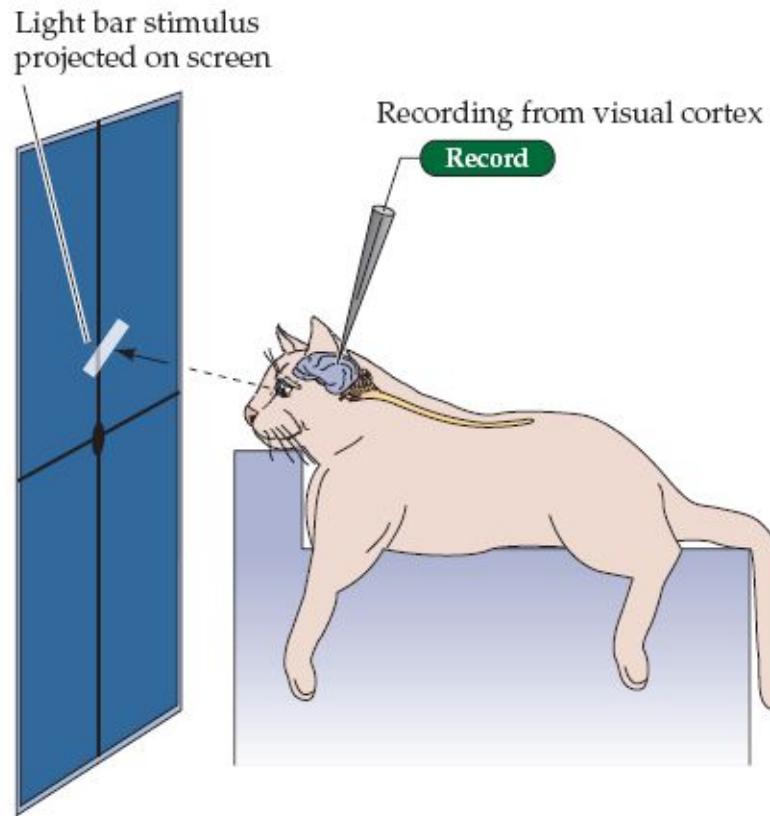
# David Hubel y Torsten Wiesel





## David Hubel y Torsten Wiesel

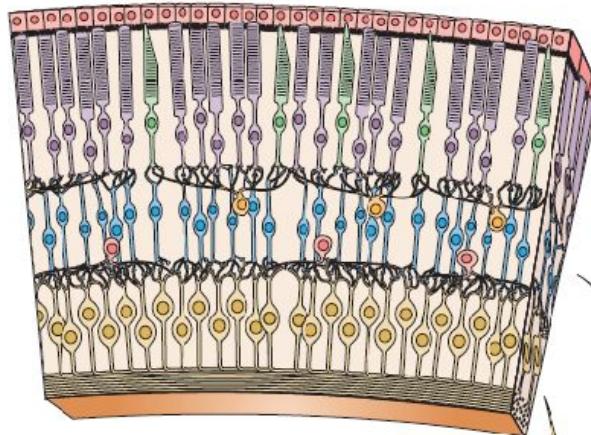
(A) Experimental setup



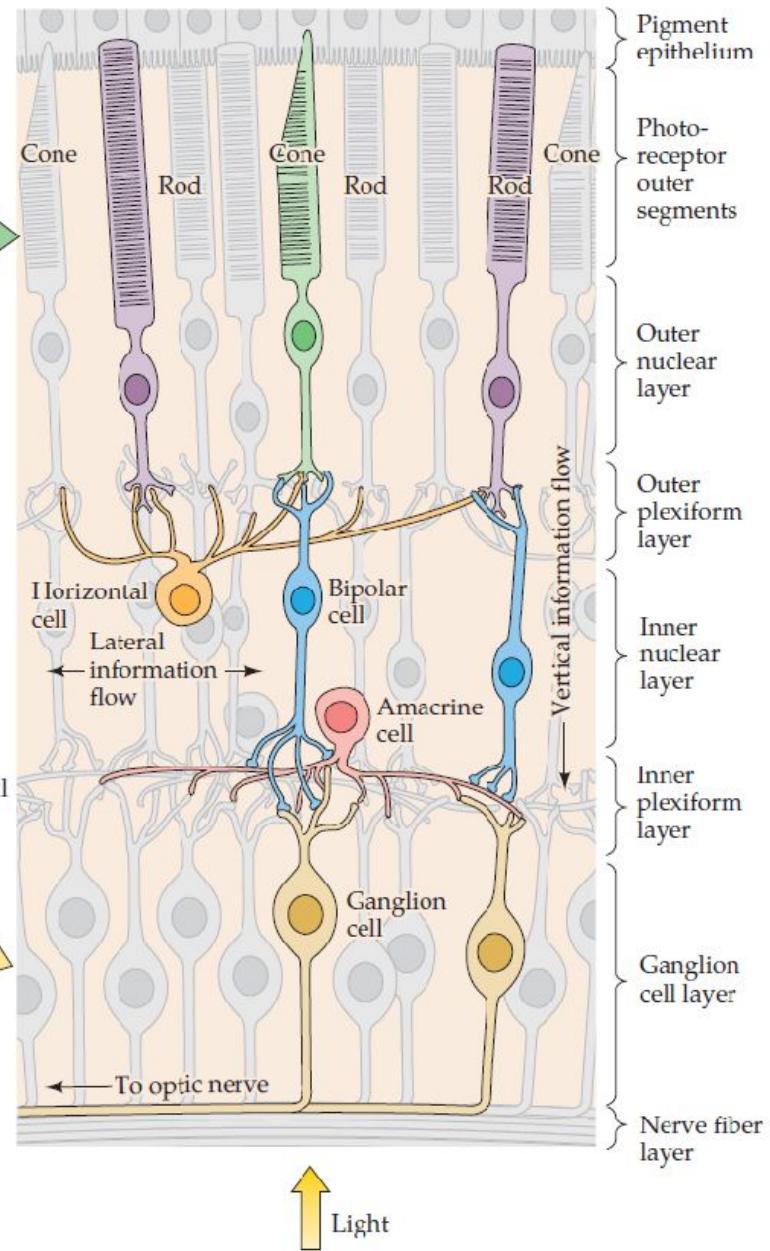
# Los experimentos de Hubel y Wiesel



(A) Section of retina

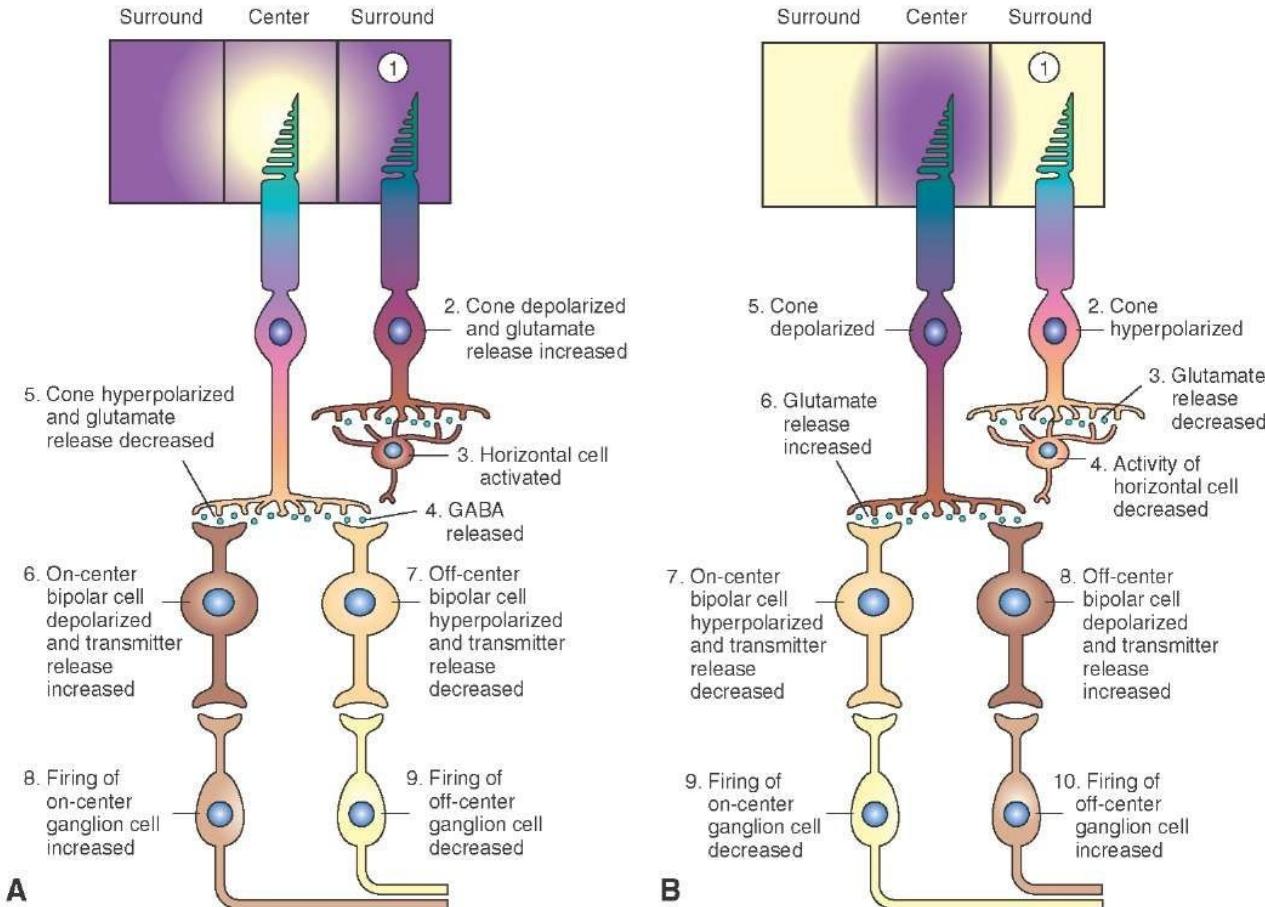


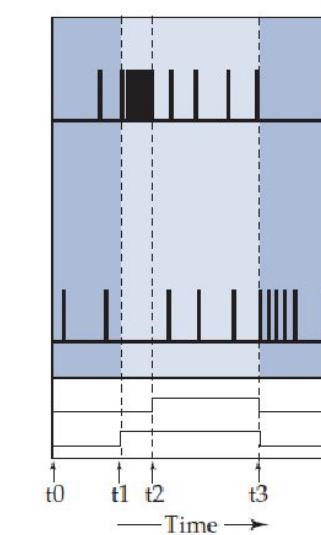
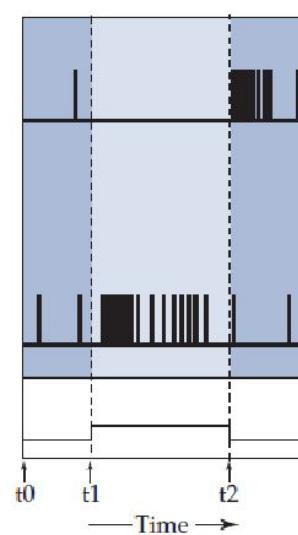
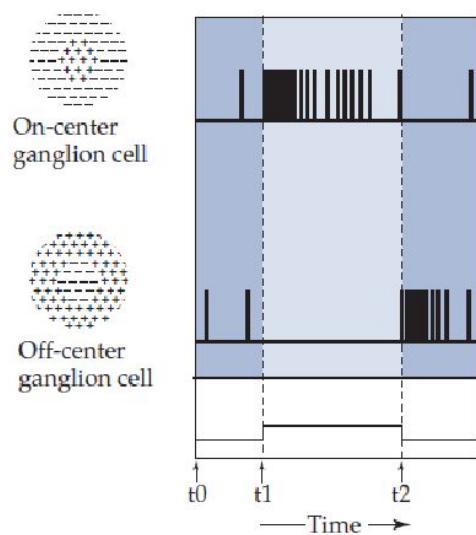
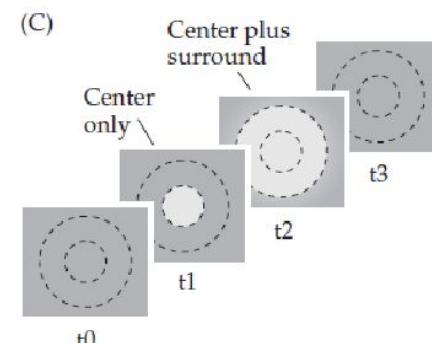
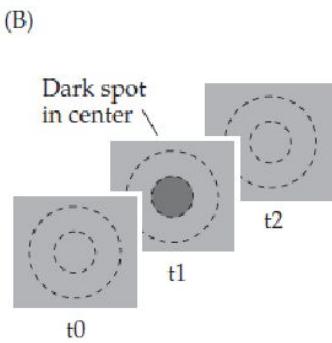
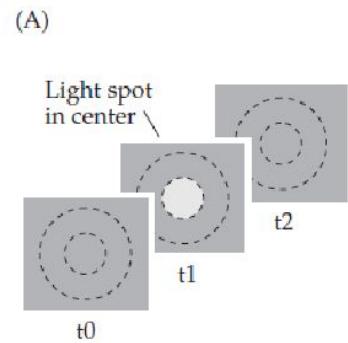
(B)



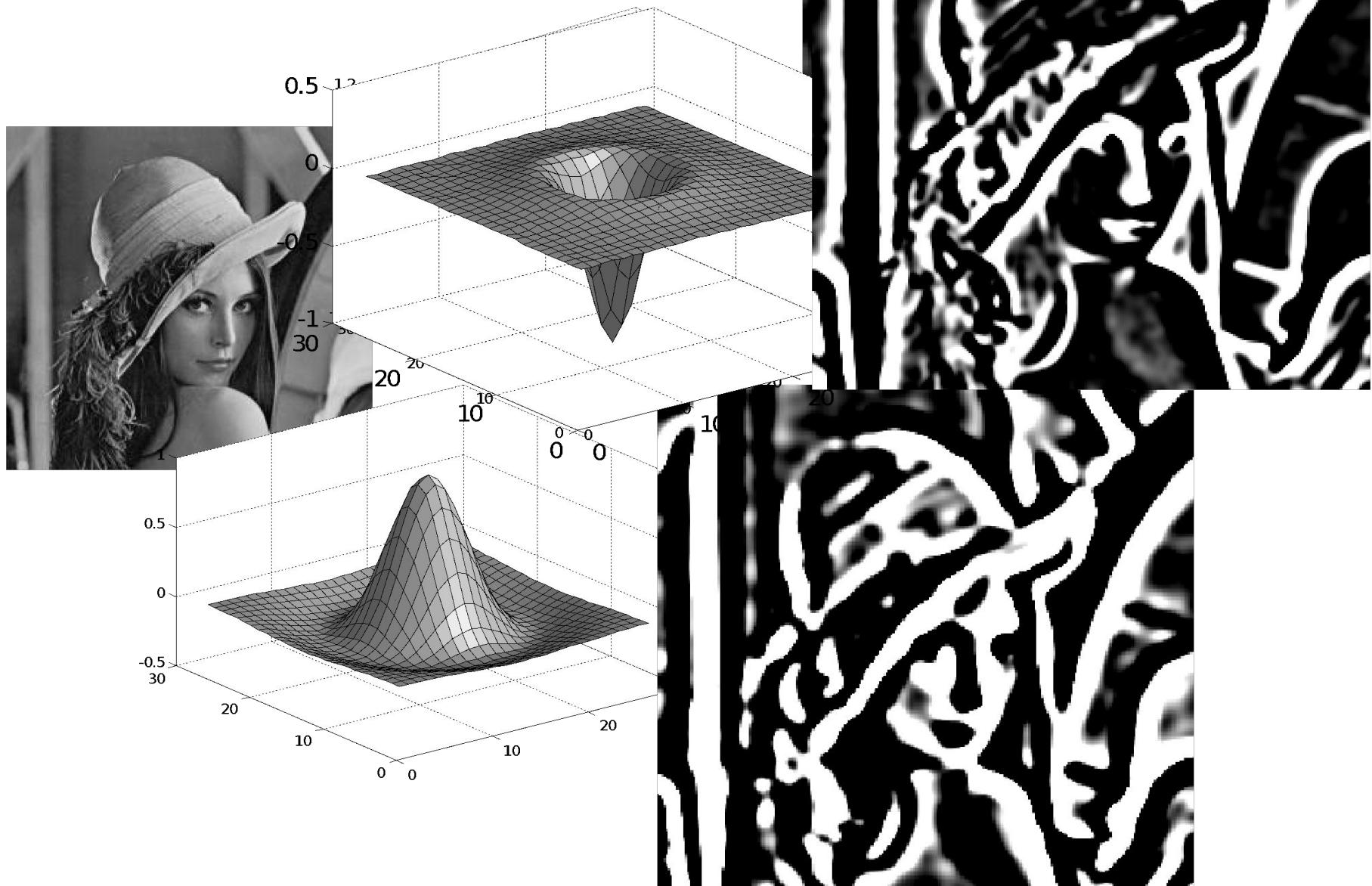
**Figure 10.4** Structure of the retina. (A) Section of the retina showing overall arrangement of retinal layers. (B) Diagram of the basic circuitry of the retina. A three-neuron chain—photoreceptor, bipolar cell, and ganglion cell—provides the most direct route for transmitting visual information to the brain. Horizontal cells and amacrine cells mediate lateral interactions in the outer and inner plexiform layers, respectively. The terms *inner* and *outer* designate relative distances from the center of the eye (inner, near the center of the eye; outer, away from the center, or toward the pigment epithelium).

# Construcción de los campos receptivos de las células ganglionares de la retina



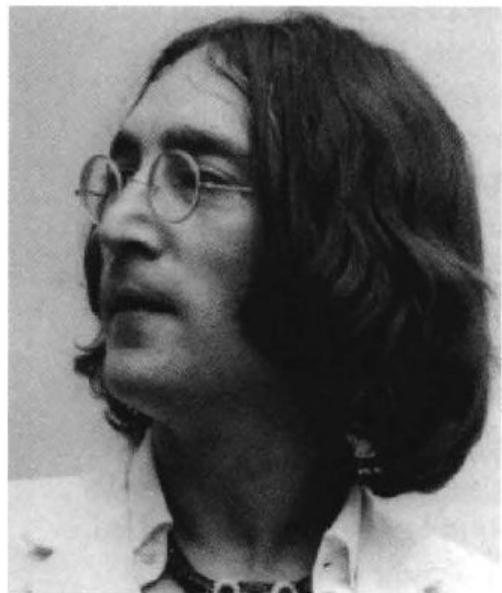


# Retina y geniculado



# Teoría computacional para la tarea de medir gradientes de imagen

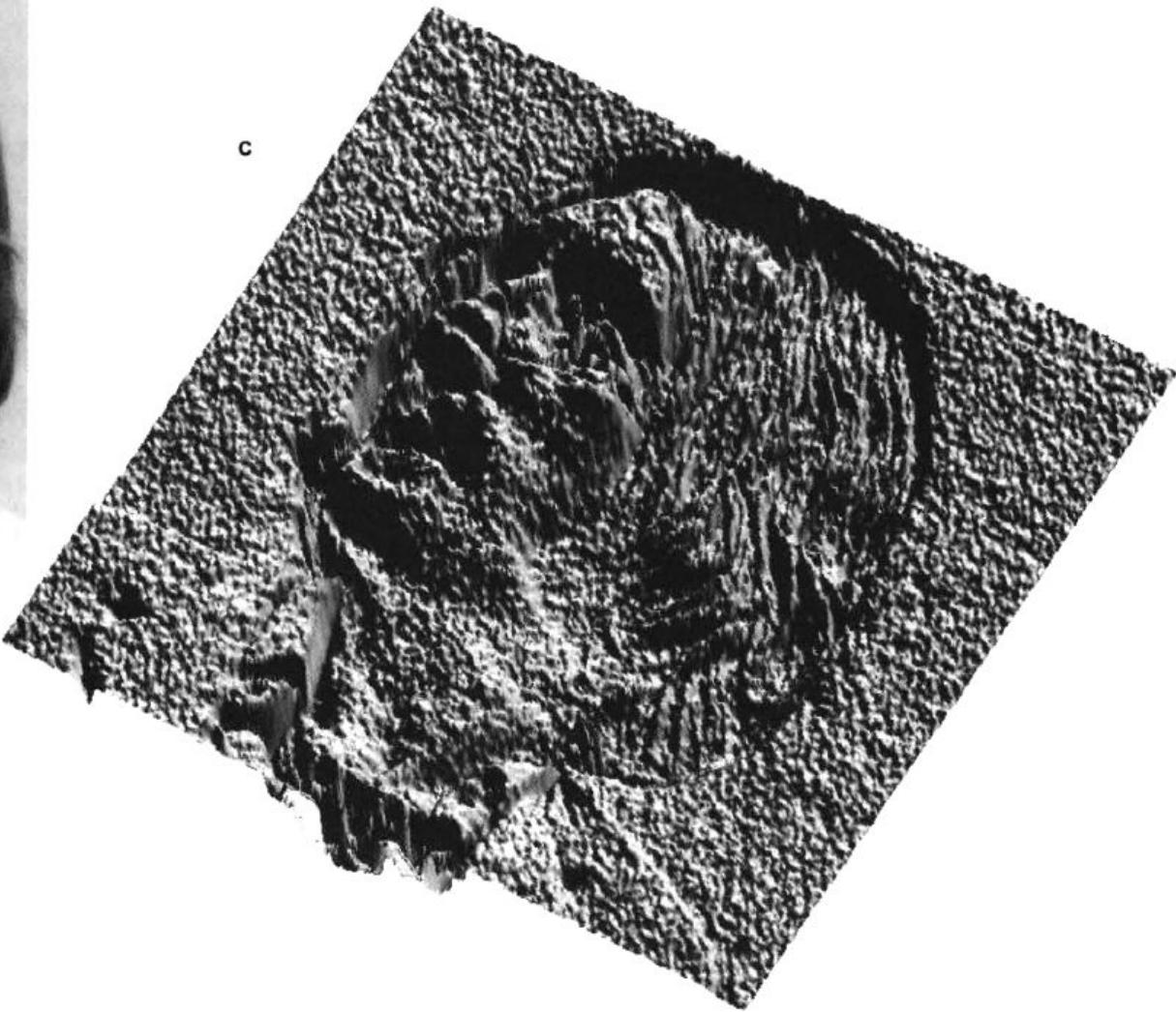
a



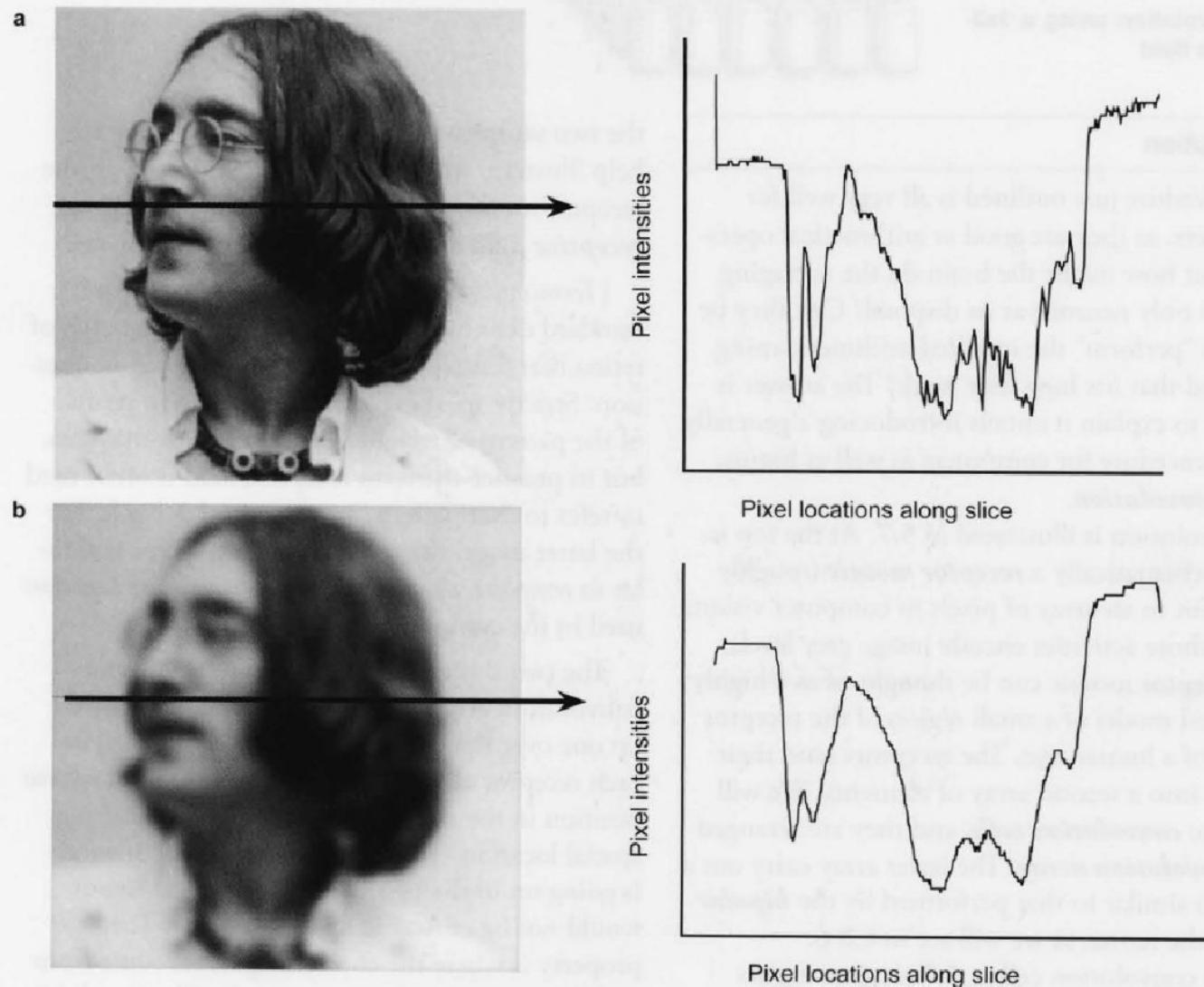
b

159	139	116	93	63
167	146	120	95	101
167	127	110	116	94
136	117	143	129	93
118	138	159	141	107

c



# Gradientes en una imagen



# Suavizado

a



b



c

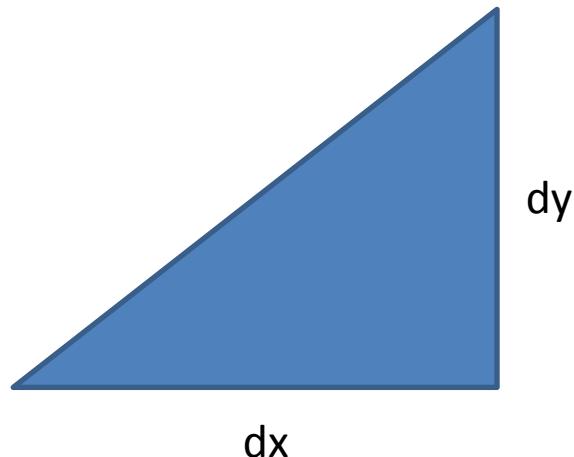


d



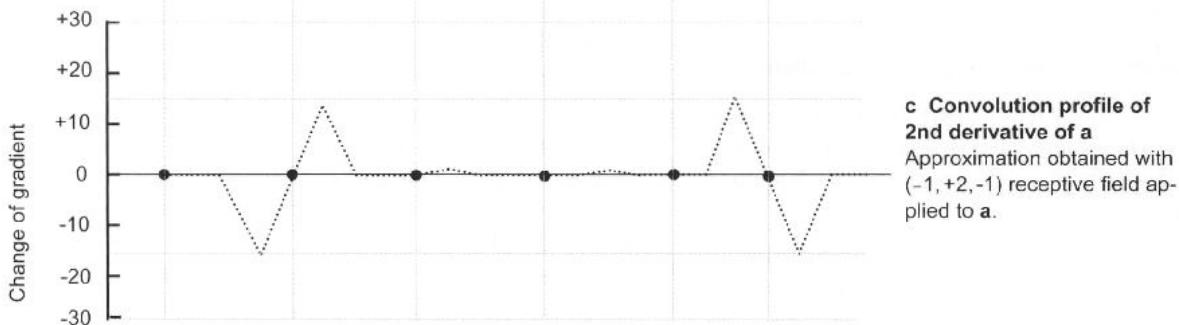
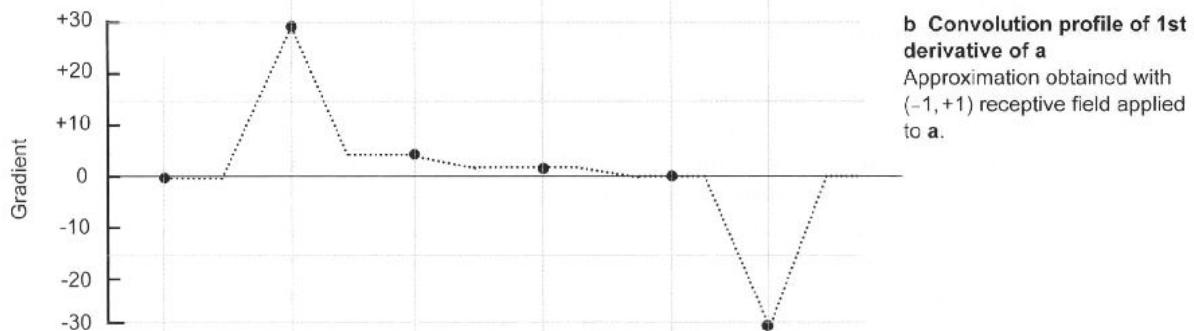
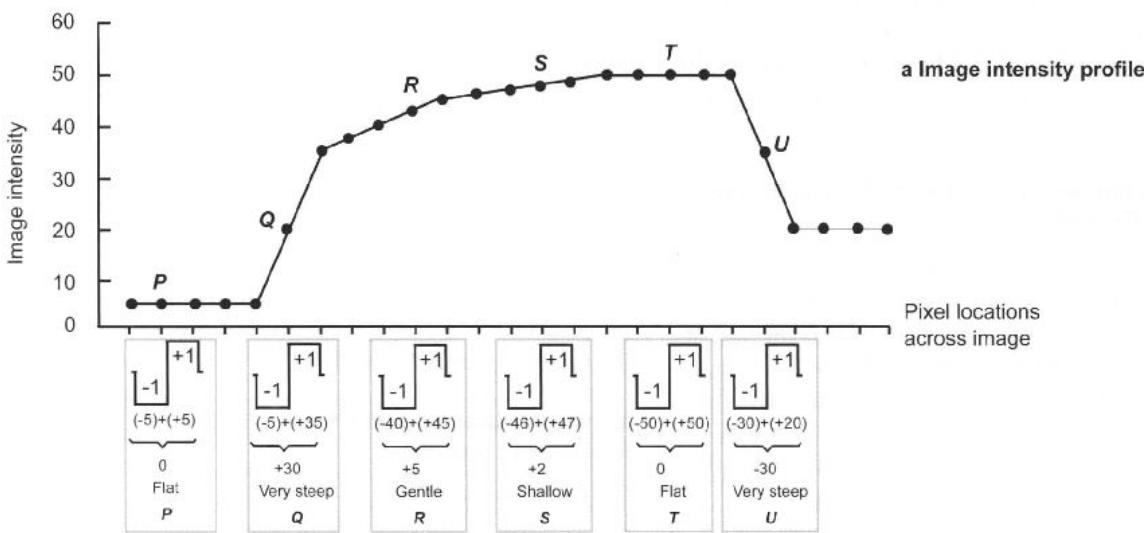
# Algoritmo para medir gradientes de imagen

- Paso 1 Mida la diferencia de intensidad (altura) entre los puntos de la imagen.
- Paso 2 Divida esta diferencia por la distancia entre los puntos, para obtener el gradiente.



# Procedimiento equivalente

- Paso 1 Para cada par de píxeles vecinos, pondera las entradas de una como positiva y la otra como negativa, multiplicando por + 1 y -1 respectivamente.
- Paso 2 Sumar las entradas ponderadas.



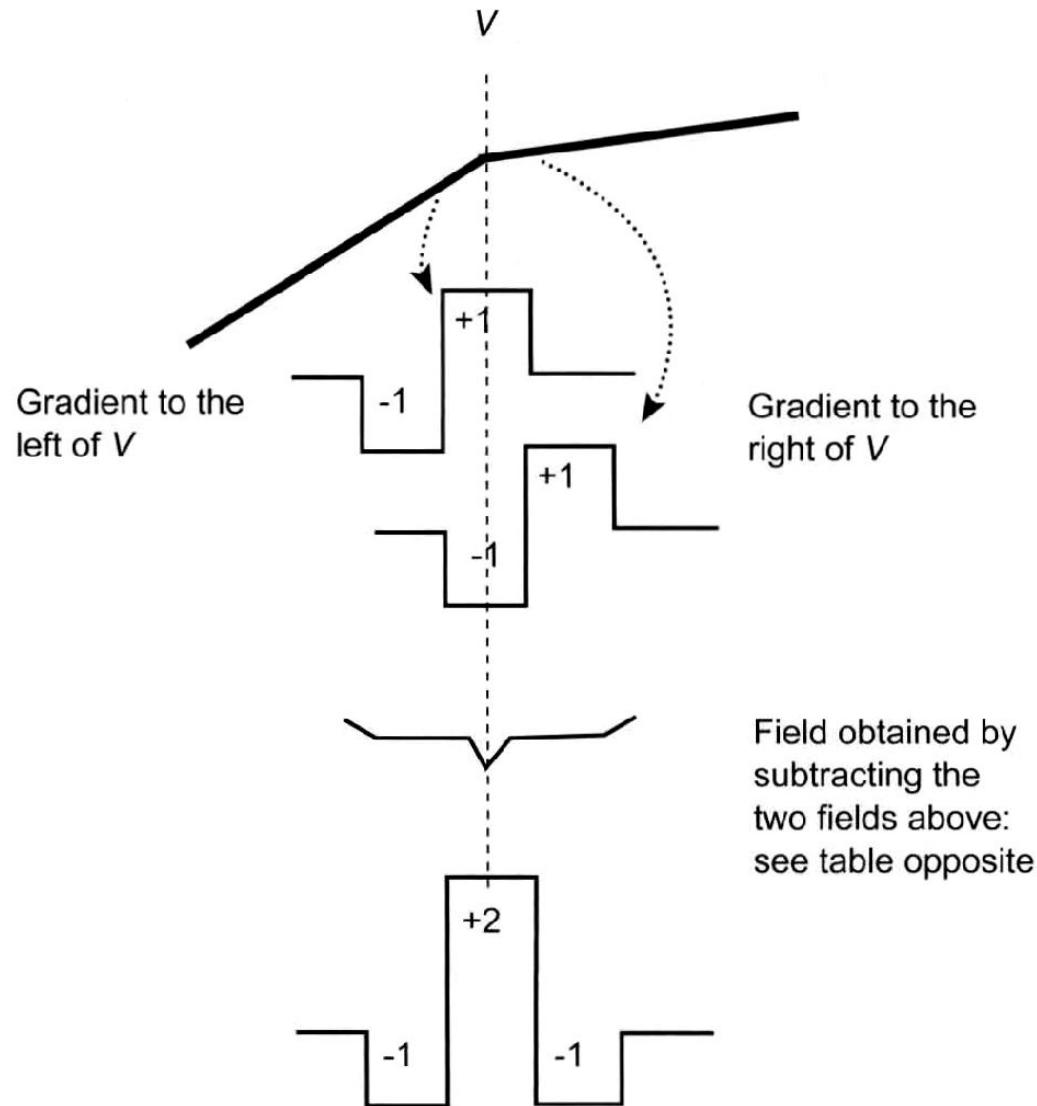
### 5.11 Measuring gradients and changes of gradients using receptive fields

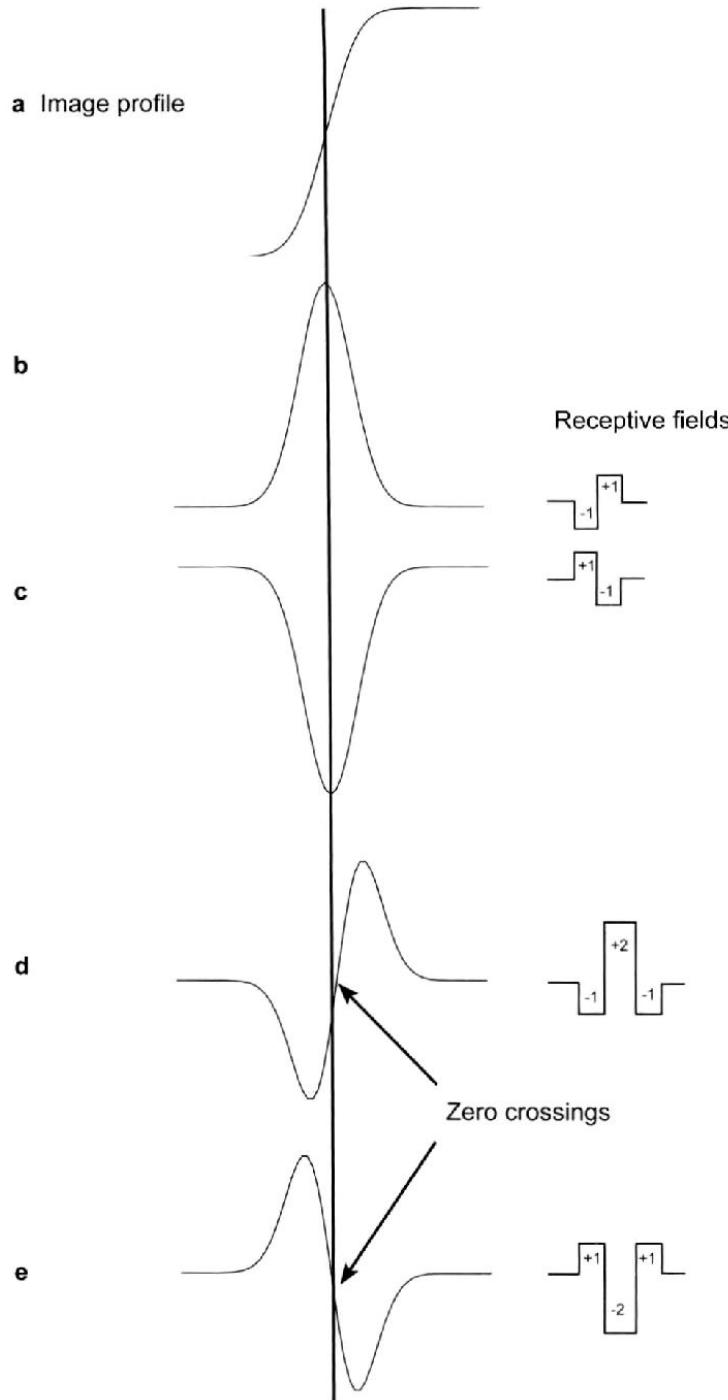
The units of intensity plotted in a are arbitrary but chosen to convey the main ideas.

# Procedimiento de un solo paso para medir los cambios en los gradientes

- Paso 1 Medir gradientes utilizando campos receptivos del tipo  $(-1, +1)$  o  $(+1, -1)$ .
- Paso 2 Medir gradientes en la salida del paso 1, de nuevo usando campos de  $(-1, +1)$  o  $(+1, -1)$ .

# Medir los cambios en los gradientes





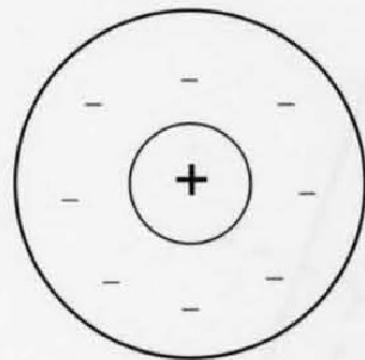
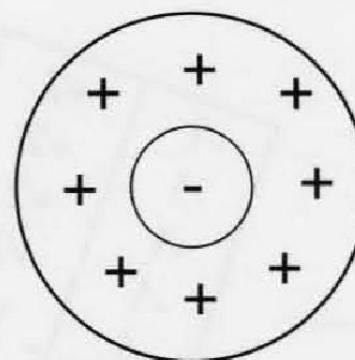
- Las células centro-ON y centro-OFF actúan como filtros complementarios que responden a incrementos o decrementos locales de luz.
- Cuando sus respuestas se comparan en niveles posteriores, los puntos donde ambas se equilibran —los *zero-crossings*— marcan con precisión los bordes del estímulo.
- Este mecanismo permite transformar gradientes de luminancia en una representación discreta de contornos.
- Así, el sistema visual refuerza la detección de límites y formas a partir de señales redundantes de la retina.

**a**

-1	-1	-1
-1	+8	-1
-1	-1	-1

**b**

+1	+1	+1
+1	-8	+1
+1	+1	+1

**c****d**

a

-1	+1
-1	+1
-1	+1
-1	+1
-1	+1
-1	+1
-1	+1

-

-1

+1

-1	+1
-1	+1
-1	+1
-1	+1
-1	+1
-1	+1
-1	+1

-

-

b

-1	+2	-1
-1	+2	-1
-1	+2	-1
-1	+2	-1
-1	+2	-1
-1	+2	-1
-1	+2	-1

-

-1

+2

-1

-1

-1	+2	-1	-1
-1	+2	-1	-1
-1	+2	-1	-1
-1	+2	-1	-1
-1	+2	-1	-1
-1	+2	-1	-1
-1	+2	-1	-1

-

-

# Convolución

0	0	0	0	0	0	0
0	105	102	100	97	96	100
0	103	99	103	101	102	104
0	101	98	104	102	100	104
0	99	101	106	104	99	104
0	104	104	104	100	98	104

Image Matrix

Kernel Matrix

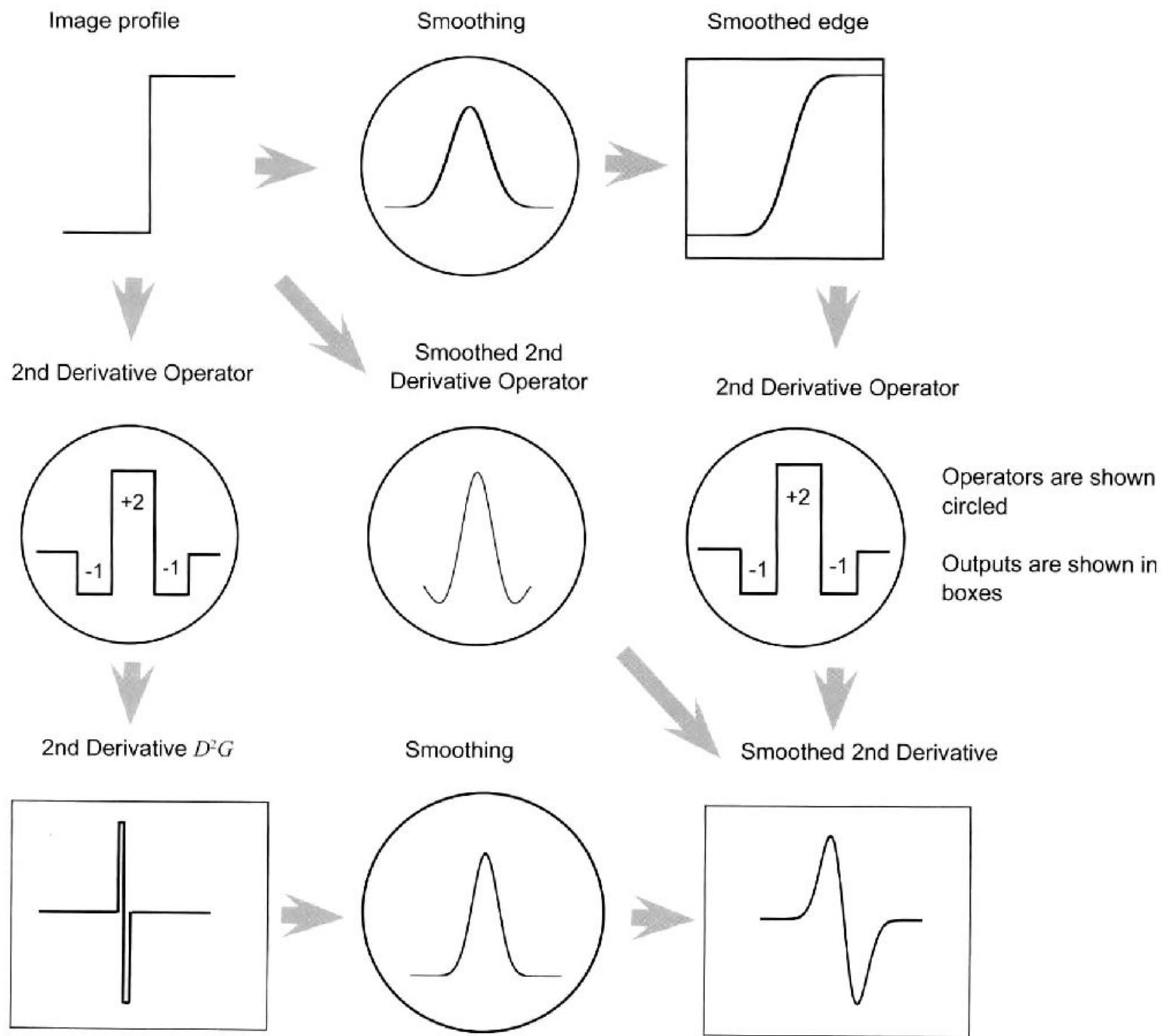
0	-1	0
-1	5	-1
0	-1	0

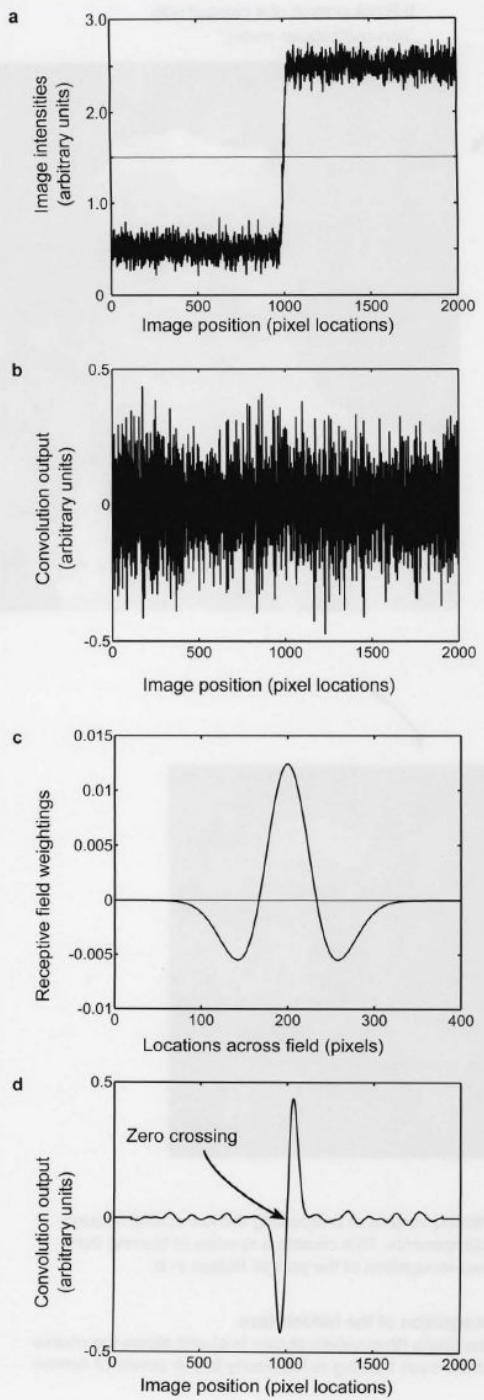
320				

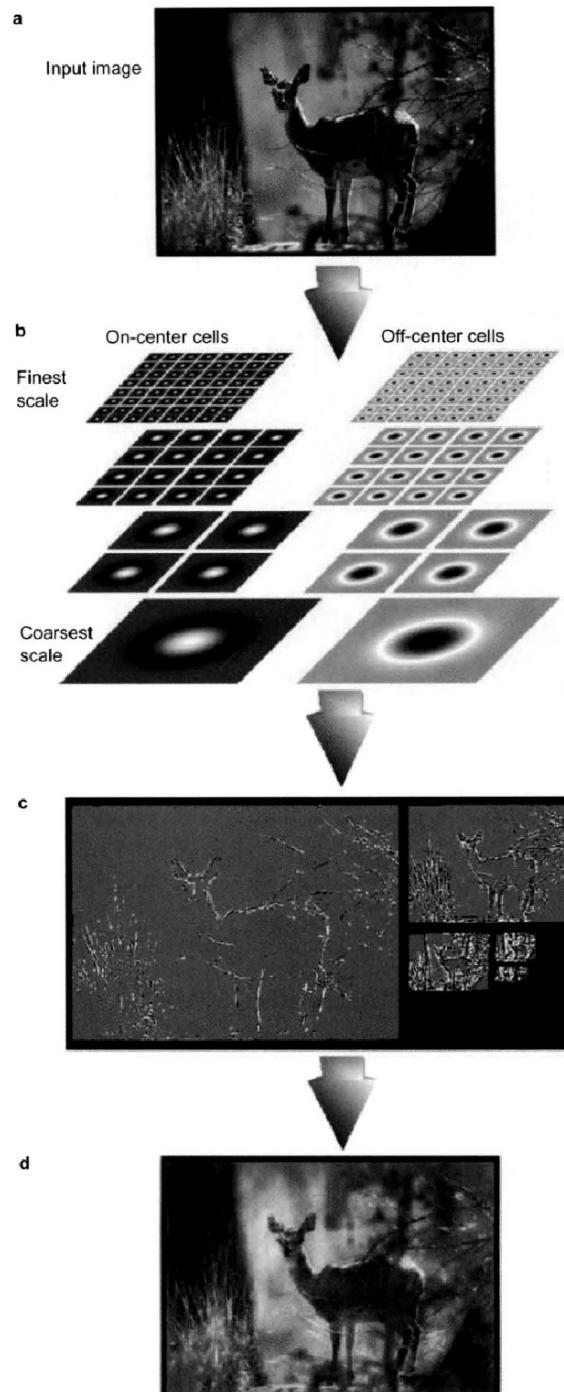
Output Matrix

$$\begin{aligned} & 0 * 0 + 0 * -1 + 0 * 0 \\ & + 0 * -1 + 105 * 5 + 102 * -1 \\ & + 0 * 0 + 103 * -1 + 99 * 0 = 320 \end{aligned}$$

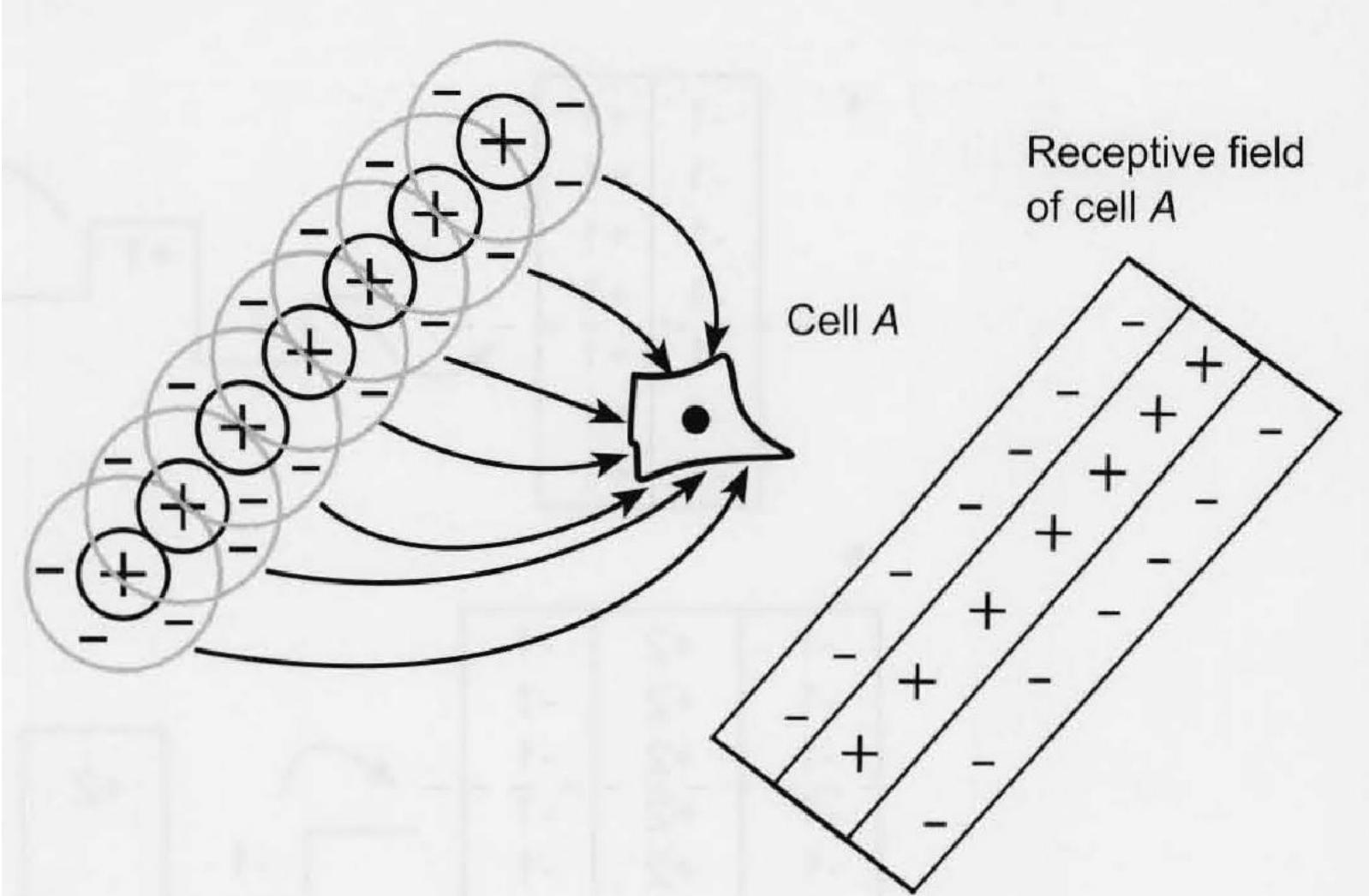
**Convolution with horizontal and vertical strides = 2**







Van Rullen and Thorpe (2001).



**a**

+	-	+
+	-	+
+	-	+
+	-	+
+	-	+

**b**

+	-	+
+	-	+
+	-	+
+	-	+

Narrow-width line

+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+

+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+

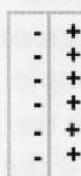
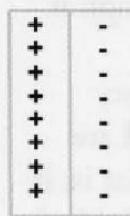
Medium-width  
line

+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+

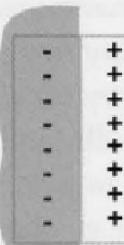
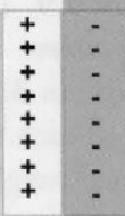
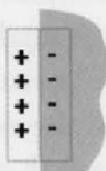
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+
+	-	+

Large-width  
line

**Receptive fields plotted  
with small spots of light**



**Optimal stimuli**



Narrow-width  
flanks

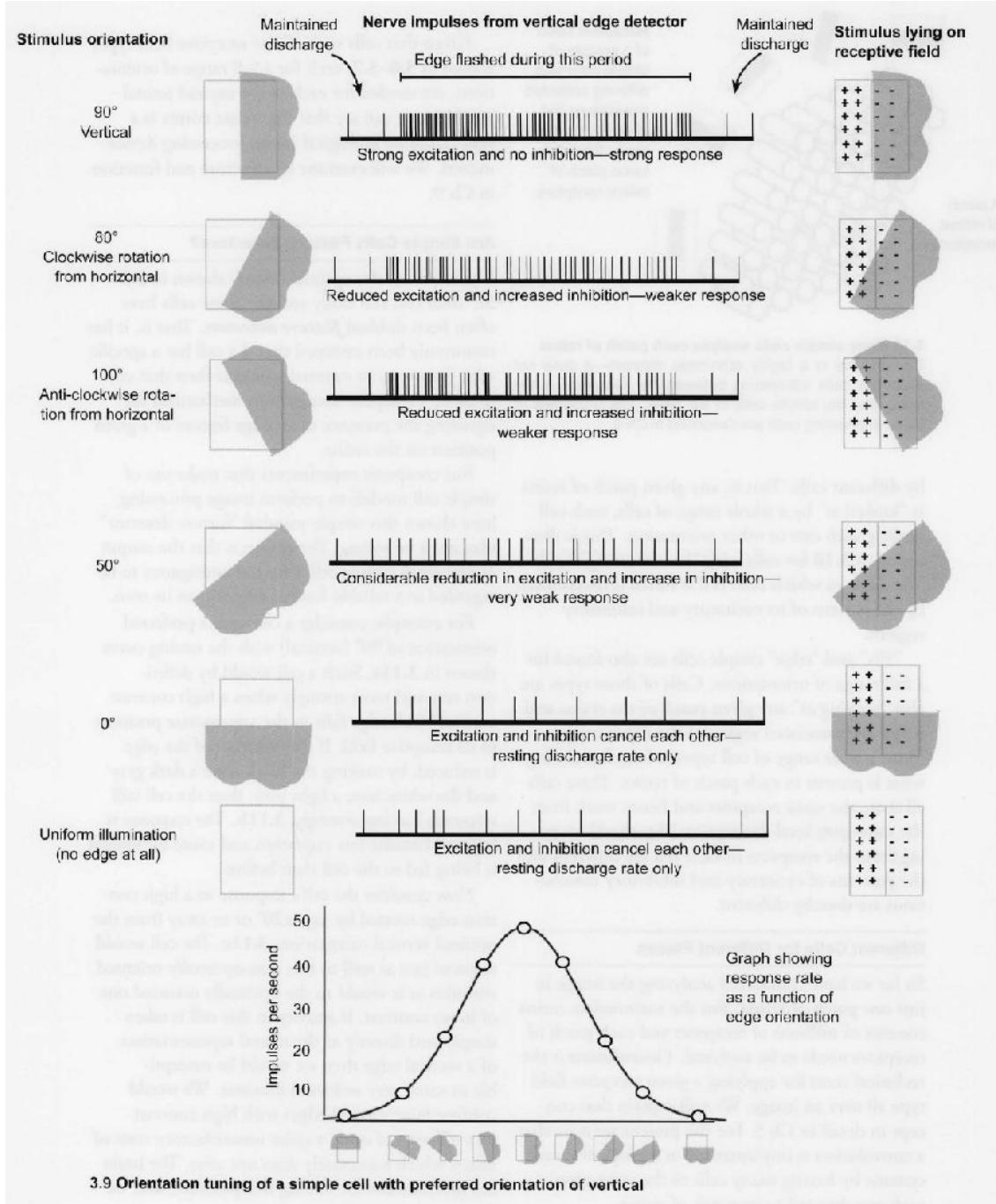
Medium-width  
flanks

Large-width  
flanks

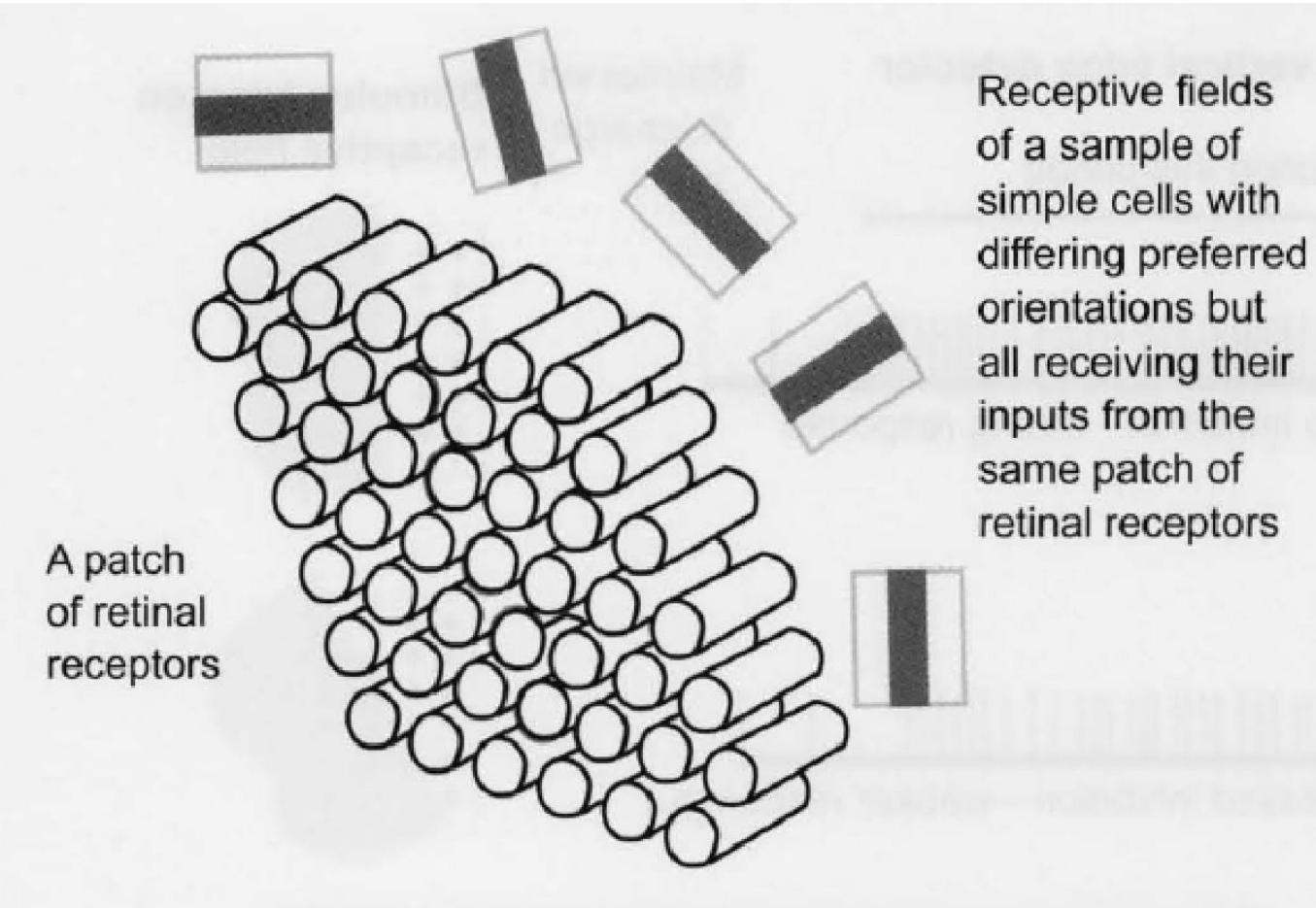
Narrow-width  
flanks

Medium-width  
flanks

Large-width  
flanks



- Una amplia gama de tipos de células se encuentra analizando lo que está presente en cada parche de la retina. Estas células comparten todos los mismos receptores y, por lo tanto, trabajan desde la misma descripción de nivel de gris. Pero las fibras que van desde los receptores a cada célula son diferentes y los patrones de conexiones excitatorias e inhibitorias son por lo tanto diferentes.



Different stimuli  
on same  
receptive field

Maintained  
discharge

Stimuli on during  
this period

Maintained  
discharge



Preferred stimulus:  
very brisk response

Some -s showing more  
inhibition than in a

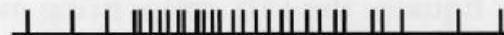
Fewer +s showing less  
excitation than in a



Preferred stimulus orientation but  
low contrast: weaker response  
than in a



Balance of excitation and  
inhibition the same as in b



Not preferred stimulus orientation  
but high contrast: similar  
response to that in b

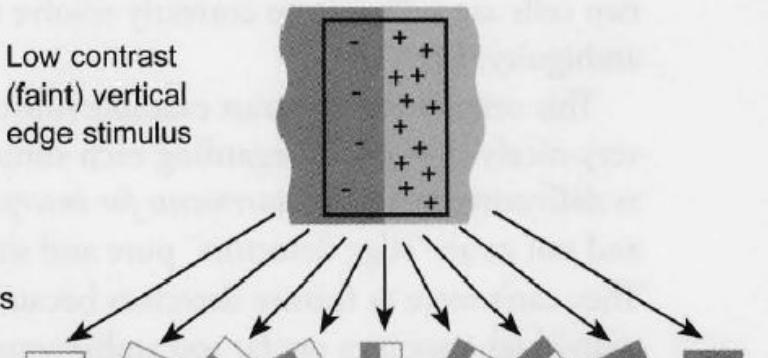


All responses are from the same simple striate cell  
which has a vertical preferred stimulus orientation

a

Low contrast  
(faint) vertical  
edge stimulus

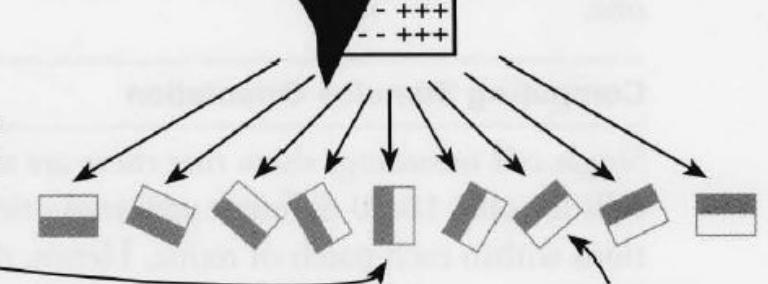
Row of cells  
varying  
in their  
preferred  
orientations



b

High contrast  
just-off-vertical  
edge stimulus

Same row  
of cells as  
above

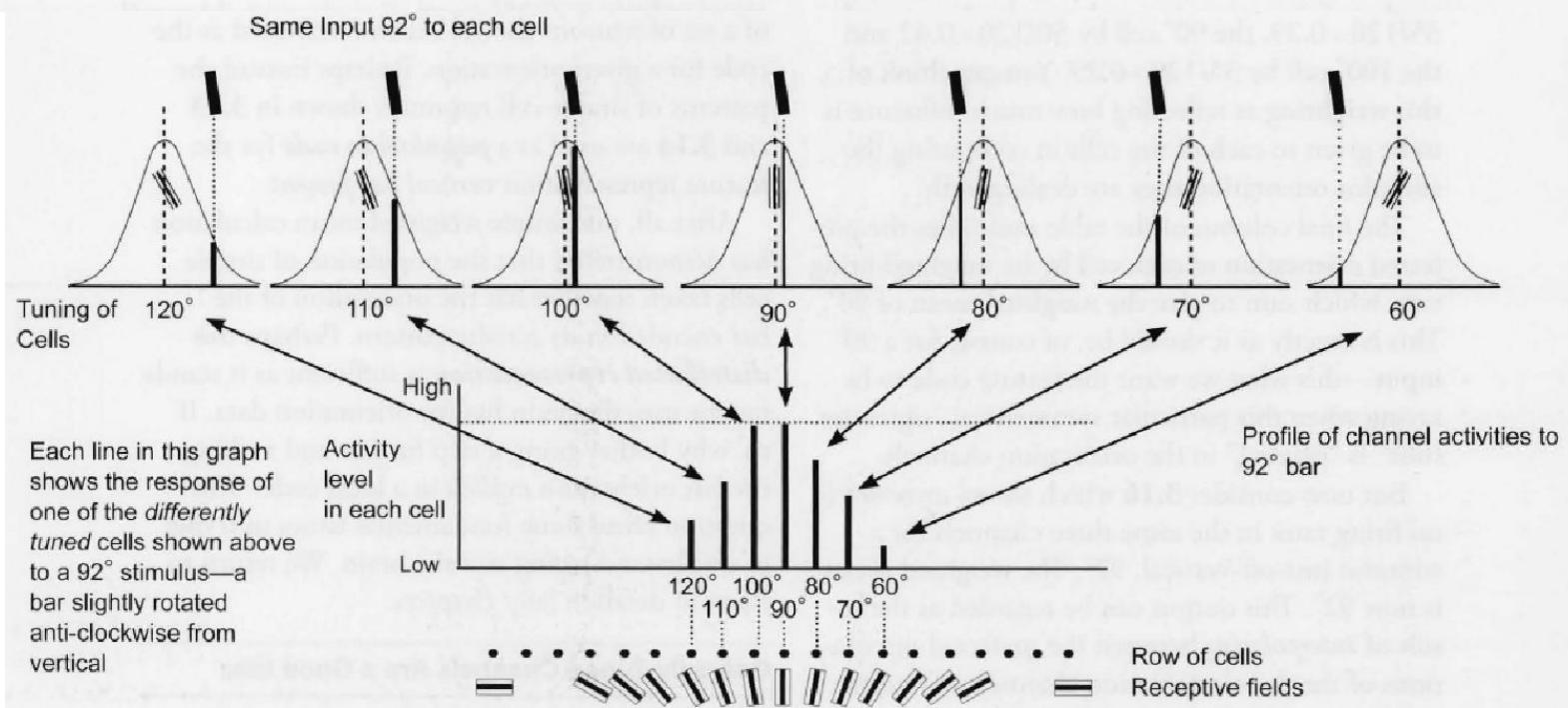
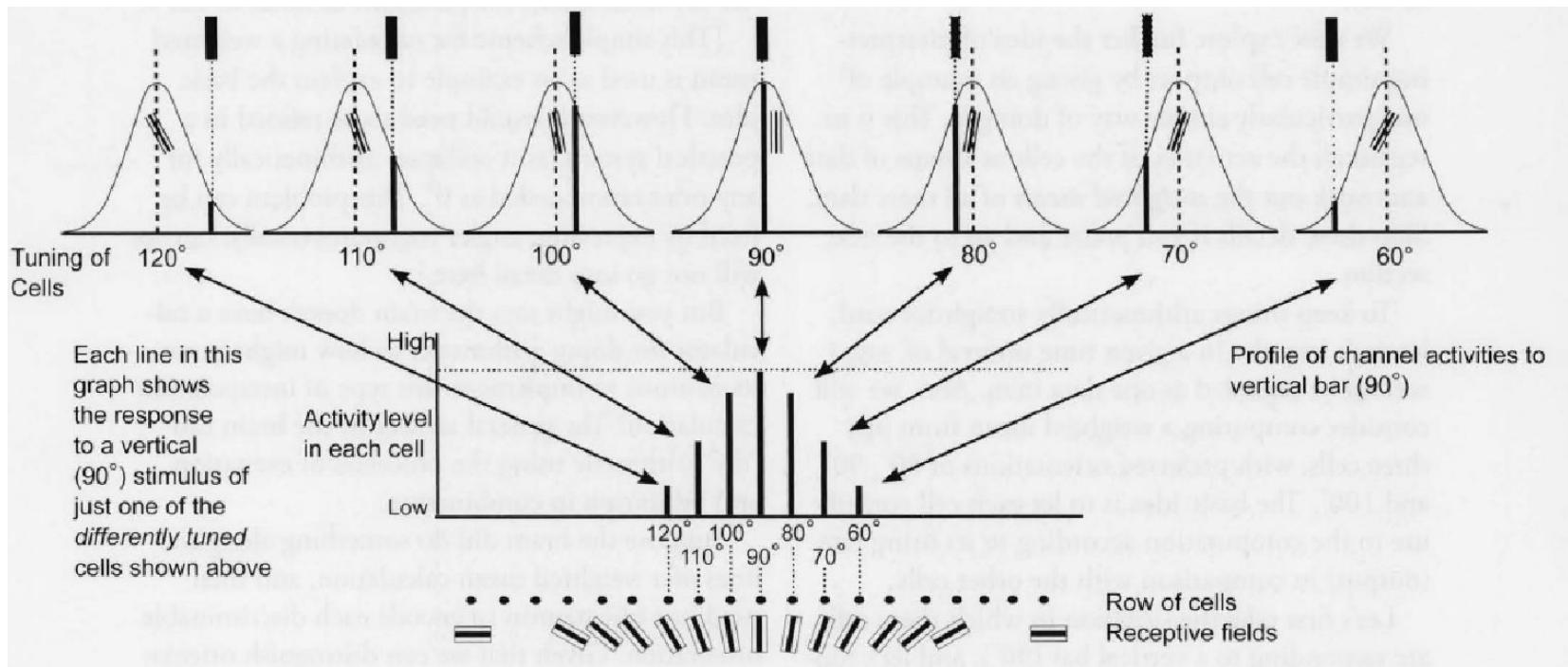


Active at 20 impulses per second, just as in  
response to the faint edge above, but still much less  
active than the neighboring cell which is tuned to  
prefer the just-off-vertical edge orientation

The most  
active cell (say,  
100 impulses  
per second)

### 3.11 Simple cell ambiguity

### 3.12 Interpreting simple cell responses in context



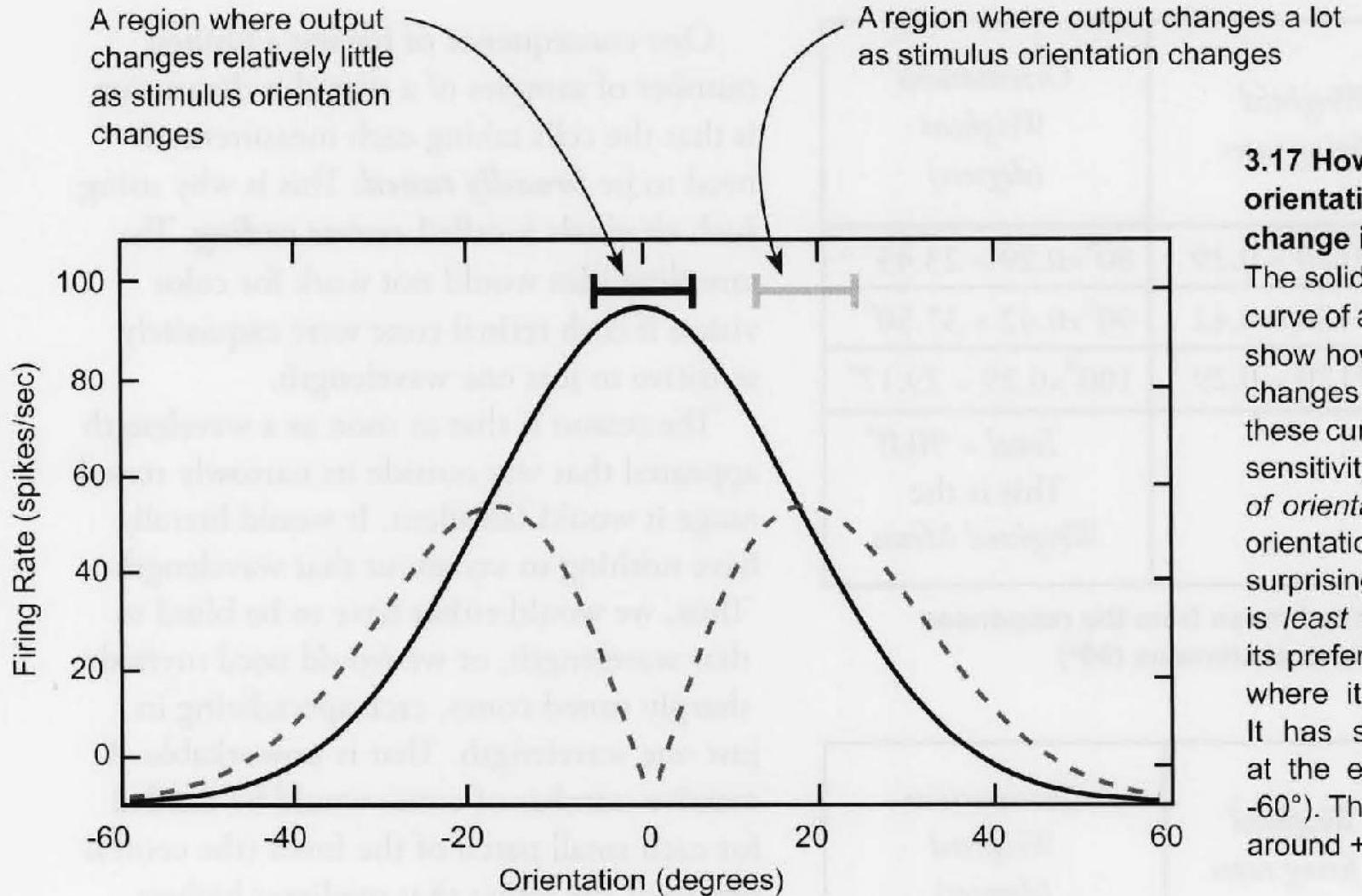
<i>Preferred Orientation of Cells</i>	<i>Cell firing rates (impulses per sec)</i>	<i>Weighted firing rates</i>	<i>Orientations Weighted (degrees)</i>
80°	35	$35/120 = 0.29$	$80^\circ \times 0.29 = 23.43^\circ$
90°	50	$50/120 = 0.42$	$90^\circ \times 0.42 = 37.50^\circ$
100°	35	$35/120 = 0.29$	$100^\circ \times 0.29 = 29.17^\circ$
	<i>Total = 120</i>		<i>Total = 90.0°</i> This is the <i>Weighted Mean</i>

**3.15 Calculating a weighted mean from the responses of three channels to a vertical stimulus (90°)**

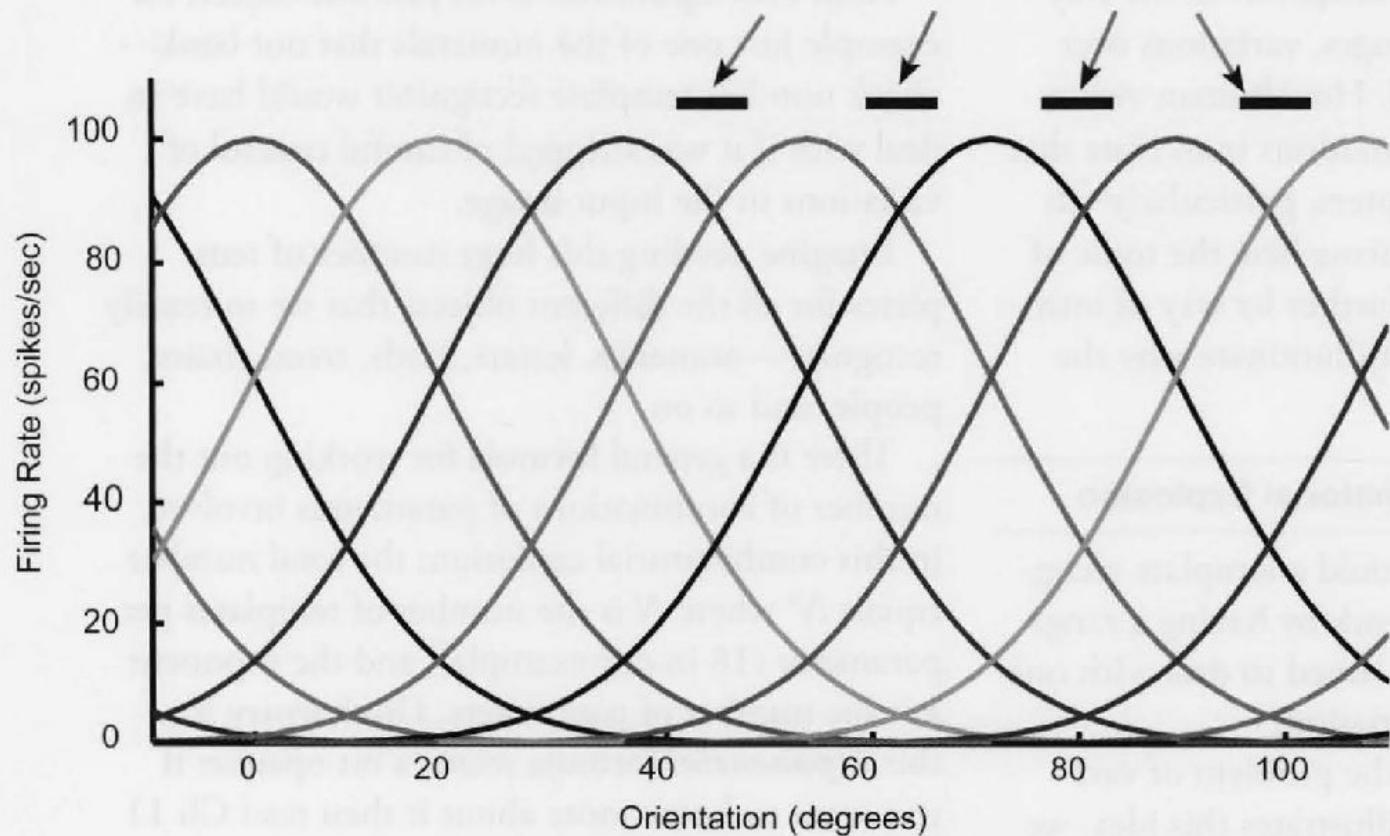
<i>Preferred Orientation of Cells</i>	<i>Cell firing rates (impulses per sec)</i>	<i>Weighted firing rates</i>	<i>Orientations Weighted (degrees)</i>
80°	24	$24/118 = 0.20$	$80^\circ \times 0.20 = 16.27^\circ$
90°	46	$46/118 = 0.39$	$90^\circ \times 0.39 = 35.08^\circ$
100°	48	$48/118 = 0.40$	$100^\circ \times 0.40 = 40.68^\circ$
	<i>Total = 118</i>		<i>Total = 92.0°</i> This is the <i>Weighted Mean</i>

**3.16 Calculating a weighted mean from the responses of three channels to a 92° stimulus**

# Problema de la resolución de los parámetros



Examples of regions of peak sensitivity to stimulus orientation changes because these are regions with steepest changes in channel outputs

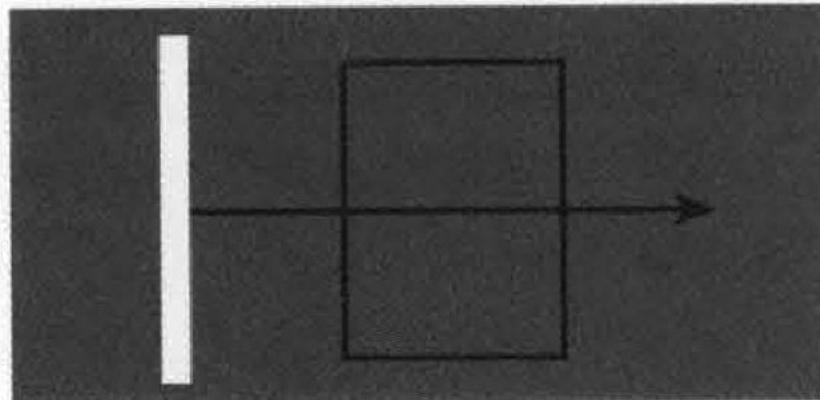


**3.18 Using many simple cells with different preferred orientations to “tile” the full stimulus orientation range**

# Células simples y complejas

a

Moving stimulus with preferred orientation for the cell



Receptive field of complex cell

Recordings

Complex cell

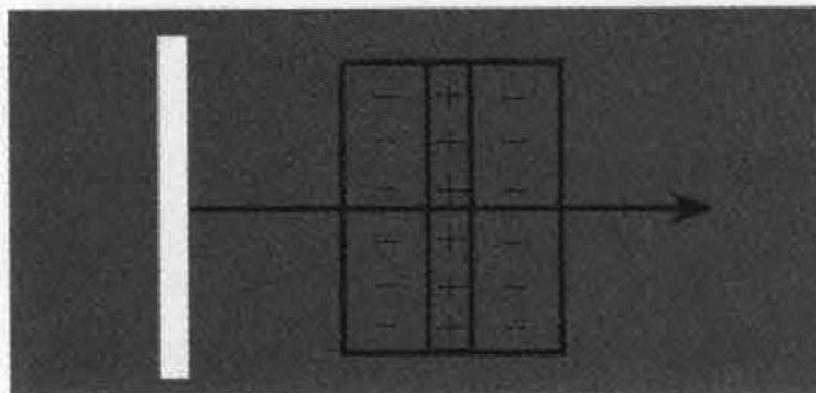


Simple cell

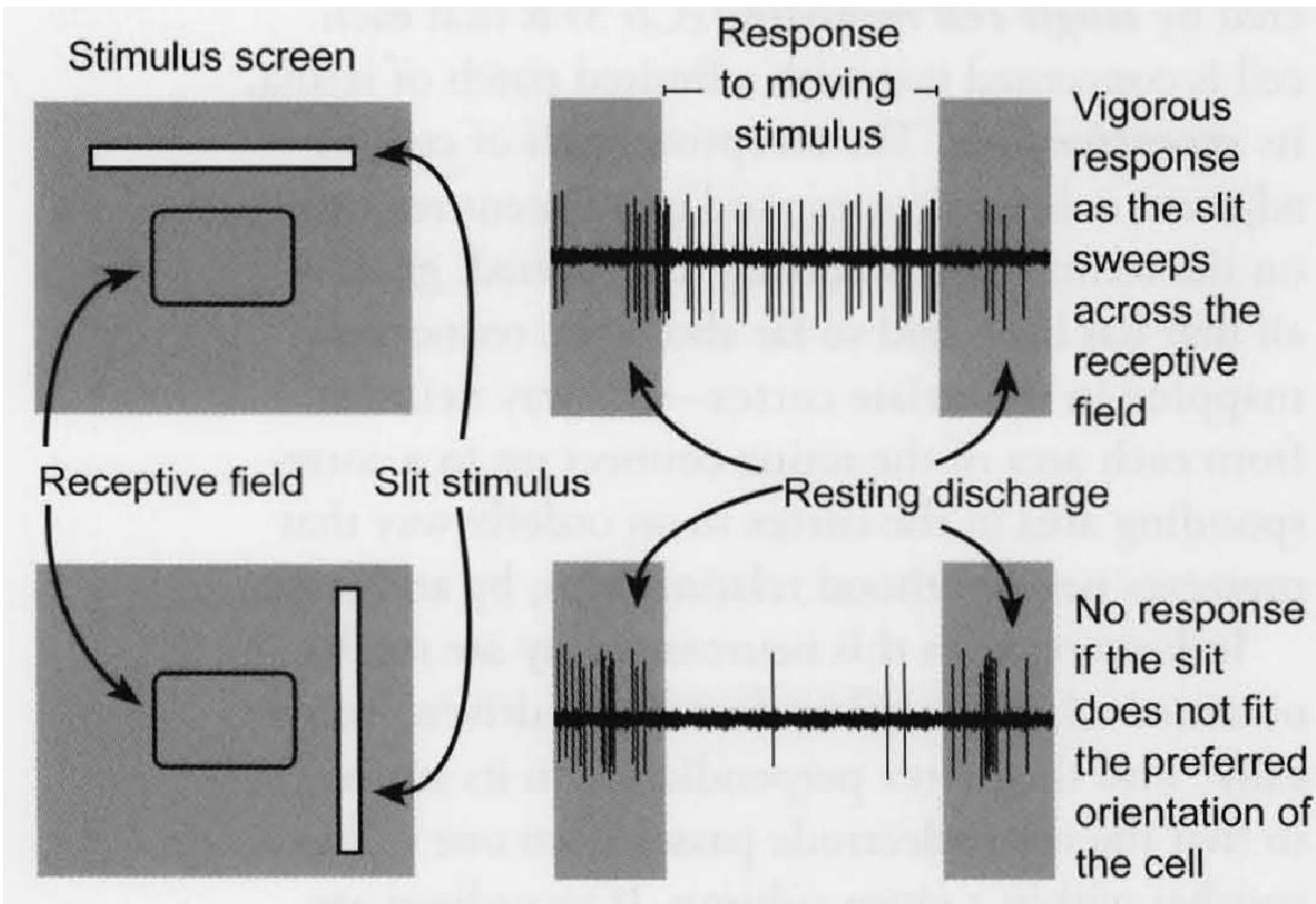


b

Moving stimulus with preferred orientation for the cell

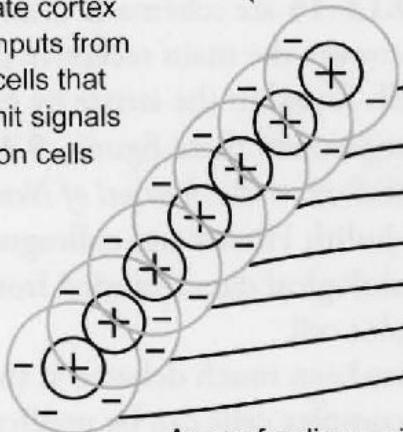


Receptive field of simple cell



a

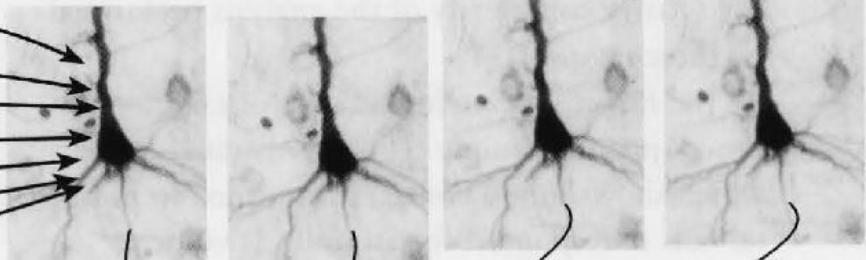
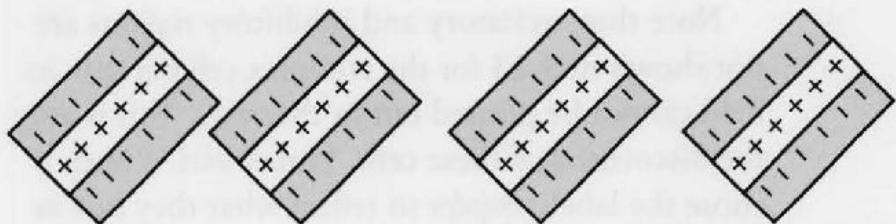
On-center receptive fields of certain cells in layer 4 (9.11, 9.20) of striate cortex that receive their inputs from lateral geniculate cells that themselves transmit signals from retinal ganglion cells



Axons feeding a simple cell in striate cortex, explaining how it becomes preferentially responsive to a particular line orientation

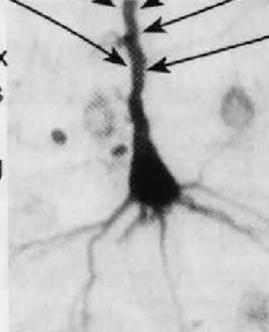
b

Receptive fields of simple cells can be mapped into excitatory and inhibitory areas



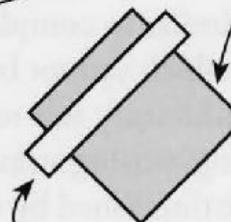
Axons

Dendrite of complex cell receiving inputs from axons of simple cells sharing same orientation tuning

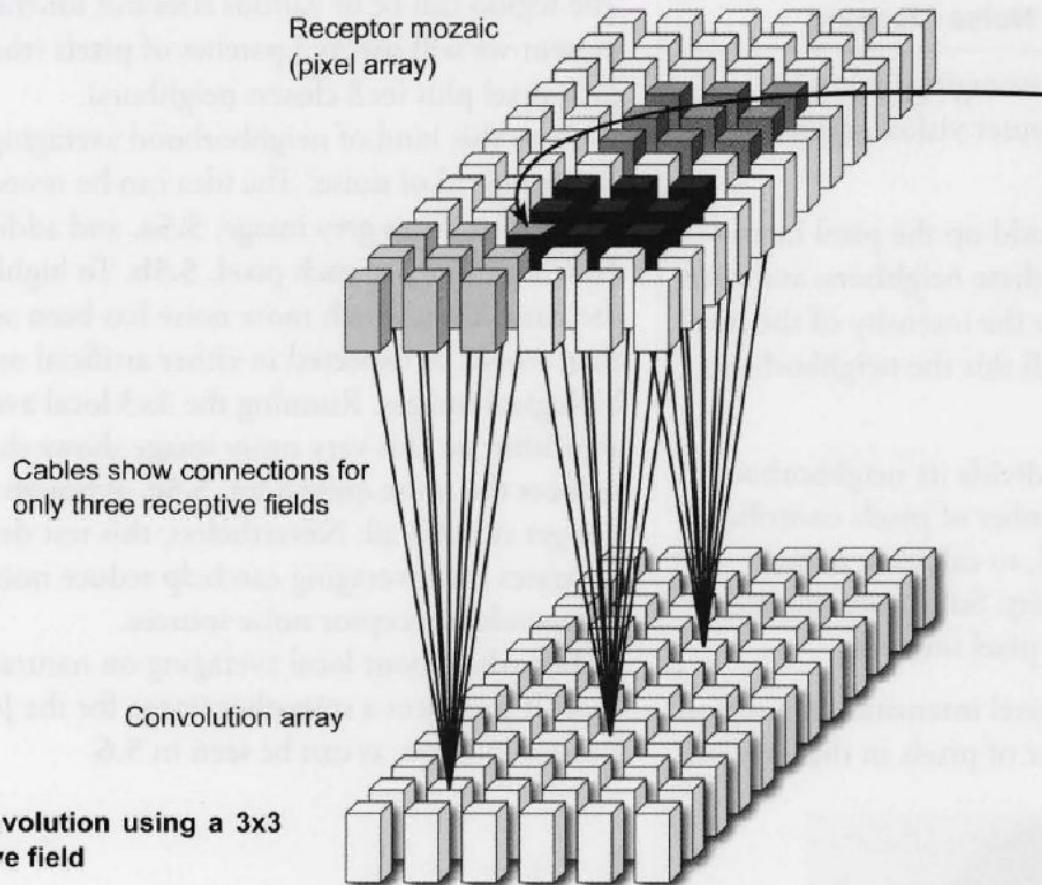


Stimulus with preferred orientation of this complex cell

Receptive field of a complex cell cannot be mapped into excitatory and inhibitory areas



reduced resolution resulting from downsampling  
and much slower learning by the network.  
A convolutional layer has a receptive field of size  $3 \times 3$   
and hence can be represented by a  $3 \times 3$  convolution array.  
The receptive field of each convolution cell is a  $3 \times 3$  block of pixels.  
The connections between the convolution array and the  
receptor mozaic are shown below.

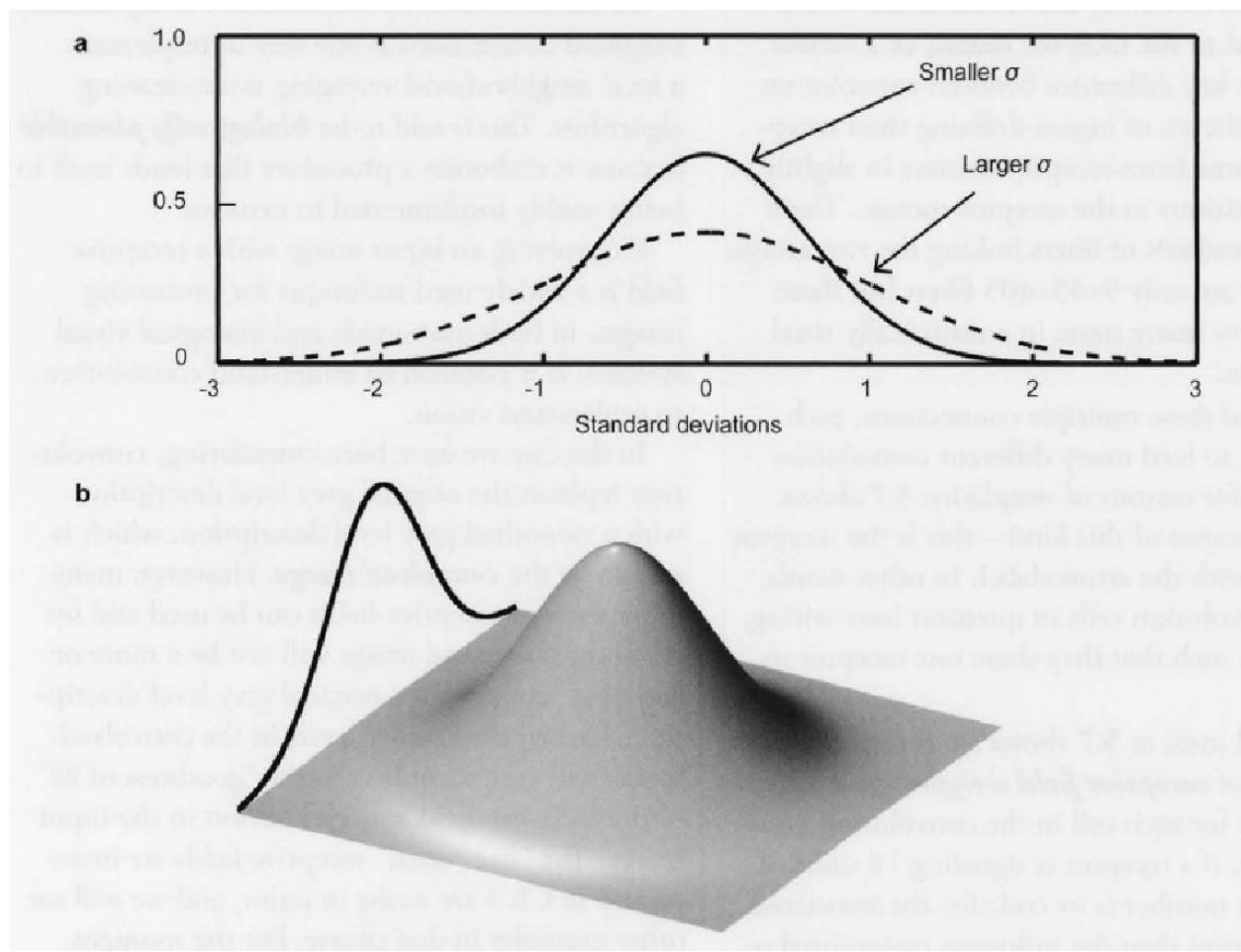


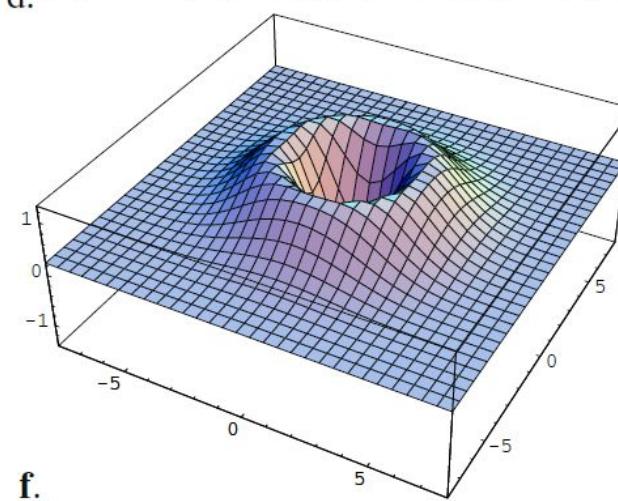
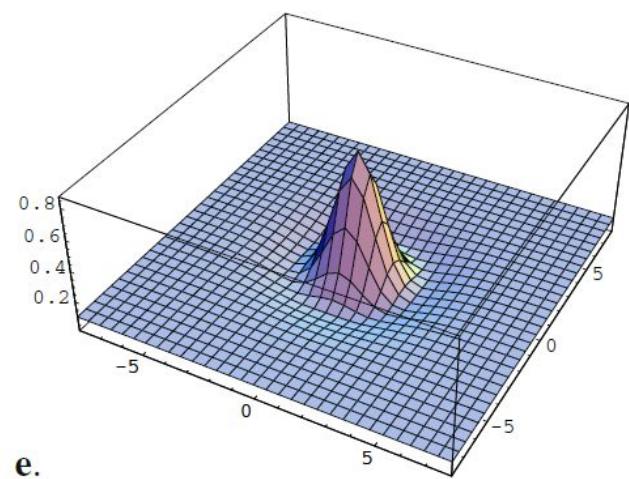
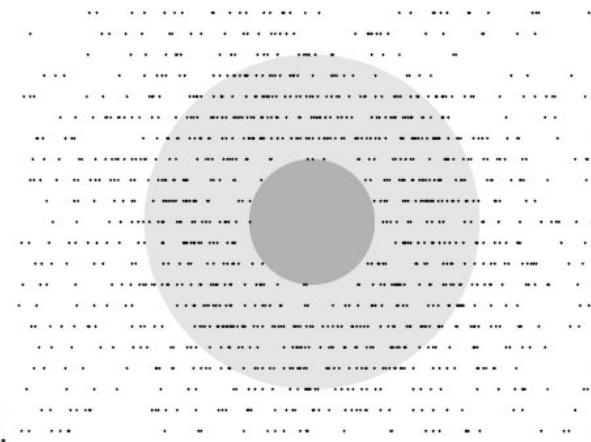
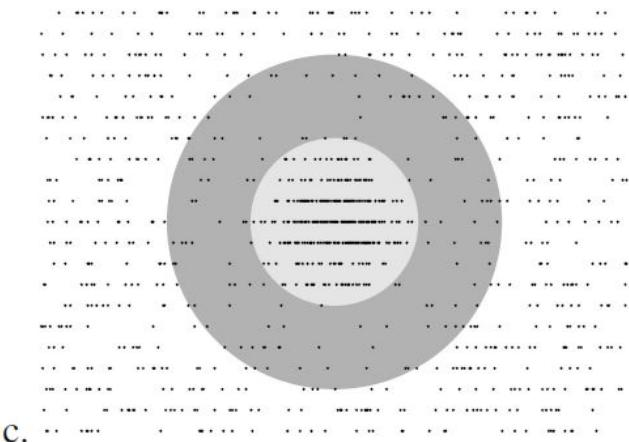
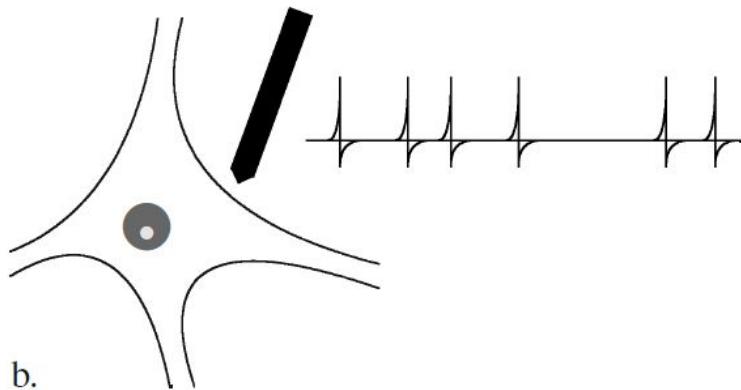
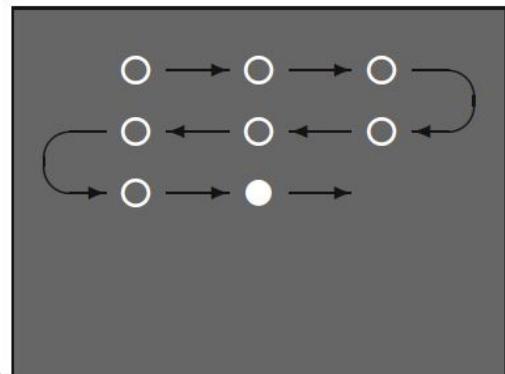
### 5.7 Convolution using a $3 \times 3$ receptive field

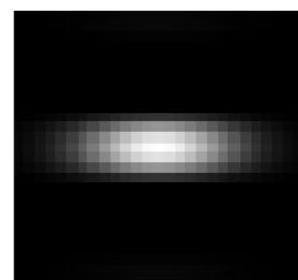
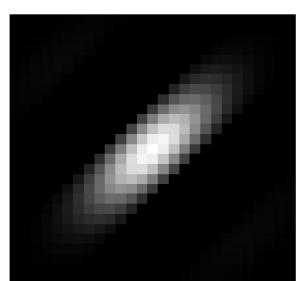
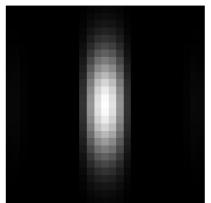
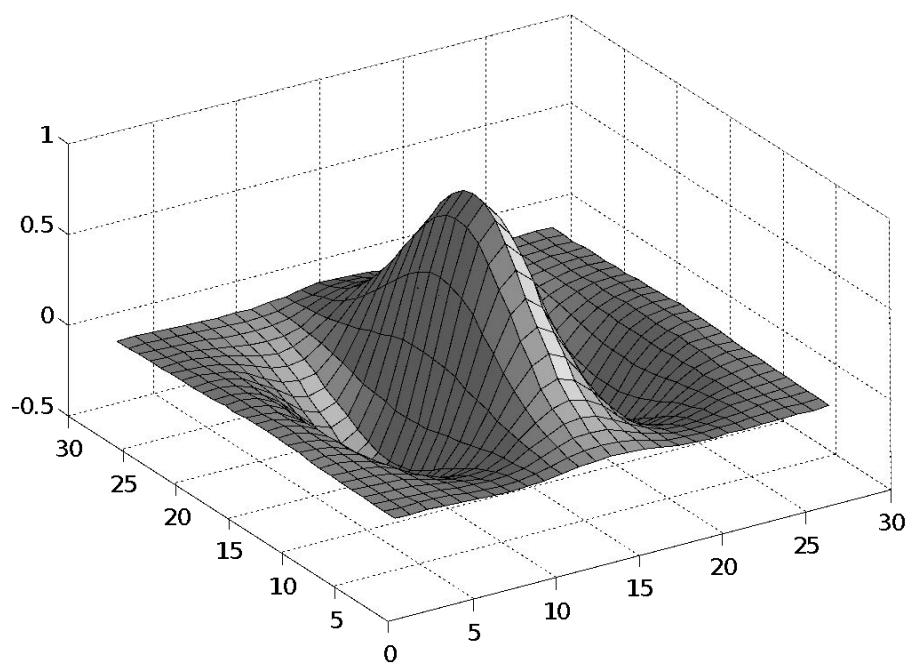
Schematic diagram of  
3x3 receptive field to  
show the  $\frac{1}{9}$  weightings  
applied to each receptor  
input to the convolution  
cells

$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

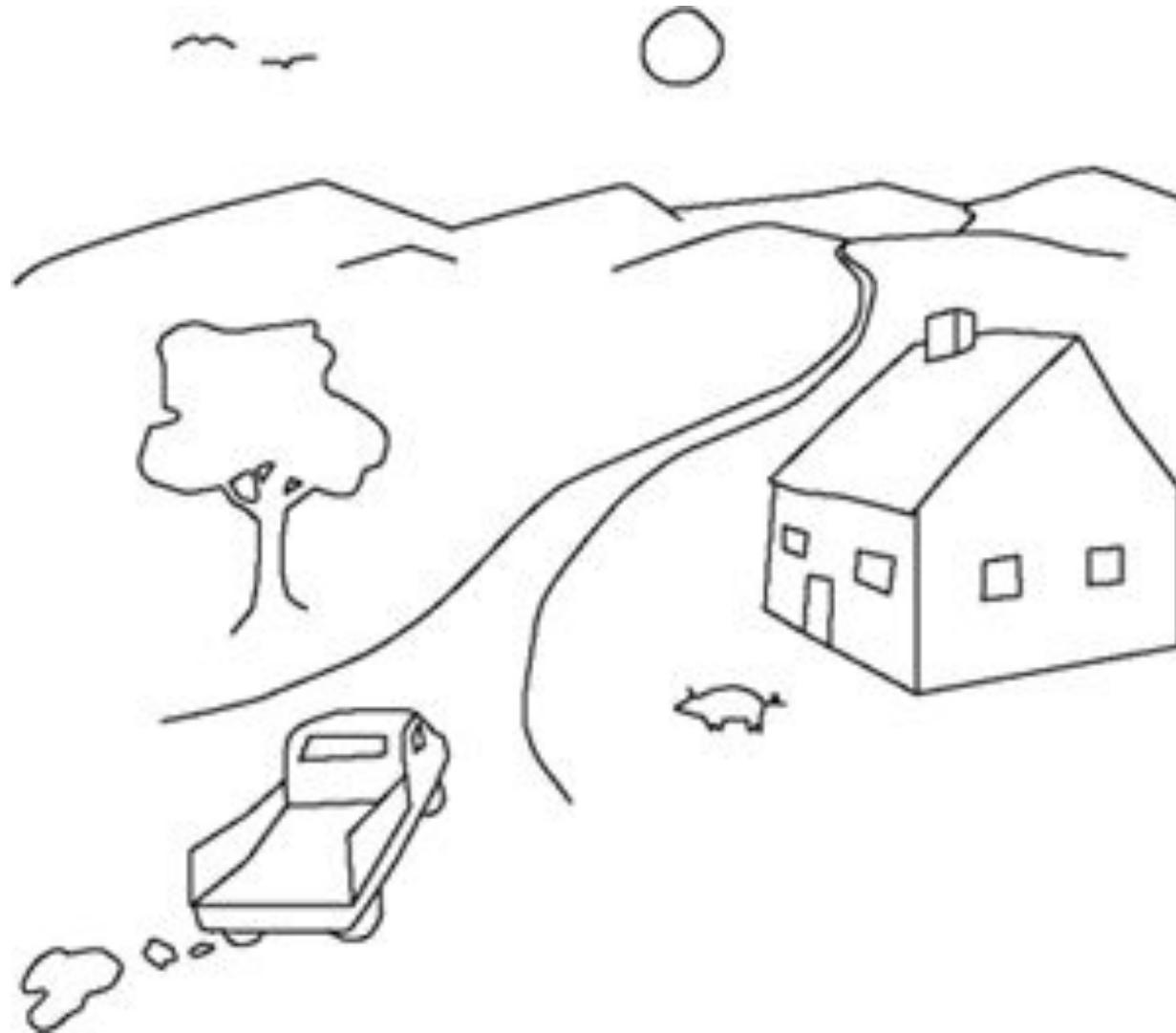
- La forma particular de la distribución de ponderación tiene que optimizar dos objetivos contradictorios: suavizar el ruido y no alterar demasiado los bordes que se encuentren en la imagen convolucionada.
- Hay un teorema (ver Marr y Hildreth, 1980) que indica que la forma de la distribución de ponderación para lograr un equilibrio óptimo entre esos dos objetivos es en forma de campana.







# Donde hay mas información?

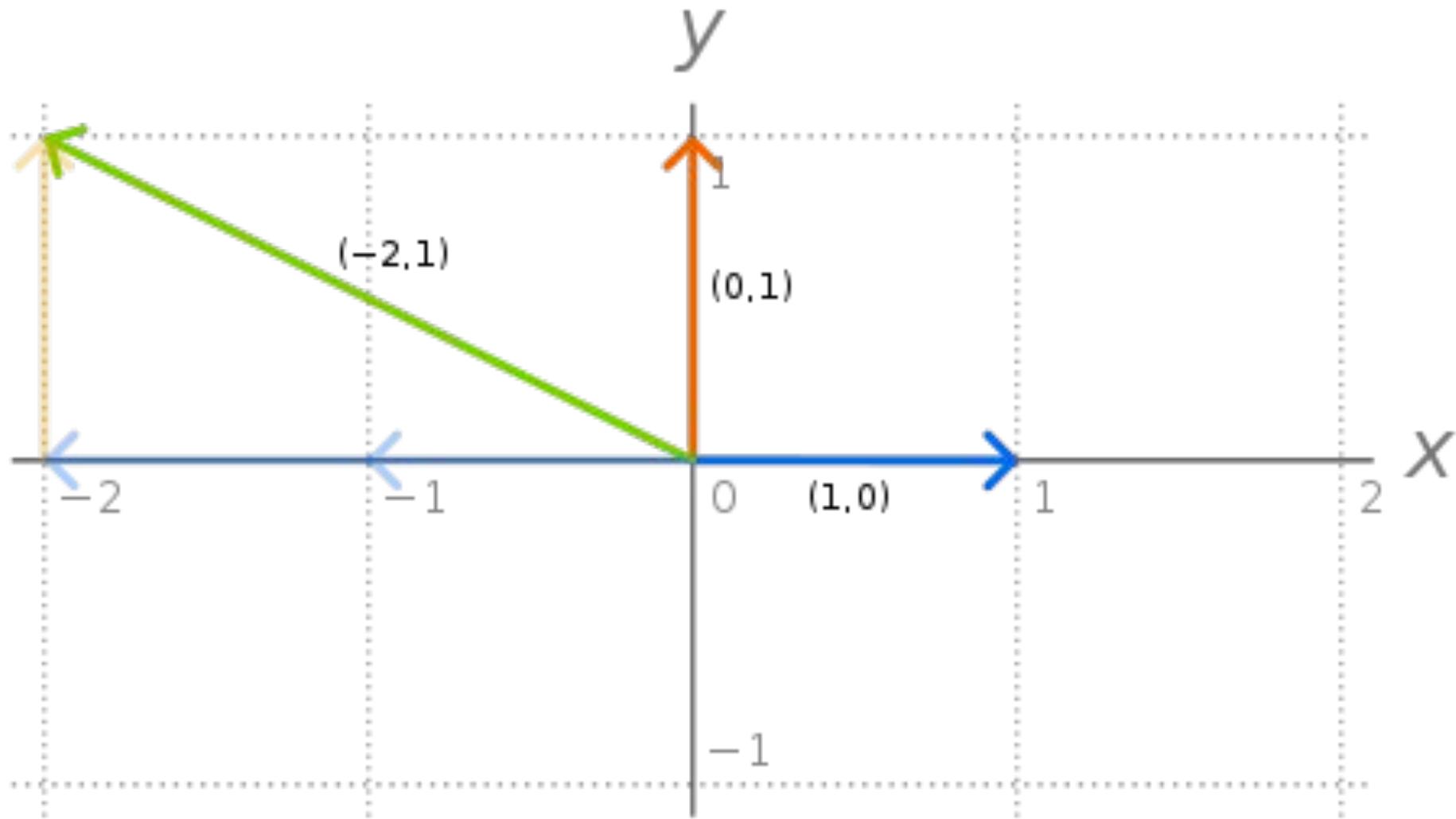


# Teoría psicofísica de las frecuencias espaciales

C. BLAKEMORE AND F. W. CAMPBELL



# Base vectorial en R<sub>2</sub>

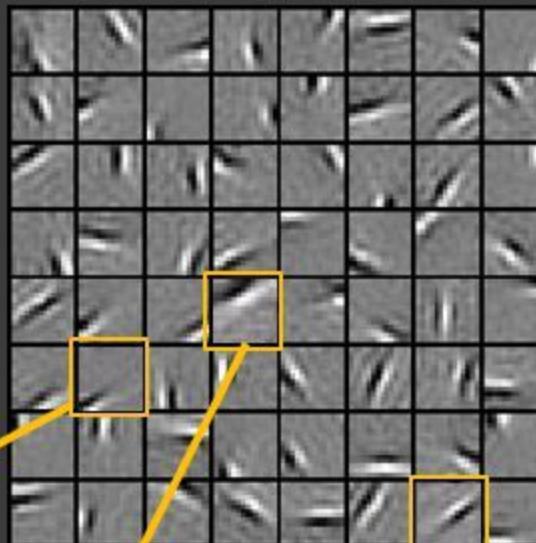


# Sparse Coding Illustration

Natural Images



Bases  $[d_1, \dots, d_{64}]$



Test example



$x$

$$\approx 0.8 *$$



$d_{36}$

$$+ 0.3 *$$



$d_{42}$

$$+ 0.5 *$$



$d_{63}$

$[0, 0, \dots, 0, \mathbf{0.8}, 0, \dots, 0, \mathbf{0.3}, 0, \dots, 0, \mathbf{0.5}, \dots]$    Compact & easily  
 $= [\gamma_1, \dots, \gamma_{64}]$    (feature representation)   interpretable

- Un enfoque muy prometedor para resolver los difíciles problemas en la visión se basa en la adaptación a las estadísticas de la entrada.
- Una representación adaptativa no trata de representar todos los tipos posibles de datos sino que se adapta a un tipo particular de datos.

# El espacio de las imágenes

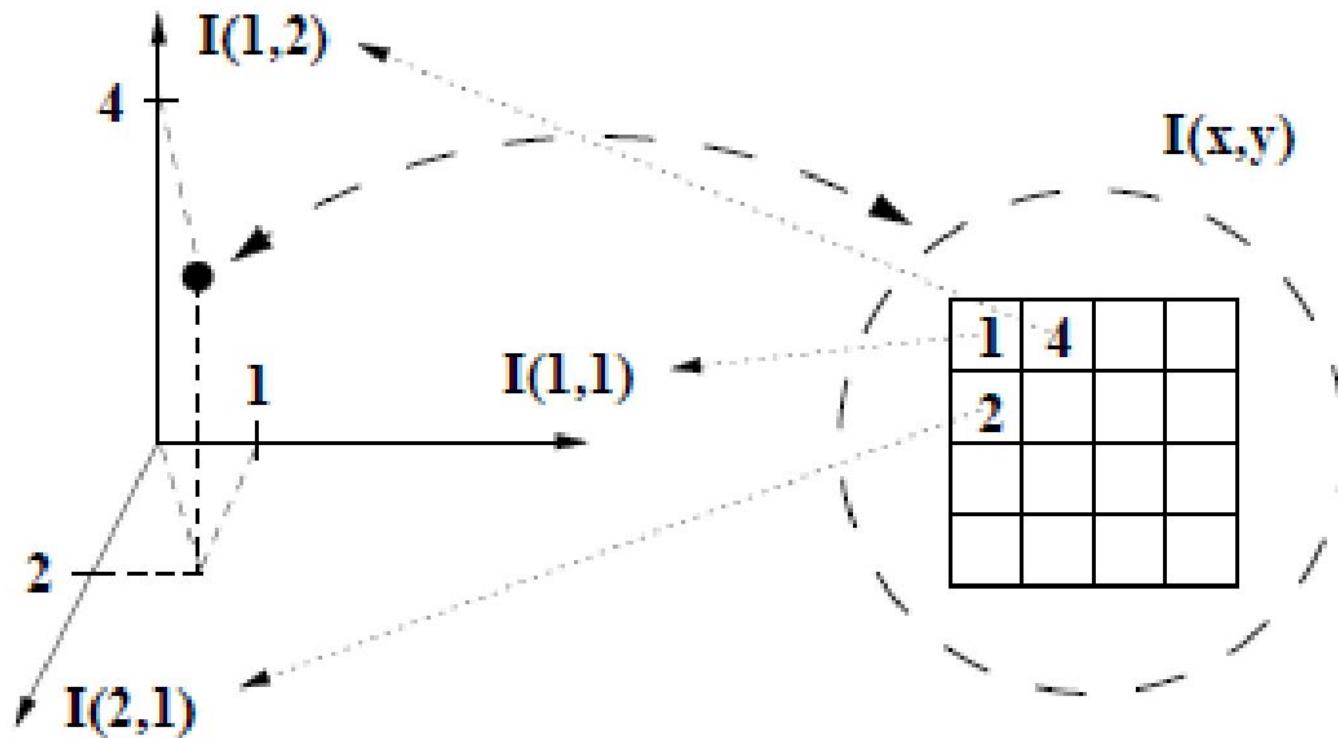
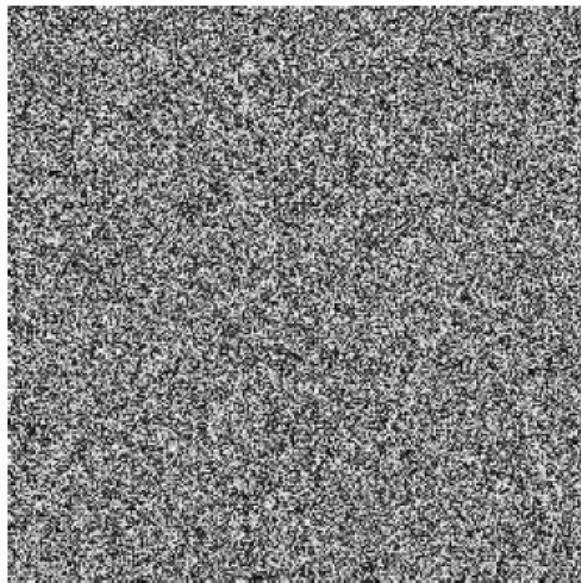
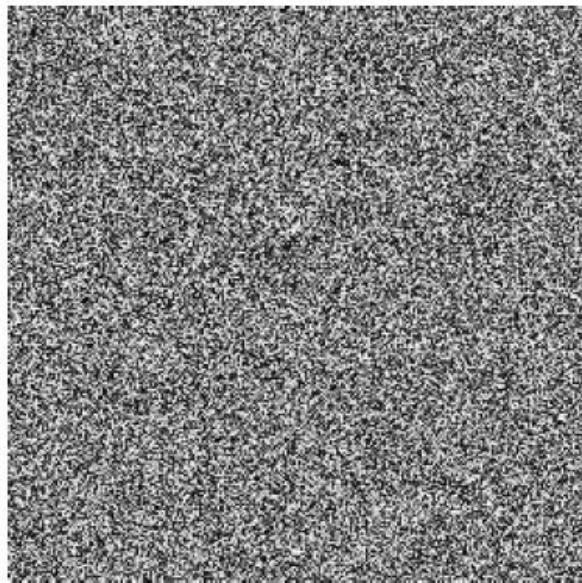
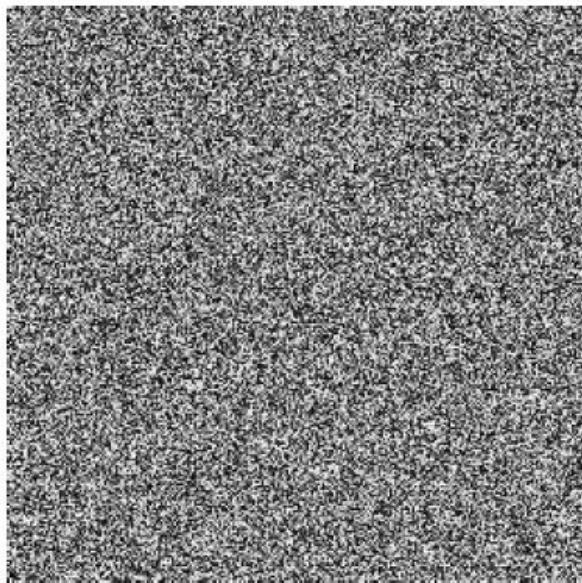


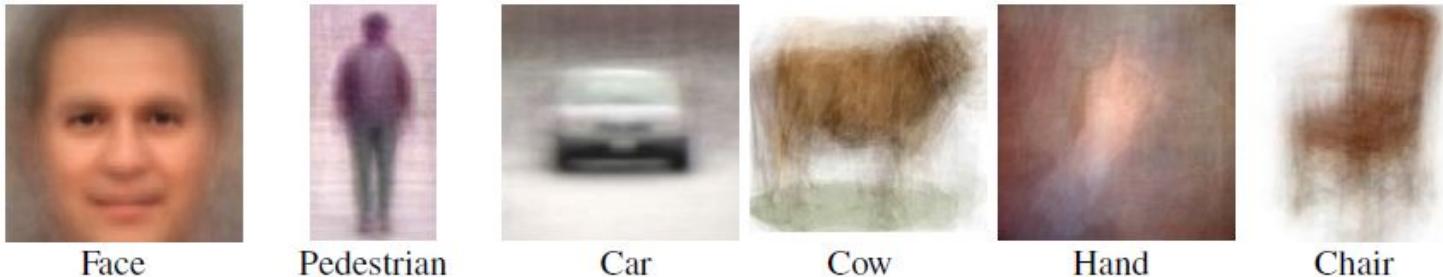
Image space

Image pixels



# Statistics of natural images

Objects



Face

Pedestrian

Car

Cow

Hand

Chair

Scenes



Mountain

Beach

Forest

Highway

Street

Indoor

Objects in scenes



Animal  
in natural scene

Tree  
in urban scene

Close-up person  
in urban scene

Far pedestrian  
in urban scene

Car in  
urban scene

Lamp in  
indoor scene

# Motion perception

- The goal of motion perception is to infer the true motion of objects in the world from the changing pattern of retinal stimulation.
- Involves solving many challenging problems
- The corresponding neural processing must be done within strong time constraints.

# Complexity of motion perception

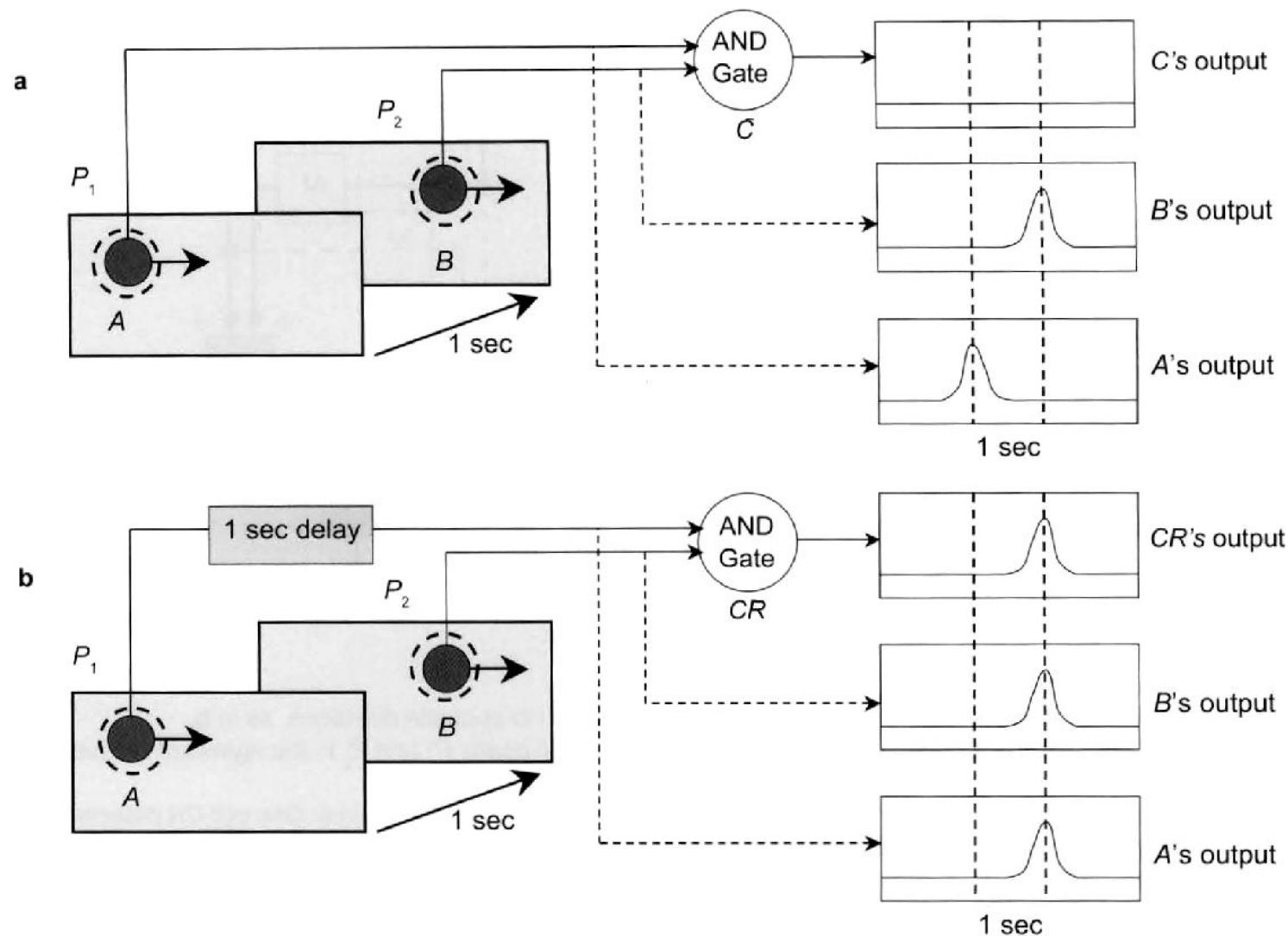


Hernan Badino and Takeo Kanade. A Head-Wearable Short-Baseline Stereo System for the Simultaneous Estimation of Structure and Motion. In IAPR Conference on Machine Vision Applications (MVA), Nara, Japan, June 2011.

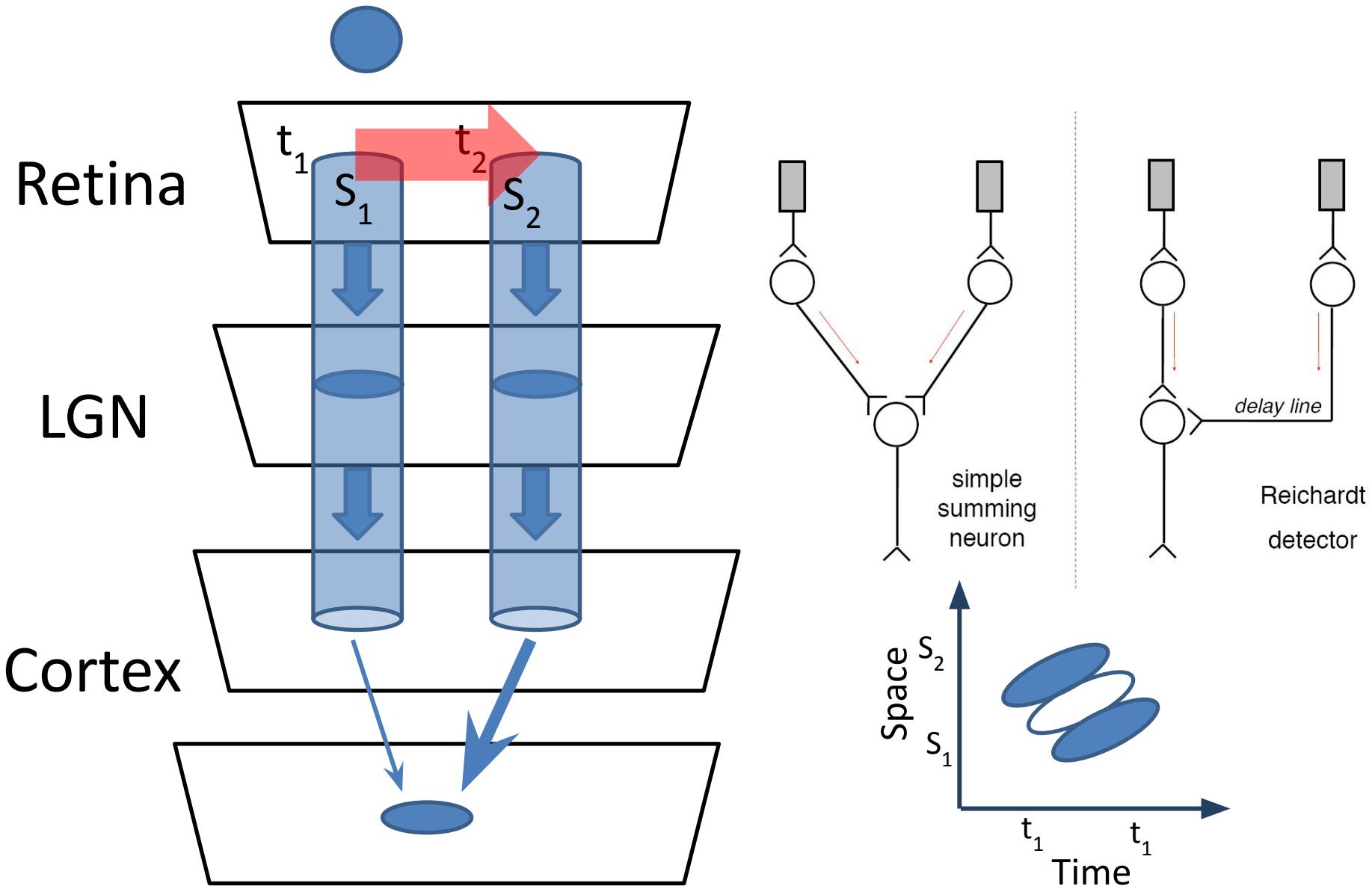
# Sensores elementales de movimiento

- La detección de movimiento comienza con "sensores de movimiento" simples en el sistema visual que se han desarrollado para detectar un cambio en la luminancia en un punto de la retina y que se correlacionan con un cambio en la luminancia en un punto vecino en la retina después de un breve retraso.

# Módulo detector de movimiento: principio básico de un detector Reichardt

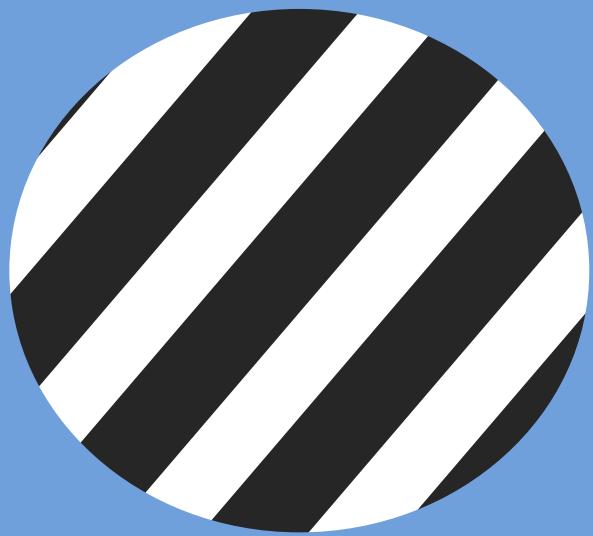


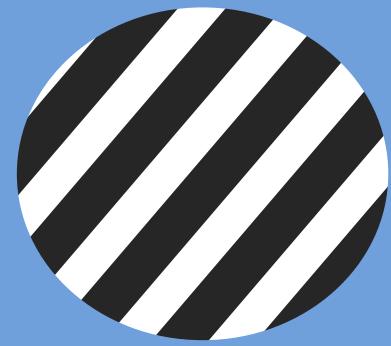
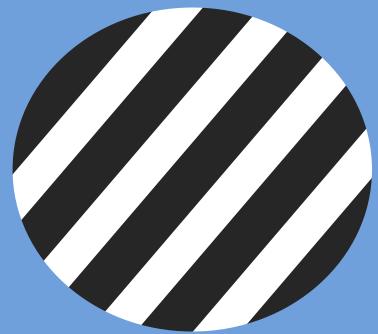
# Motion detection

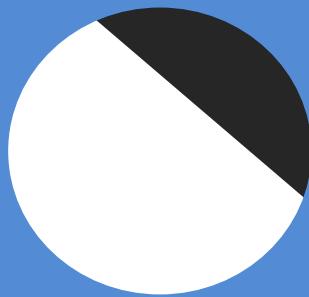


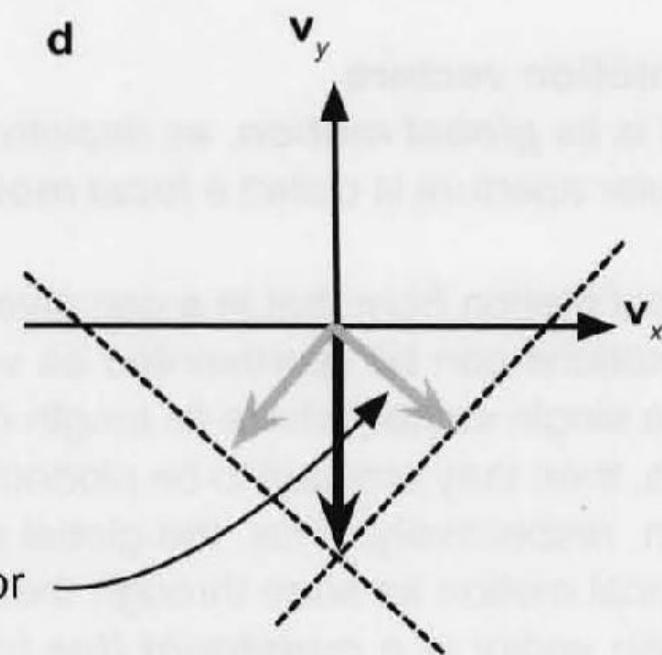
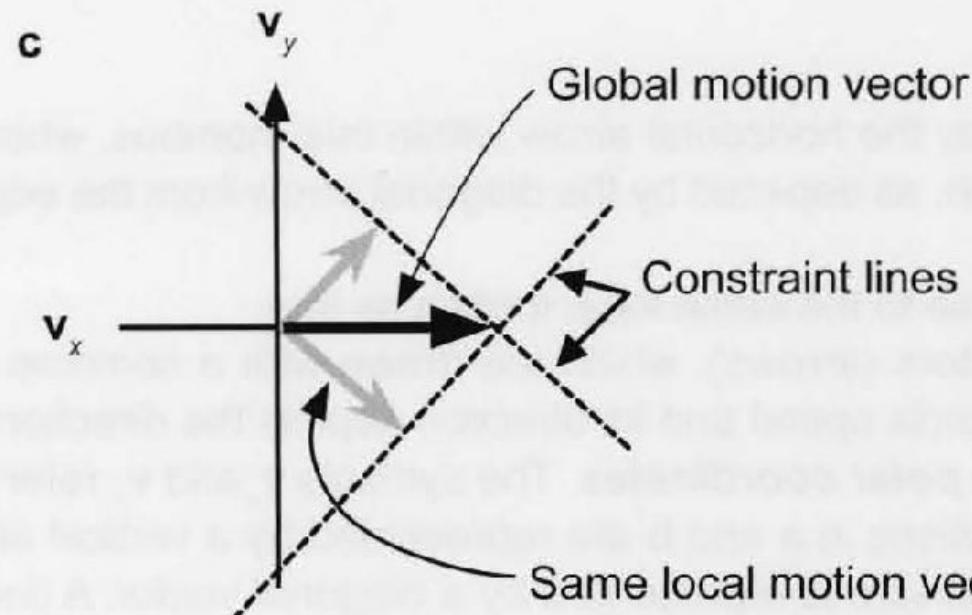
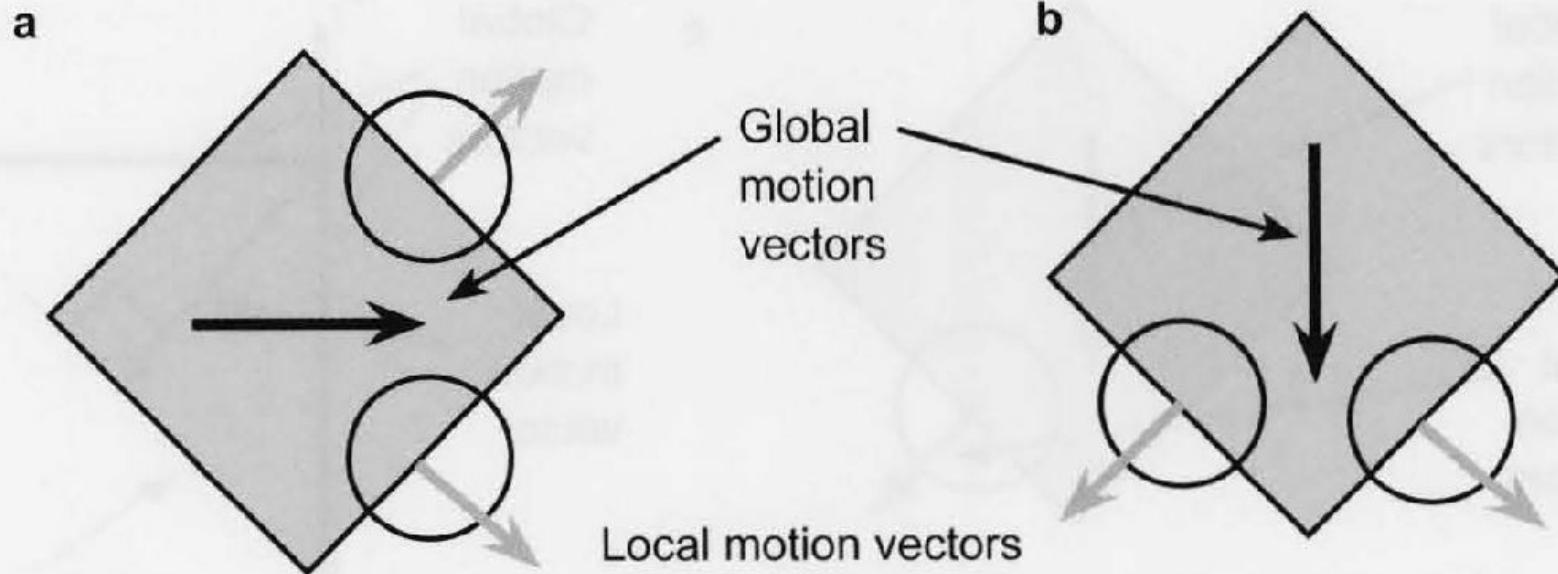
# Classic view of motion processing

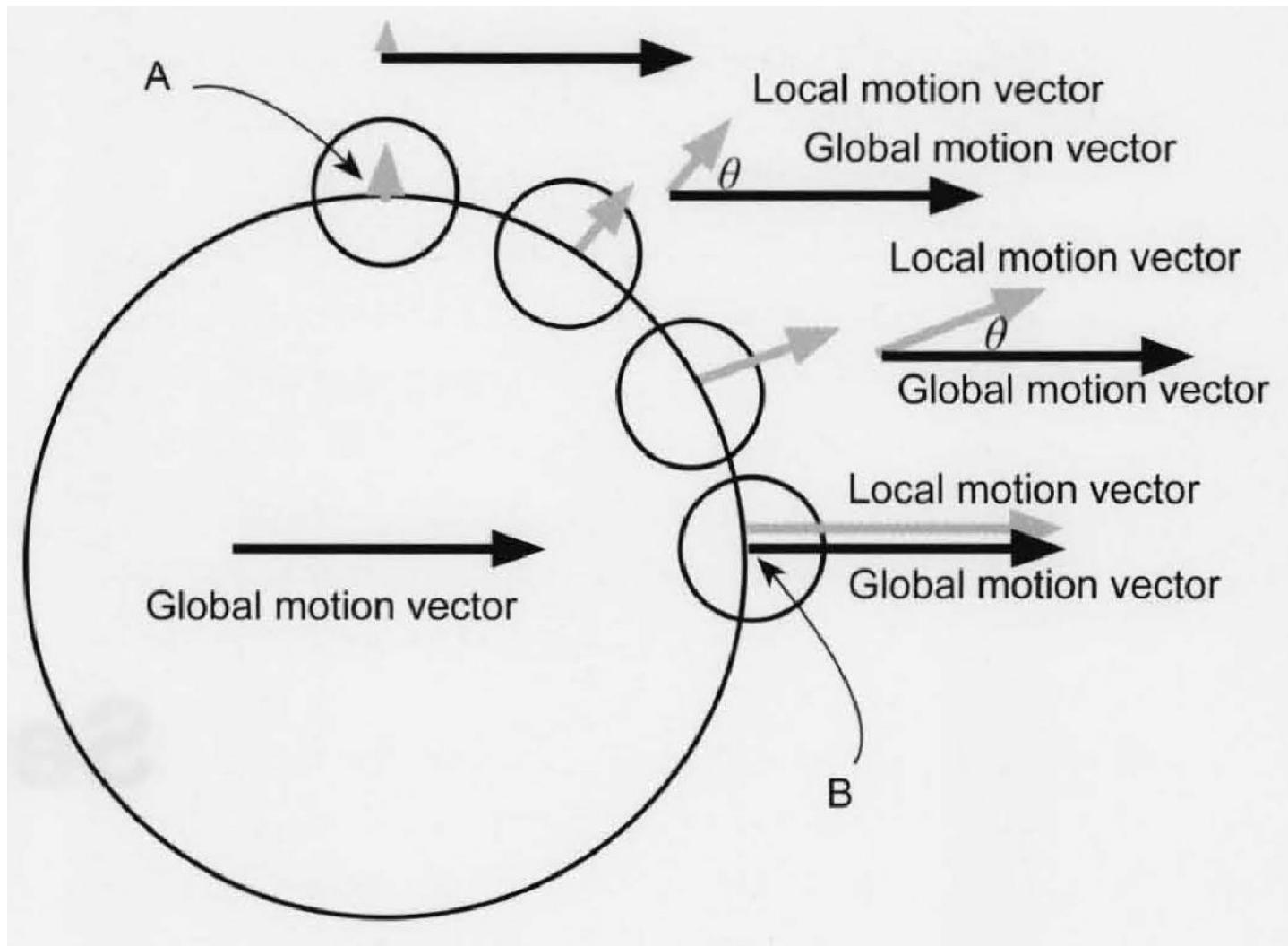
- Primary visual cortex (V1) computes an initial step in making inferences about motion.
- Direction-selective neurons in V1 encode very local measurements of image motion.
- A later visual area, designated MT, contains a second-stage representation.
- MT neurons receive input from a subset of V1 neurons and combine their inputs in such a way as to encode a more global “pattern” or object-based motion.

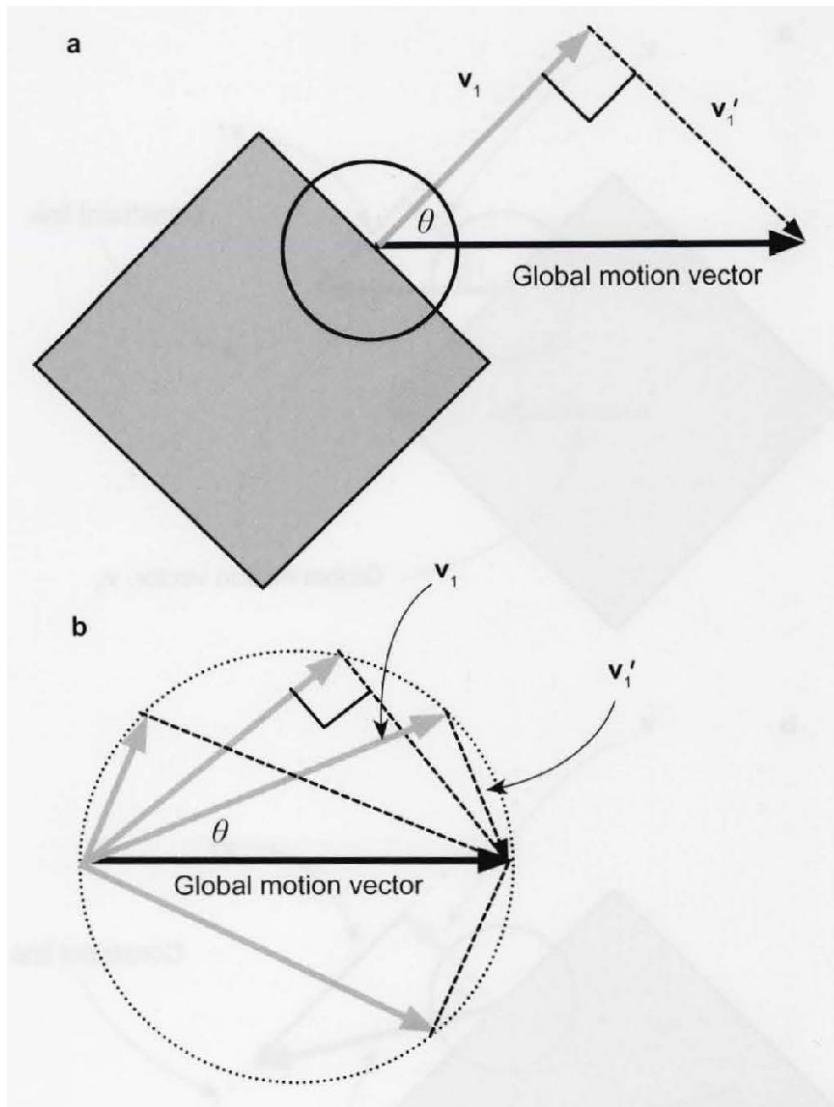




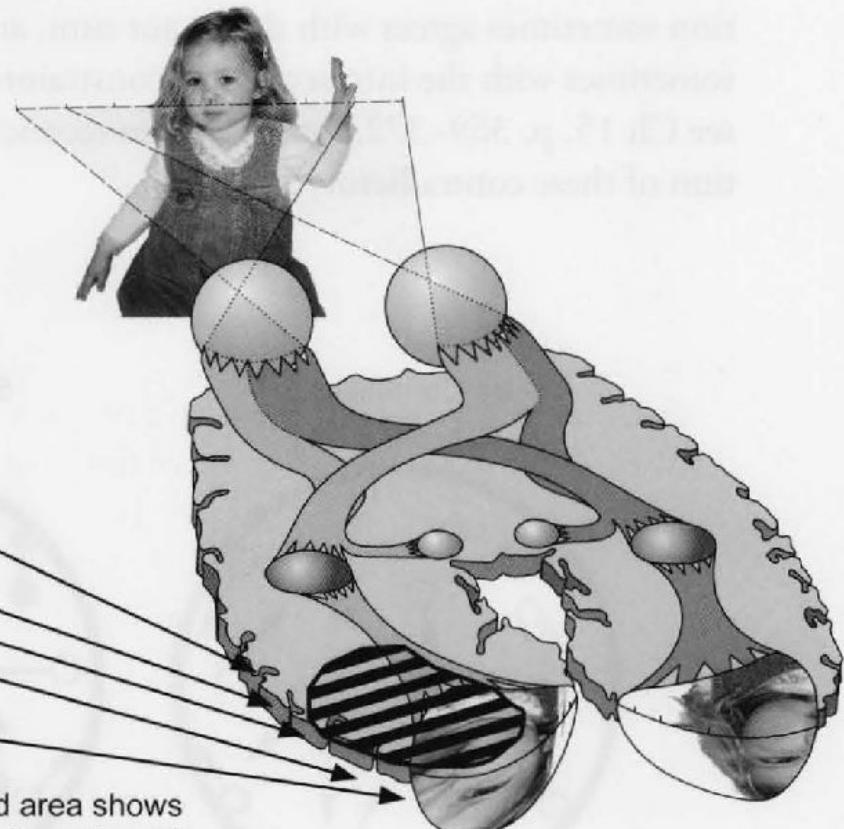
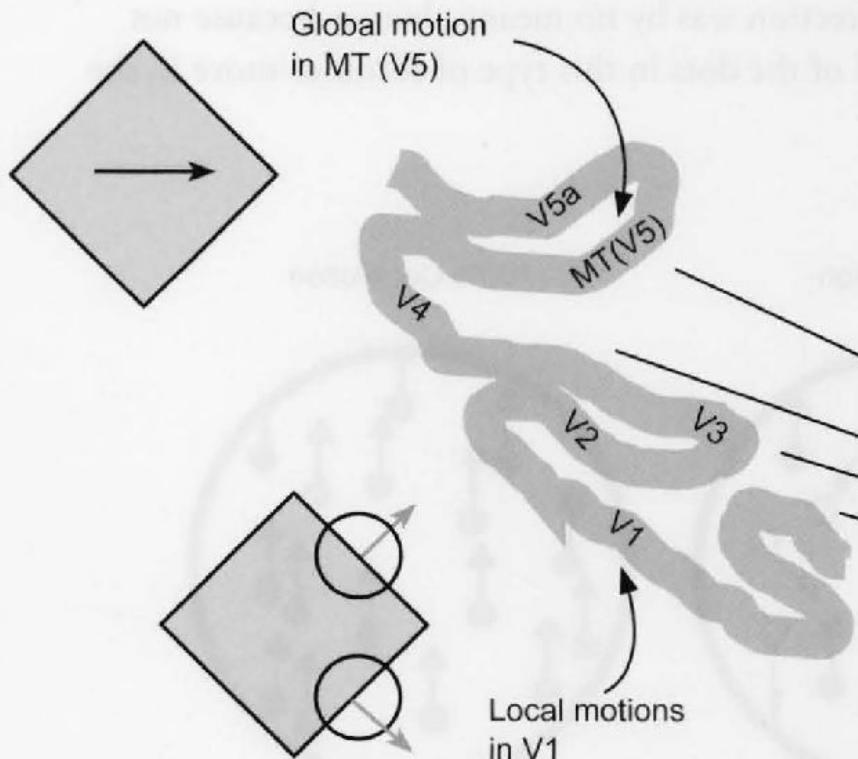






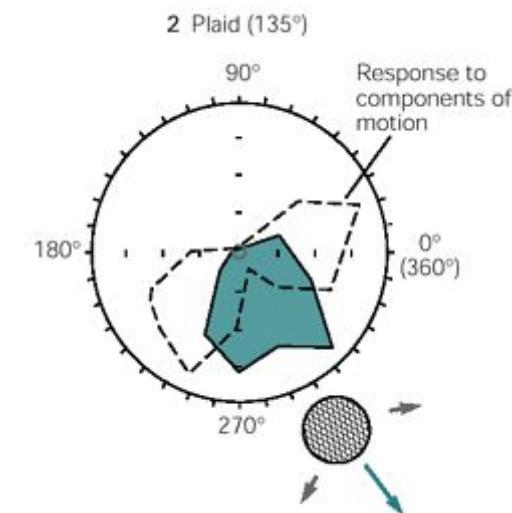
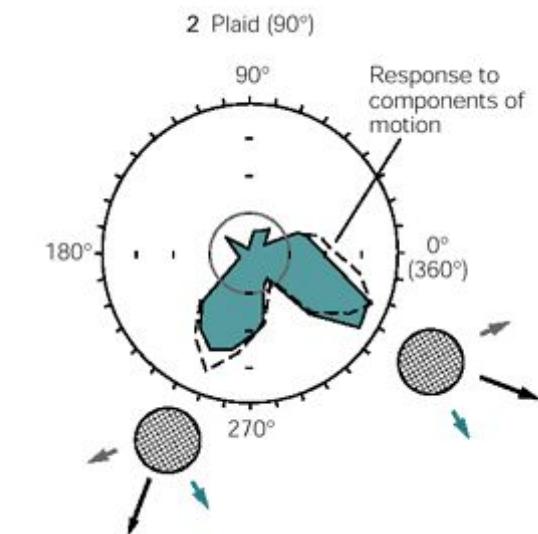
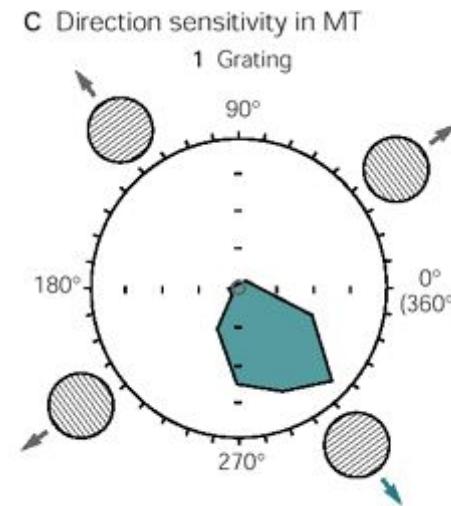
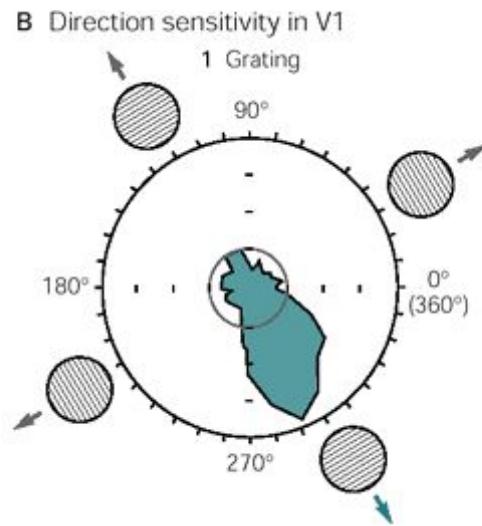
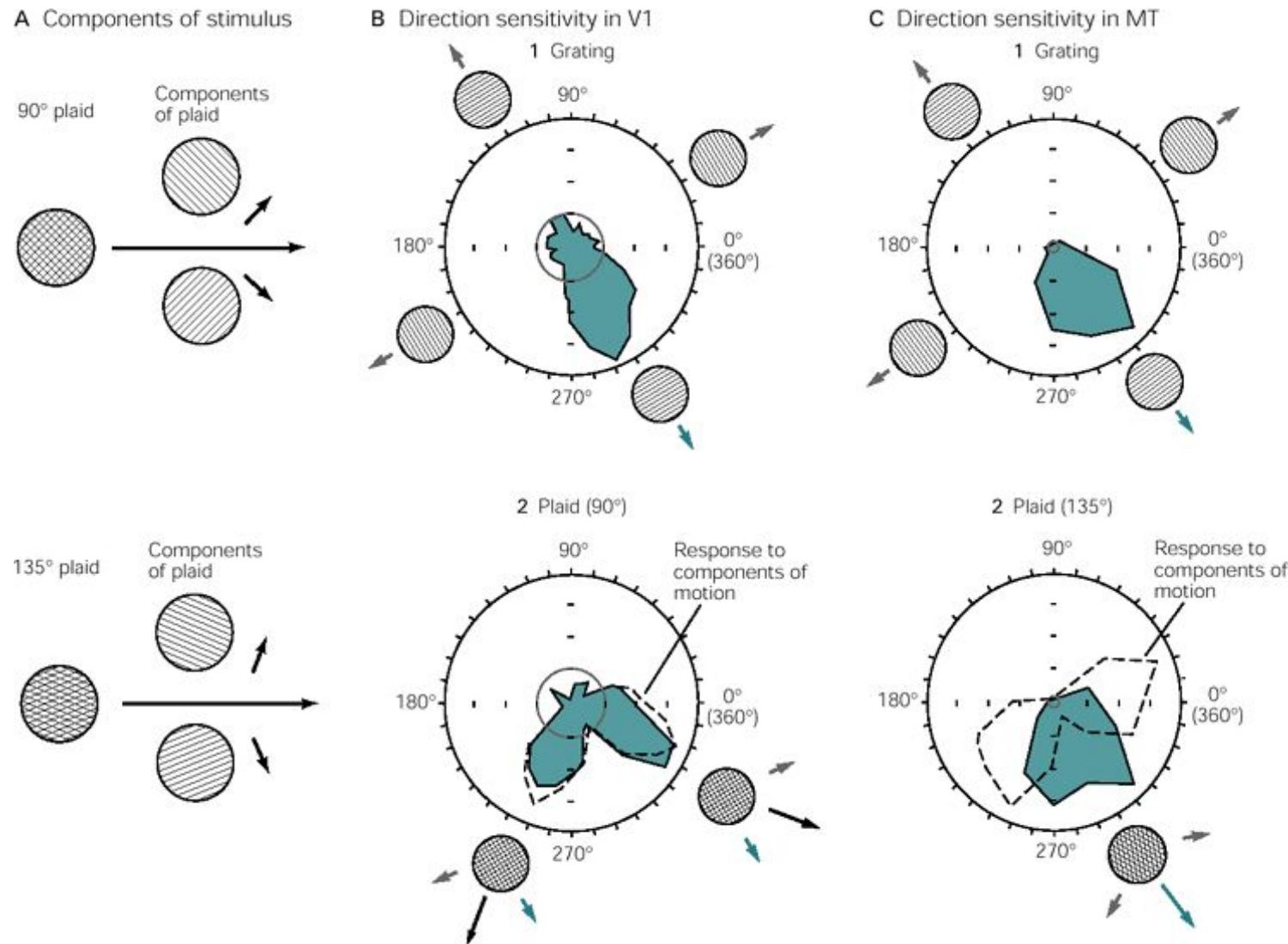


# Movimiento local y global en el cerebro



Striped area shows  
where the enlarged  
cross-section on left is  
located

# Cells in MT Solve the Aperture Problem

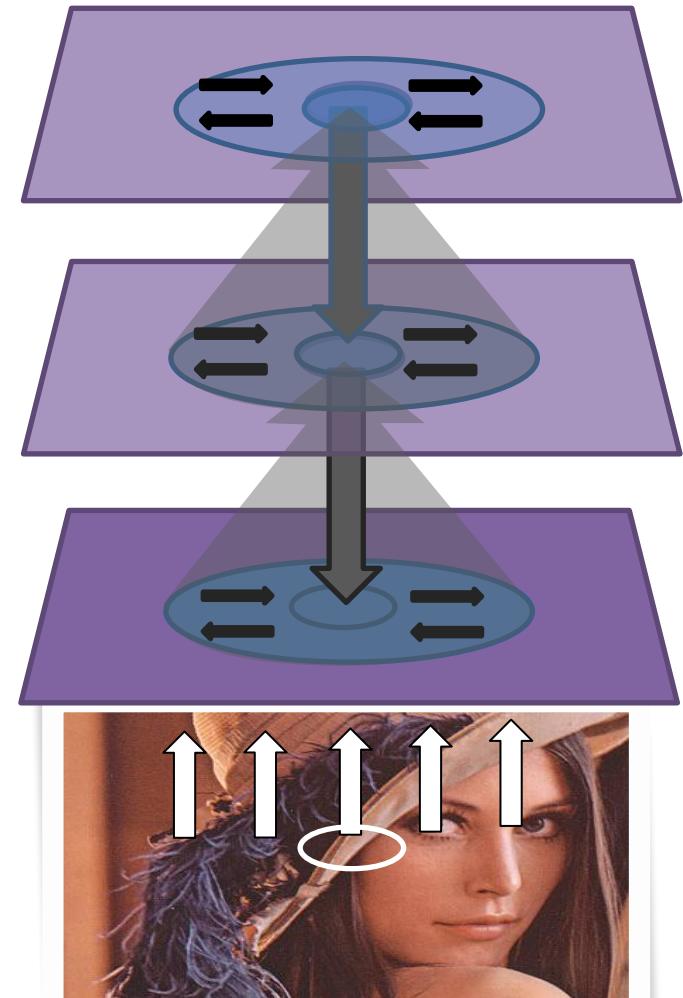


# Sensación y percepción

- Helmholtz (1860) distingue entre percepciones, que son los productos de reconocimiento de las causas de la entrada sensorial, y la sensación en sí.
- Reconocimiento (es decir, inferir las causas de la sensación) es la inversa de la generación de los datos sensoriales a partir de sus causas.
- De ello se desprende que el reconocimiento descansa en modelos aprendidos a partir de la experiencia de cómo las sensaciones son causadas.
- La Inferencia perceptiva emerge de la interacción, de la información mutua, entre procesos de arriba hacia abajo (“top down”) y de abajo hacia arriba (“bottom up”) que permiten que la sensación constriñe la percepción.

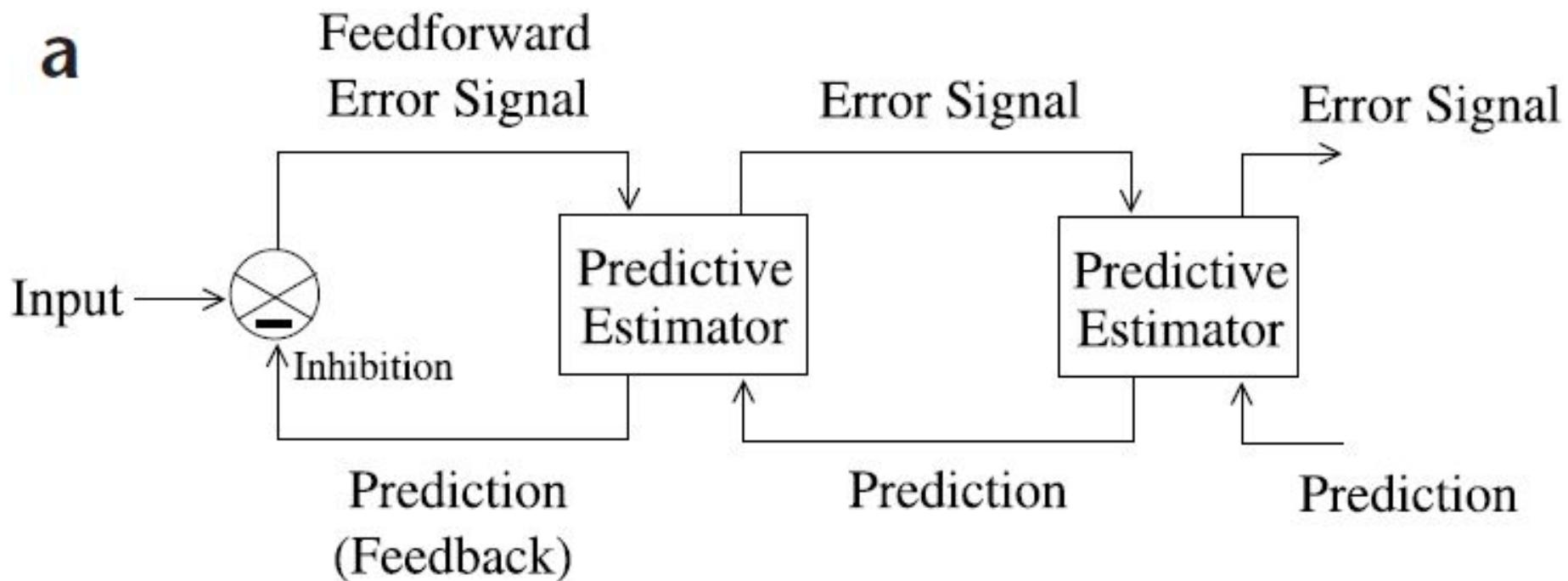
# Sensory inference and Predictive coding

- At the lowest level the current visual scene patterns the light that stimulate sensory receptors
- These signals are then processed via a multilevel cascade in which each level attempts to predict the activity at the level below it via backward connections.
- The backward connections allow the activity at one stage of the processing to return as another input at the previous stage.
- If this successfully predicts the lower level activity no further action needs to ensue but where there is a mismatch, “prediction error” occurs and error-indicating activity is propagated to the higher level.
- This automatically adjusts probabilistic representations at the higher level so that top-down predictions cancel prediction errors at the lower level (yielding rapid perceptual inference).
- At the same time, prediction error is used to adjust the structure of the model so as to reduce any discrepancy next time around (yielding slower timescale perceptual learning).



Rao, Rajesh PN, and Dana H. Ballard. "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects." *Nature neuroscience* 2.1 (1999): 79-87.

# Predictive coding

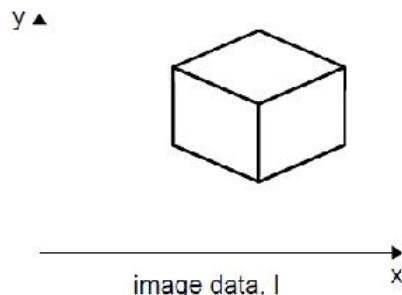


Rao, Rajesh PN, and Dana H. Ballard. "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects." *Nature neuroscience* 2.1 (1999): 79-87.

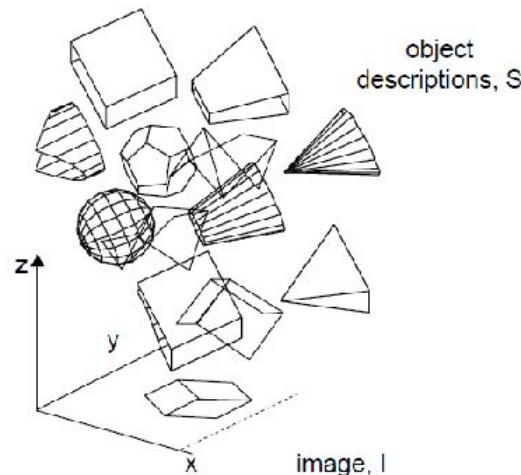
# Percepción Bayesiana

Which object description generated  
the image data?

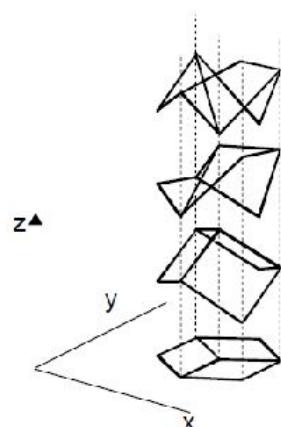
(a)



(b)

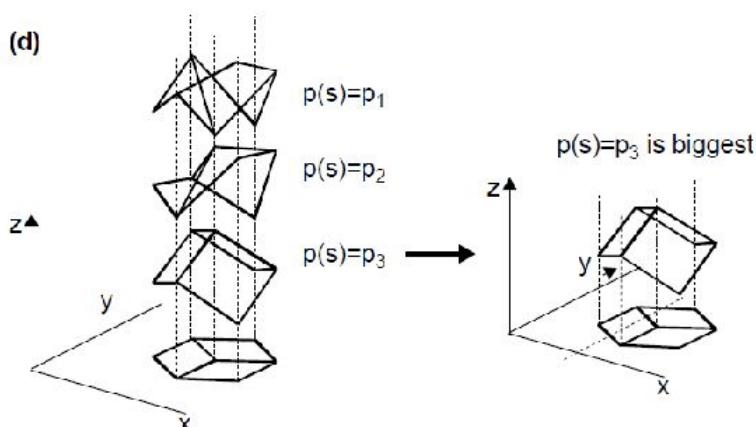


(c)



Bayesian solution

(d)



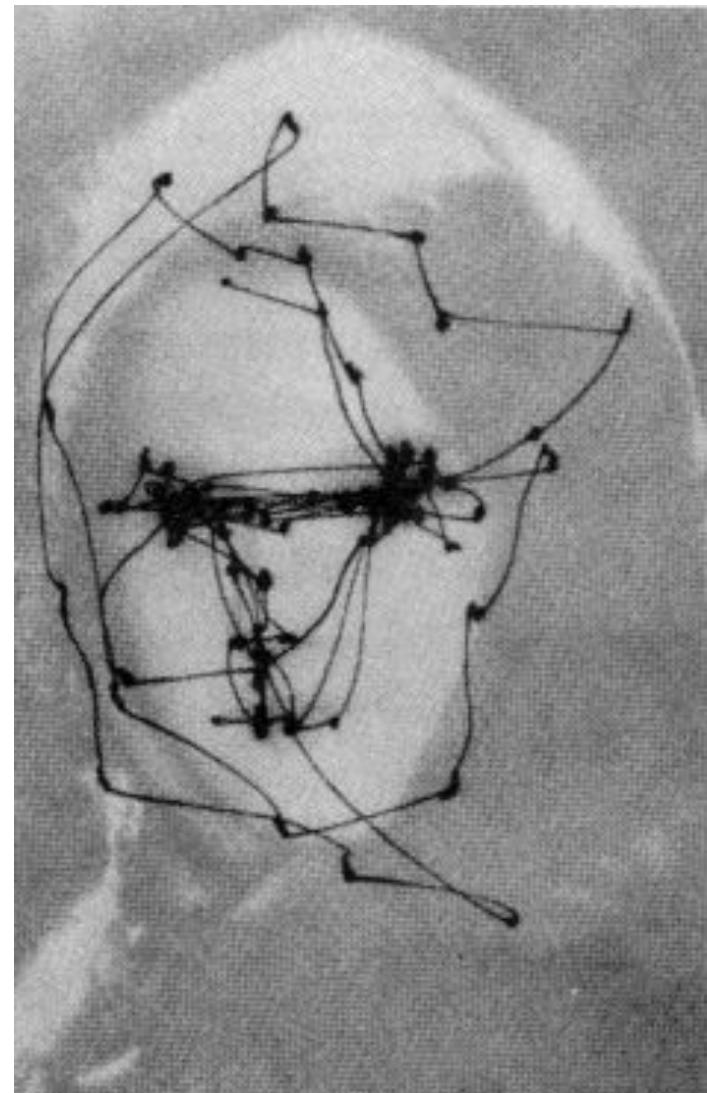
likelihood,  $p(I|S)$ , narrows selection  
consistent with projection

prior,  $p(S)$ , further narrows selection

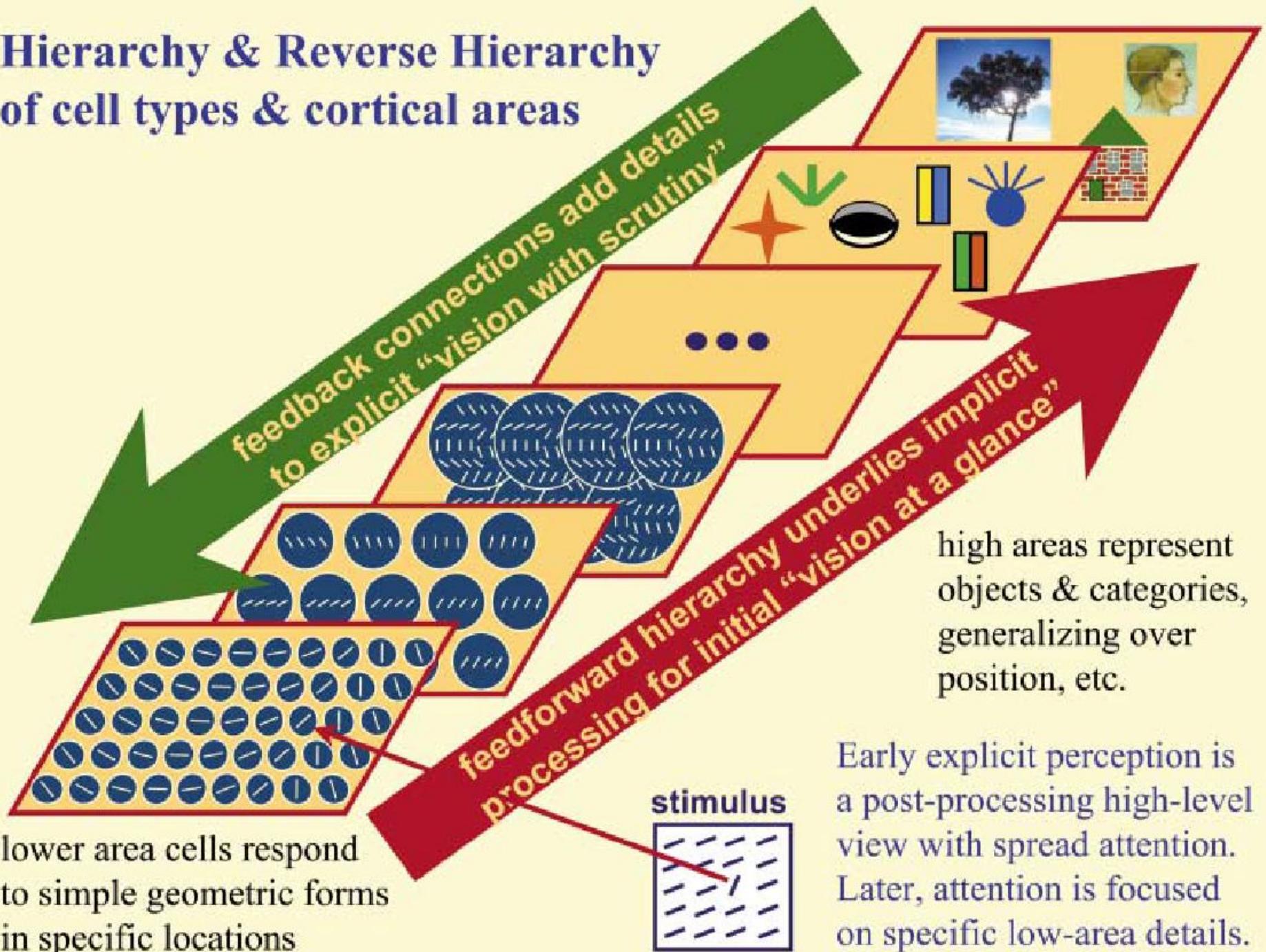
**FIN**

- El dogma era que una neurona codifica la configuración del estímulo que maximiza su frecuencia de disparo

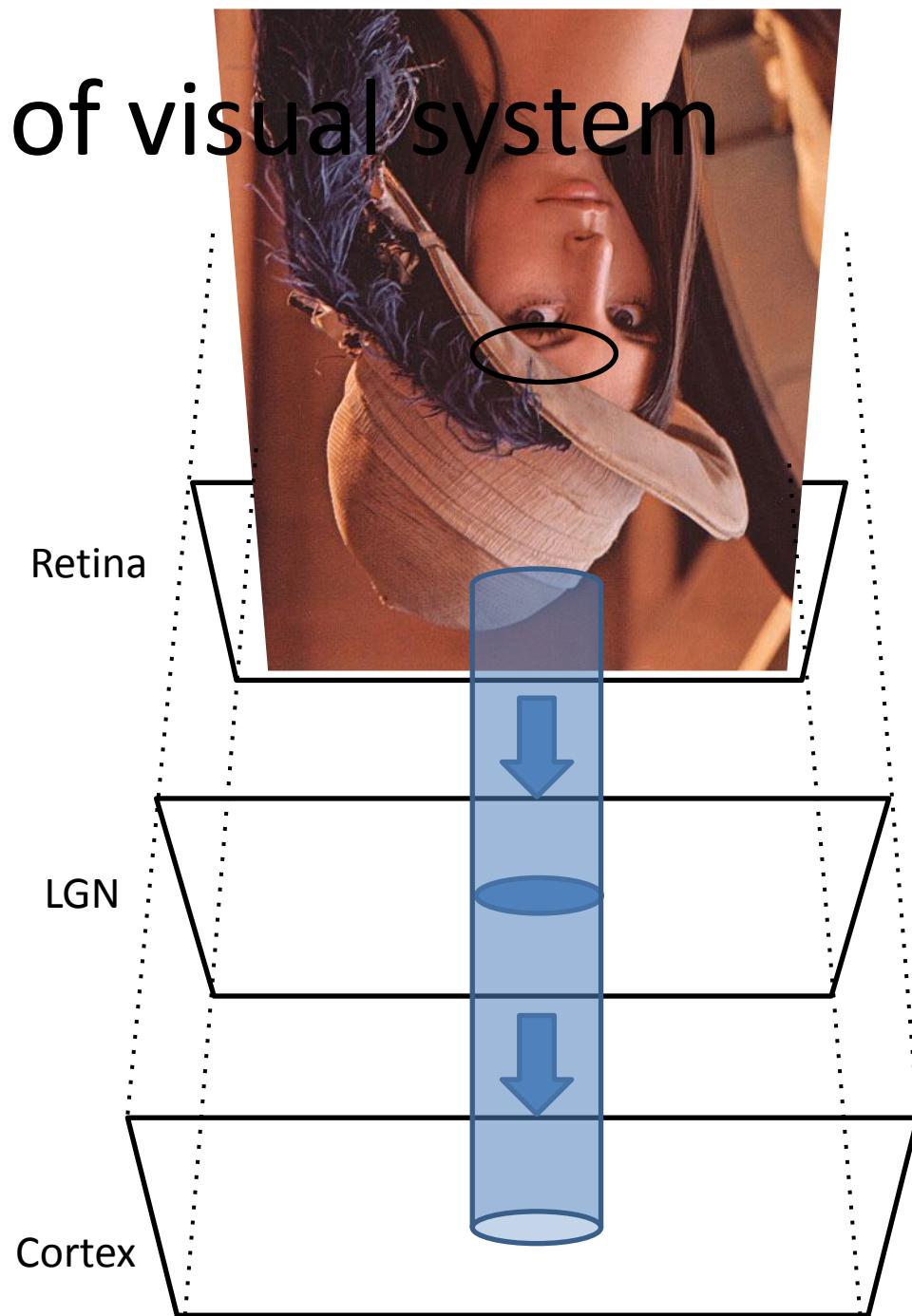
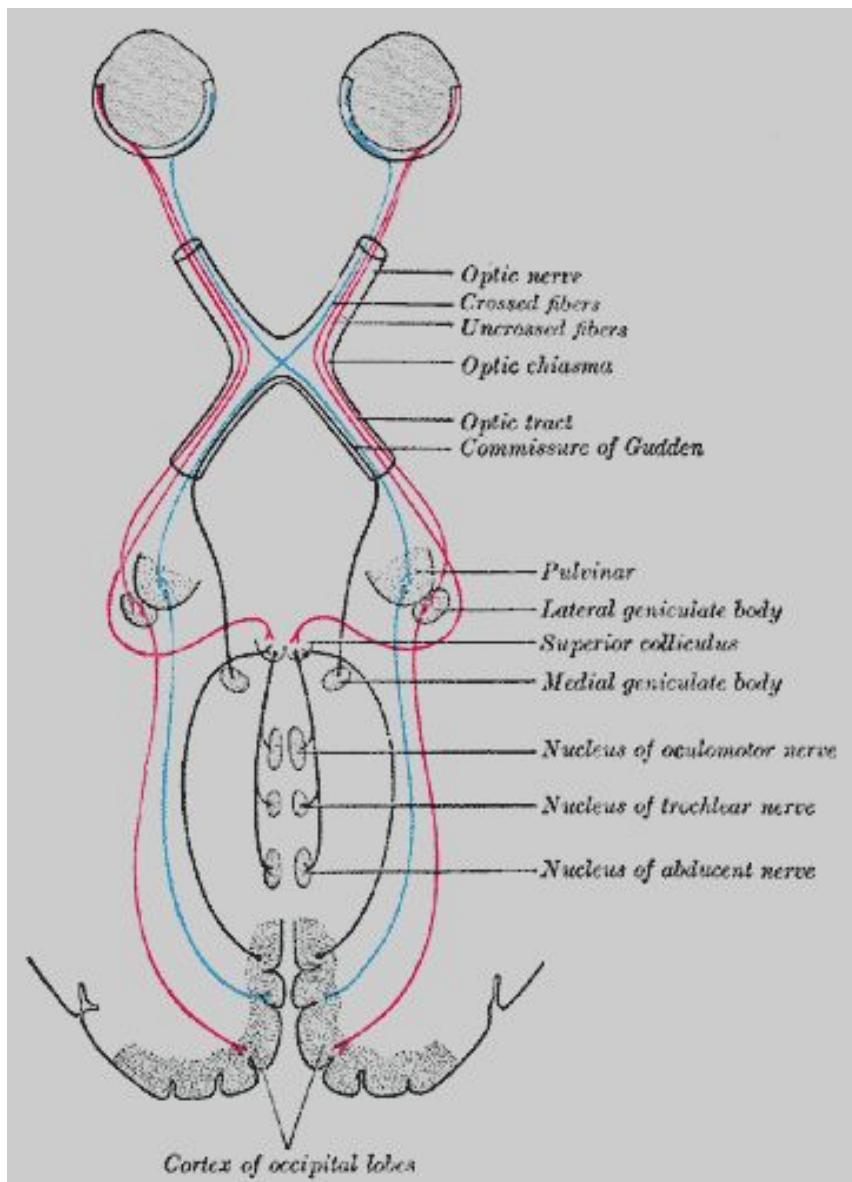
# Visión activa. Yarbus



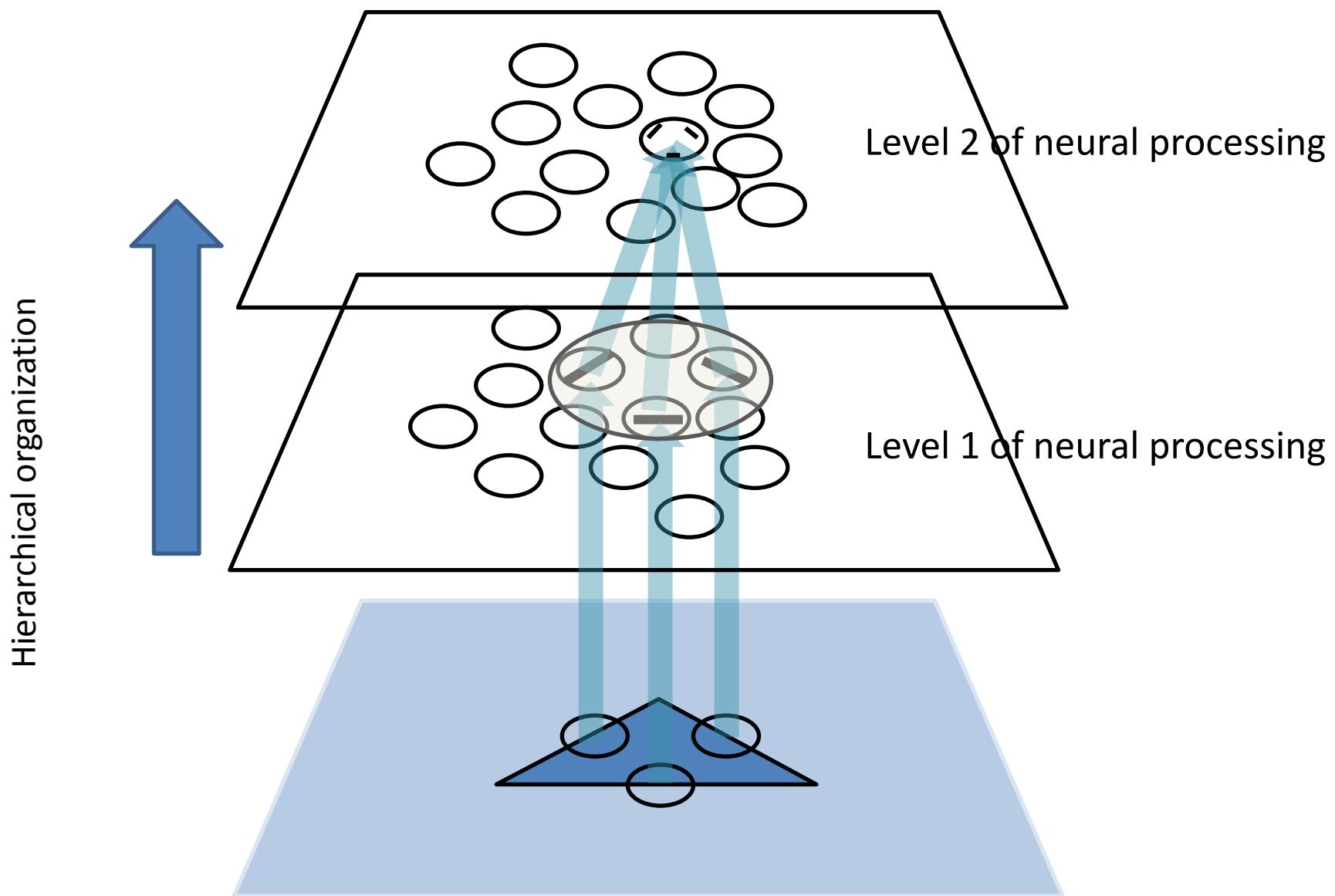
# Hierarchy & Reverse Hierarchy of cell types & cortical areas



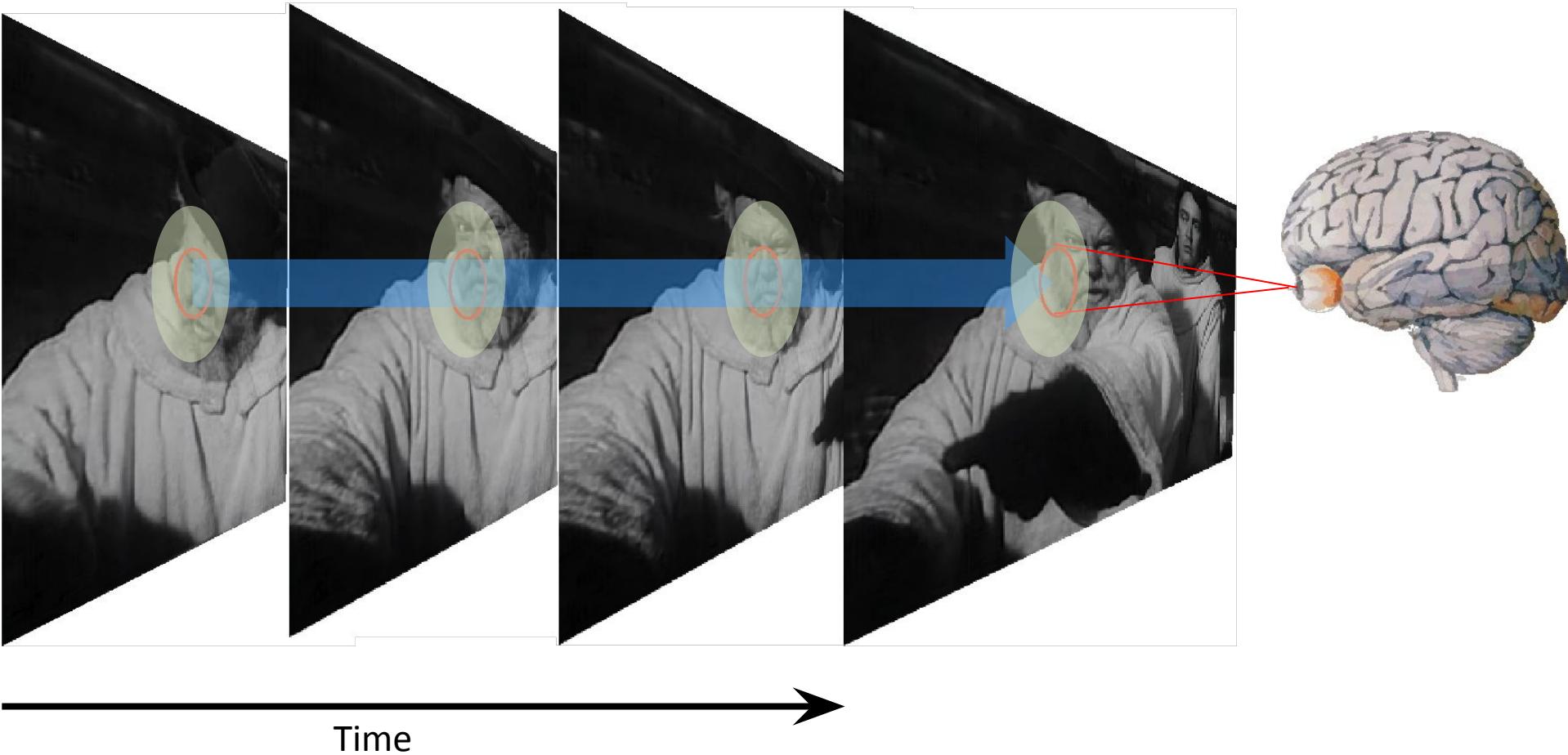
# Classic view of visual system



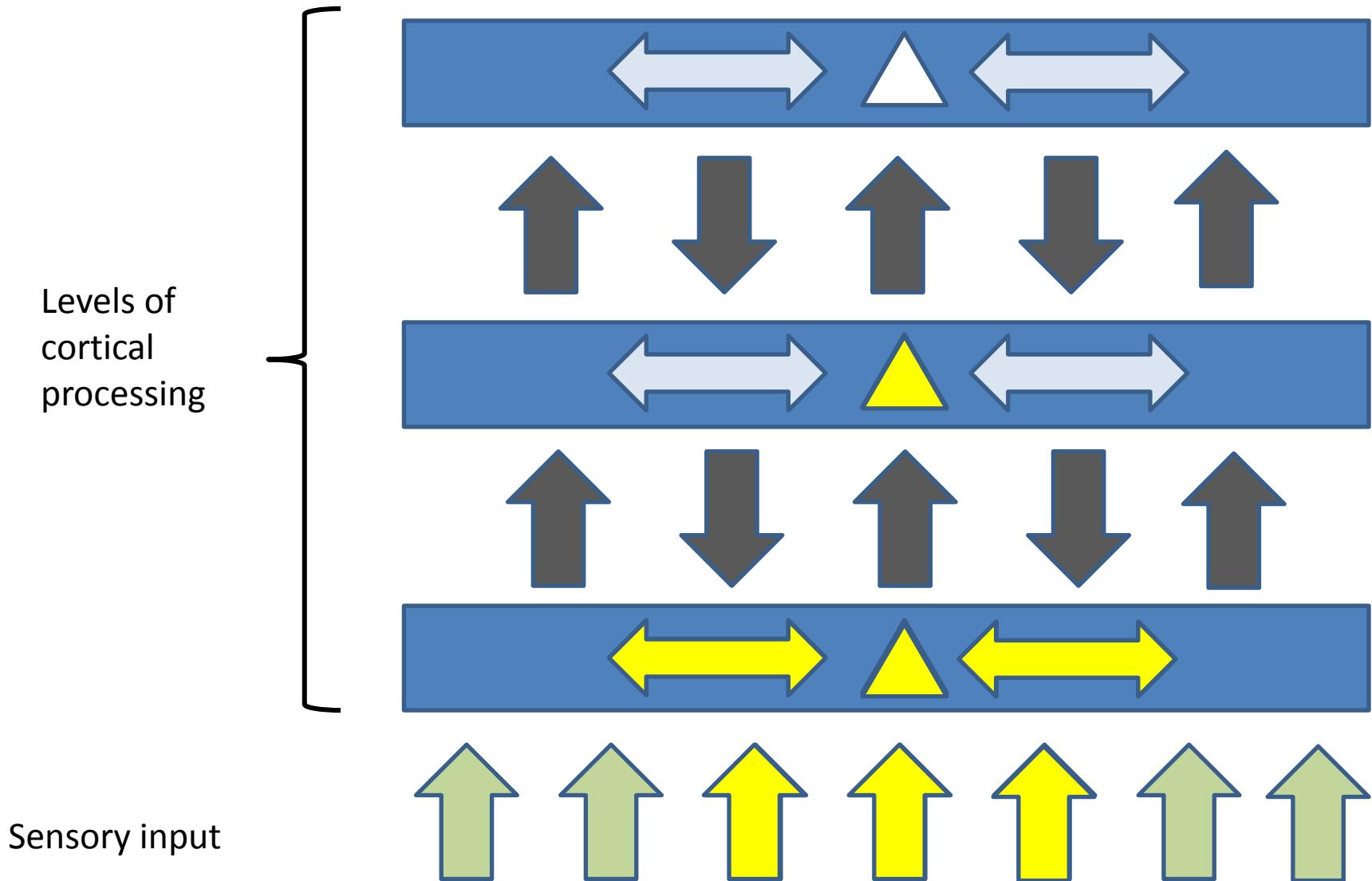
# Classic view of sensory functioning



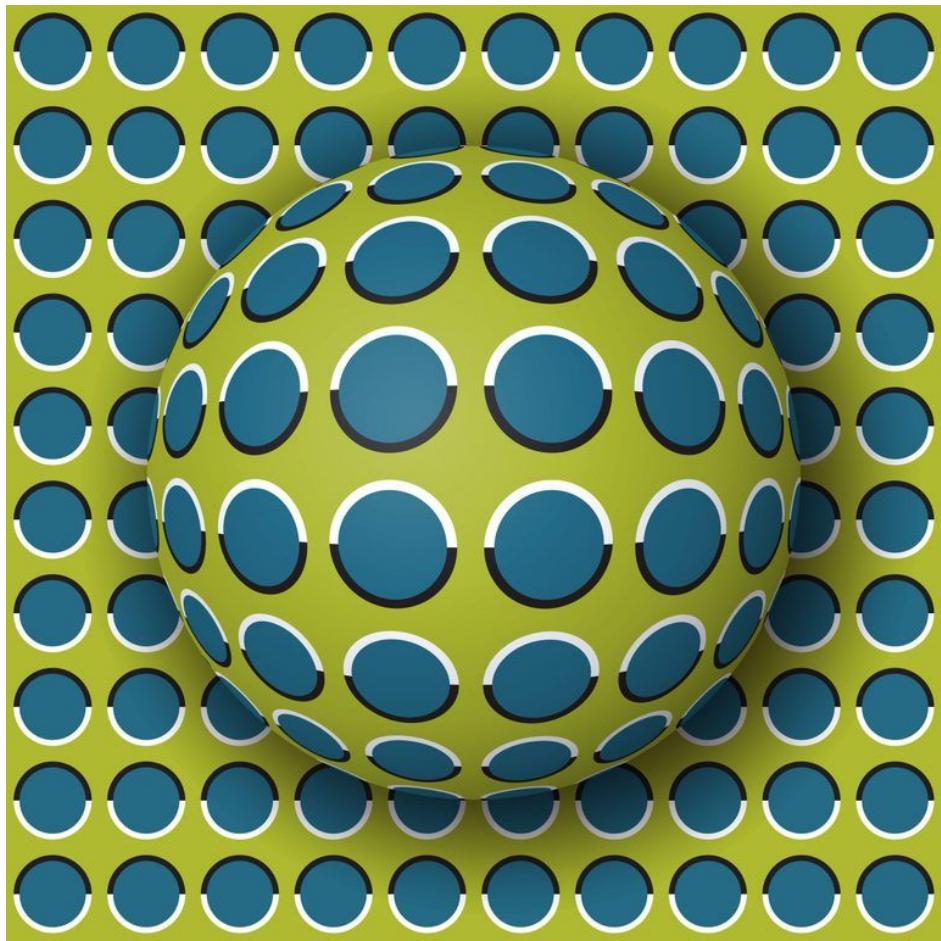
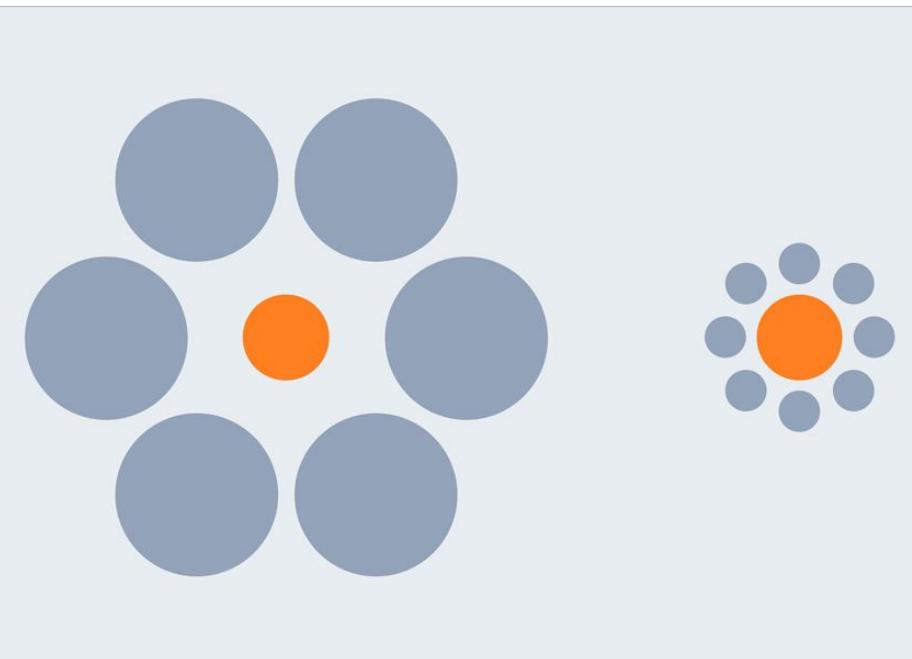
# Taking full advantage of the information available



# Current model of cortical processing



# There is no direct access to reality and the sensory report can fail: Visual illusions



Like in Matrix the activity of some neurons in our brains makes us believe something about the world that is not true.

# Los niveles de Marr

David Marr (Marr, 1982) distingue tres niveles de análisis en el estudio de sistemas cognitivos:

- (1) el proceso computacional realizado por el cerebro,
- (2) su instancia algorítmica, y
- (3) el sustrato físico de la implementación