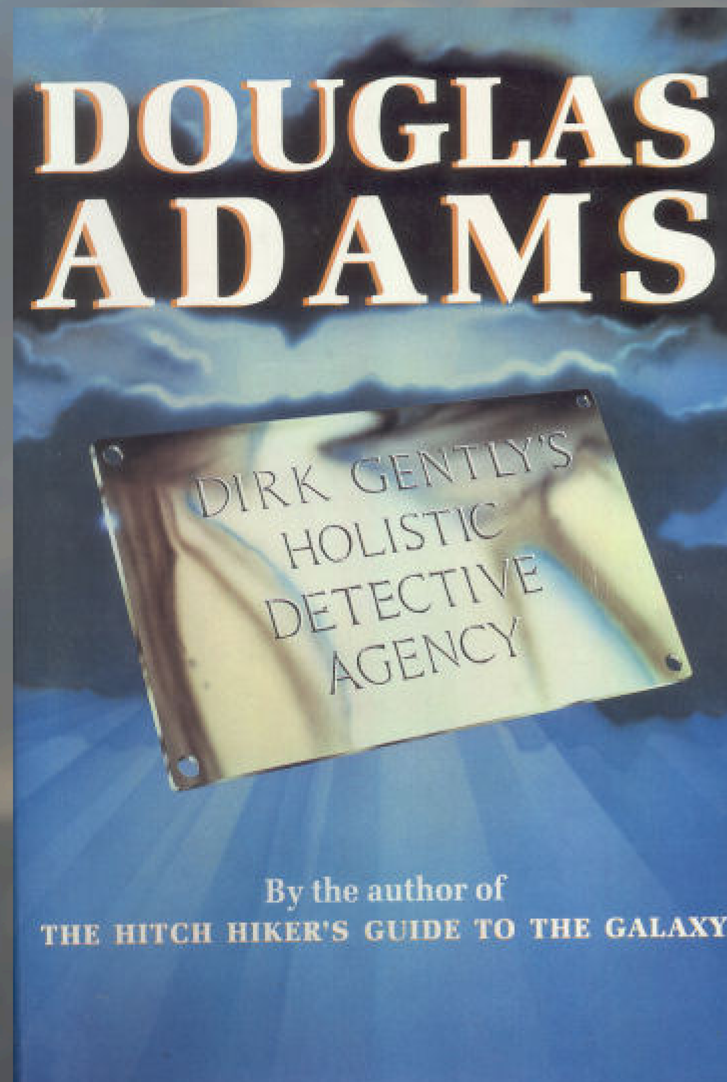
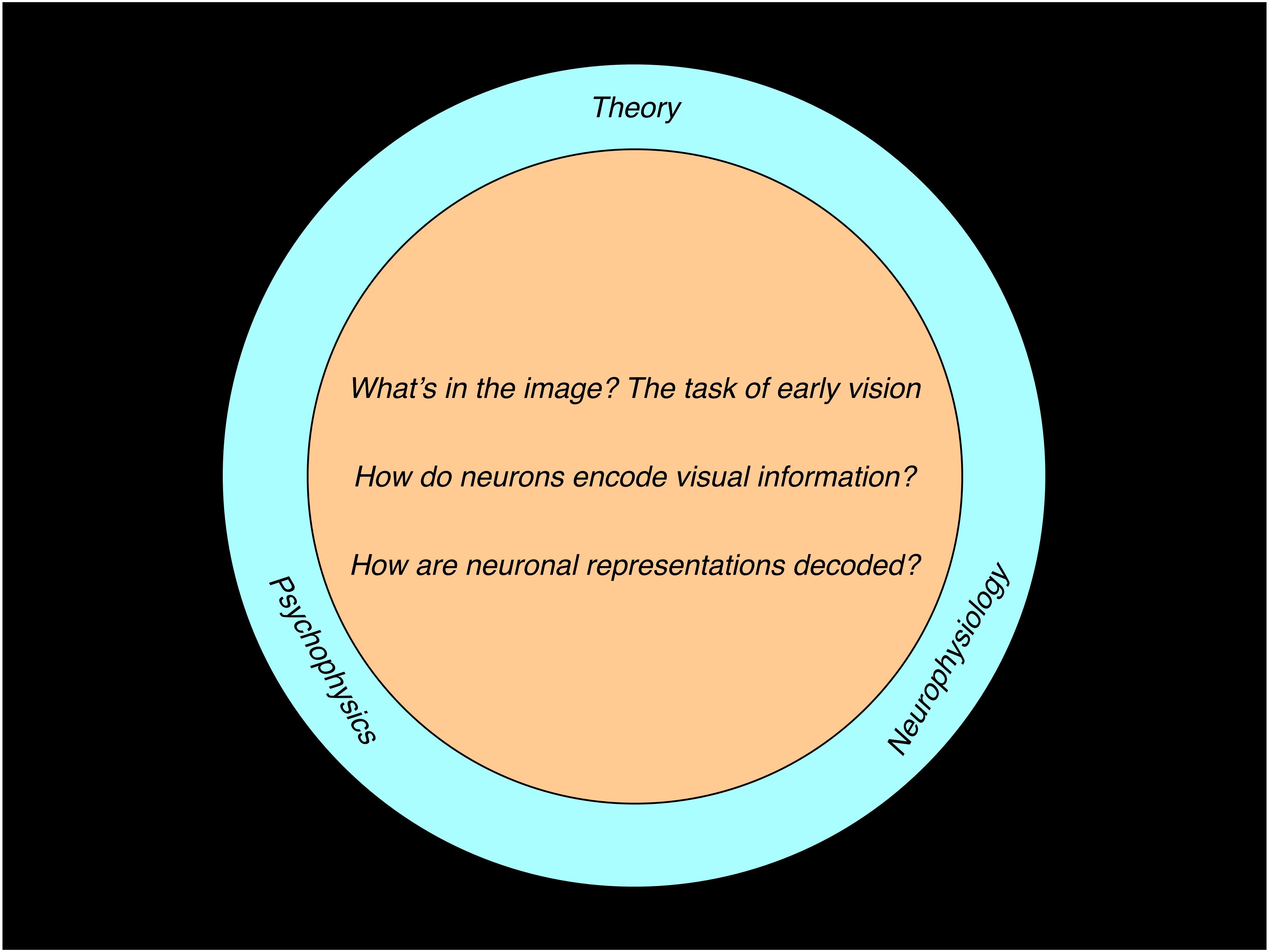
The elements of early vision

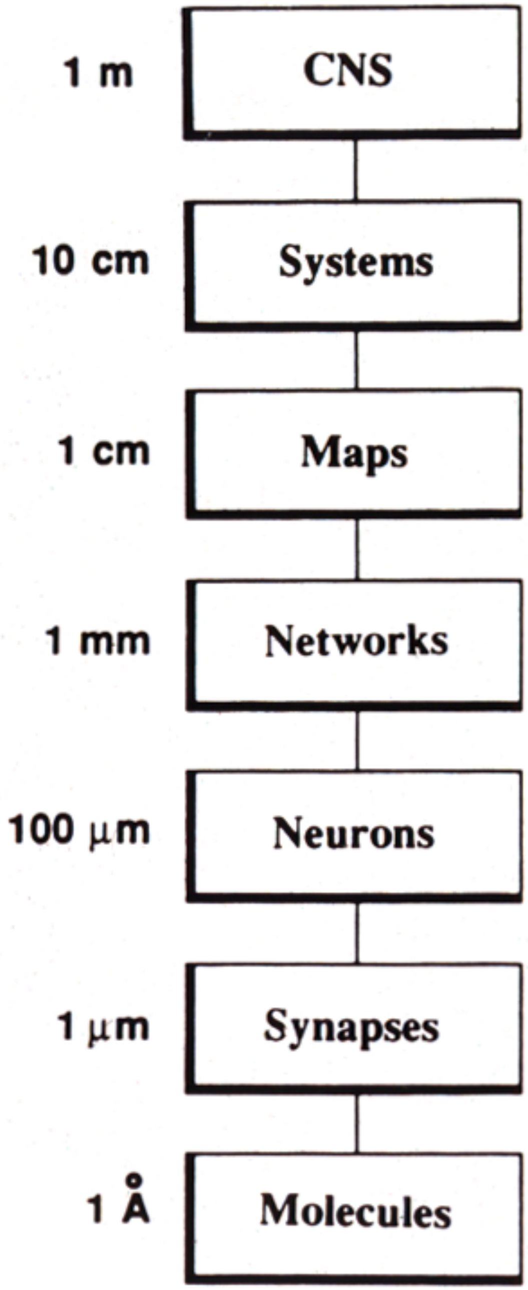
or, what vision (and this course) is all about

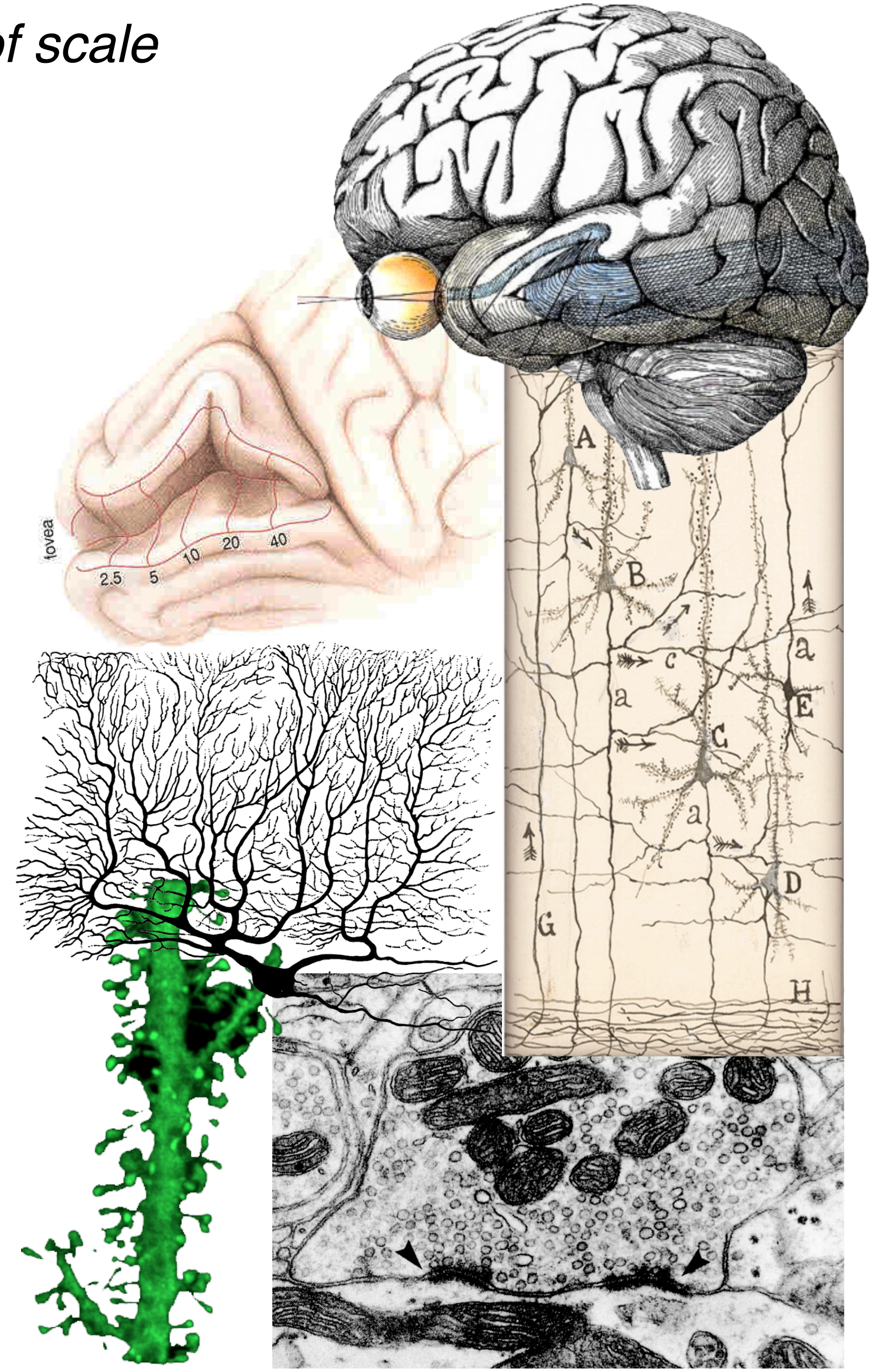
The Electric Monk was a labor saving device, like a dishwasher or a video recorder. Dishwashers washed tedious dishes for you, thus saving you the bother of washing them yourself, video recorders watched tedious television for you, thus saving you the bother of looking at it yourself; Electric Monks believed things for you, thus saving you what was becoming an increasingly onerous task, that of believing all the things the world expected you to believe.



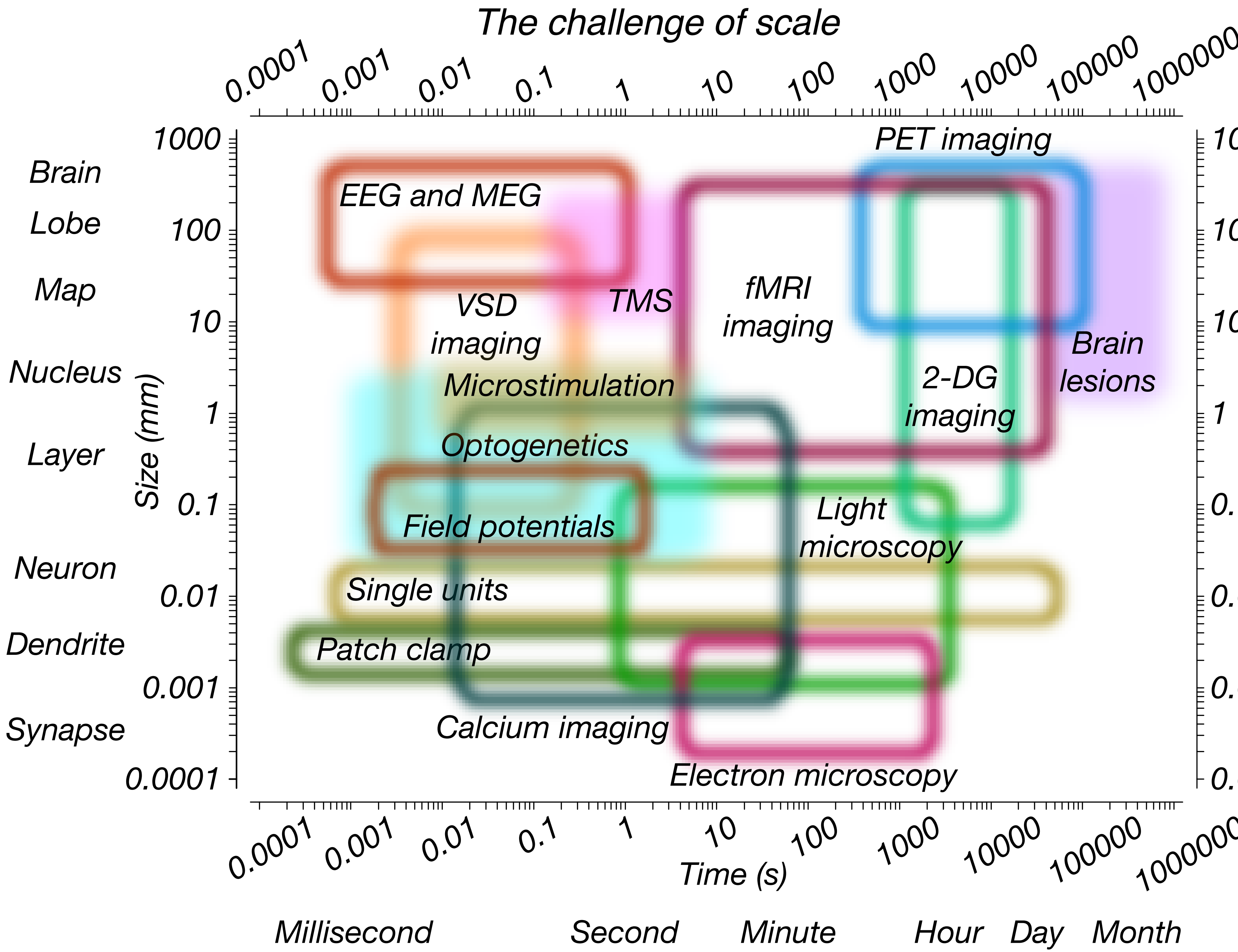

Theory
What's in the image? The task of early vision
How do neurons encode visual information?
How are neuronal representations decoded?
Psychophysics
Neurophysiology

The challenge of scale





after Churchland and Sejnowski, 1988

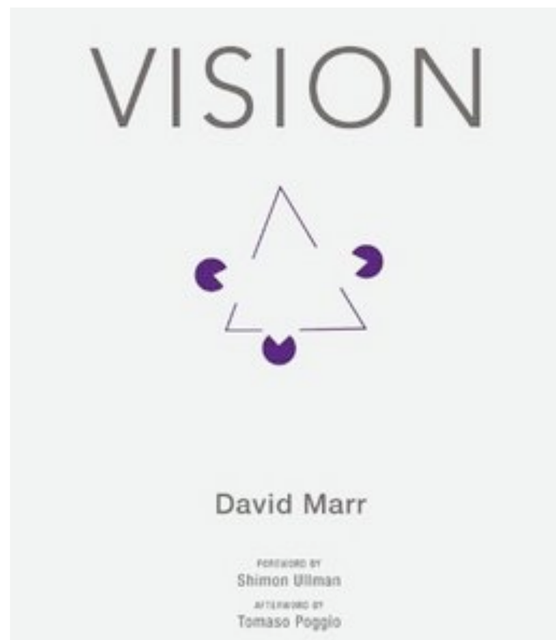


Sejnowski, Churchland & Movshon (2014) after Churchland & Sejnowski (1988)

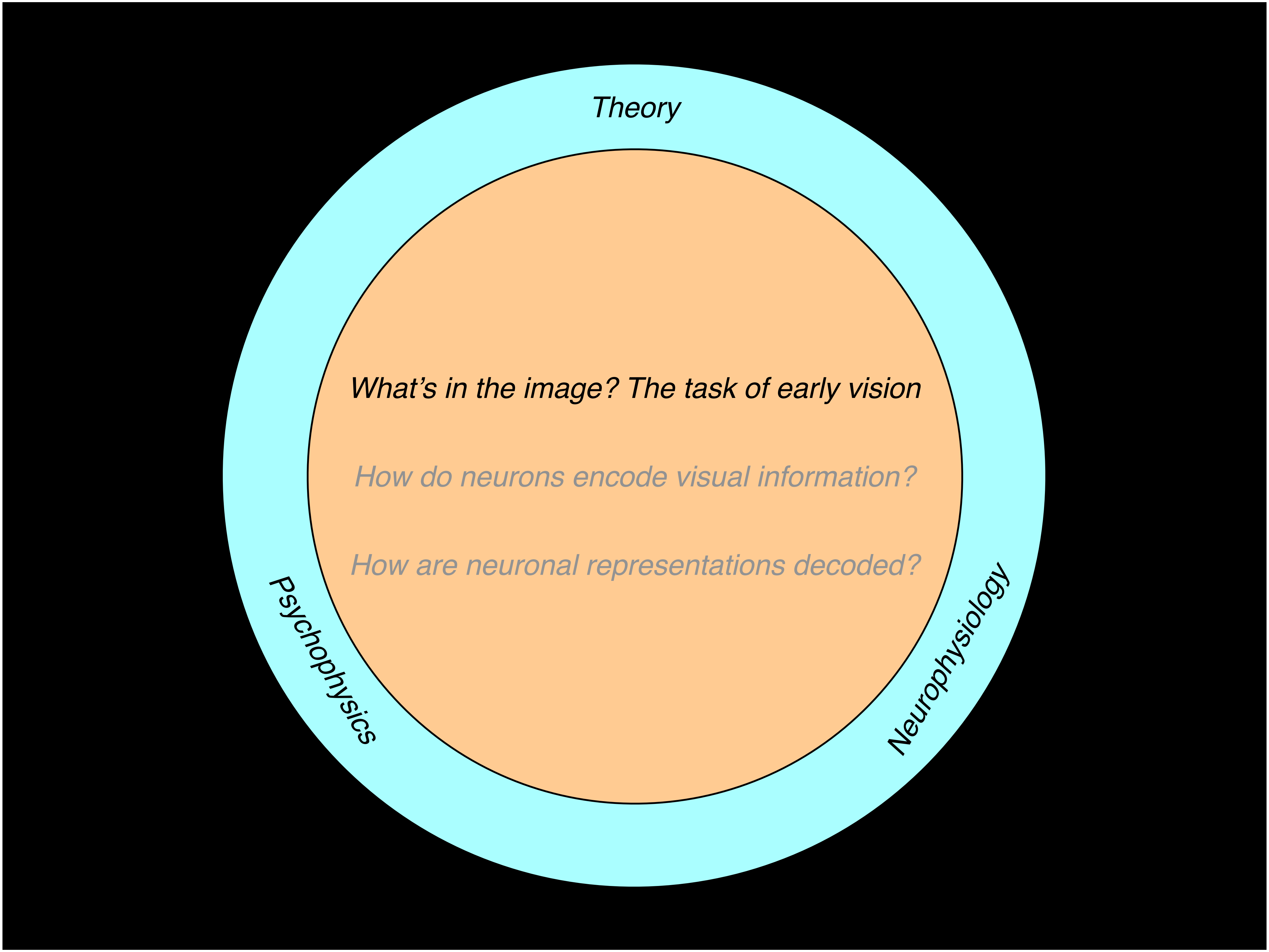
The challenge of level

Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

David Courtnay Marr
(1946-1980)



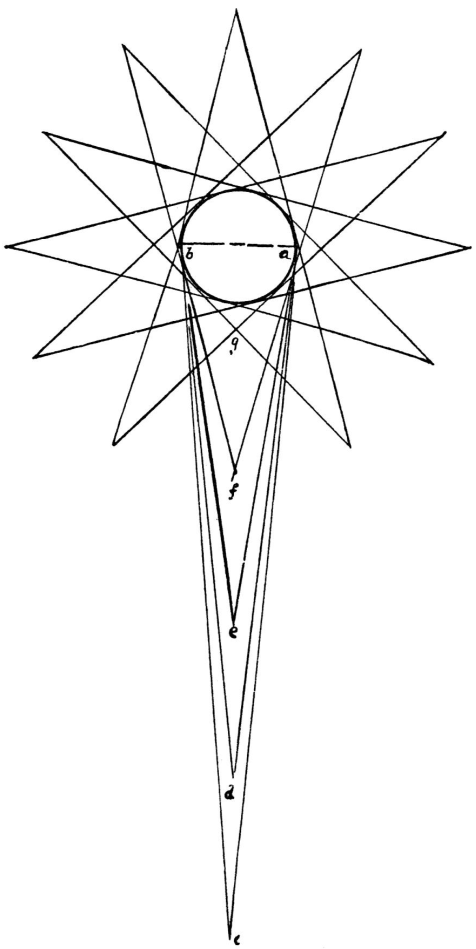
Marr, 1982


Theory
What's in the image? The task of early vision
How do neurons encode visual information?
How are neuronal representations decoded?
Psychophysics
Neurophysiology

The Plenoptic Function and the Elements of Early Vision

Edward H. Adelson and James R. Bergen

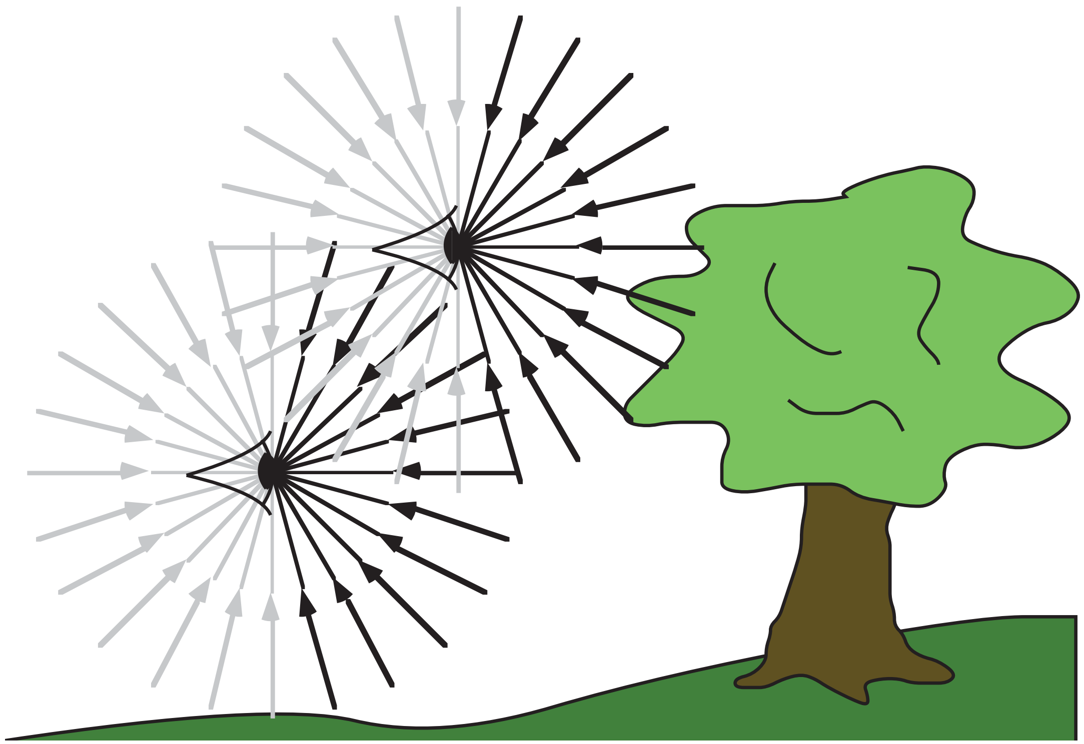
In M. Landy and J. A. Movshon (eds), *Computational Models of Visual Processing* (pp. 3-20). Cambridge, MA: MIT Press (1991).



Every body in light and shade fills the surrounding air with infinite images of itself; and these, by infinite pyramids diffused in the air, represent this body throughout space and on every side. Each pyramid that is composed of a long assemblage of rays includes within itself an infinite number of pyramids and each has the same power as all, and all as each.

– The Notebooks of Leonardo da Vinci

The plenoptic function describes all the information available to an observer anywhere in space and time



What is there to see?

Black and white photo: x, y
Black and white movie: x, y, t
Color movie: x, y, t, λ
Holographic movie: x, y, t, λ, V_x, V_y, V_z

Practical applications

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 14, NO. 2, FEBRUARY 1992

Single Lens Stereo with a Plenoptic Camera

Edward H. Adelson and John Y.A. Wang

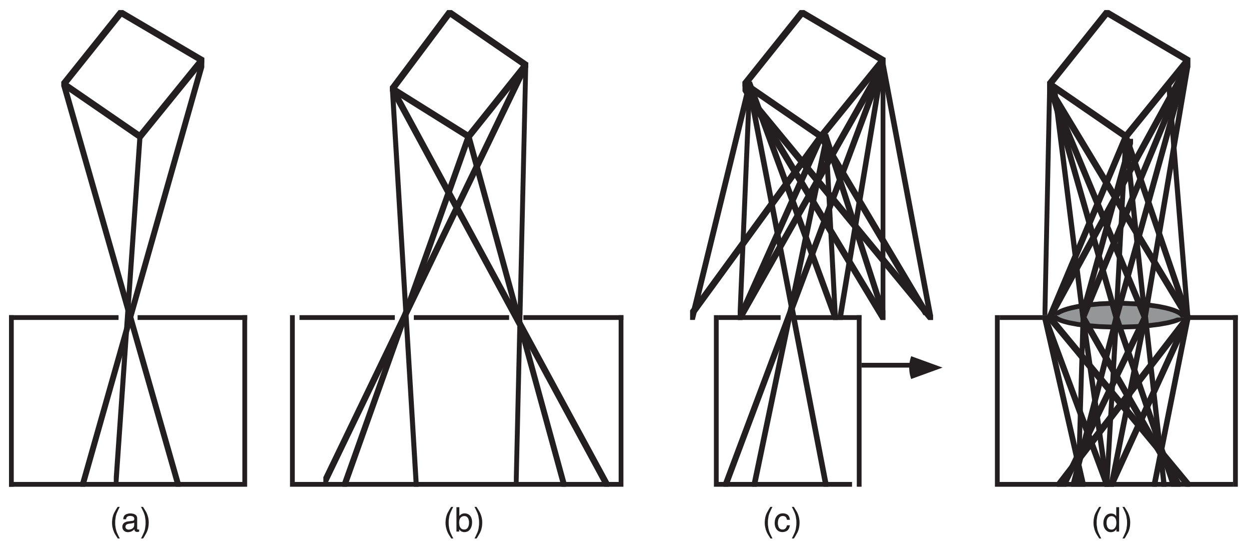


Fig. 2. (a) Pinhole camera forms an image from a single viewpoint; (b) in a stereo system, two images are formed from different viewpoints; (c) in a motion parallax system, a sequence of images are captured from many adjacent viewpoints; (d) a lens gathers light from a continuum of viewpoints; in an ordinary camera these images are averaged at the sensor plane.

Practical applications

LYTRO

ILLUM | Apps | Gallery | Learn | Shop Lytro

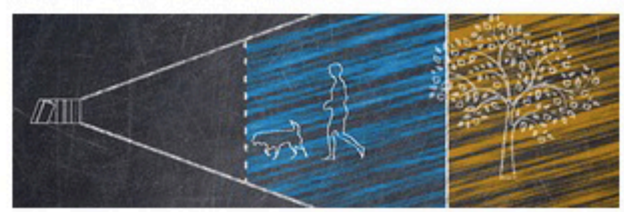
The invention of photography gave us the power to capture our most beautiful, tragic and thought-provoking moments. But traditional cameras only capture part of a story. What if we could look deeper? What if we could tell the full story of a moment captured in time?

Now, with a single exposure from Lytro ILLUM, you can adjust aperture, use selective focus, tilt, focus spread, animation and even 3D to differentiate yourself with multiple deliverables.

Lytro ILLUM Technical Specs

Lytro ILLUM in action

Explore Lytro ILLUM tutorials



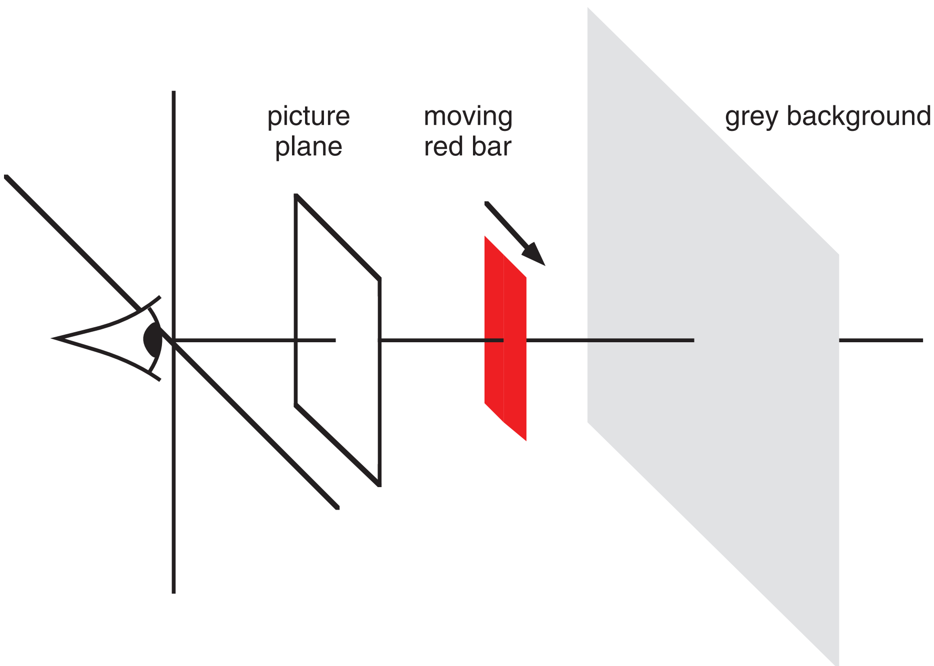
What is there to see?

Black and white photo: x, y
Black and white movie: x, y, t
Color movie: x, y, t, λ
Holographic movie: x, y, t, λ, V_x, V_y, V_z

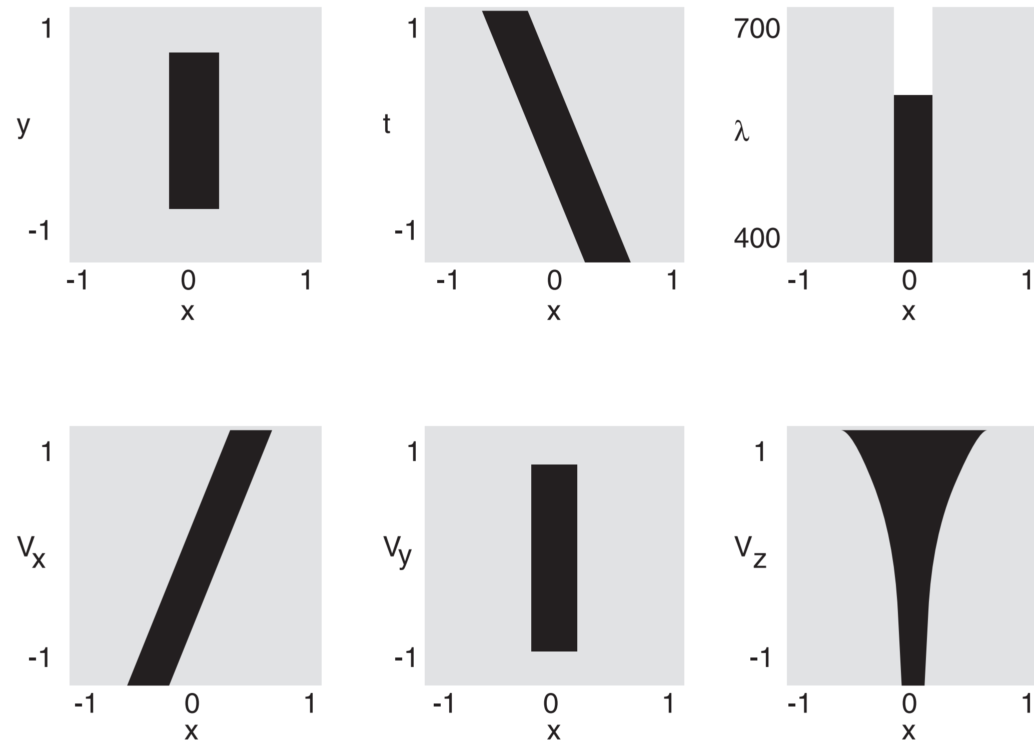
Hence, the plenoptic function: $\boldsymbol{P(x, y, t, \lambda, V_x, V_y, V_z)}$

Proposition 1. *The task of early vision is to deliver a small set of useful measurements about each observable location in the plenoptic function.*

Proposition 2. *The elemental operations of early vision measure local change along various directions within the plenoptic function.*

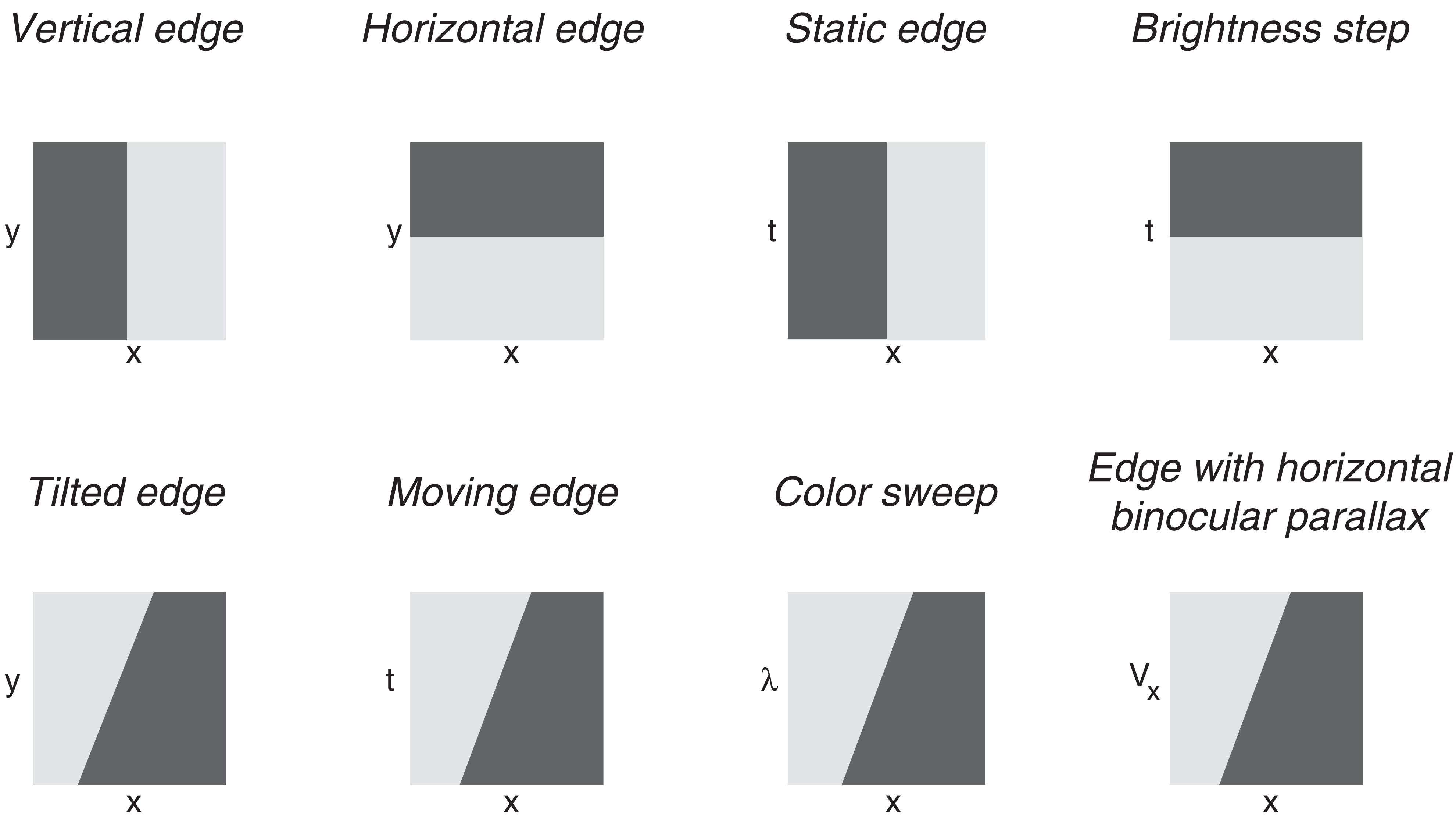


A hypothetical scene that produces a variety of simple plenoptic structures.



The plenoptic structures found along various planes. Each panel represents a slice through the plenoptic function.

Some edge-like structures found in particular planes within the plenoptic function



Local derivatives will turn out to be handy

Visual Pattern Analysis in Machines and Animals

H. B. Barlow, R. Narasimhan, and A. Rosenfeld

SCIENCE 18 August 1972, Volume 177, Number 4049

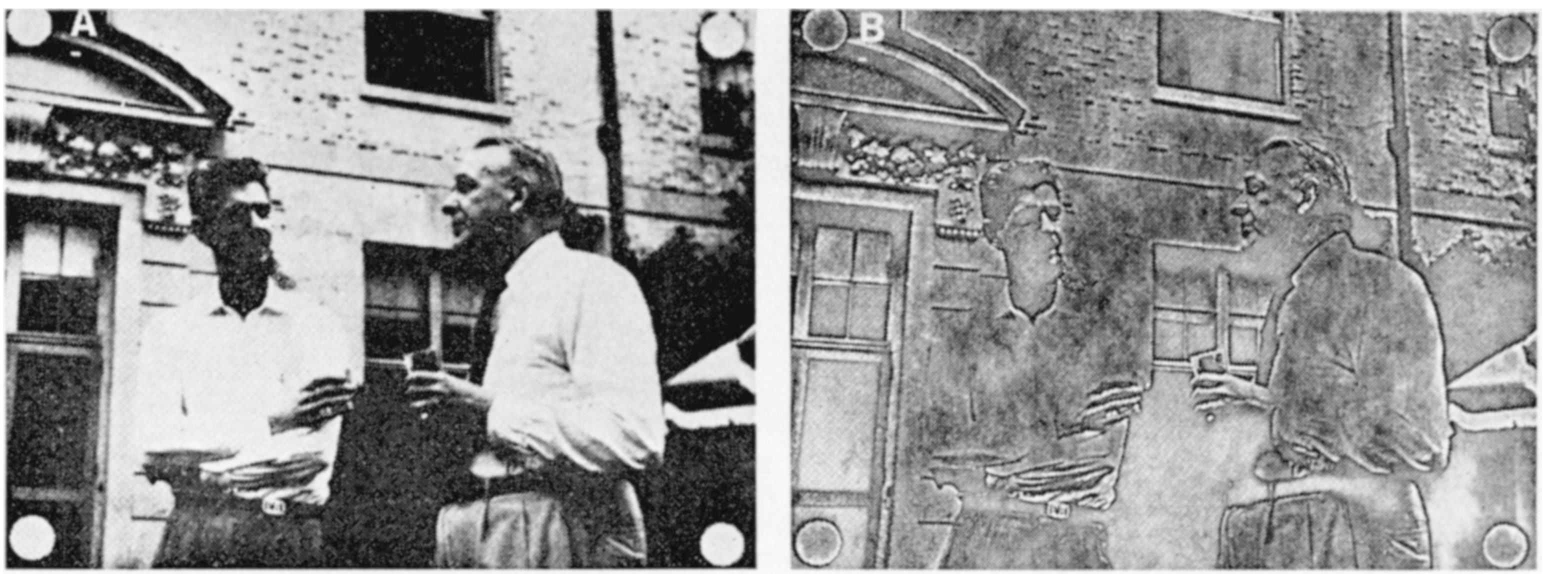
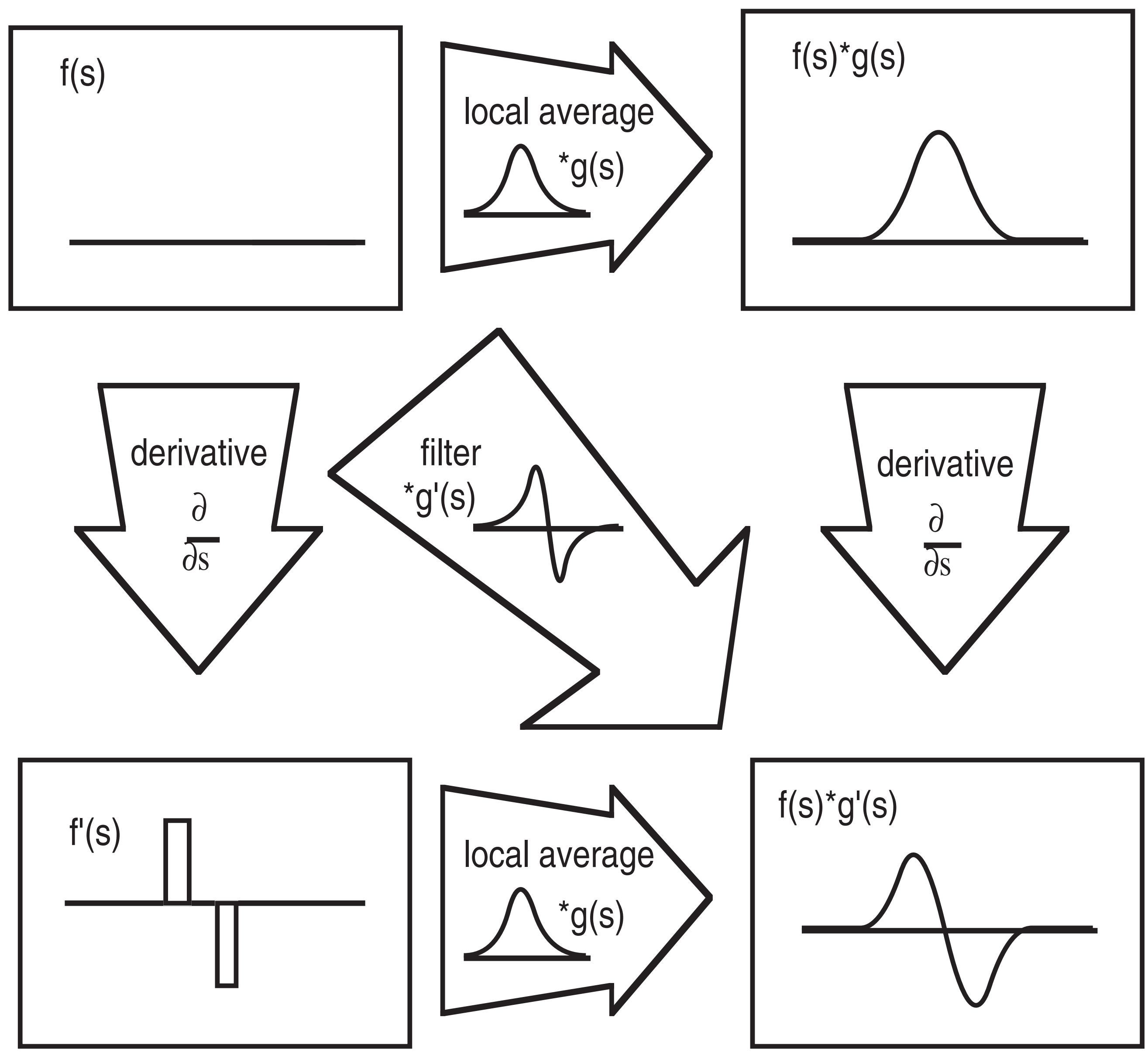
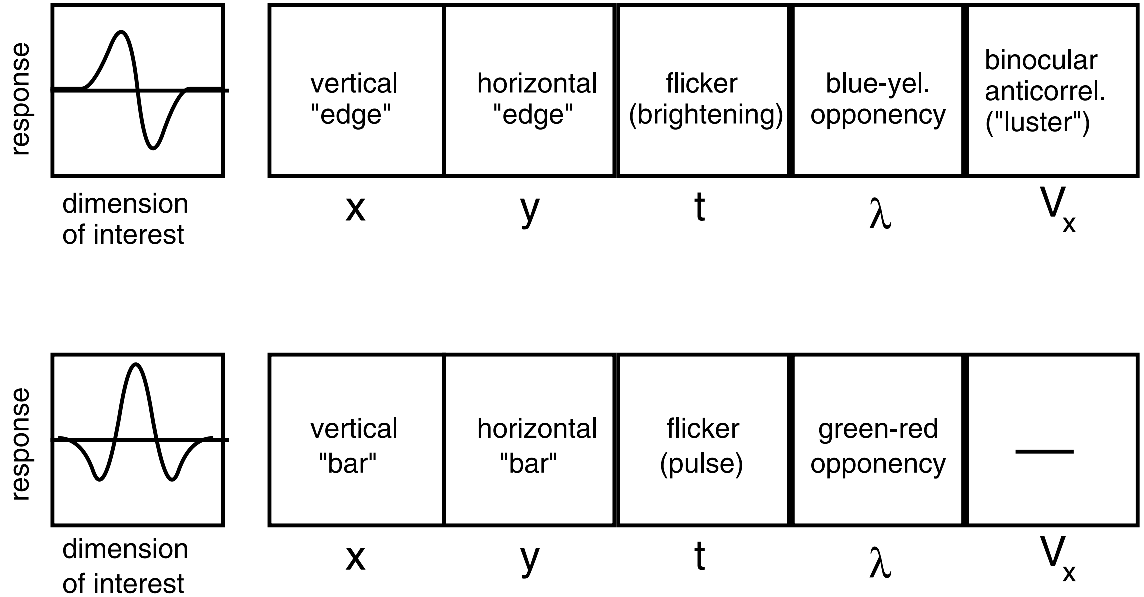


Fig. 2. Example of unsharp masking. (Left) Original scene. (Right) The same scene after superimposing a defocused negative. The person on the left is F. Attneave, who first pointed out the relation between Gestalt principles and redundancy reduction (*34*). [Courtesy of M.I.T. Press (*26*)]

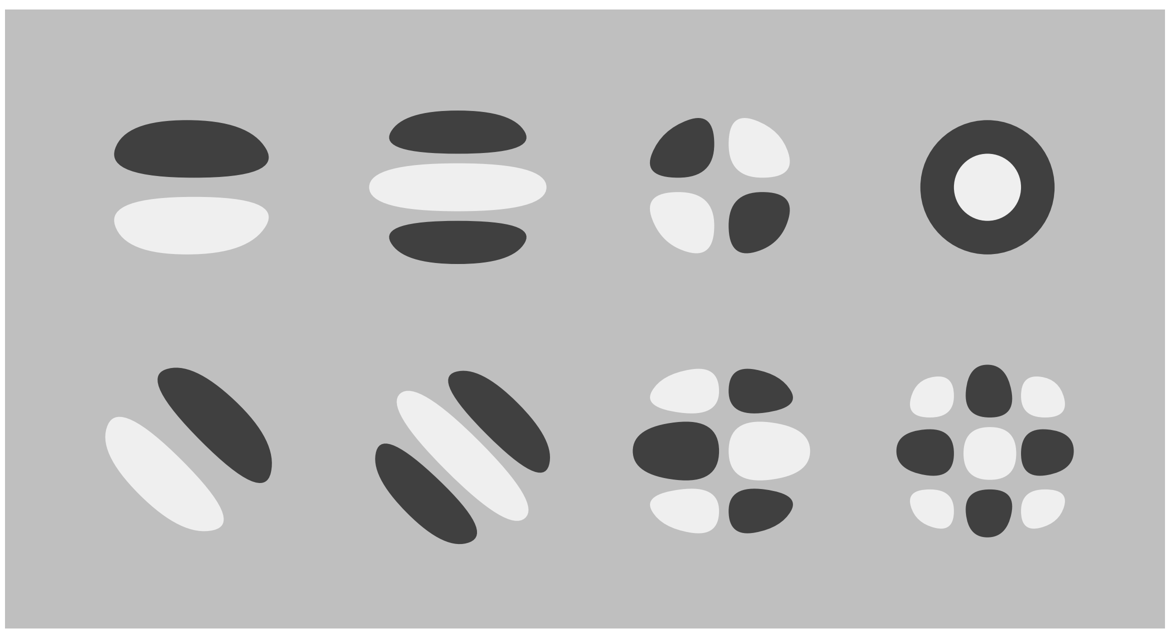
Local derivatives will turn out to be handy



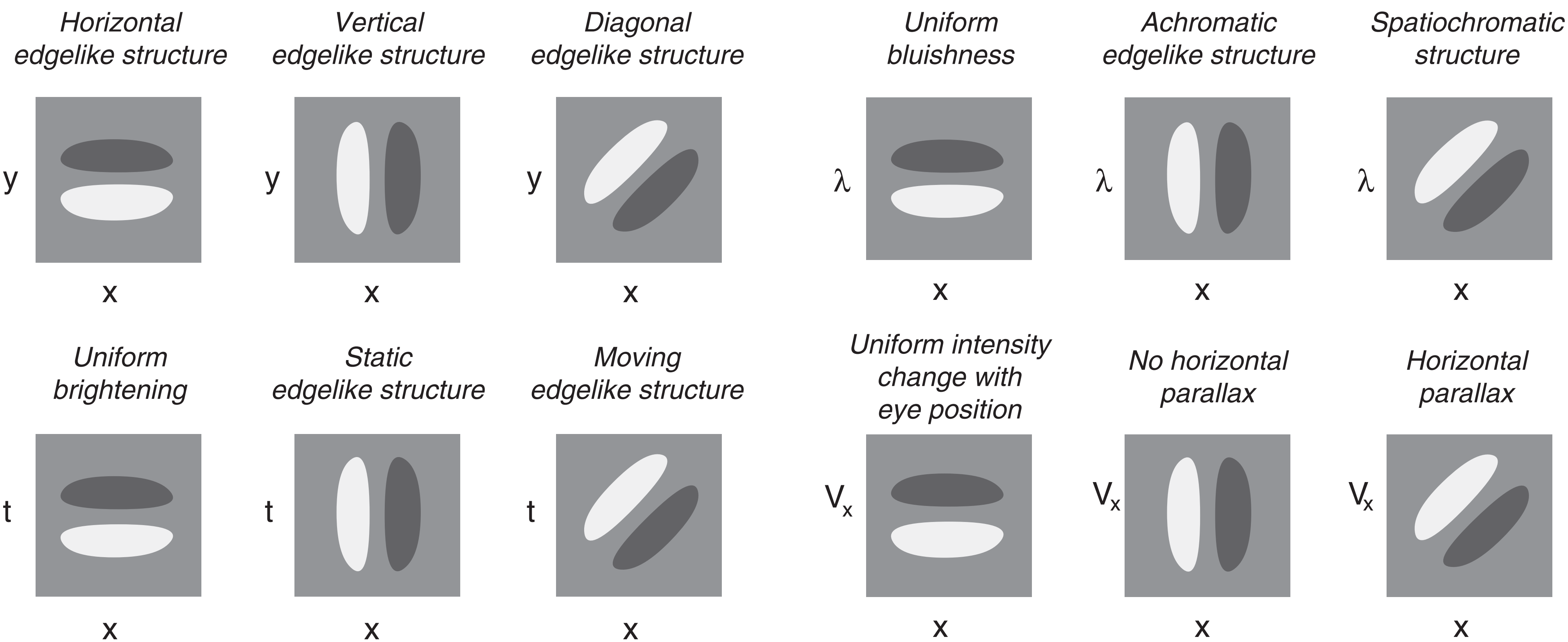
Derivatives along single dimensions yield some basic visual measurements



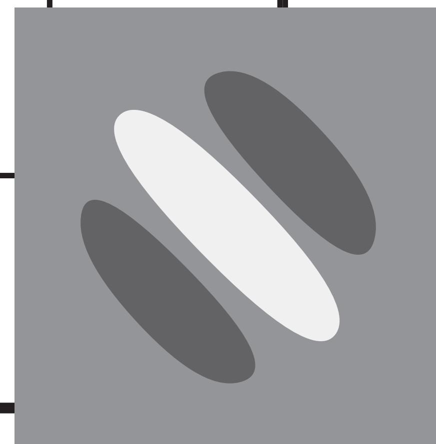
Low-order derivatives lead to a few two-dimensional operators (receptive fields)



The same receptive field structures produce different measurements when placed along different planes in plenoptic space

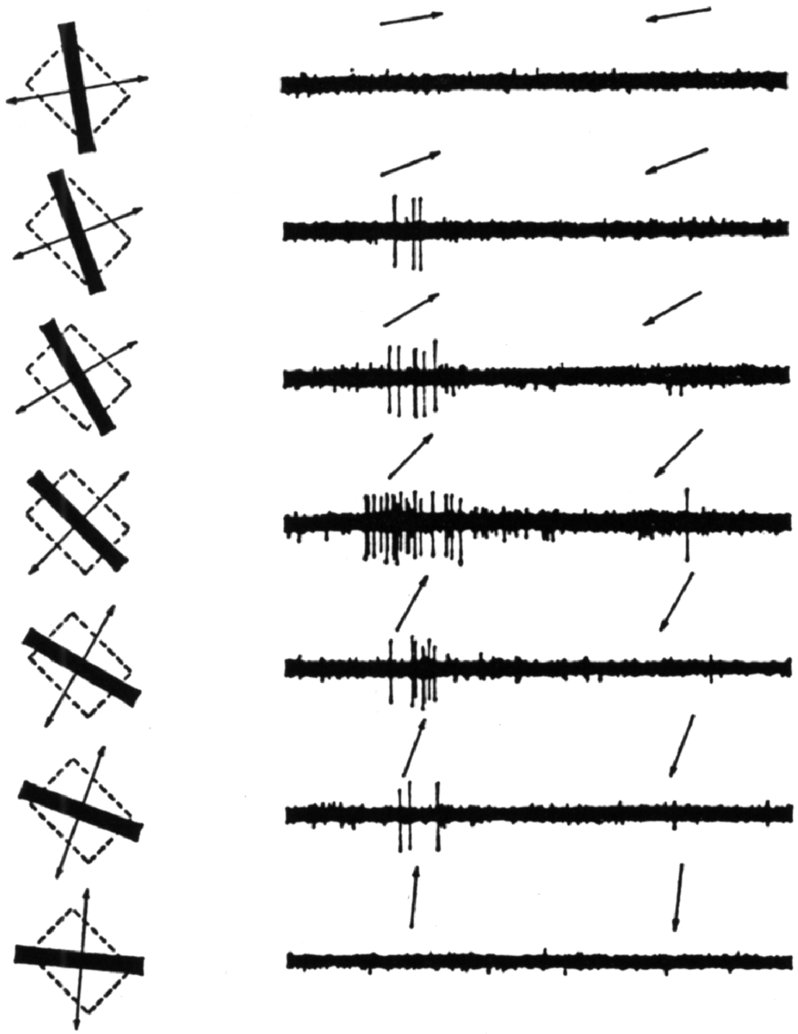


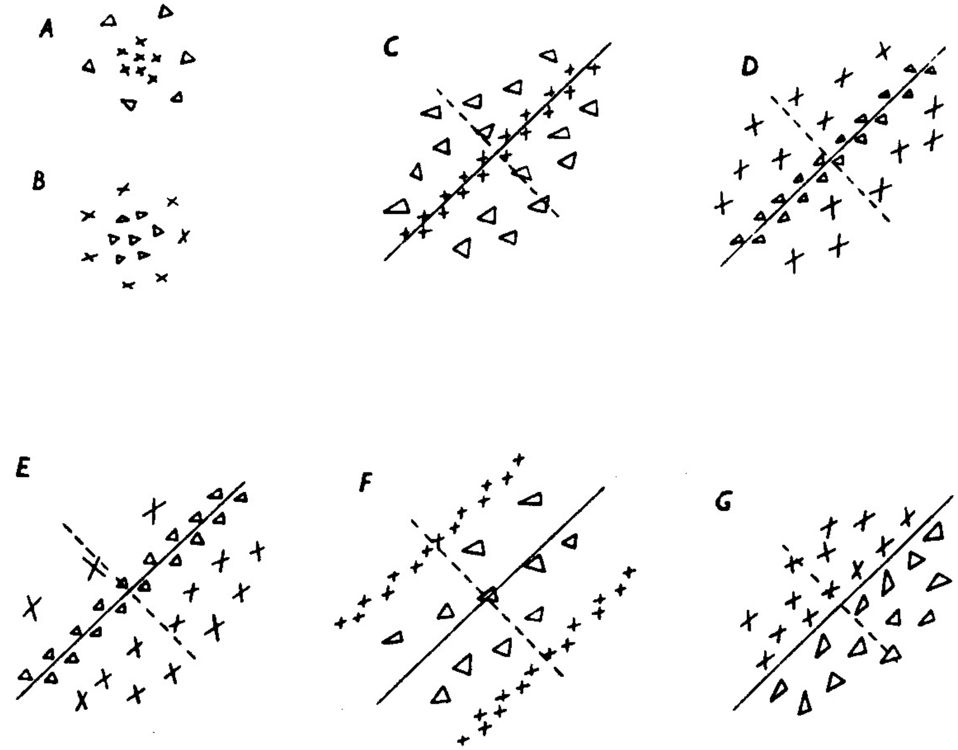
What can you measure with a tilted second derivative?

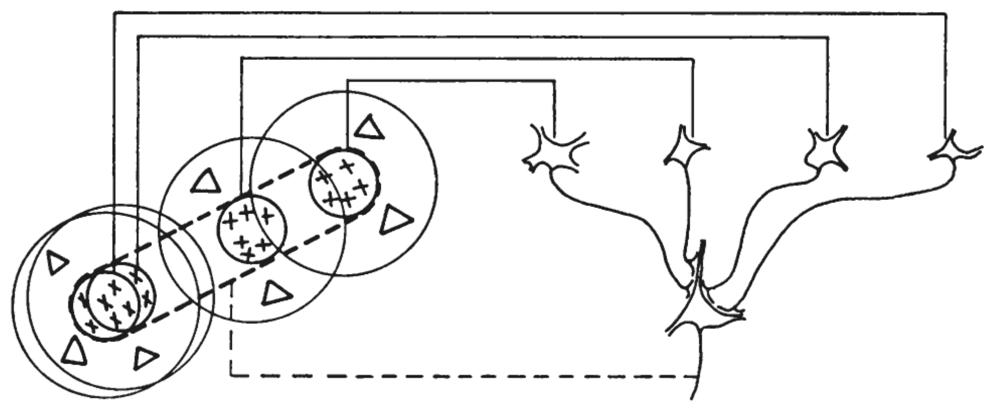


	x	y	t	λ	V_x
x	—				
y	**diag. "bar"** static achromatic no dispar	—			
t	**vert. "bar"** **leftward** achromatic no dispar	**hor. "bar"** **downward** achromatic no dispar	—		
λ	**vertical** static **hue-sweep** no dispar	**horiz.** static **hue-sweep** no dispar	full-field **sequential** **hue-sweep** no dispar	—	
V_x	**vert. "bar"** static achromatic **hor. dispar**	**hor. "bar"** static achromatic **vert. dispar**	full-field **sequential** achromatic **eye-order**	full-field static **hue-shift** **luster**	—

Orientation selectivity in V1

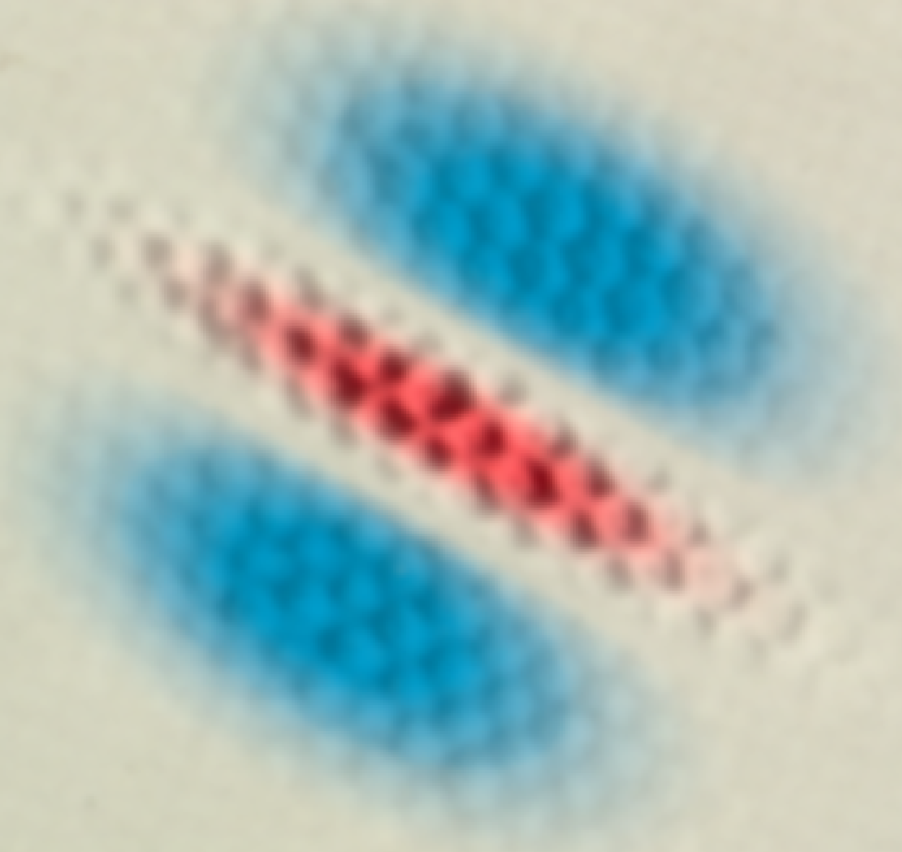






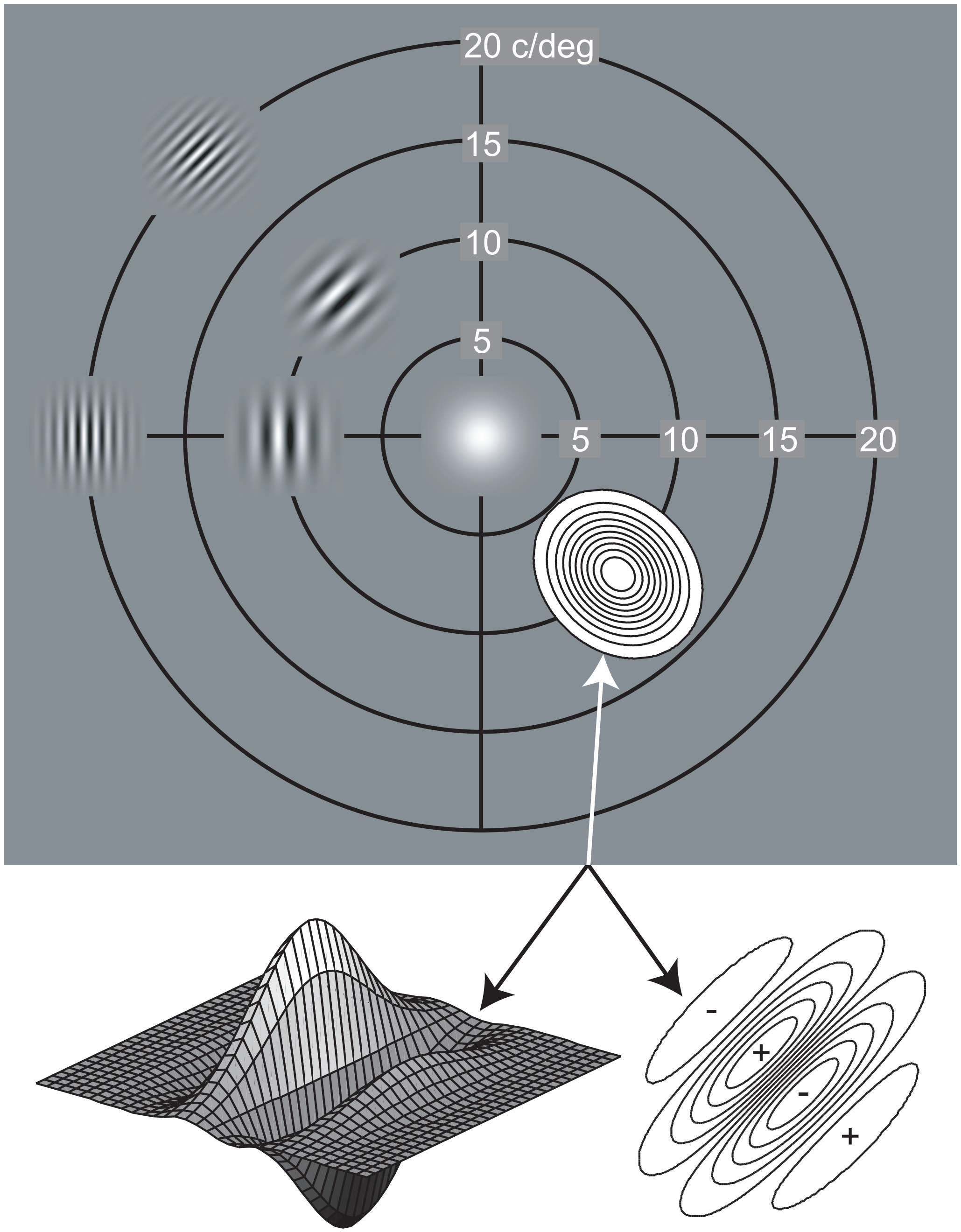
Hubel & Wiesel (1962, 1968)

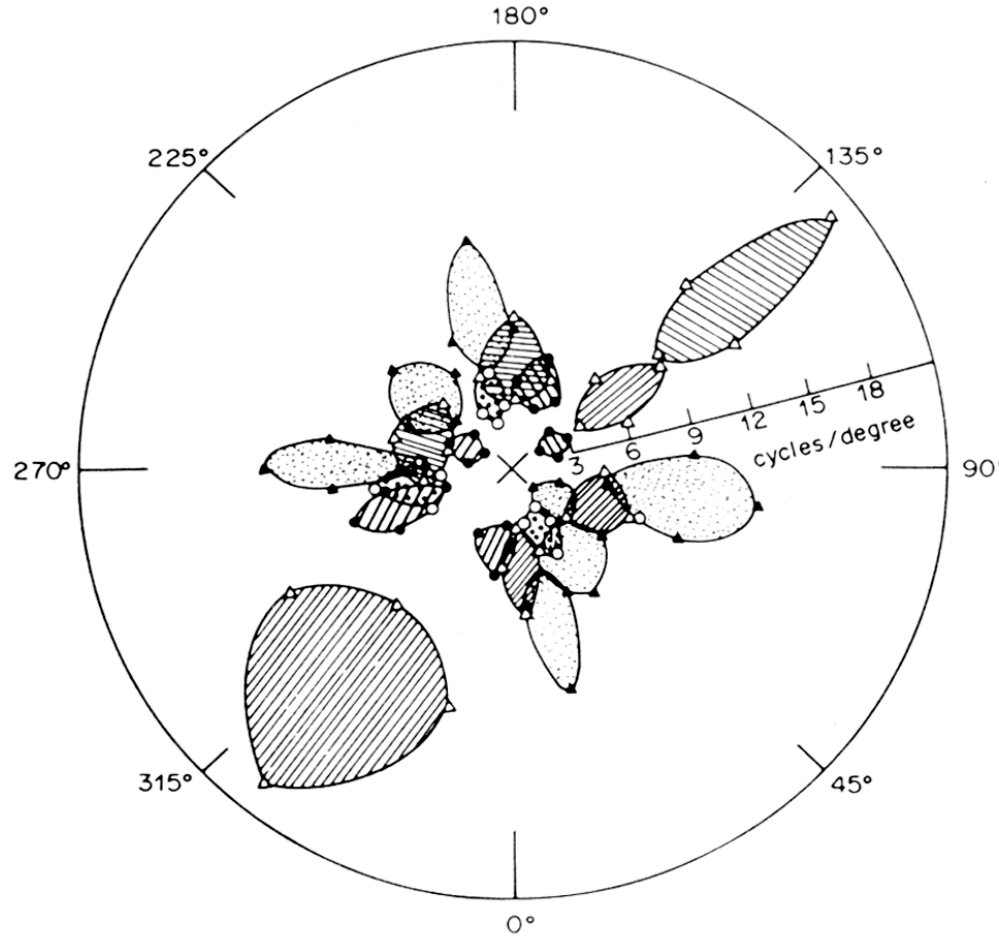
Orientation in space can be detected and measured by oriented filters



Hawken & Parker, 1991

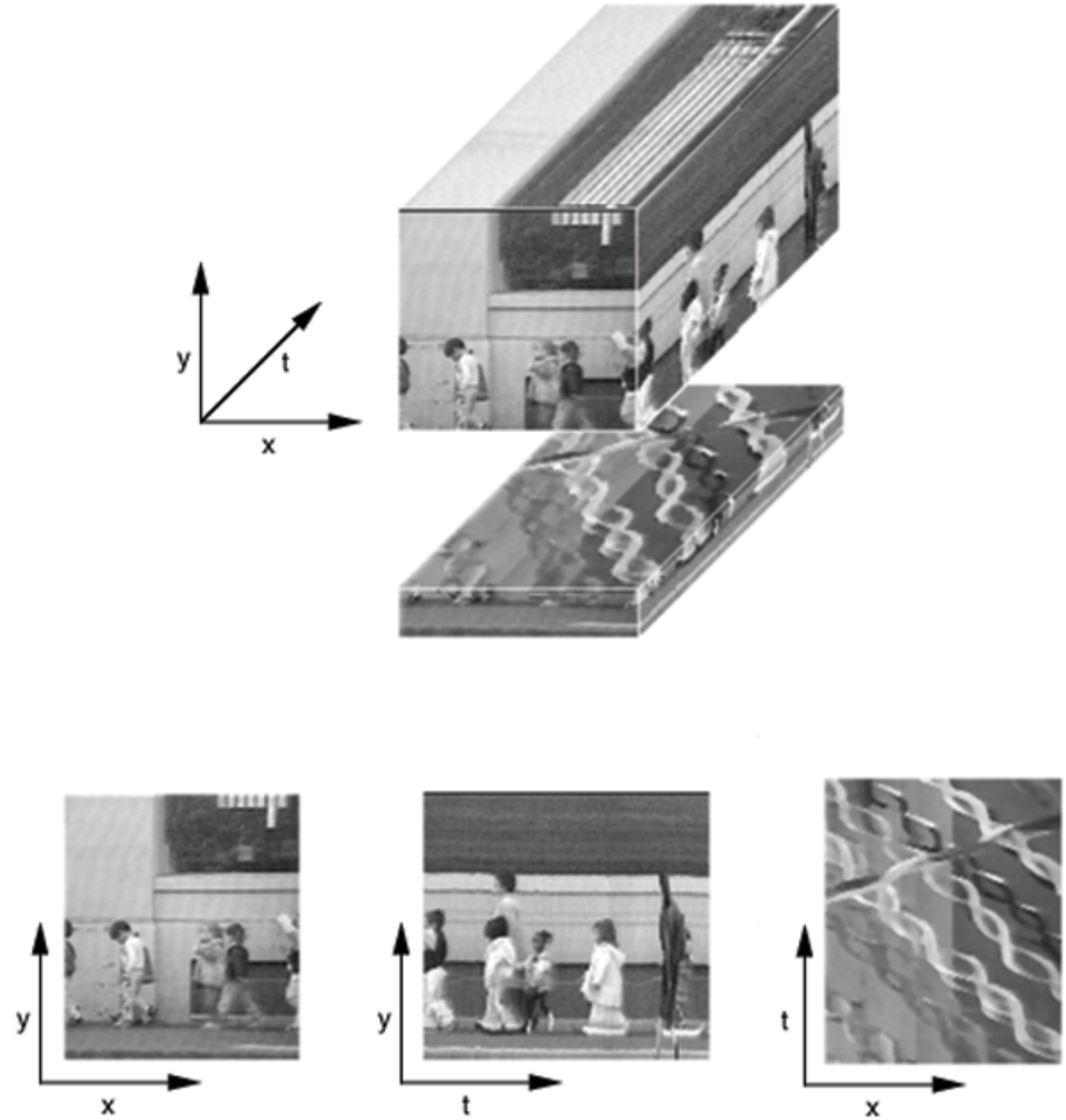
Orientation in space can be detected and measured by oriented filters



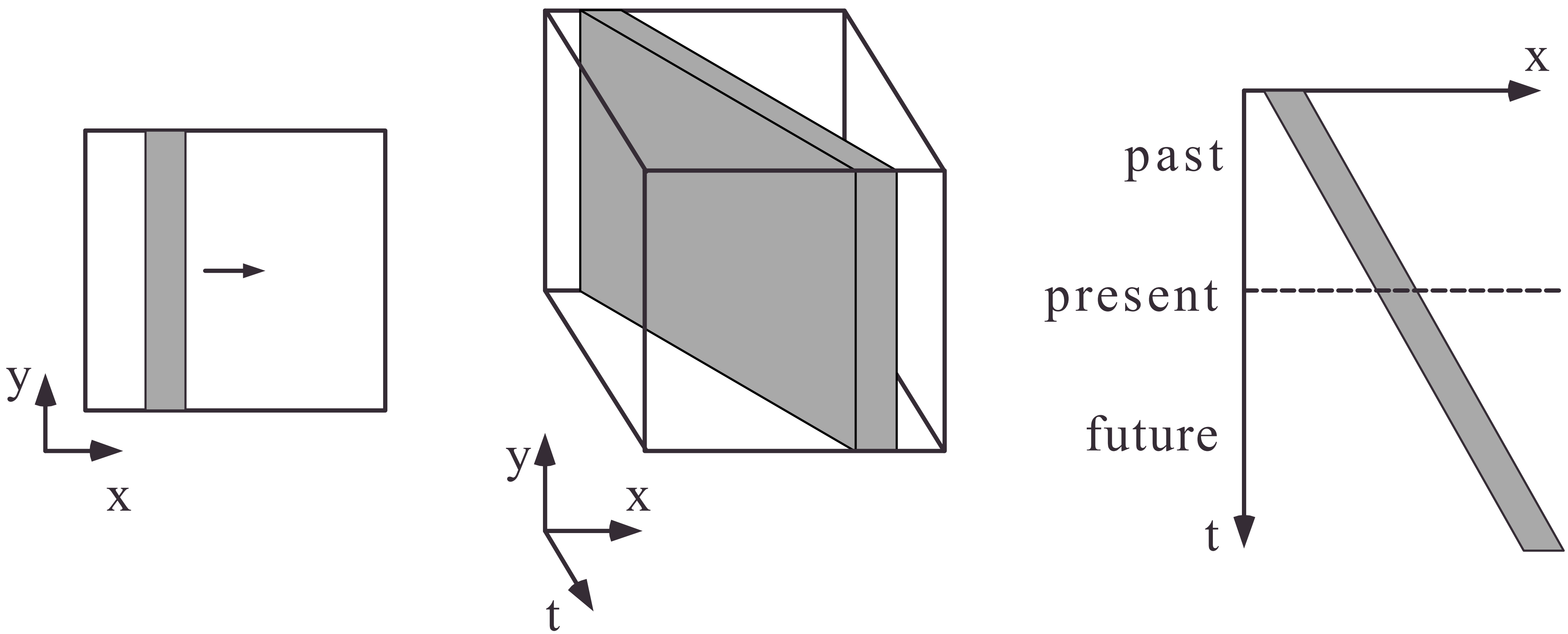


DeValois, Albrecht and Thorell, 1982

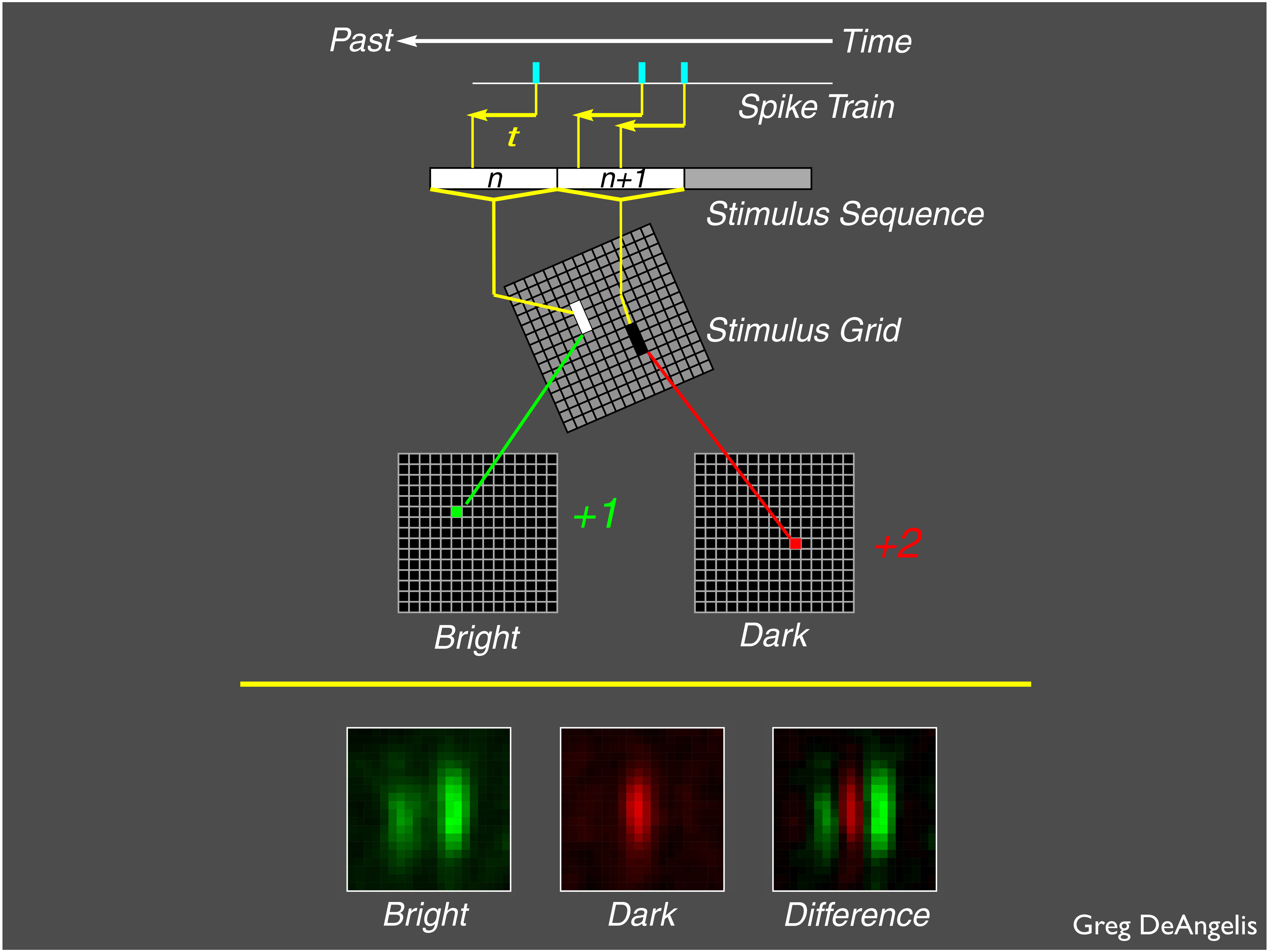
Motion is orientation in space-time

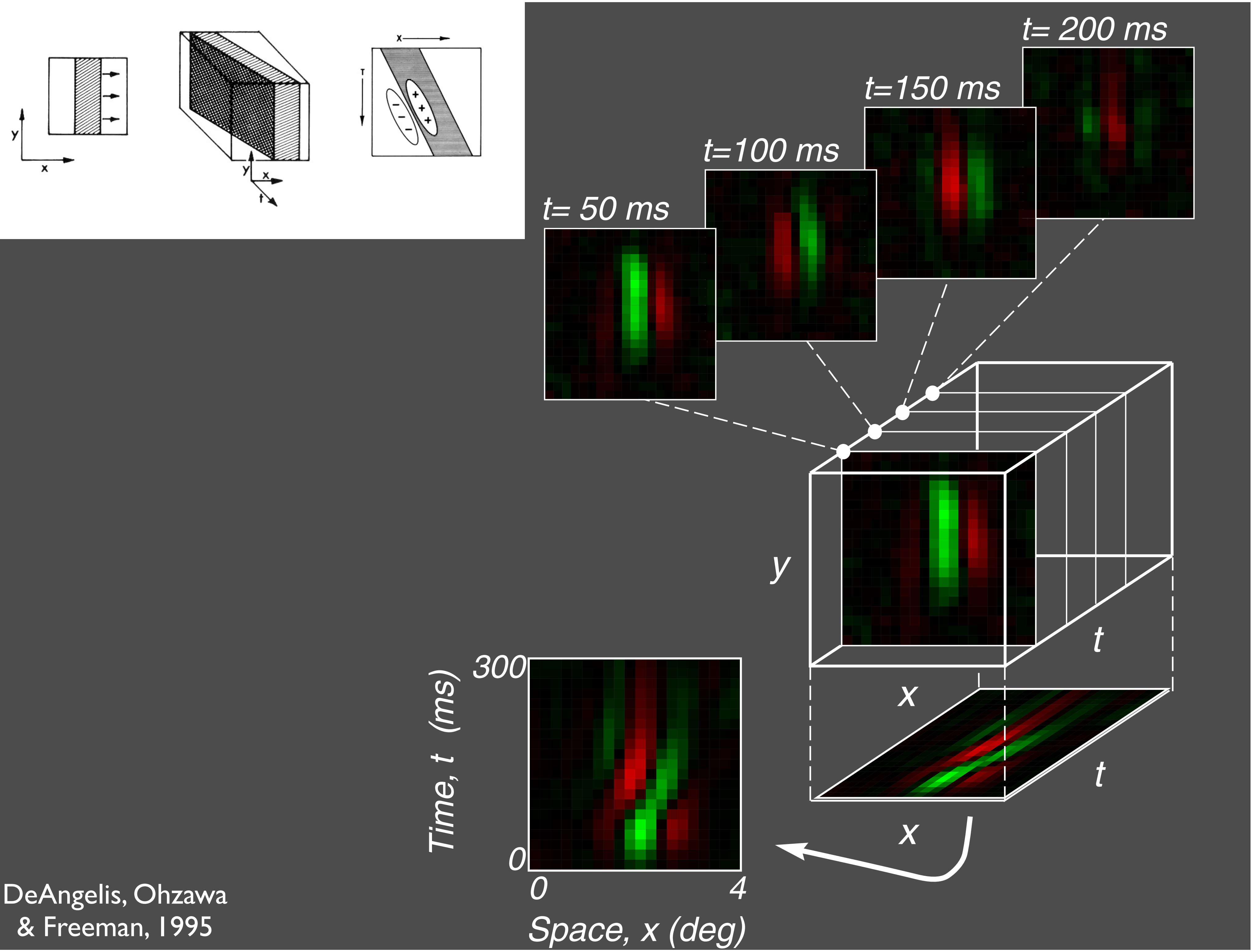


Motion is orientation in space-time
and spatiotemporally oriented filters can be used to detect and measure it



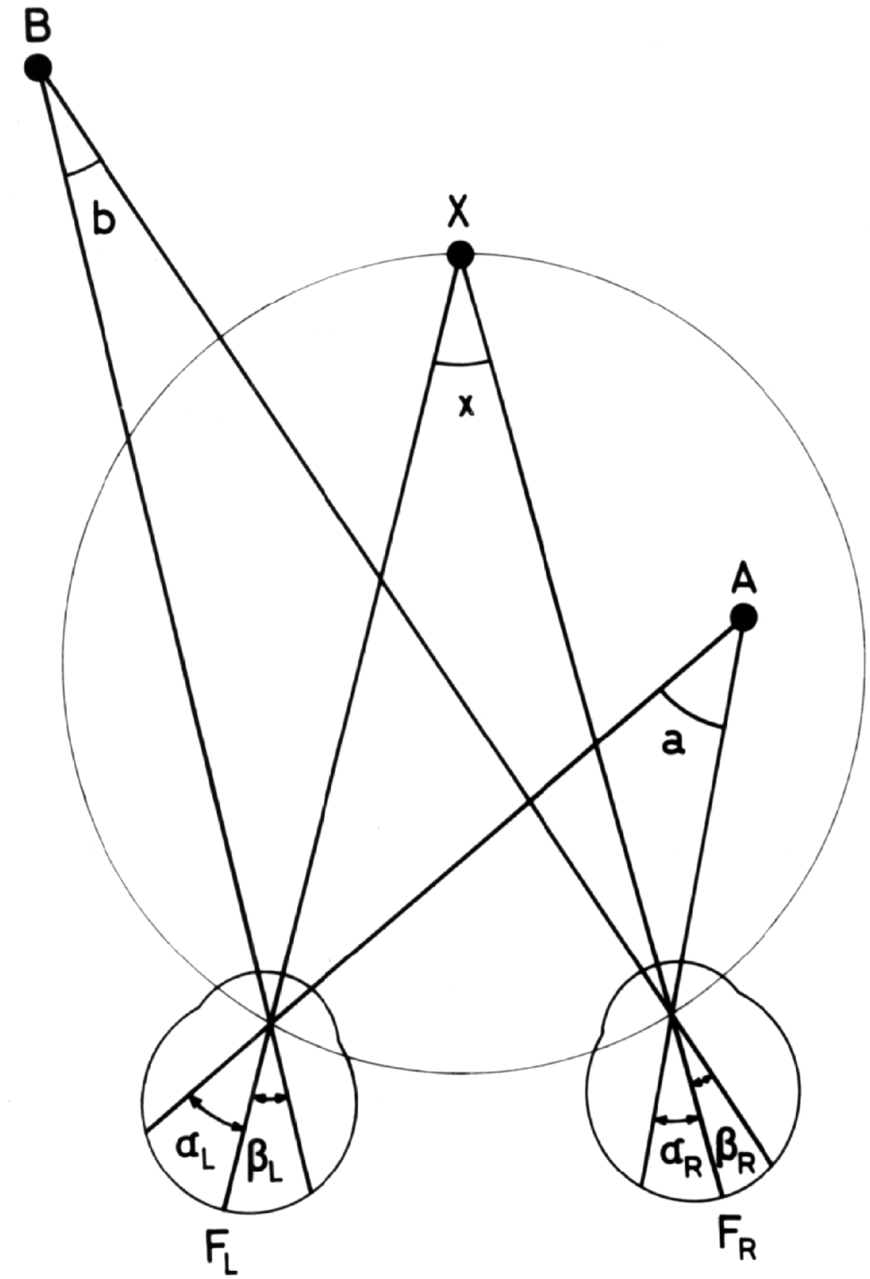
Adelson & Bergen (1985)


Past
Time
Spike Train
t
n
n+1
Stimulus Sequence
Stimulus Grid
+1
+2
Bright
Dark
Bright
Dark
Difference
Greg DeAngelis

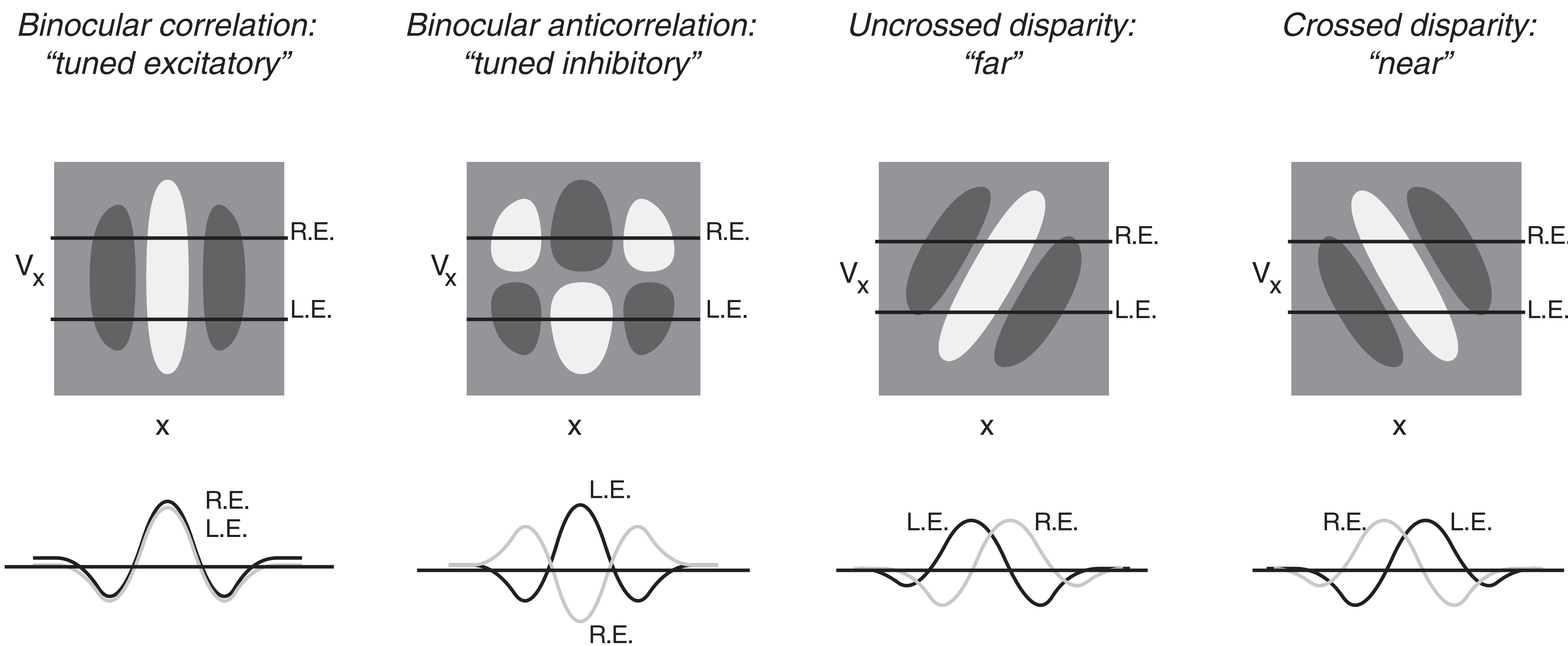


DeAngelis, Ohzawa
& Freeman, 1995

DIsparity is orientation in space-eye position

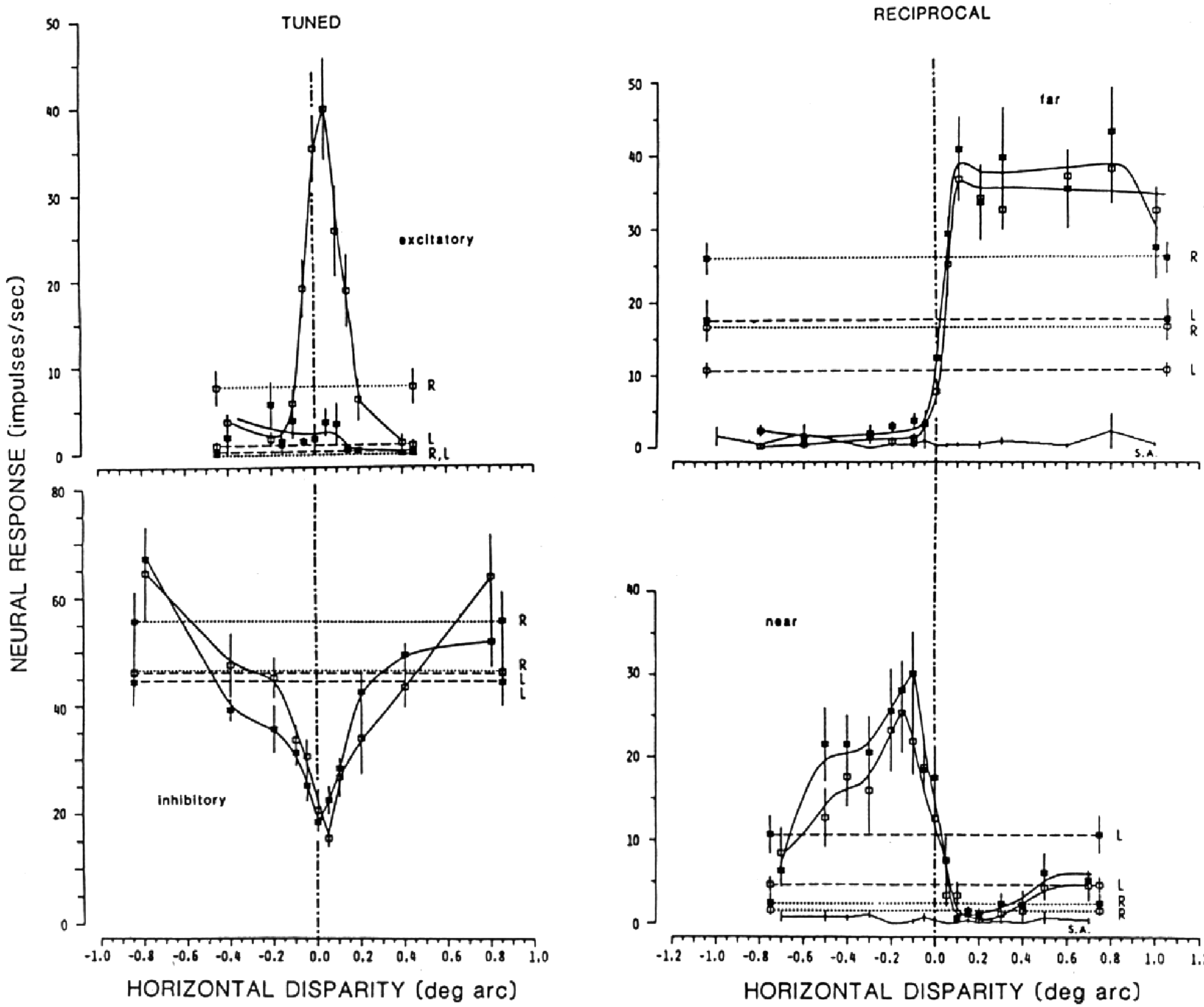


DIsparity is orientation in space-eye position



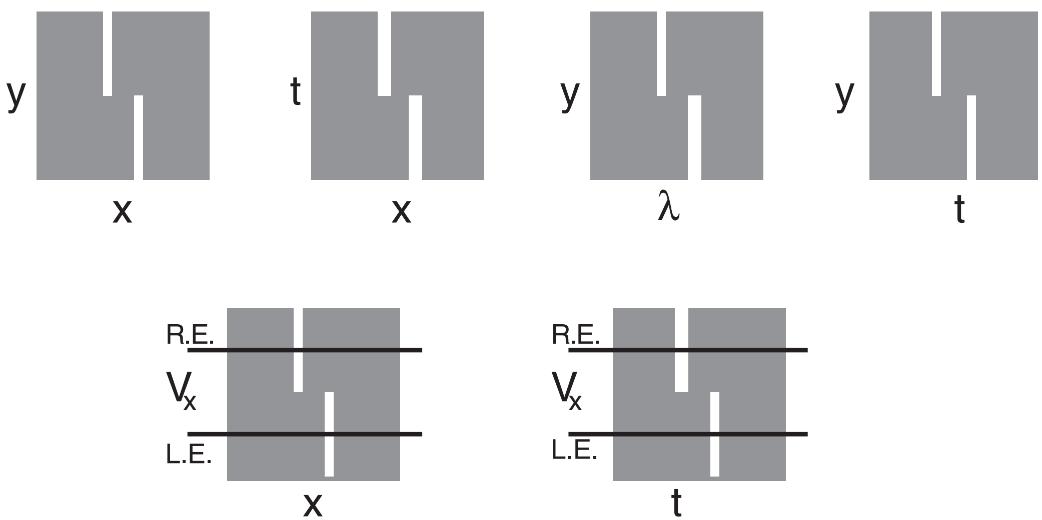
Four examples of binocular receptive fields. Humans only take two samples from the V_x axis, as shown by the two lines labeled R.E. and L.E. for right eye and left eye. The curves beneath each receptive field indicate the individual weighting functions for each eye alone.

DIsparity is orientation in space-eye position



Poggio (1981)

Many psychophysical tasks look like Vernier acuity in different planes



	x	y	t	λ	V_x
x	—	**hor. vernier** static achromatic no dispar.	**vert. bipart.** **pulse order** achromatic no dispar.	**vert. edge** static λ **diff.** no dispar.	—
y	**vert. vernier** static achromatic no dispar.	—	**hor. bipart.** **pulse order** achromatic no dispar.	**hor. edge** static λ **diff.** no dispar.	—
t	**vert. line** **jump left** achromatic no dispar.	**horiz. line** **jump down** achromatic no dispar.	—	full-field **sequential** λ **change** no dispar.	—
λ	**2 vert. lines** static λ **change** no dispar.	**2 hor. lines** static λ **change** no dispar.	full-field **pulse order** **blue-yel.** no dispar.	—	—
V_x	**vert. line** static achromatic **h disp.**	**horiz. line** static achromatic **v disp.**	full-field **pulse order** achromatic **anticorr.**	full-field static λ **diff.** **anticorr.**	—

What is there to see?

Black and white photo: x, y
Black and white movie: x, y, t
Color movie: x, y, t, λ
Holographic movie: x, y, t, λ, V_x, V_y, V_z

Hence, the plenoptic function: $\boldsymbol{P(x, y, t, \lambda, V_x, V_y, V_z)}$

Proposition 1*. The task of early vision is to deliver a small set of useful measurements about each observable location in the plenoptic function.*

Proposition 2*. The elemental operations of early vision measure local change along various directions within the plenoptic function.*

What is there to see?

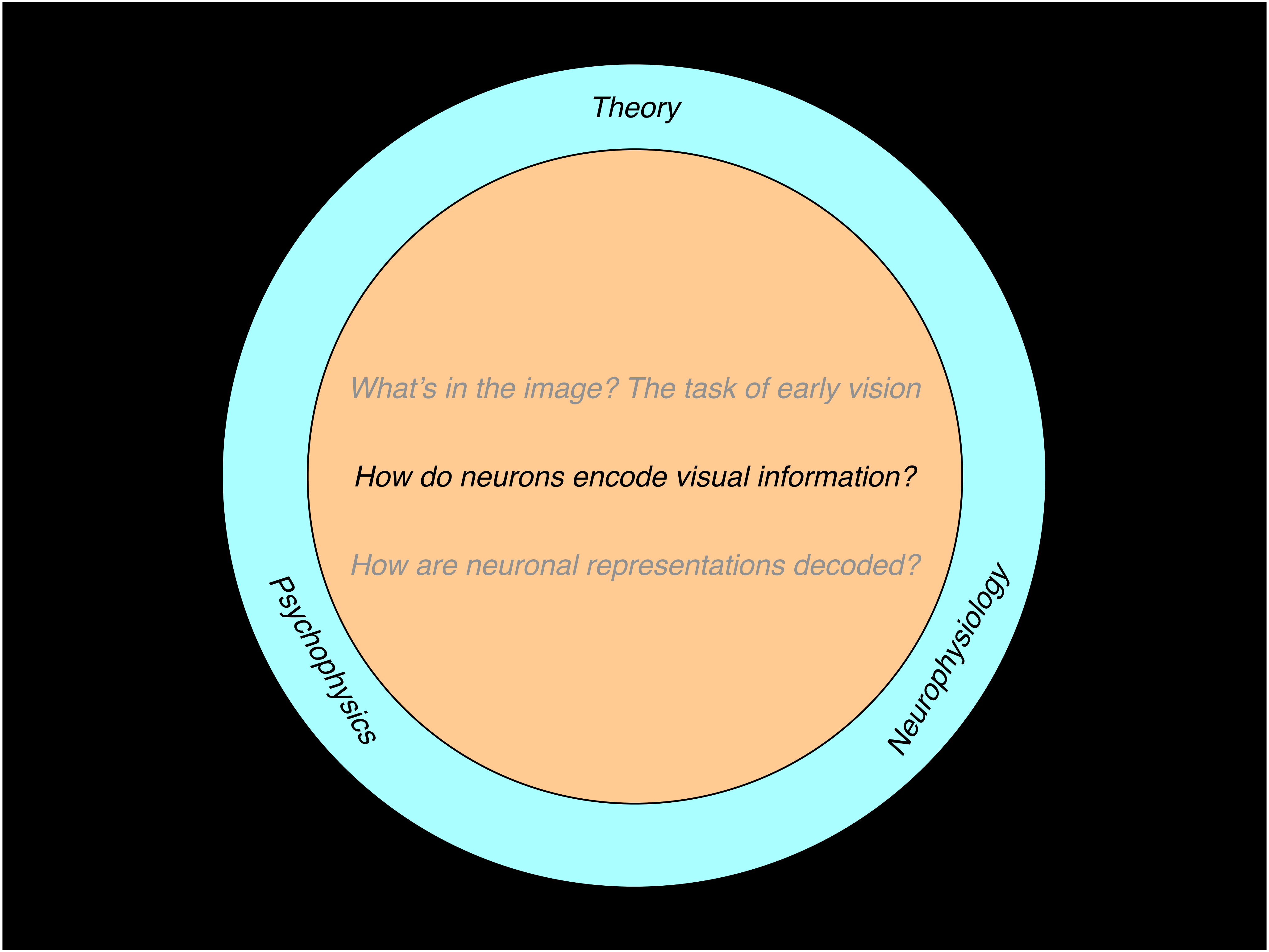
Black and white photo: x, y
Black and white movie: x, y, t
Color movie: x, y, t, λ
Holographic movie: $x, y, t, \lambda, V_x, V_y, V_z$

Hence, the plenoptic function: $\boldsymbol{P(x, y, t, \lambda, V_x, V_y, V_z)}$

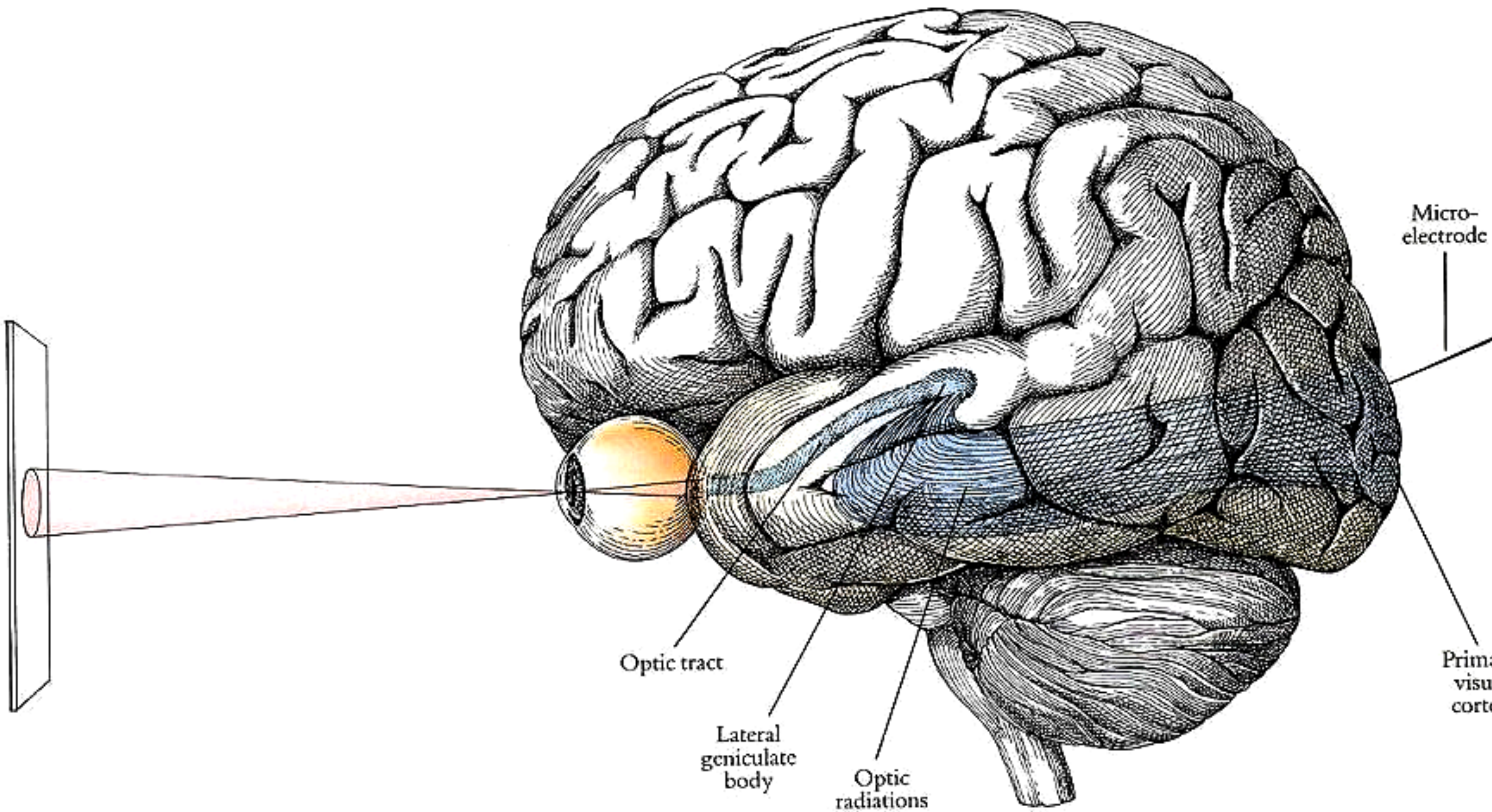
The analysis of the plenoptic function tells us what elementary measurements **can be** *computed, but not which ones* **are** *computed or which ones* **should be** *computed.*

What parts of the possible information are present in the world?

What parts of the information present in the world are important to the organism?

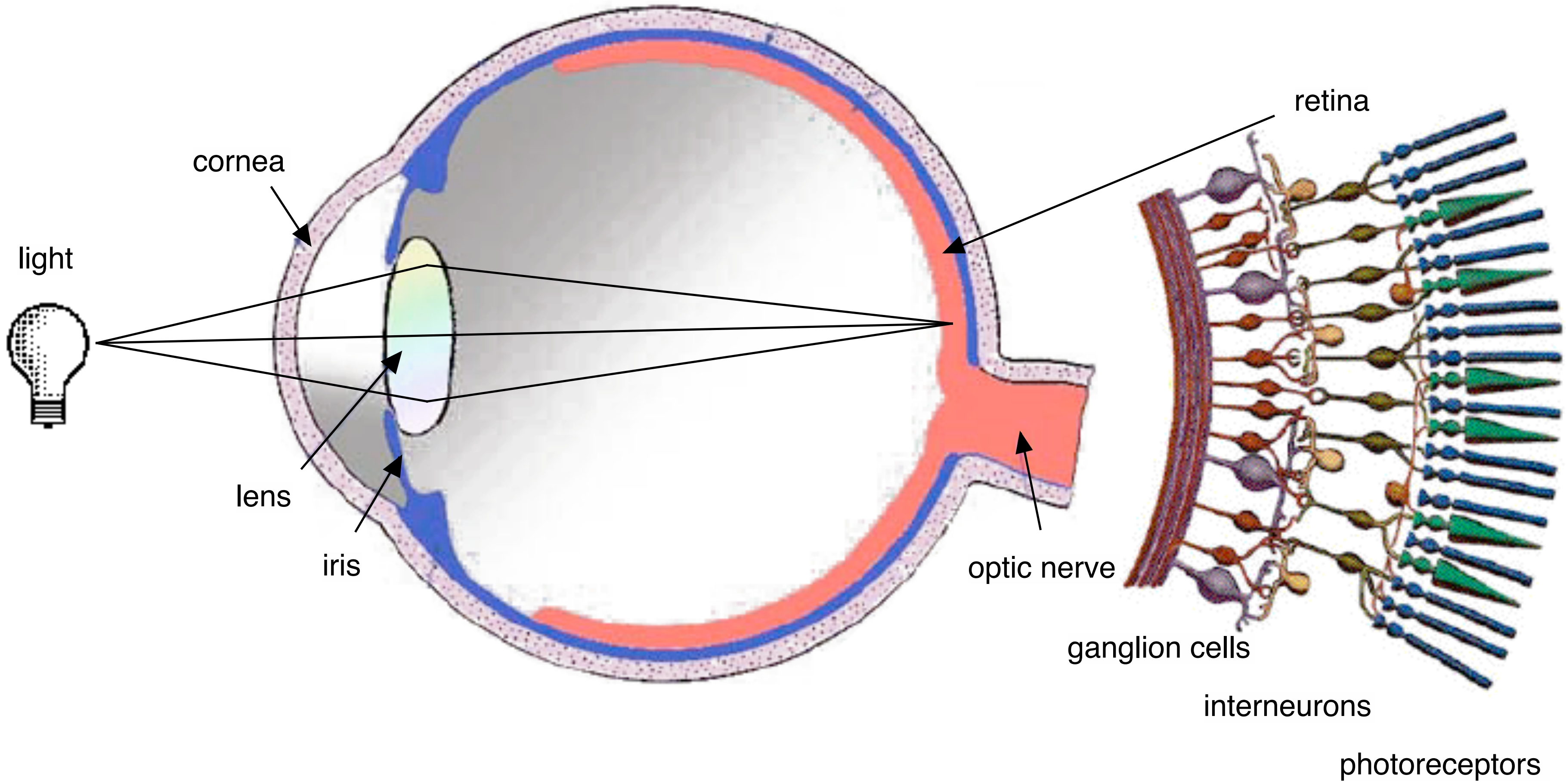

Theory
What's in the image? The task of early vision
How do neurons encode visual information?
How are neuronal representations decoded?
Psychophysics
Neurophysiology

Neural circuits perform computations

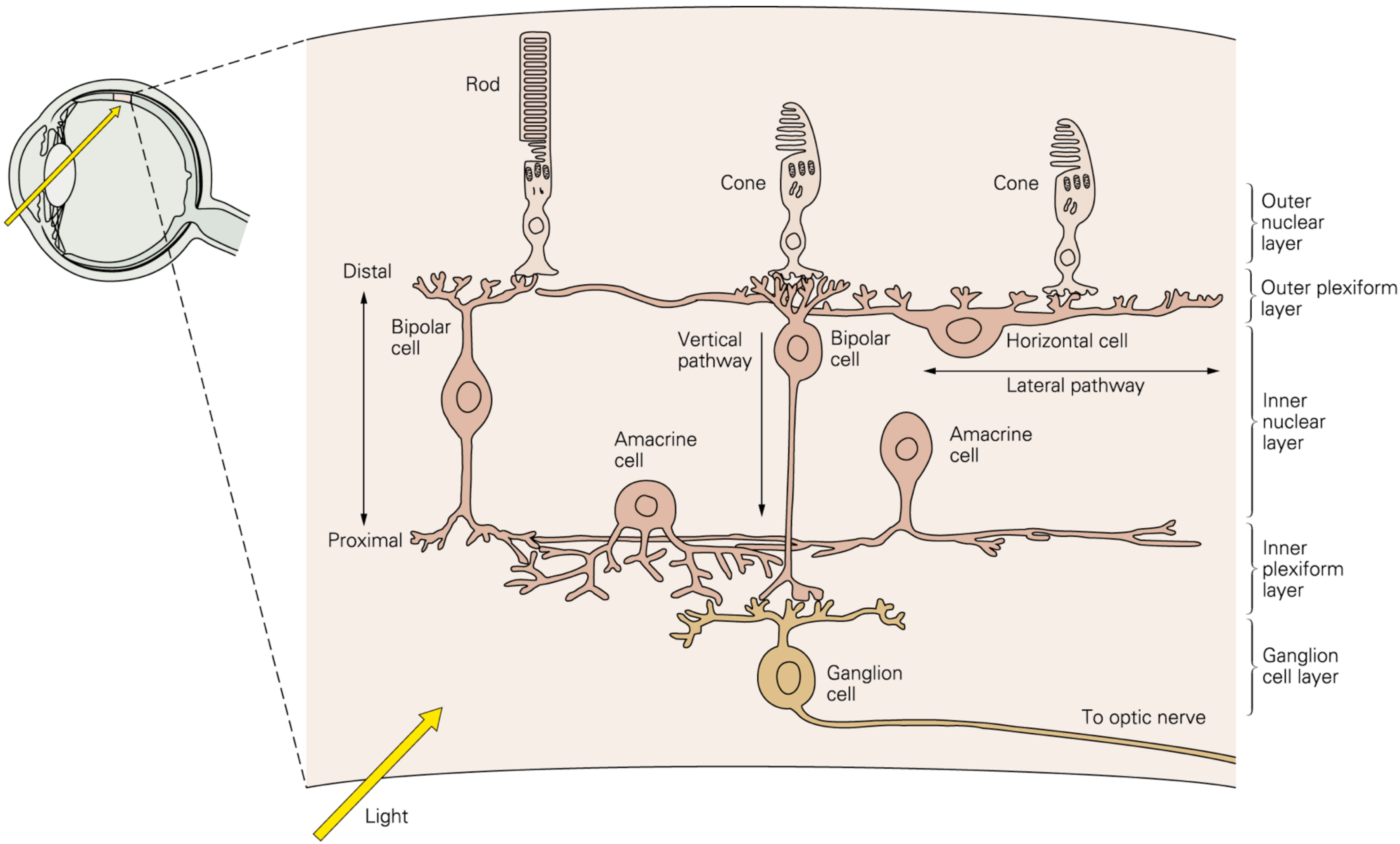


Hubel, 1988

Vertebrate retina

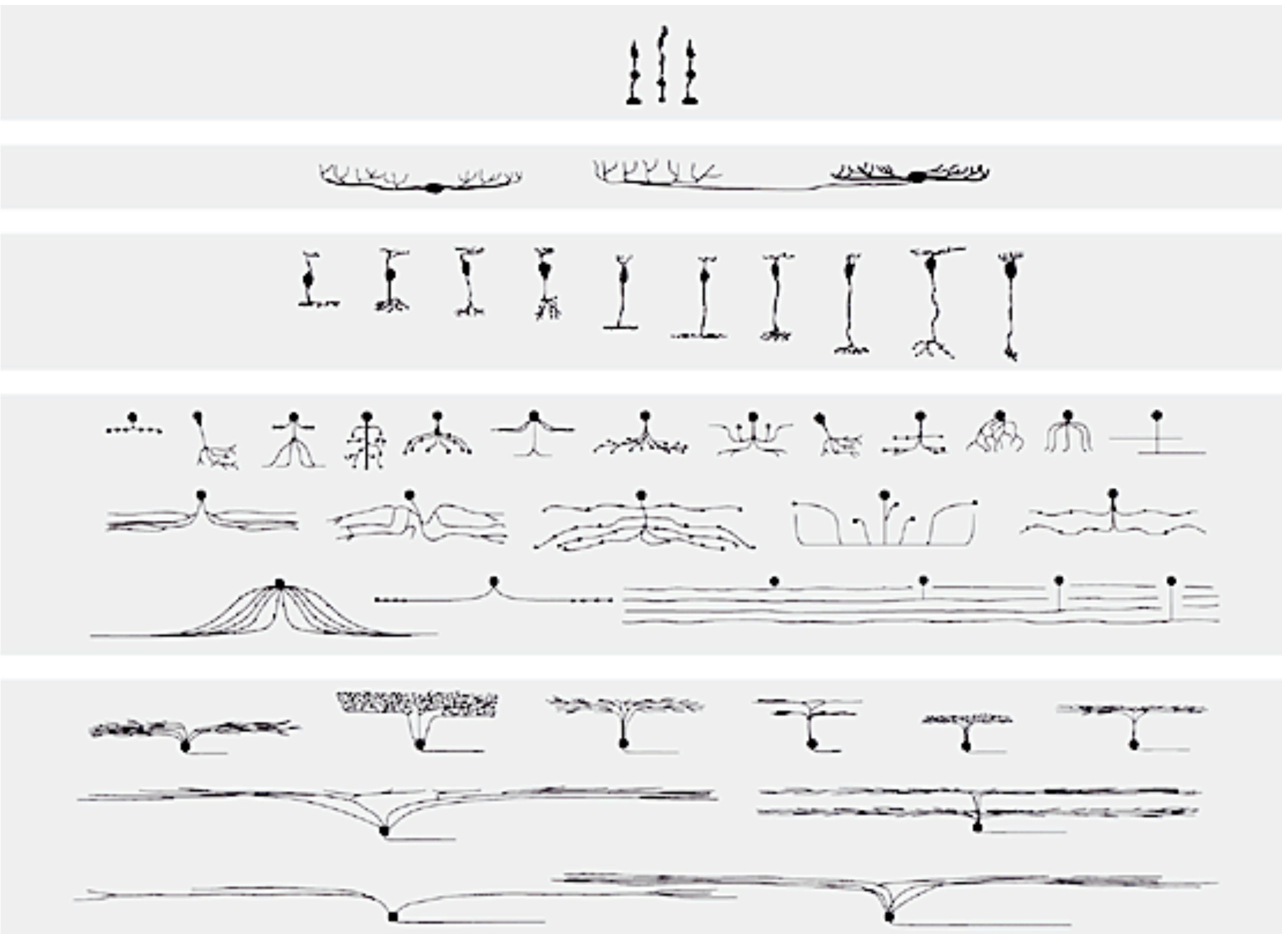


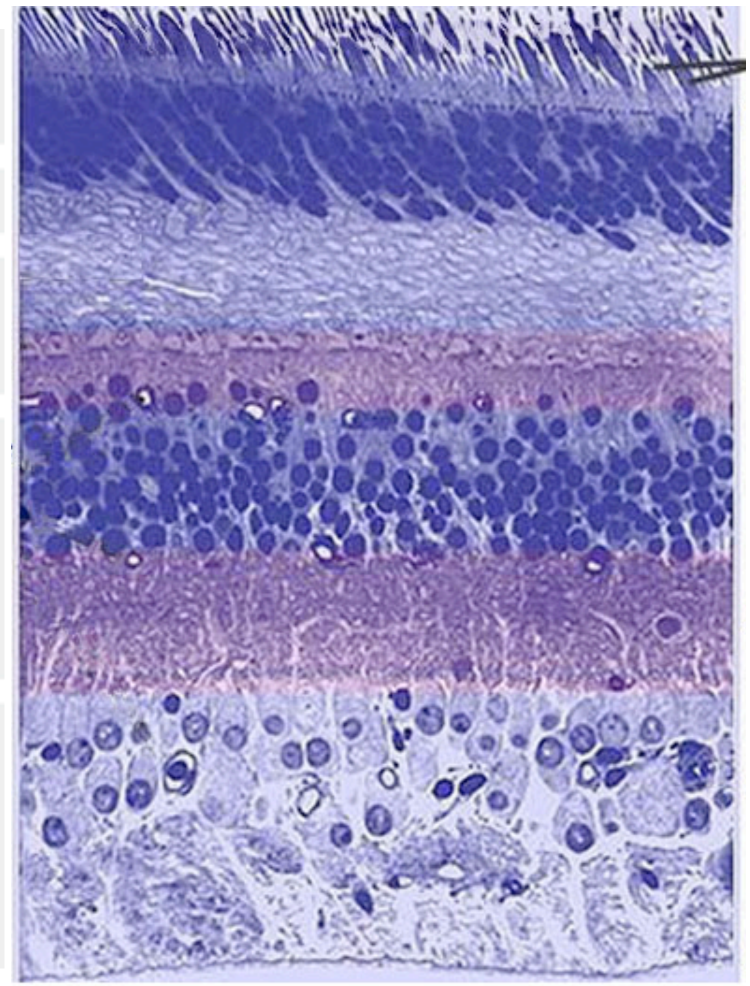
Vertical and horizontal pathways for information flow in the retina



Kandel, Schwartz & Jessell, 2001

Retinal neuronal diversity and circuit specificity





Masland, 2001

J. Physiol. (1966), **187**, *pp.* 517–552
With 17 *text-figures*
Printed in Great Britain

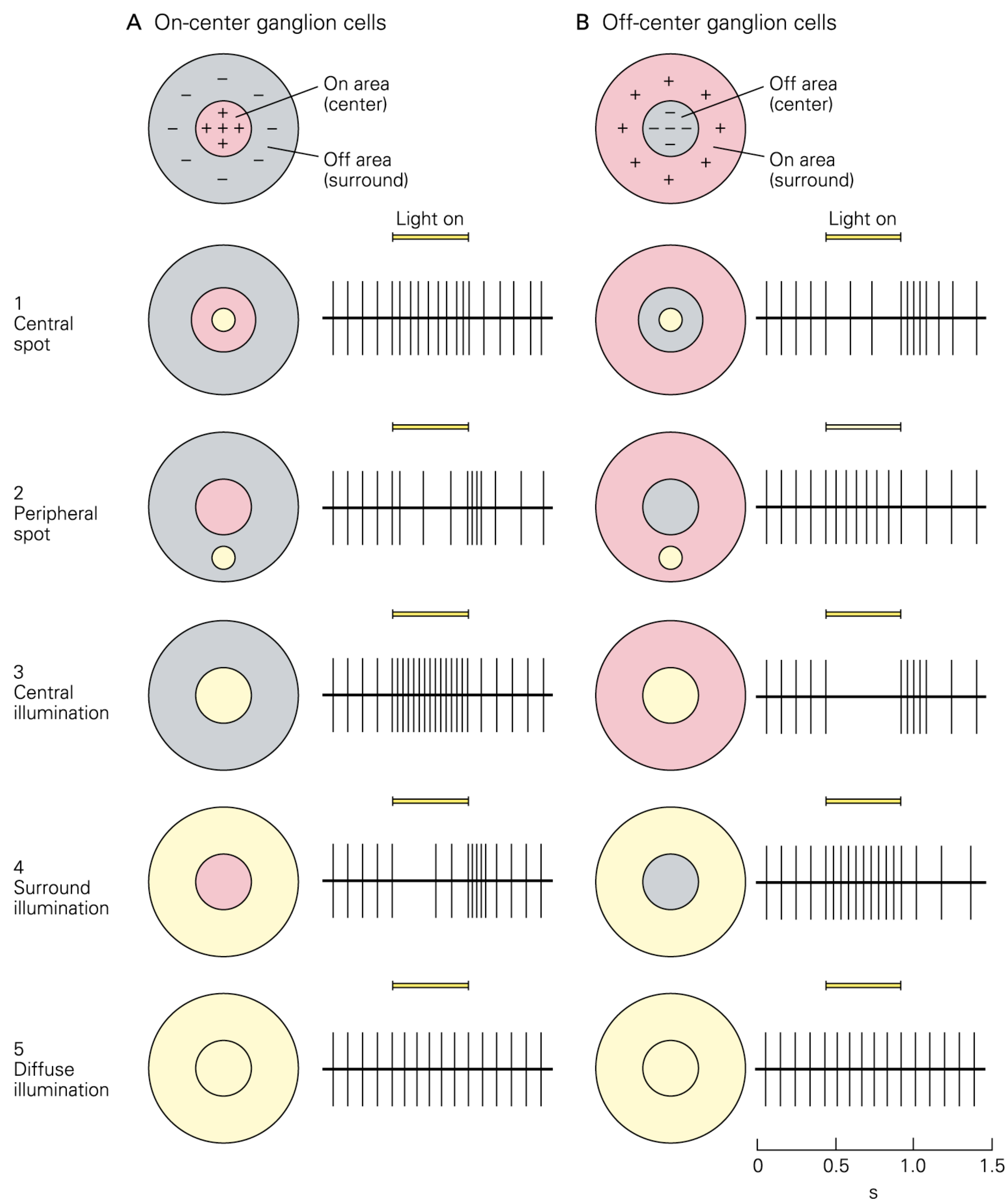
THE CONTRAST SENSITIVITY OF RETINAL GANGLION CELLS OF THE CAT

BY CHRISTINA ENROTH-CUGELL AND J. G. ROBSON*

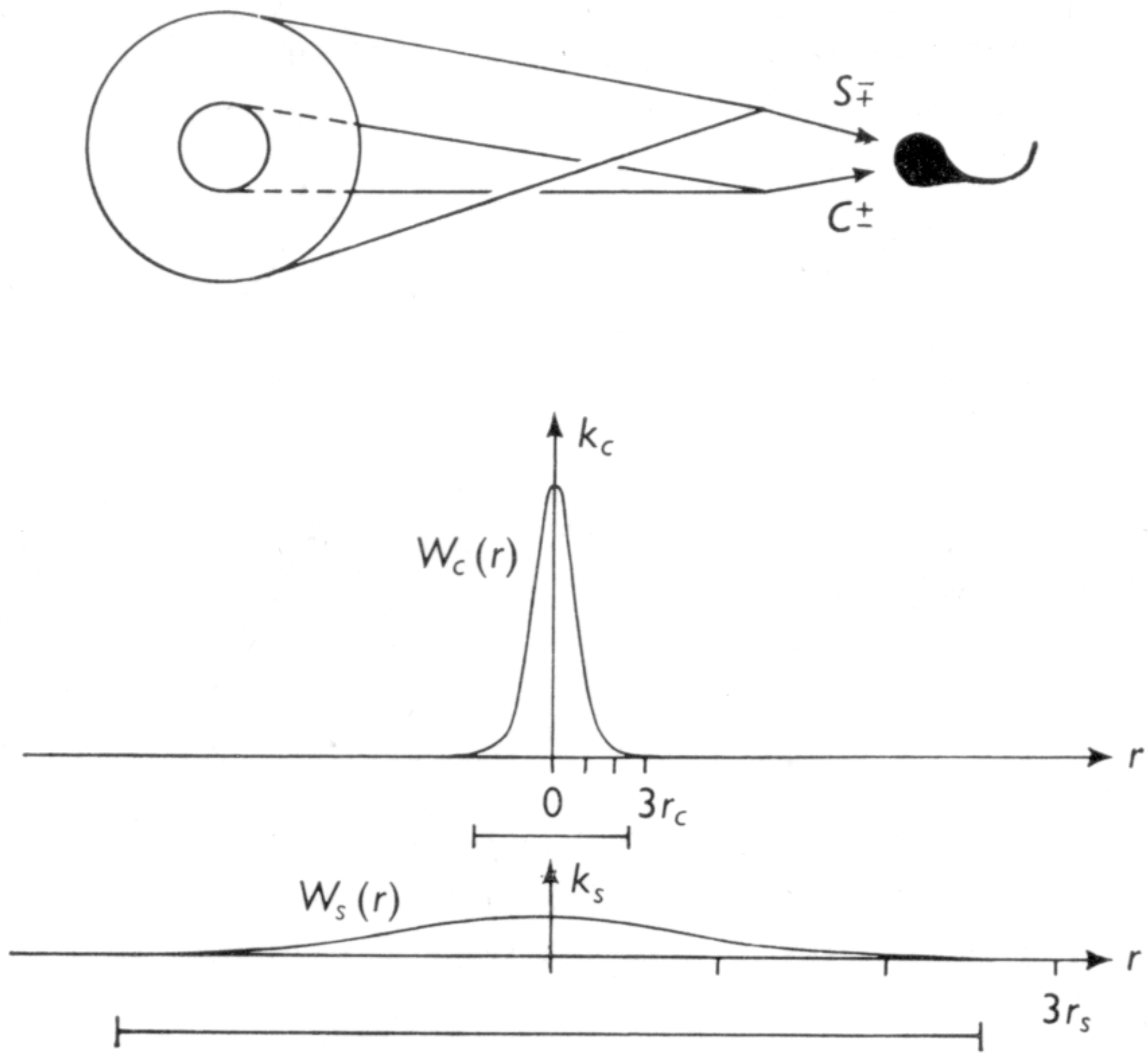
From the Biomedical Engineering Center, Technological Institute, Northwestern University, Evanston, Illinois, U.S.A.† and the Department of Physiology, Northwestern University Medical School, Chicago, U.S.A.

(*Received* 19 *April* 1966)

Spatial structure is first measured by neurons with center-surround receptive fields

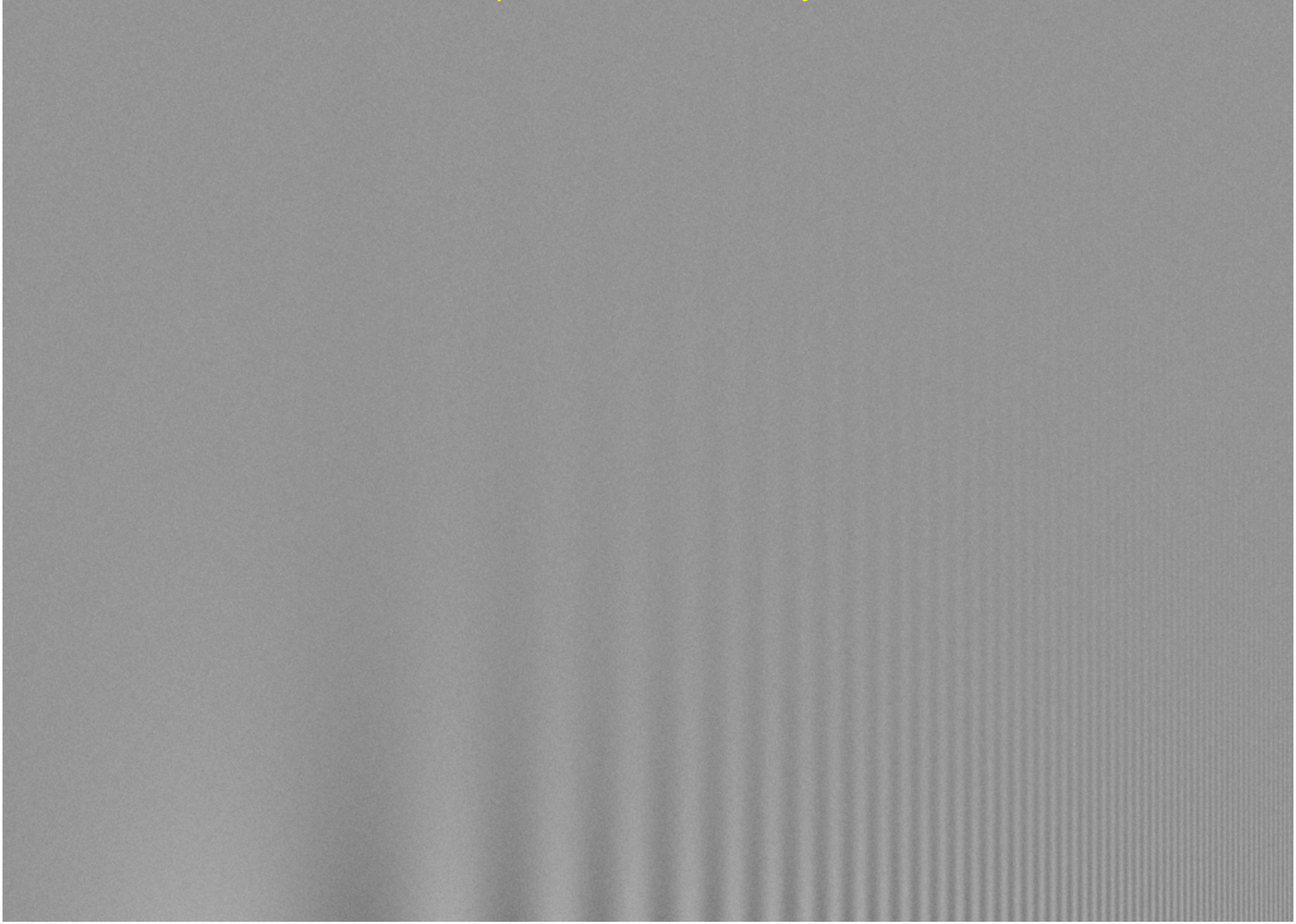


Ganglion cell receptive field modeled as difference of Gaussians

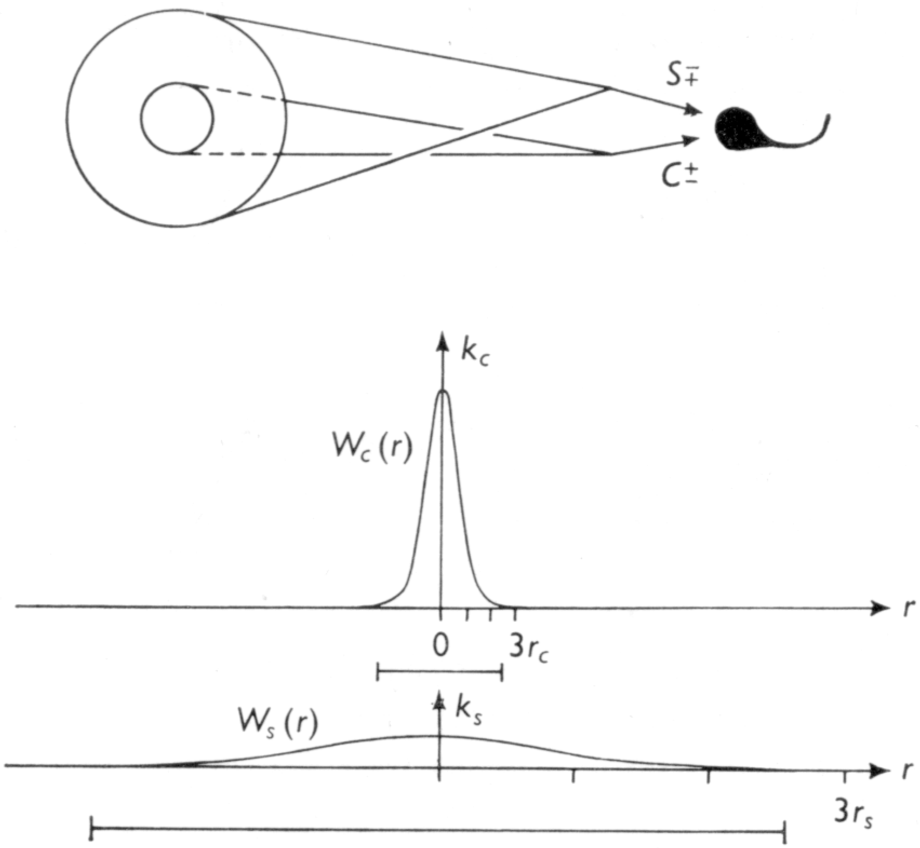


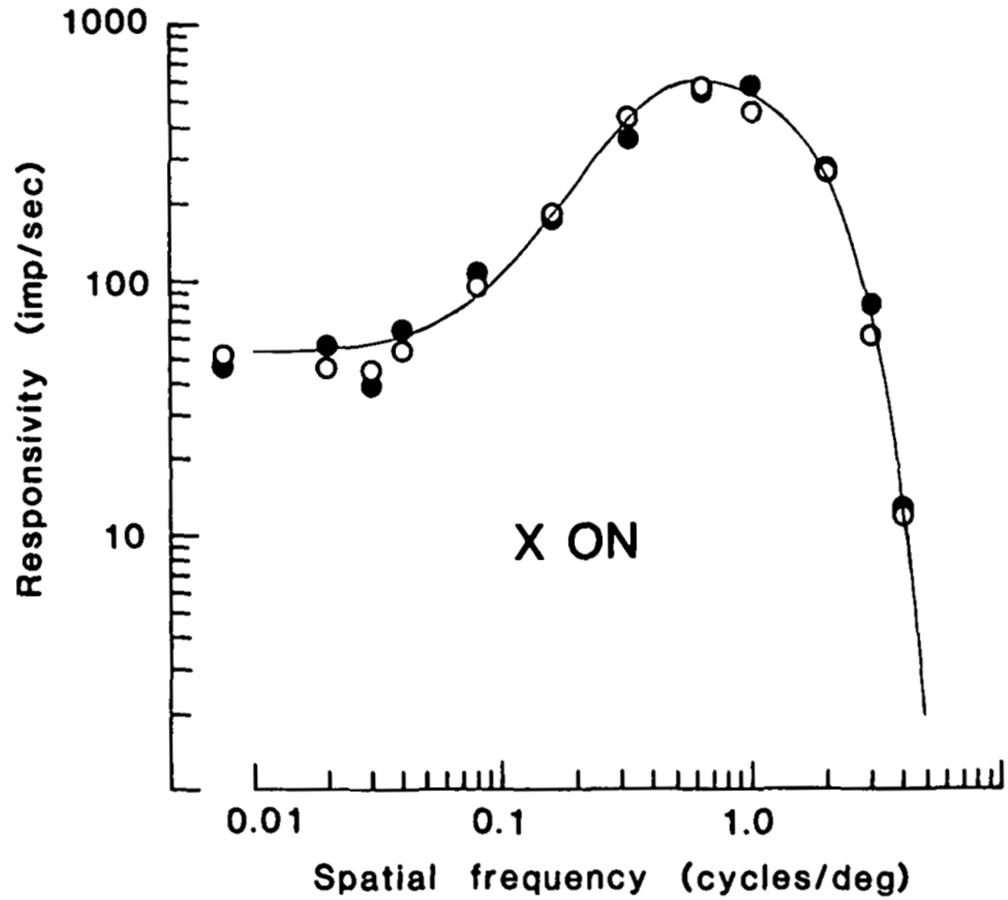
Rodieck, 1965; Enroth-Cugell & Robson, 1966

Spatial contrast sensitivity



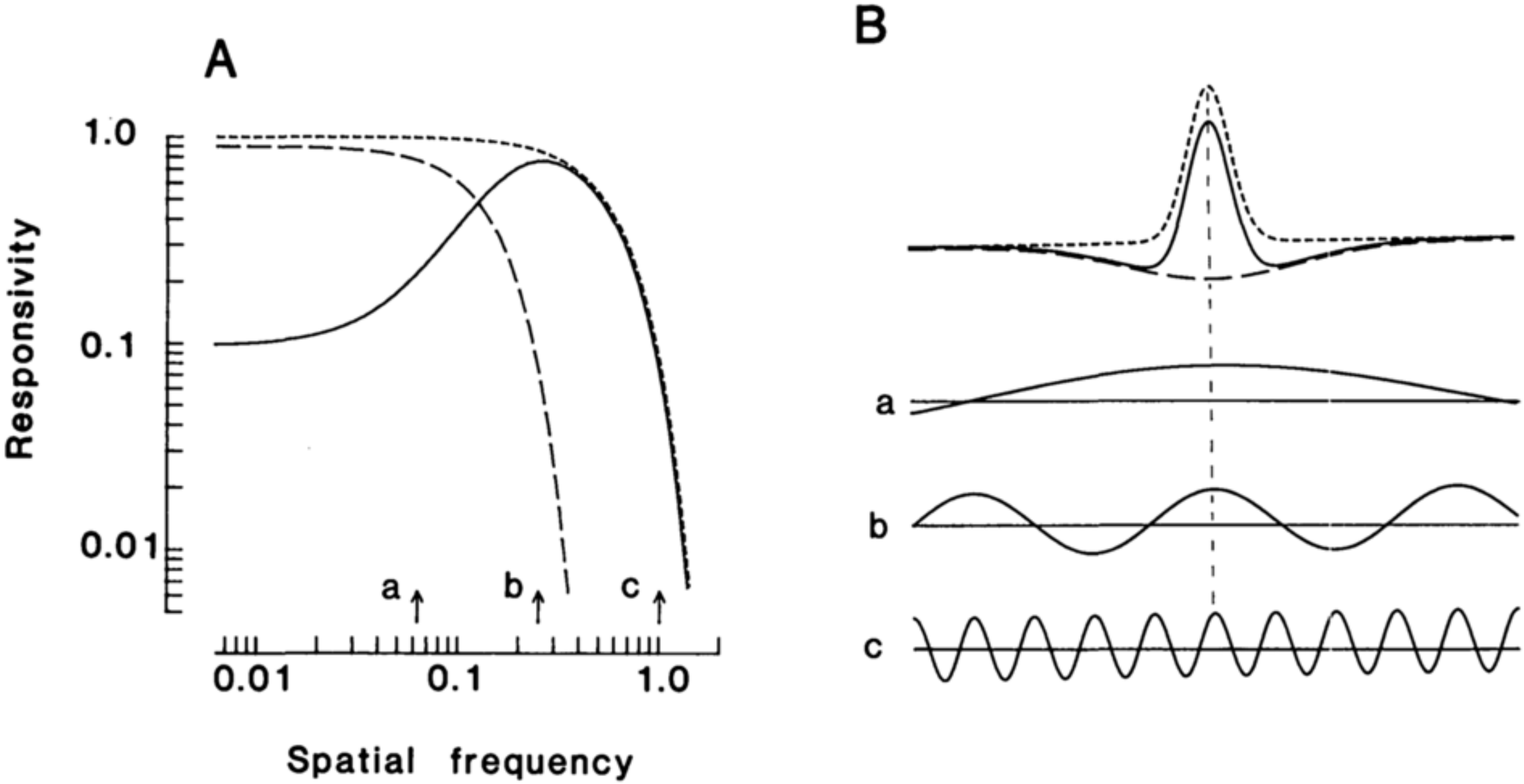
Ganglion cell receptive field modeled as difference of Gaussians





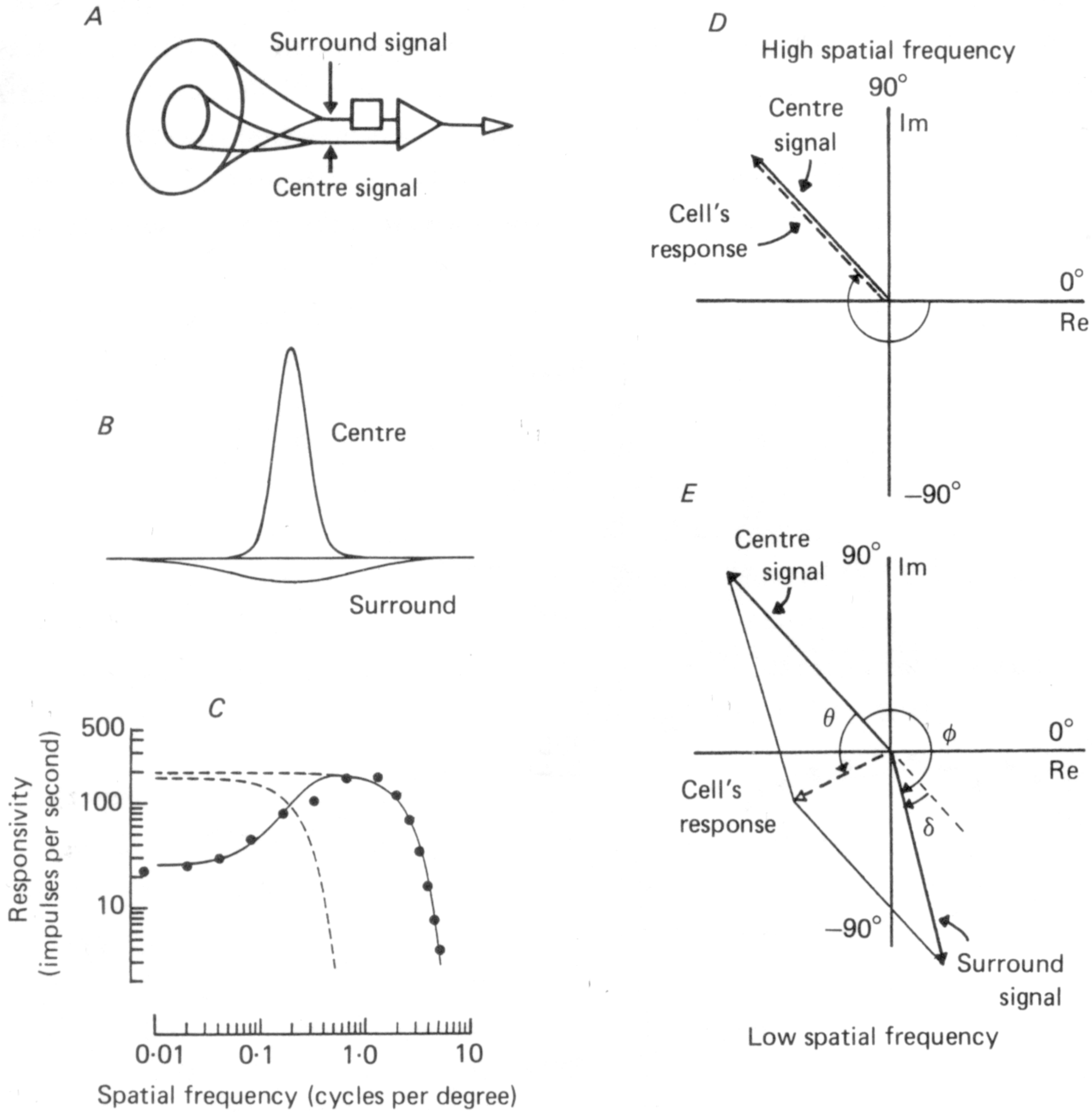
Enroth-Cugell & Robson, 1966

Frequency-domain representation of the difference of Gaussians



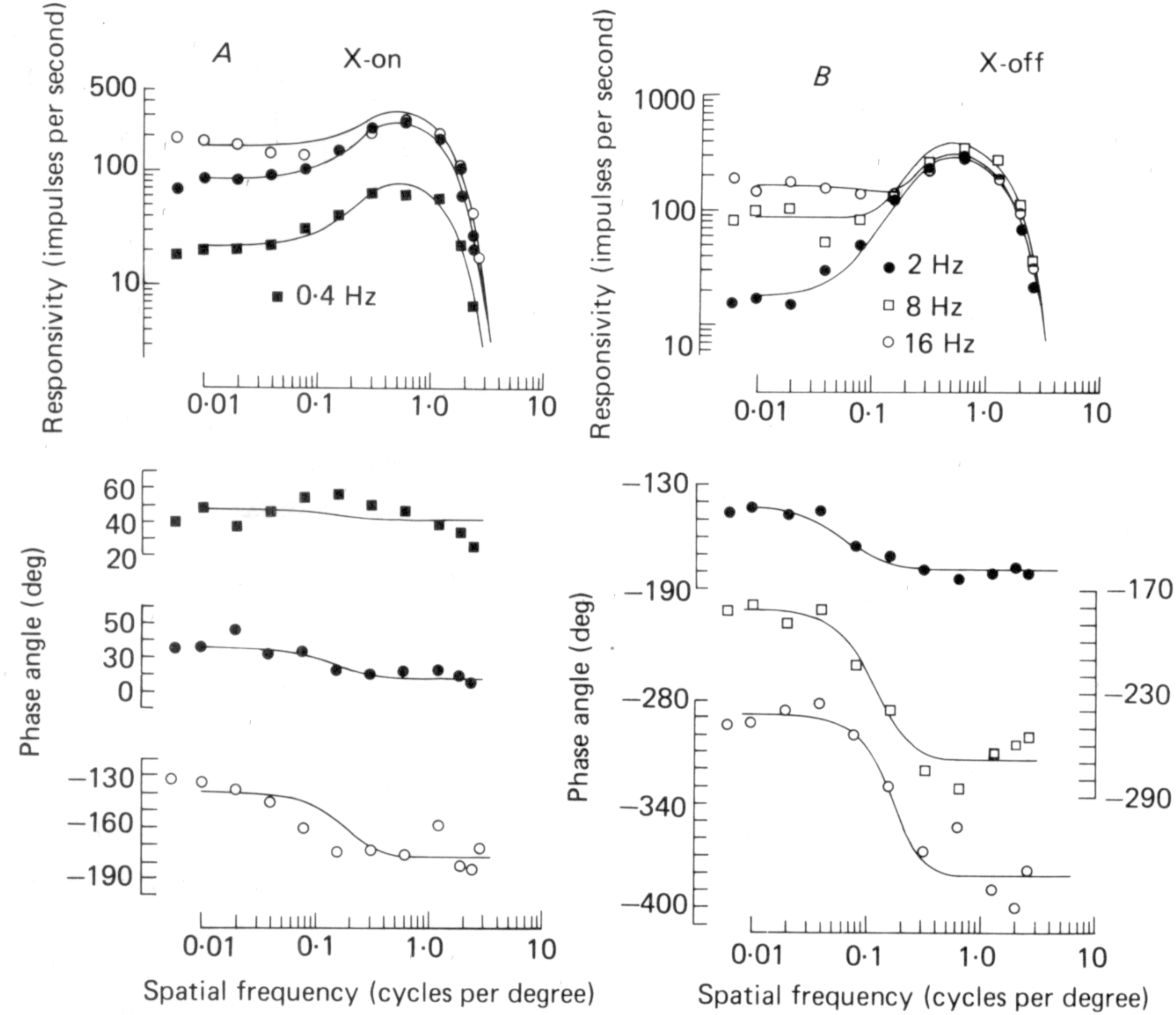
Enroth-Cugell & Robson, 1984

Temporal linearity in X cells



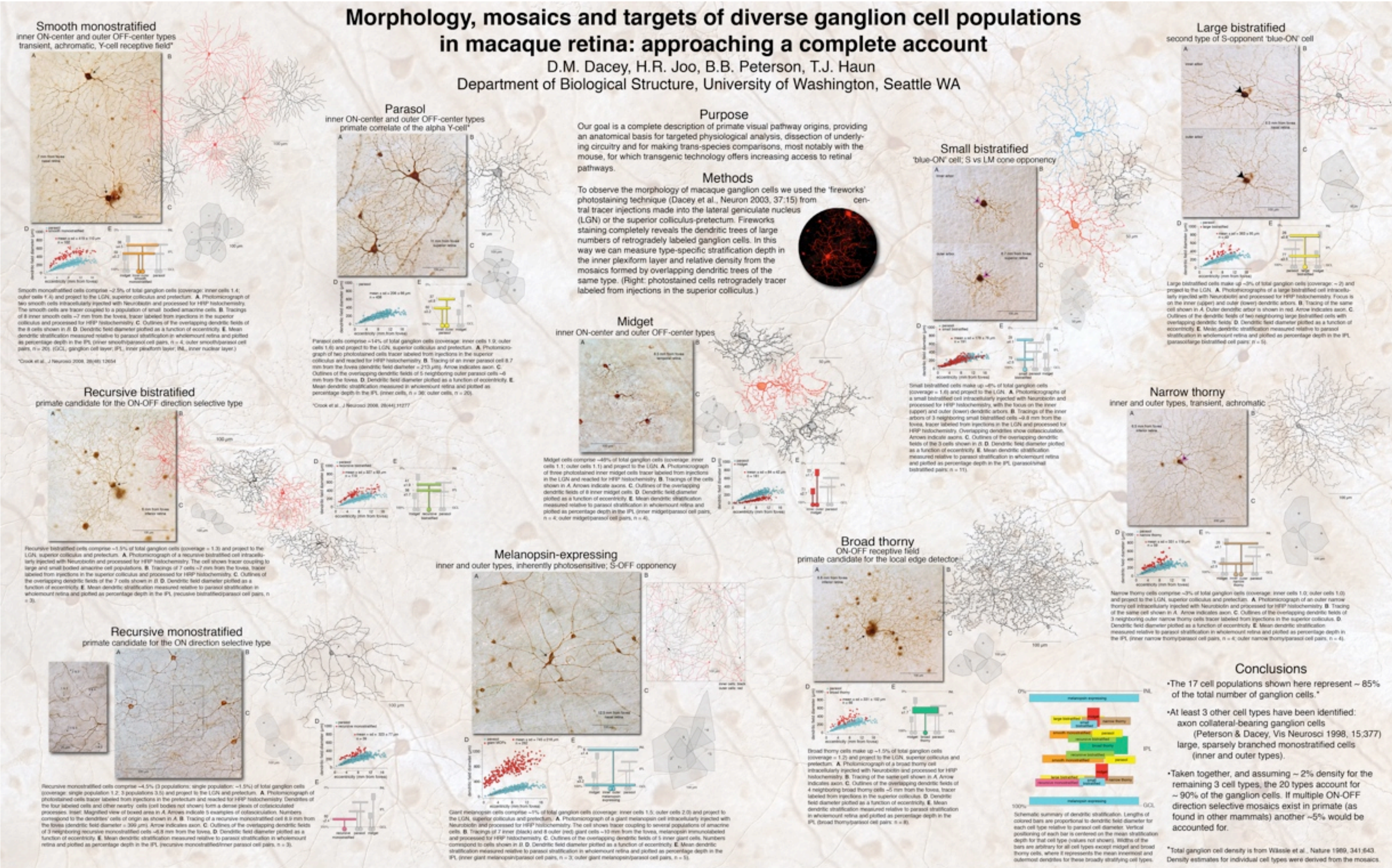
Enroth-Cugell, Robson, Schweitzer-Tong & Watson, 1983

Temporal linearity in X cells

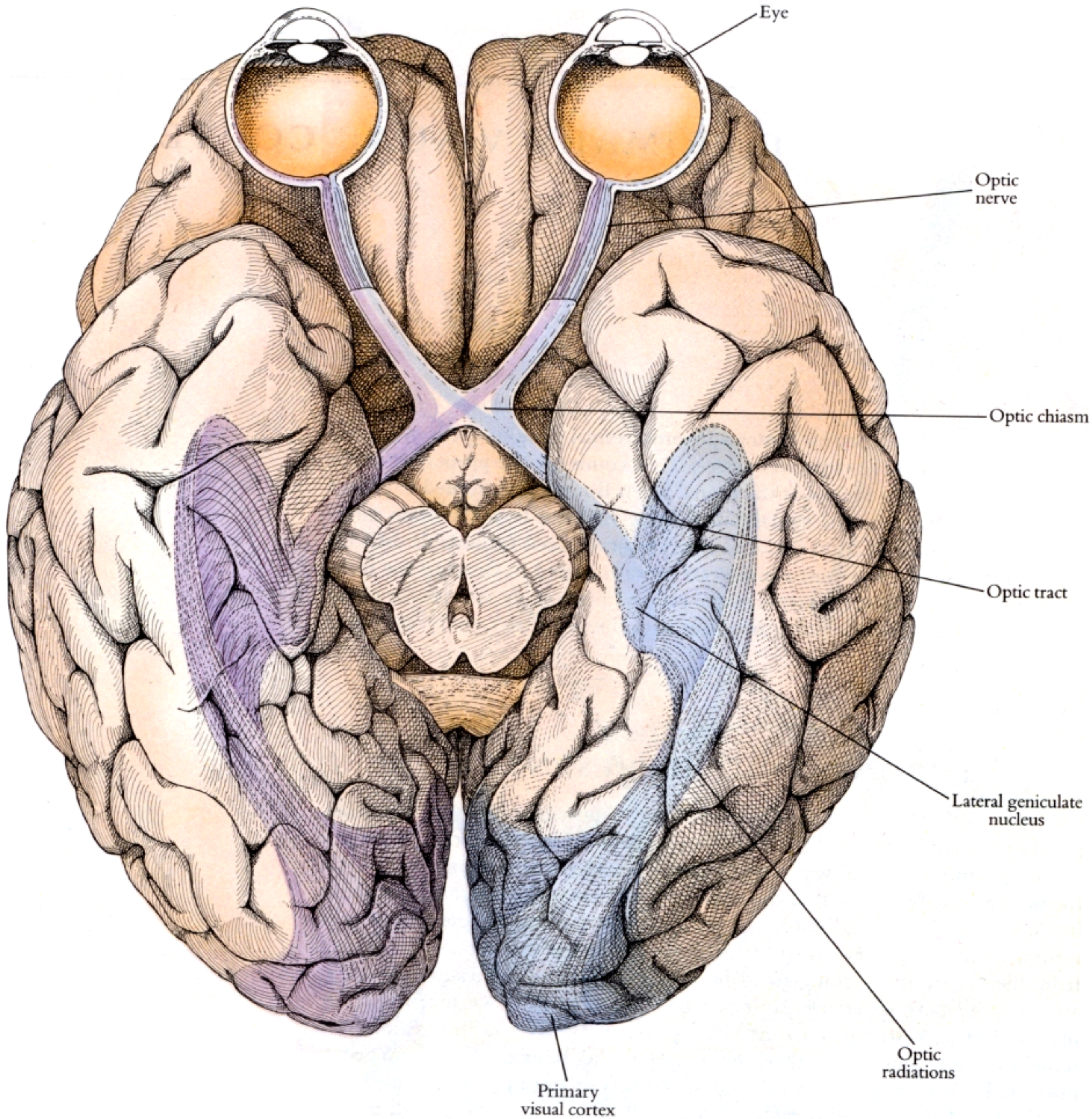


Enroth-Cugell, Robson, Schweitzer-Tong & Watson, 1983

Diversity of retinal ganglion cell types



Neural circuits perform computations



Hubel, 1988