Cortical Coding of Visual Information

Silvio P. Sabatini

silvio@dibe.unige.it

The Physical Structure of Perception and Computation

Department of Biophysical and Electronic Engineering - University of Genoa

In the last decade the opportunity emerges to exceed the formal framework of ANNs, by referring more decisively to models derived in Neurosciences

Getting physical

"Just as in analogue computation, where the physical process is the computation, so in the neocortex the physical structure provides us with constraints that are so conspicuously absent from current 'computational' approaches. It is precisely a belief in the indivisibility of structure and function in the neocortex that now drives us back to confront the cortical microcircuits and ask in what sense they can be said to transform their biology into computation."

R.J. Douglas and K.A.C. Martin, *Opening the grey box*. TINS (1991)

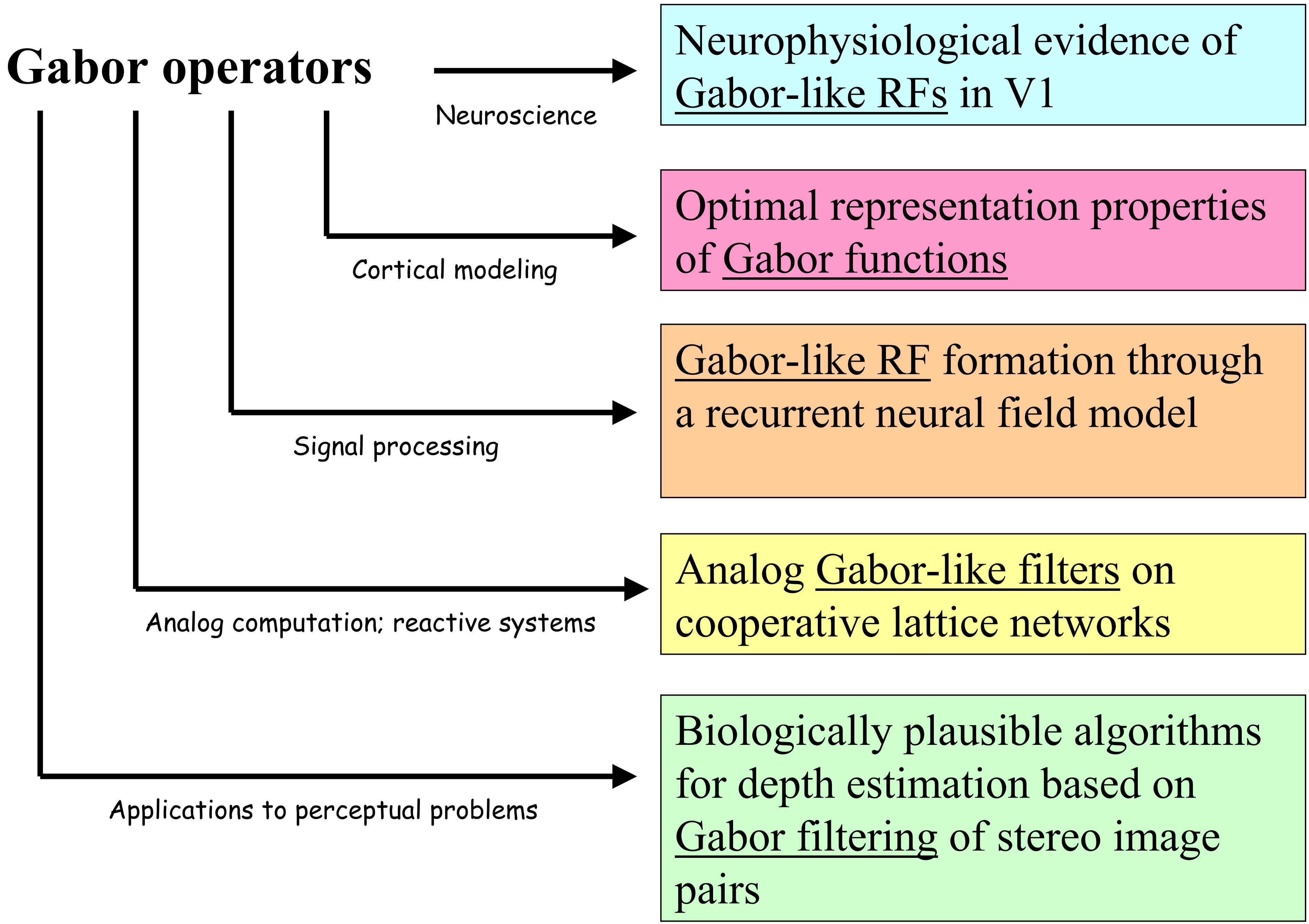
Objective:

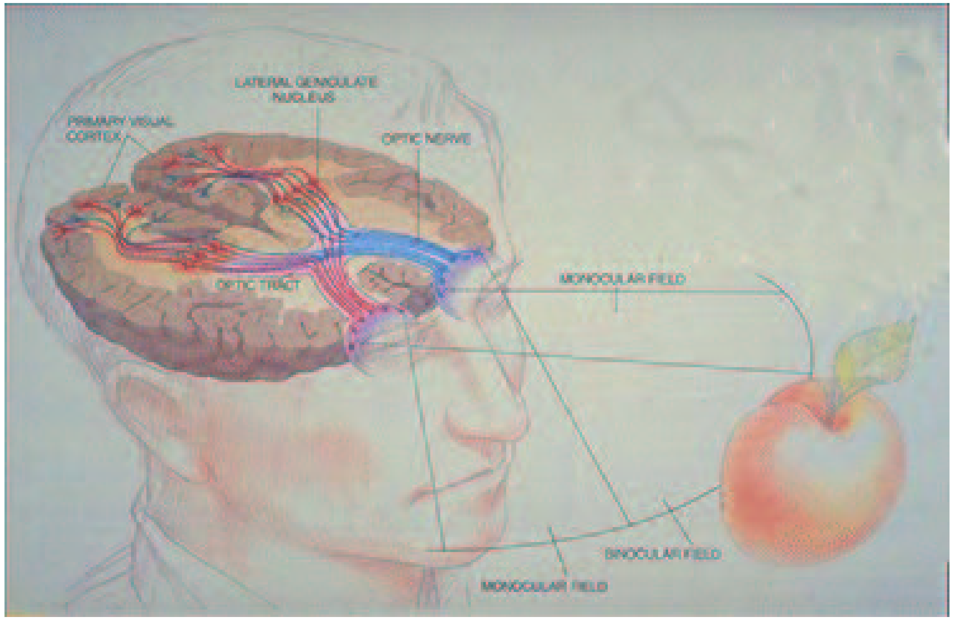
individuation of structural principles underlying visual perception.

- To steer new experimental research
- To foster modeling studies on visual perception and cortical functional architecture
- To conceive innovative hardware and software artificial systems (analog computation, reactive systems)

A case study:

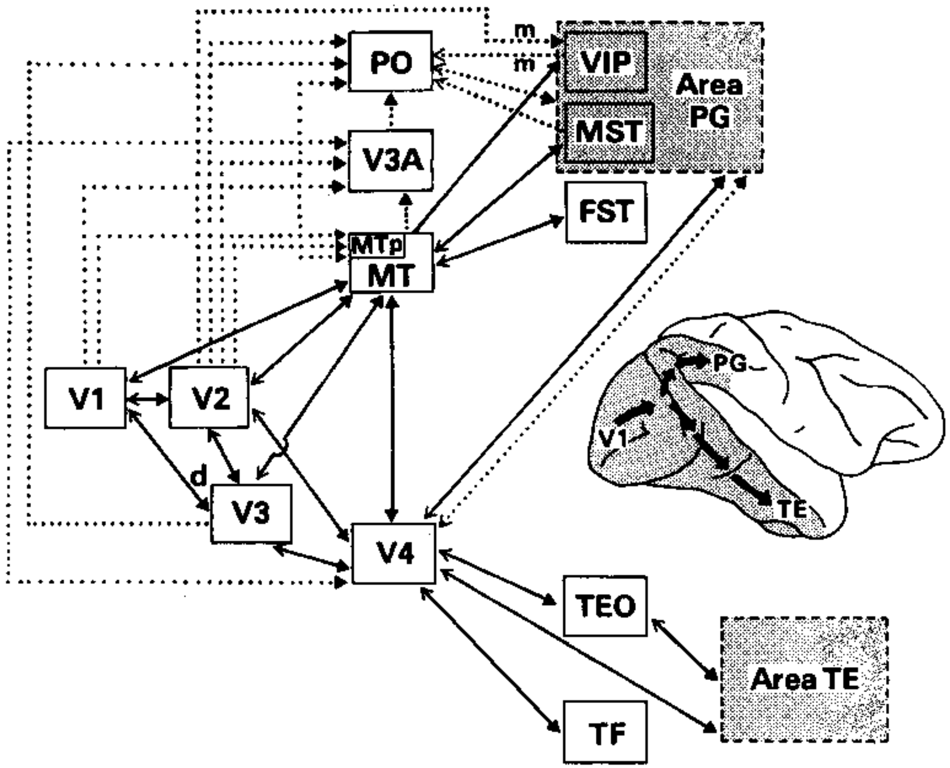
Gabor operators through neurophysiology, computational neuroscience, signal processing, neuromorphic engineering, computational vision


Gabor operators
Neuroscience
Neurophysiological evidence of Gabor-like RFs in V1
Cortical modeling
Optimal representation properties of Gabor functions
Signal processing
Gabor-like RF formation through a recurrent neural field model
Analog computation; reactive systems
Analog Gabor-like filters on cooperative lattice networks
Applications to perceptual problems
Biologically plausible algorithms for depth estimation based on Gabor filtering of stereo image pairs

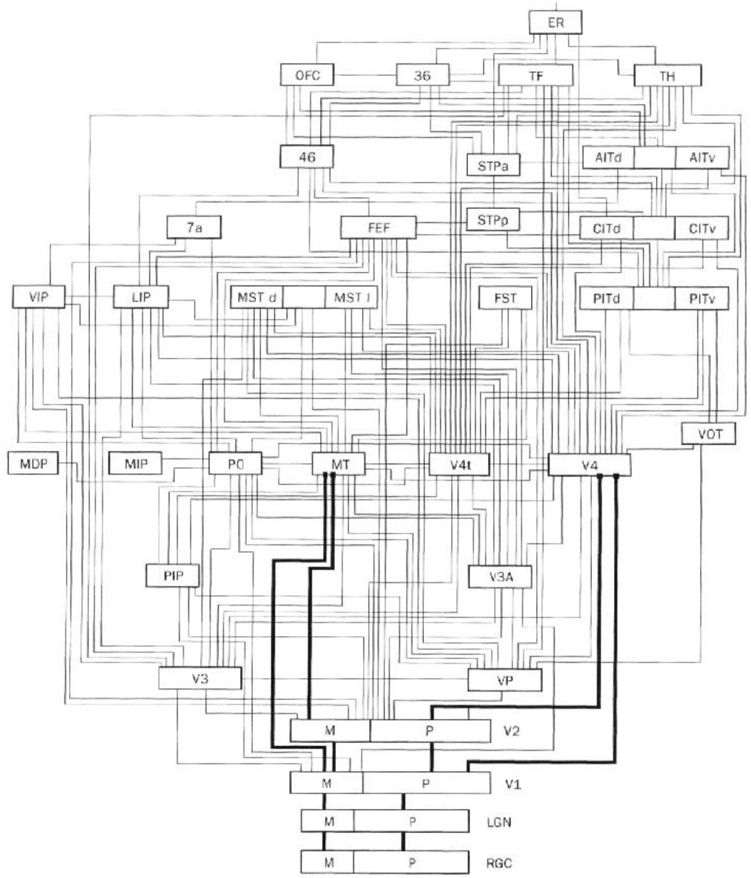


- Different neuronal nuclei (cortical and subcortical)
- More than 20 cortical areas
- 10^3 - 10^5 neurons$/mm^3$
- 10^6 - 10^9 synapses$/mm^3$

CORTICAL FLOW CHARTS



Desimone and Ungerleider, 1989



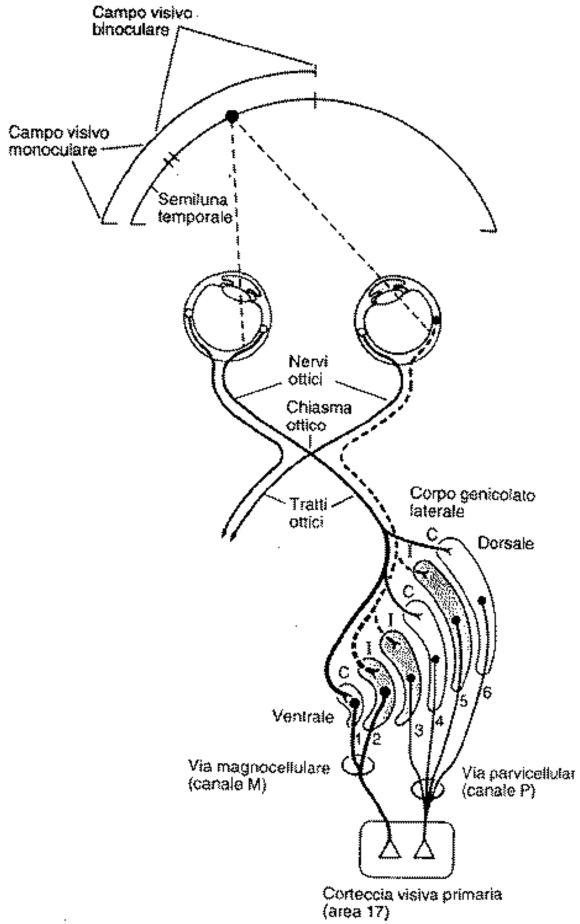
Van Essen *et al.*, 1991

GLOBAL RETINOTOPY

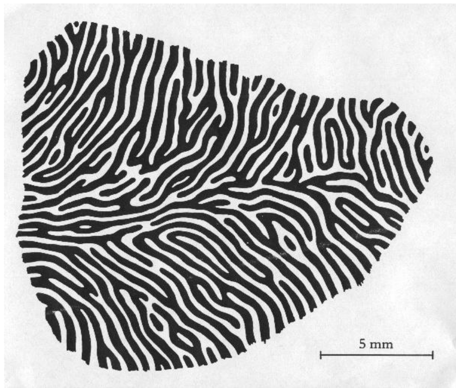
- Cortical mapping of inputs from the left and the right eyes and the ocular dominance columns

- Topographic (retinotopic) mapping

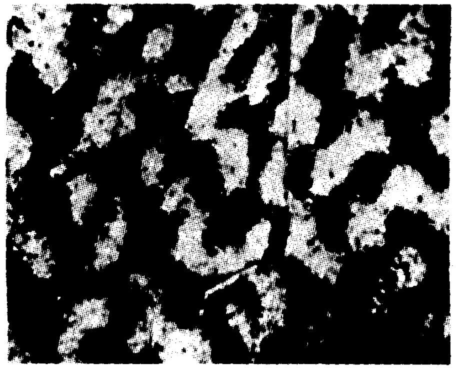
OCULAR DOMINANCE COLUMNS



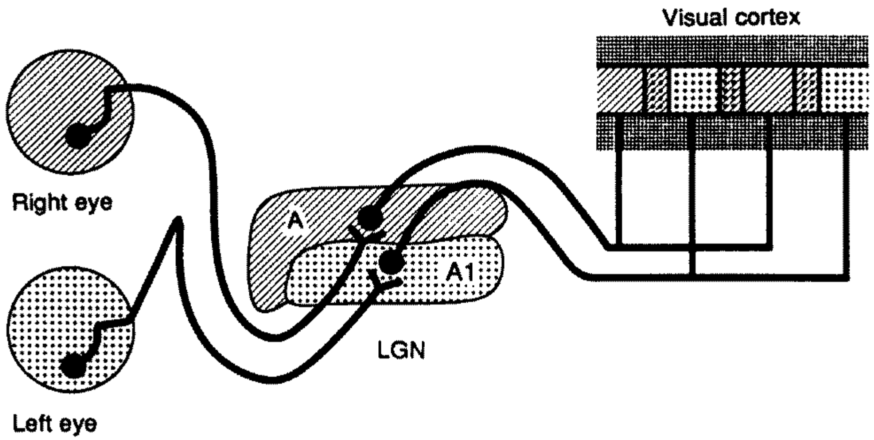
Kandel *et al.*, 1991



LeVay *et al.*, 1975

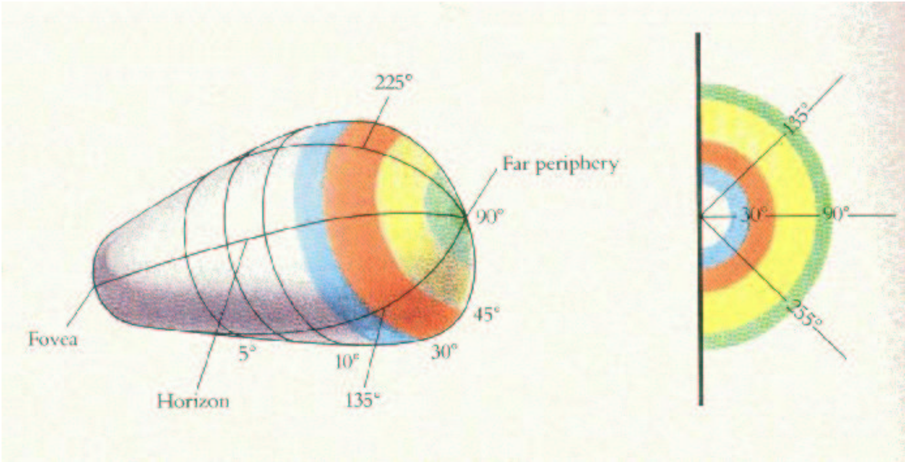


Miller *et al.*, 1989



Miller *et al.*, 1989

RETINOTOPIC MAPPING

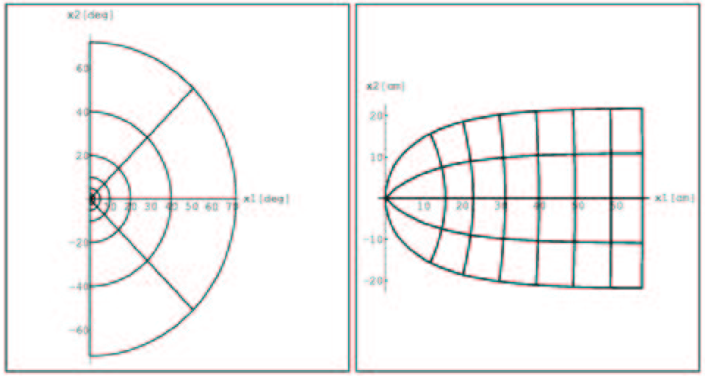


Hubel, 1988

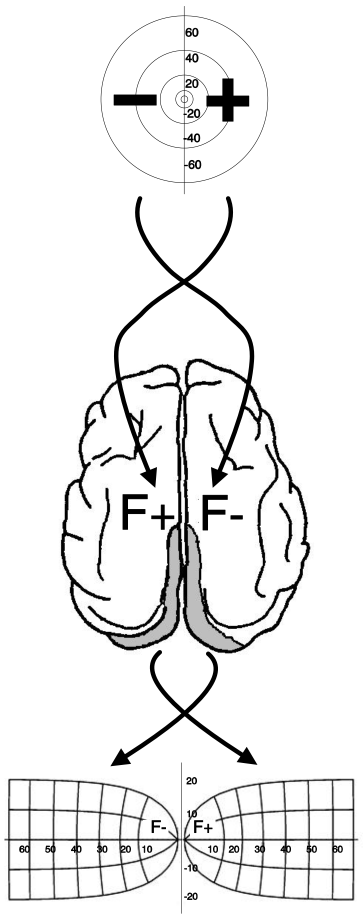
Formally:

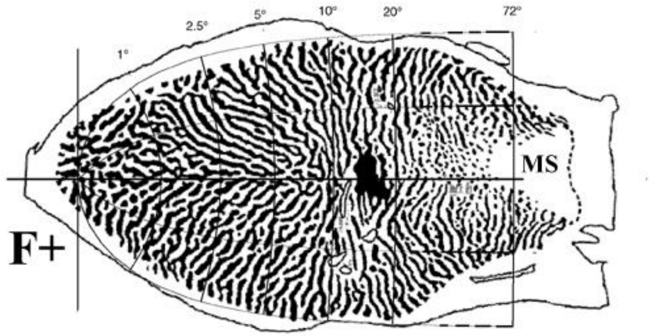
$$\mathbf{x} = \log(\mathbf{z} + \alpha)$$

Schwartz, 1977

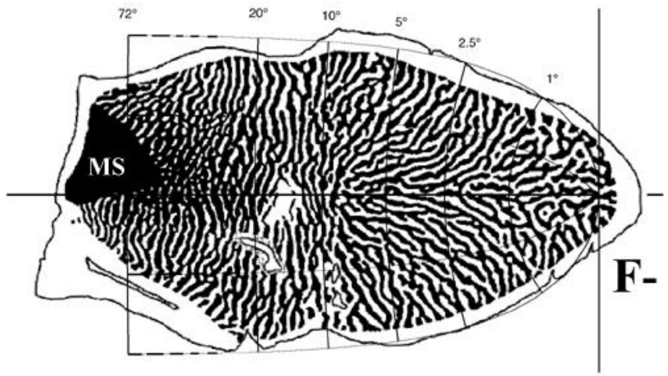


RETINOTOPIC MAPPING (cont'd)



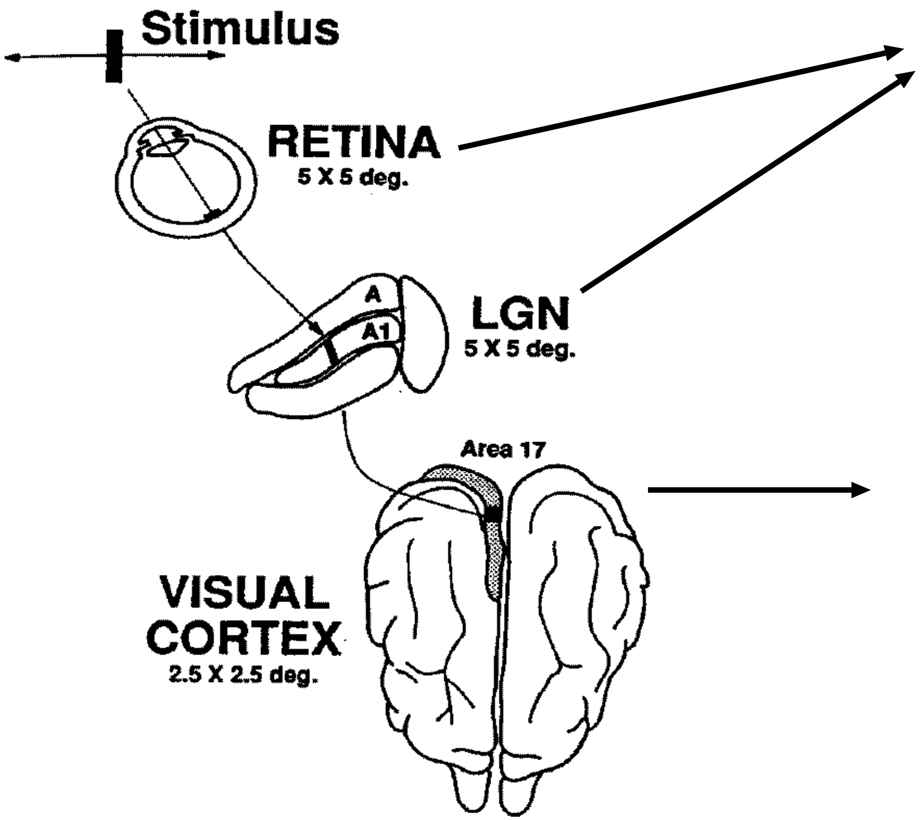


left hemisphere
right hemifield

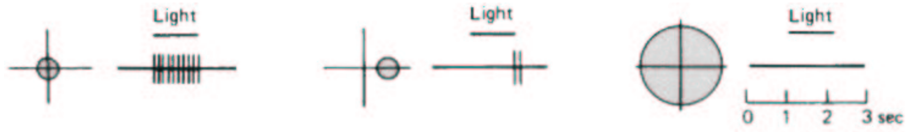


right hemisphere
left hemifield

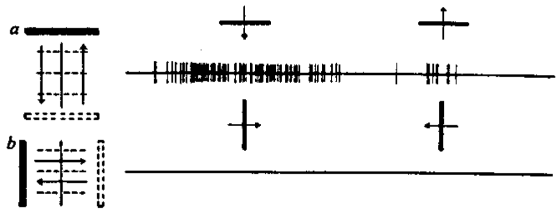
THE PRIMARY VISUAL PATHWAY



Spots

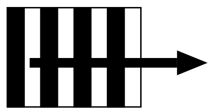


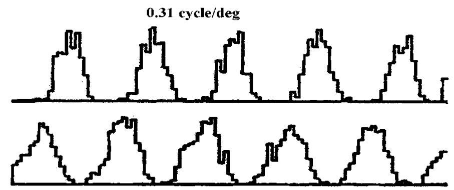
Bars



Hubel and Wiesel, 1959-62

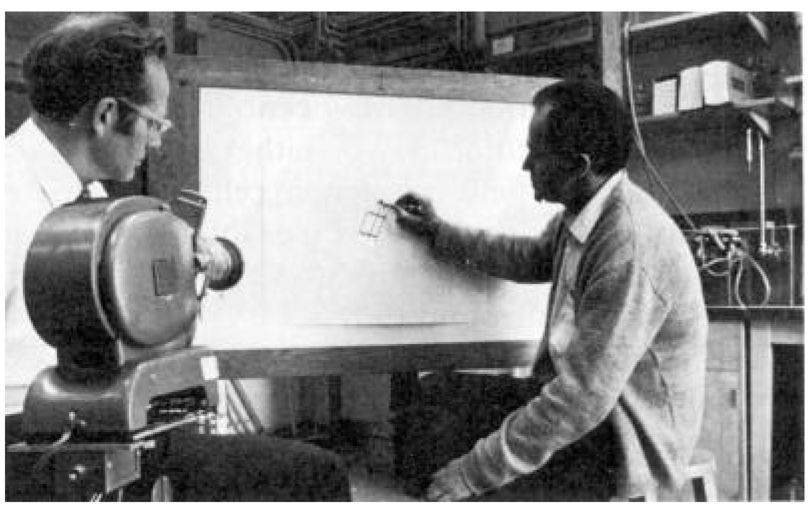
Gratings





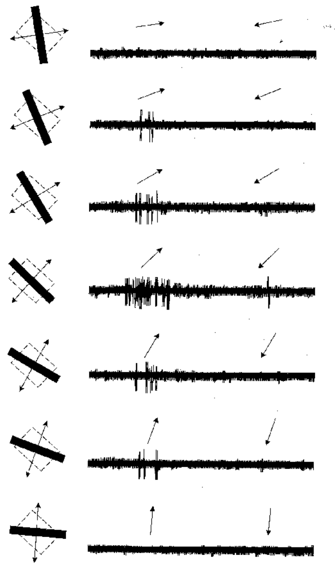
Pollen and Ronner, 1981

Responses of neurons in the primary visual cortex (V1)



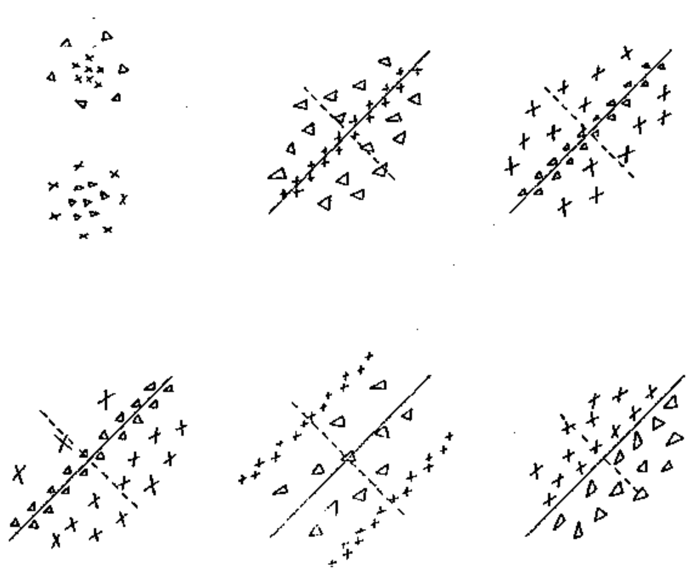
Hubel and Wiesel, circa 1969
in Nicholls et al. (1992)

Selectivity for stimulus orientation and direction



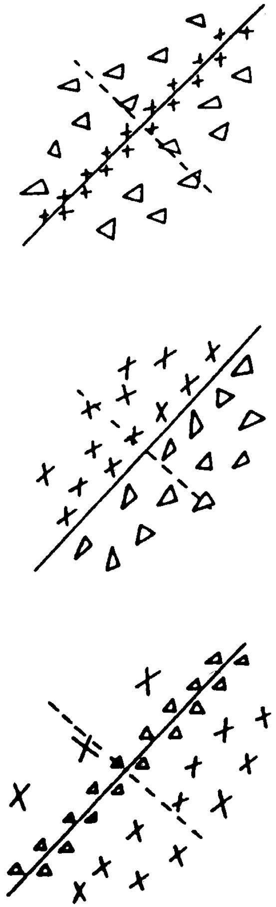
Hubel and Wiesel (1968)
in Wandell (1995)

Receptive fields of LGN and V1 simple cells

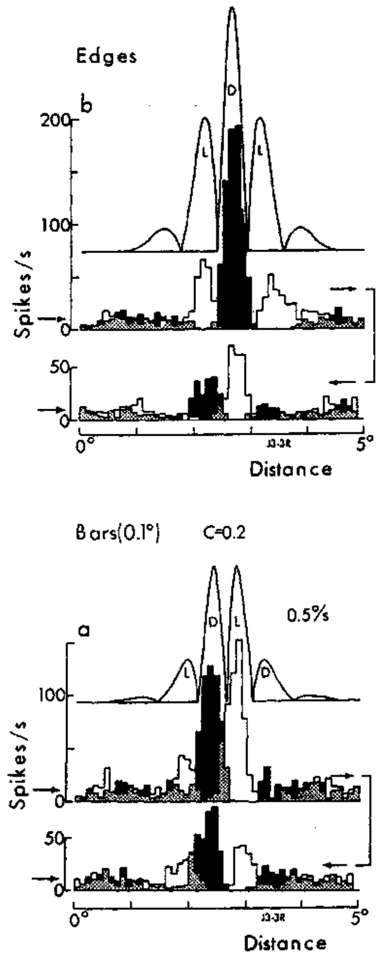


Hubel & Wiesel (1963)

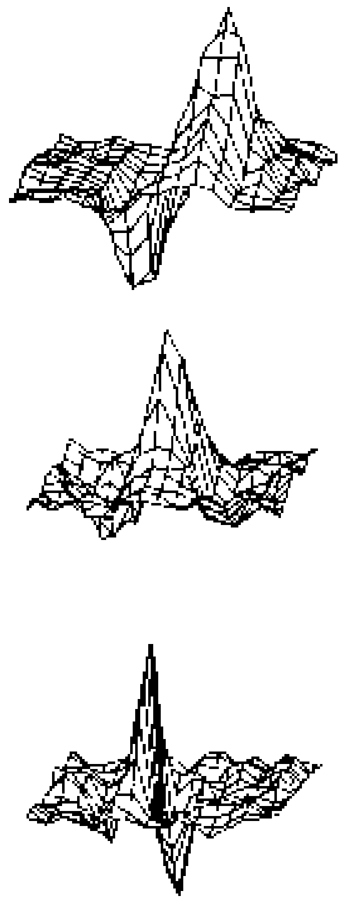
THE SPATIAL STRUCTURE OF RECEPTIVE FIELDS IN V1



Hubel and Wiesel, 1962

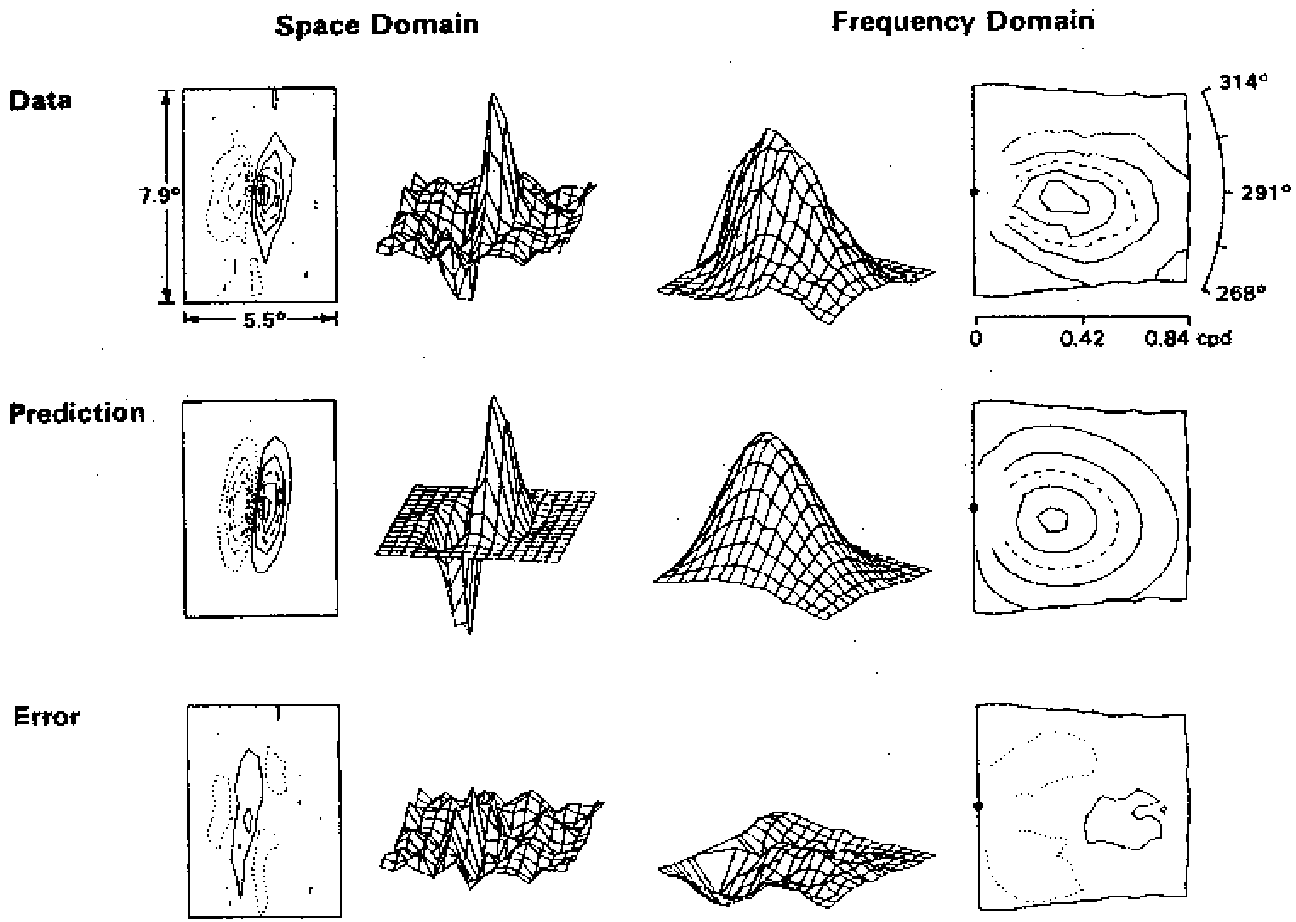


Kulikowski and Bishop, 1981

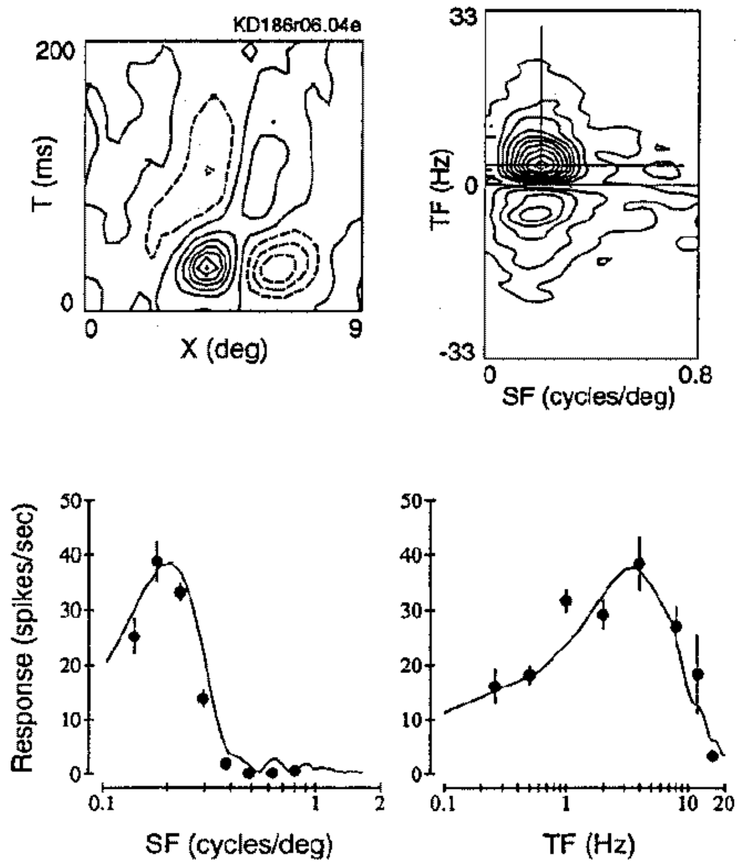


Jones and Palmer, 1987

THE LINEARITY ASSUMPTION



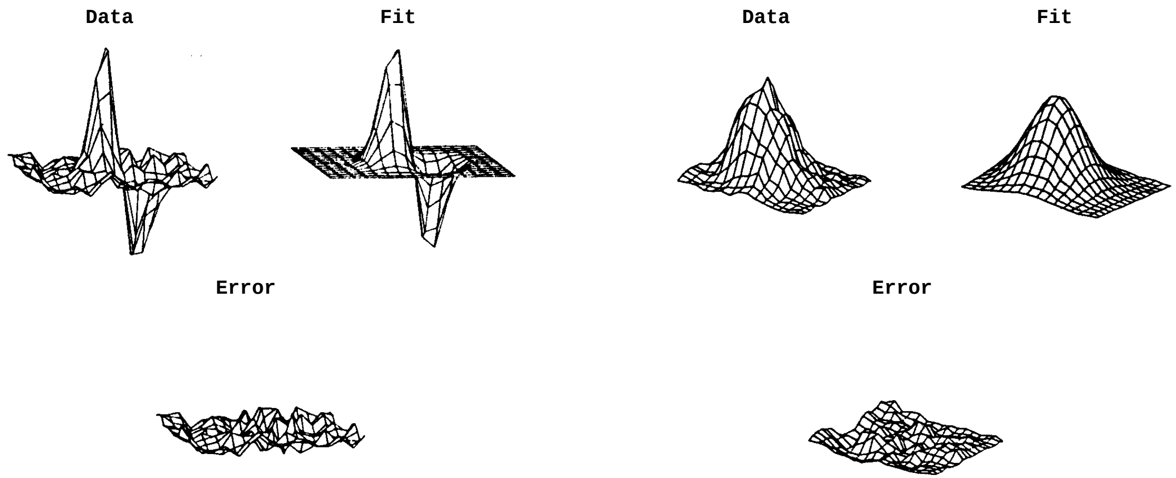
Jones and Palmer, 1987



DeAngelis *et al.*, 1993

GABOR-LIKE RECEPTIVE FIELDS

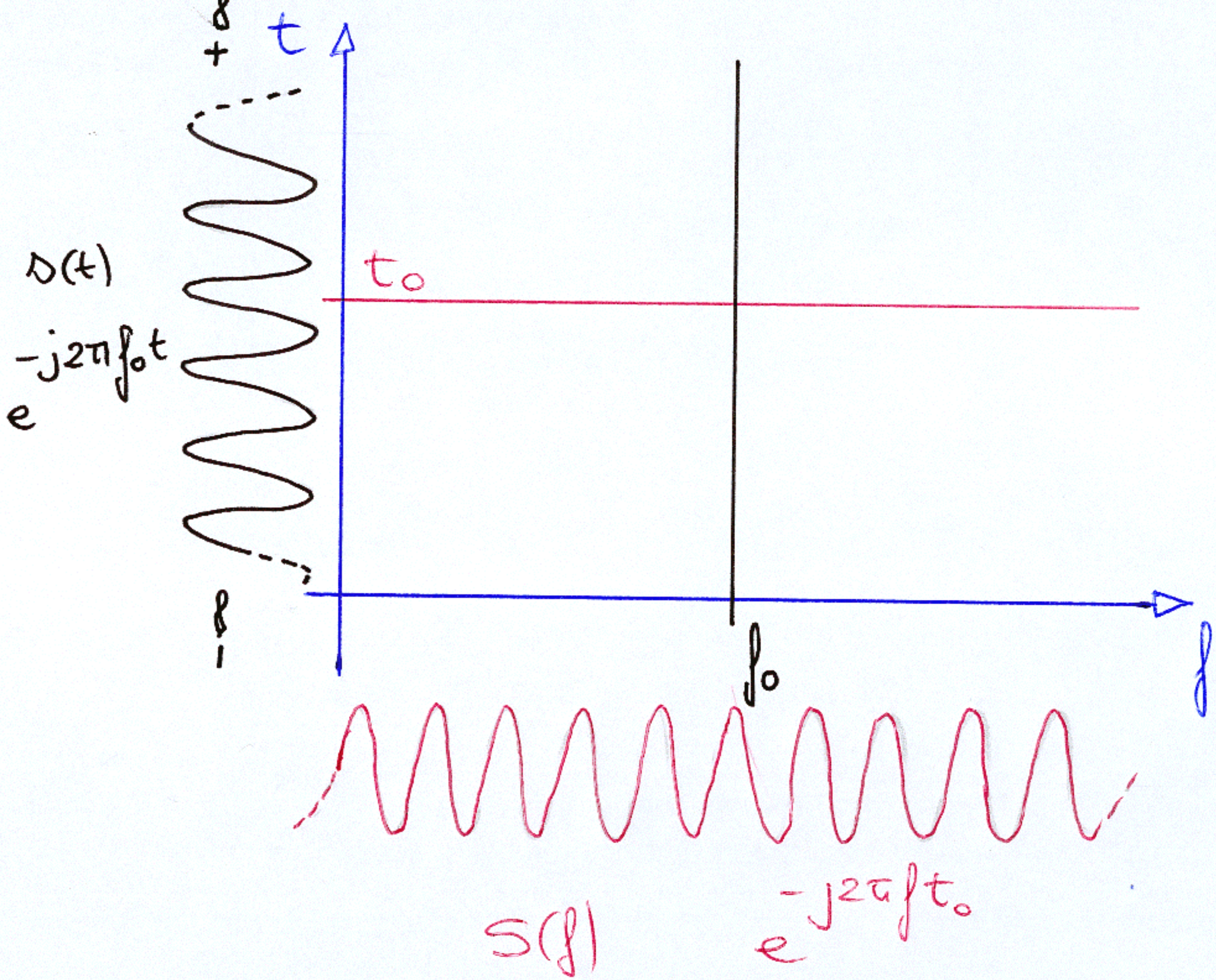
$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{(-\frac{x^2+y^2}{2\sigma^2})} cos(f_0 x + \phi) \;\circ\!\!-\!\!\bullet\; H(f_x, f_y) = \frac{1}{2} e^{-2\pi^2\sigma^2(f_x^2+f_y^2)} * [e^{j\phi}\delta(f_x - f_0) + e^{-j\phi}\delta(f_x + f_0)]$$



Jones and Palmer, 1987

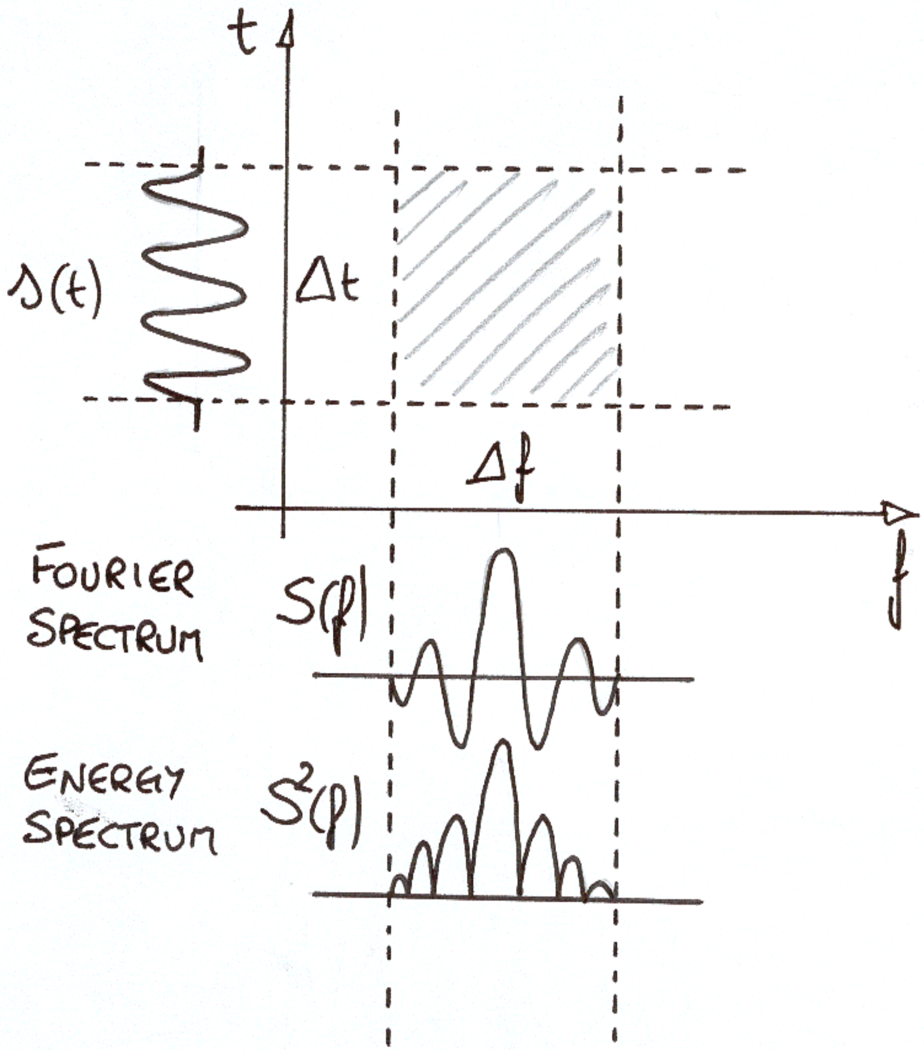
Gabor functions are characterized by optimal localization properties in both spatial and spatial-frequency domains (Daugman, 1984) $\triangle x \triangle y \triangle f_x \triangle f_y \longrightarrow 1/16\pi^2$

JOINT LOCALIZATION IN TIME (SPACE) AND FREQUENCY



INFORMATION DIAGRAMS

JOINT LOCALIZATION IN TIME (SPACE) AND FREQUENCY (CONT'D)



INFORMATION $\propto \Delta t \Delta f$

HEISENBERG PRINCIPLE

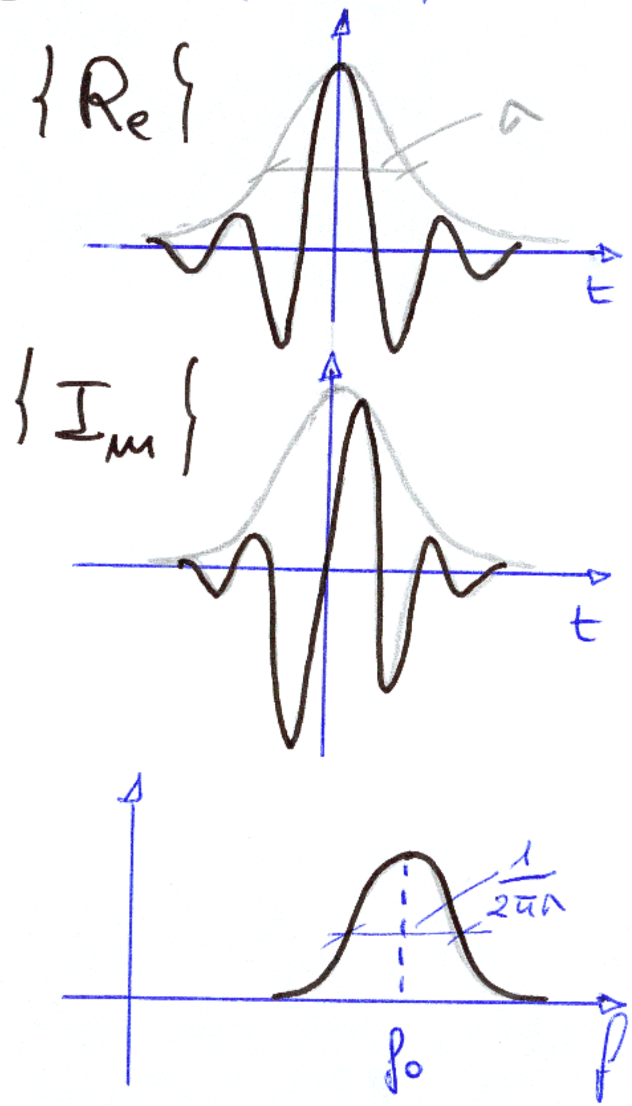
$$\Delta t \Delta f \geq \frac{1}{4\pi}$$

THE 1-D GABOR LOGON (GABOR, 1946)

$$h(t) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(t-t_0)^2}{2\sigma^2}} \cdot e^{2\pi j f_0 t}$$

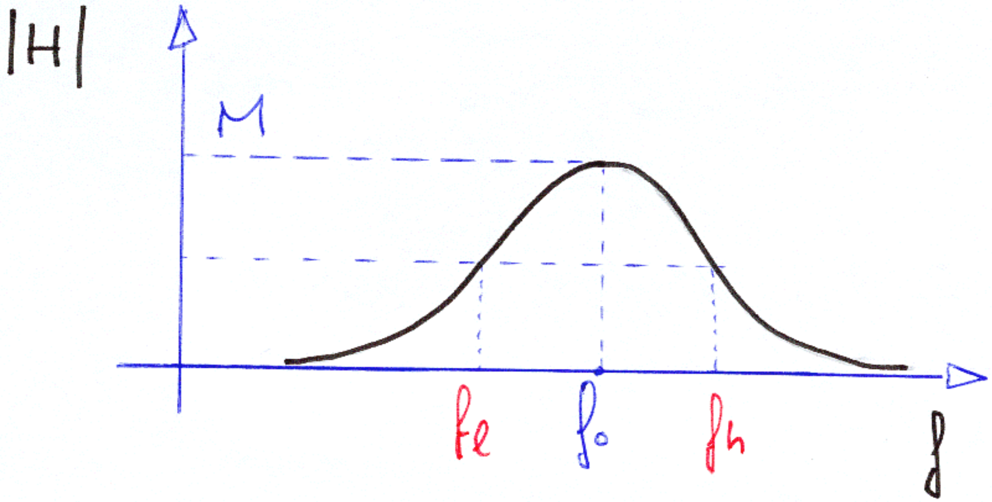
$$\Delta t \, \Delta f = \frac{1}{4\pi}$$

$$H(f) = e^{-2(f-f_0)^2 \pi^2 \sigma^2} \cdot e^{-2\pi j f t_0}$$



BAND-PASS FILTER

FREQUENCY CHARACTERIZATION



$$|H(f)| = \frac{1}{2}|H(f)|_{MAX} = \frac{M}{2}$$

$$f = f_l \cdot f_h$$

f_0 = PEAK FREQUENCY

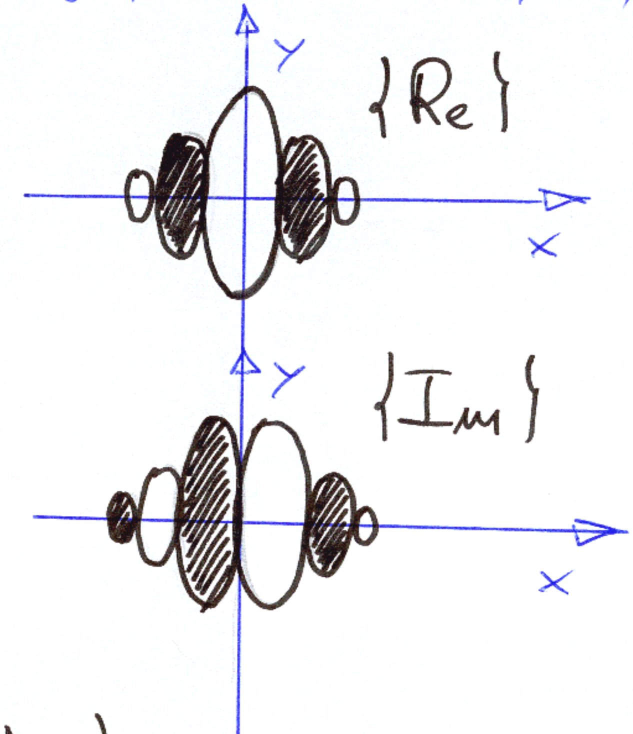
ABSOLUTE BANDWIDTH : $\Delta f = f_h - f_l$

RELATIVE BANDWIDTH : $B_{df} = \frac{\Delta f}{f_0} = f_h/f_0 - f_l/f_0$

OCTAVE RELATIVE BANDWIDTH : $B_{OCT} = \log_2 \frac{f_0}{f_l}$

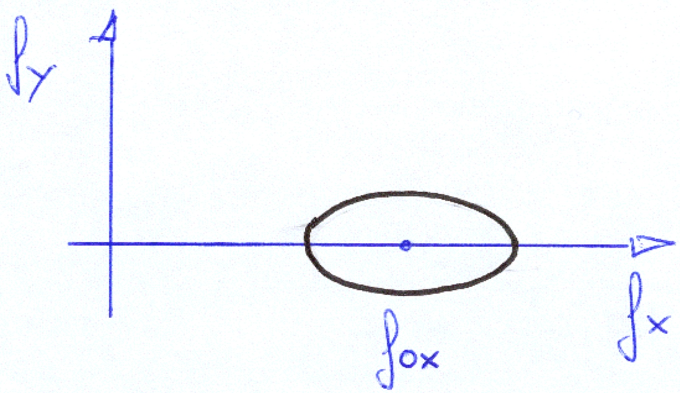
THE 2-D GABOR LOGON (MARČELJA, 1980 ; DAUGMAN, 1985)

$$h(x,y) = \frac{1}{2\pi \sigma_x \sigma_y} e^{-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}} \cdot e^{2\pi j f_{0x} x}$$



$$\Delta x \Delta y \Delta f_x \Delta f_y = \frac{1}{16\pi^2}$$

$$H(f_x,f_y) = e^{-2(f_x - f_{0x})^2 \sigma_x^2 \pi^2 - 2 f_y^2 \sigma_y^2 \pi^2} \cdot e^{-2\pi j (f_x x_0 + f_y y_0)}$$



in generale ....

THE 2-D GABOR LOGON

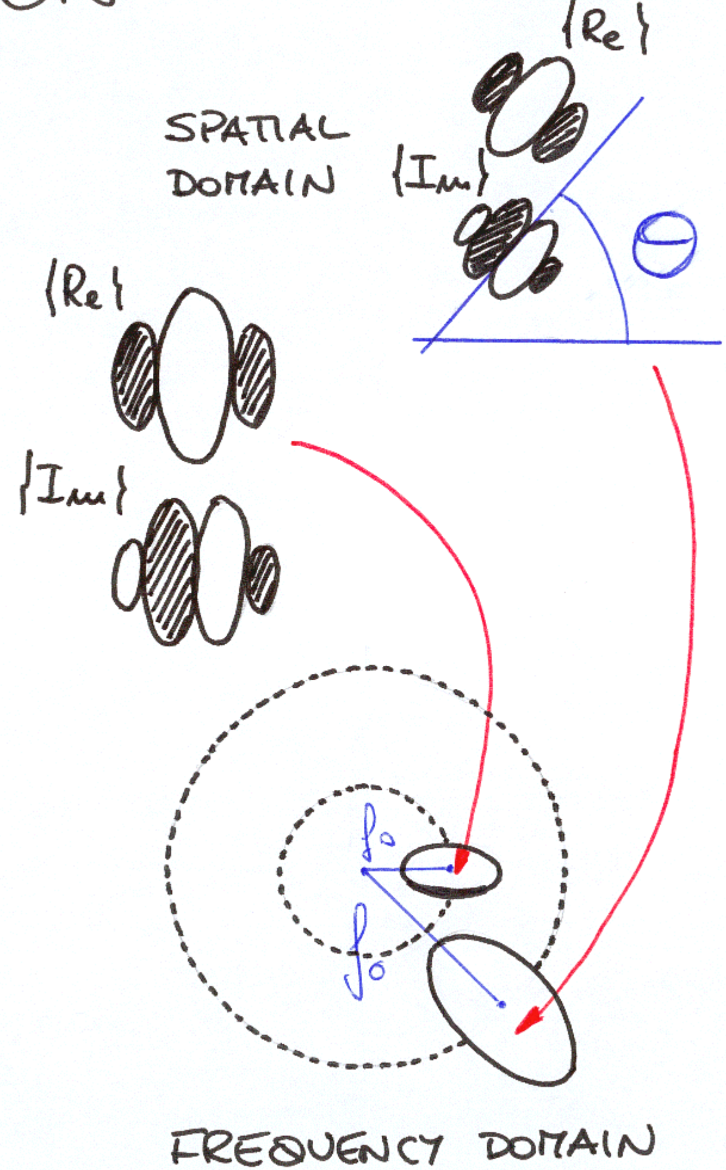
$$h(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\, e^{-\frac{(x')^2}{2\sigma_x^2} - \frac{(y')^2}{2\sigma_y^2}} \cdot e^{2\pi j f_0 x'}$$

WITH

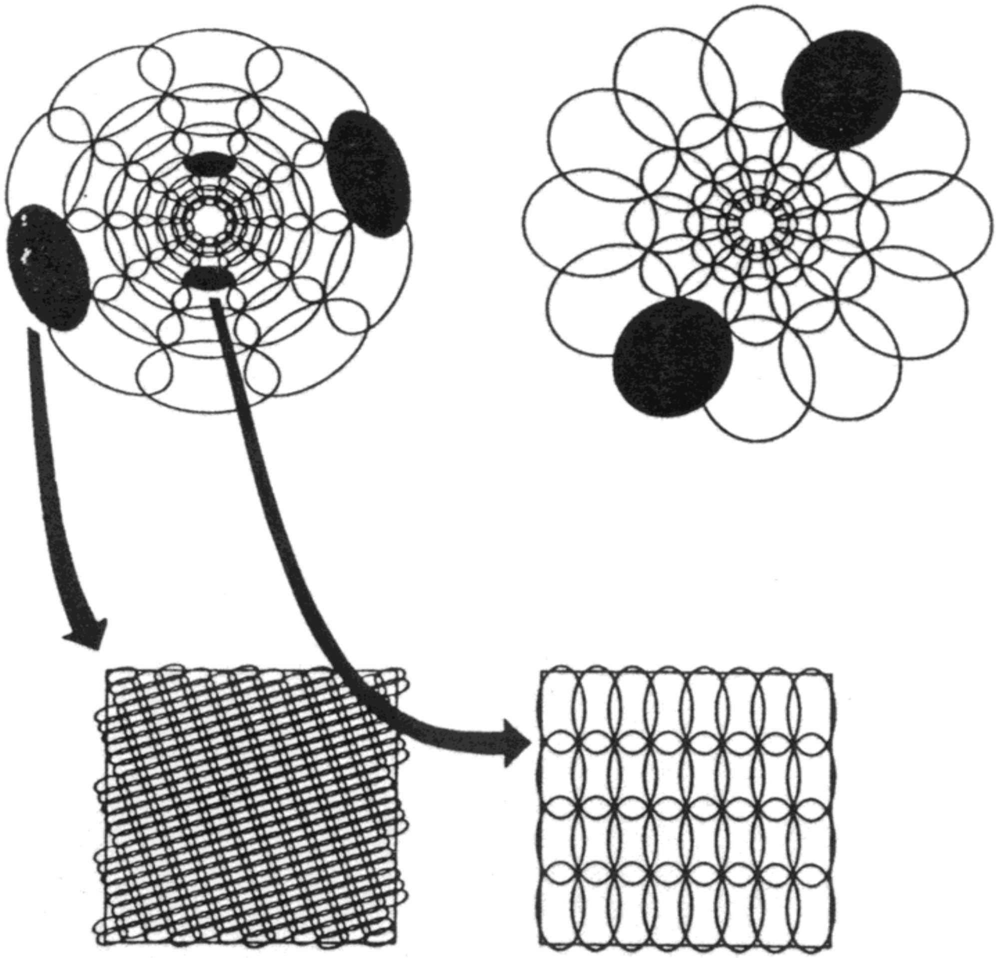
$$\begin{cases} x' = (x - x_0)\cos\theta + (y - y_0)\sin\theta \\ y' = -(x - x_0)\sin\theta + (y - y_0)\cos\theta \end{cases}$$

$$\theta = \operatorname{atan}\frac{f_{oy}}{f_{ox}}$$

$$f_0 = \sqrt{f_{ox}^2 + f_{oy}^2} =$$ RADIAL PEAK FREQUENCY



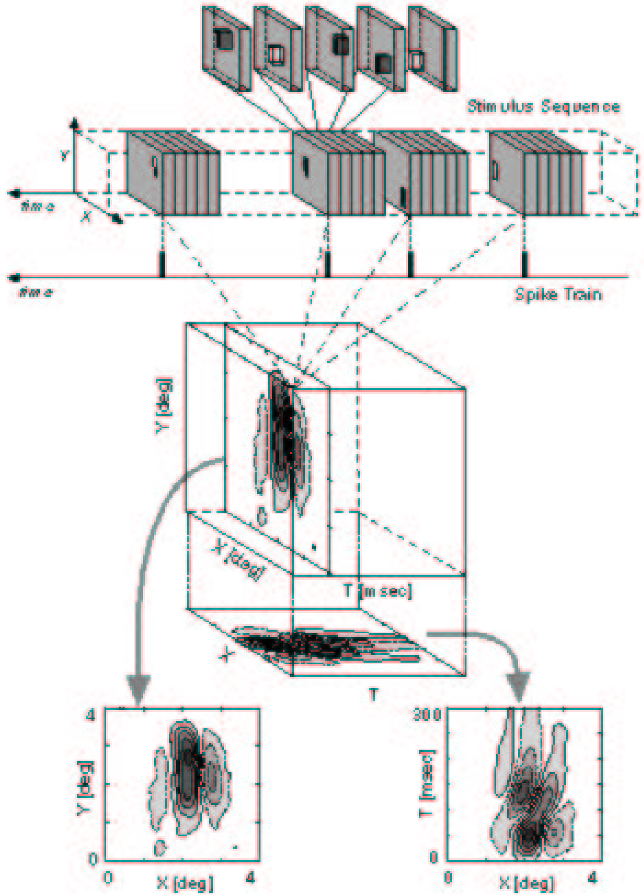
DAISY DIAGRAMS



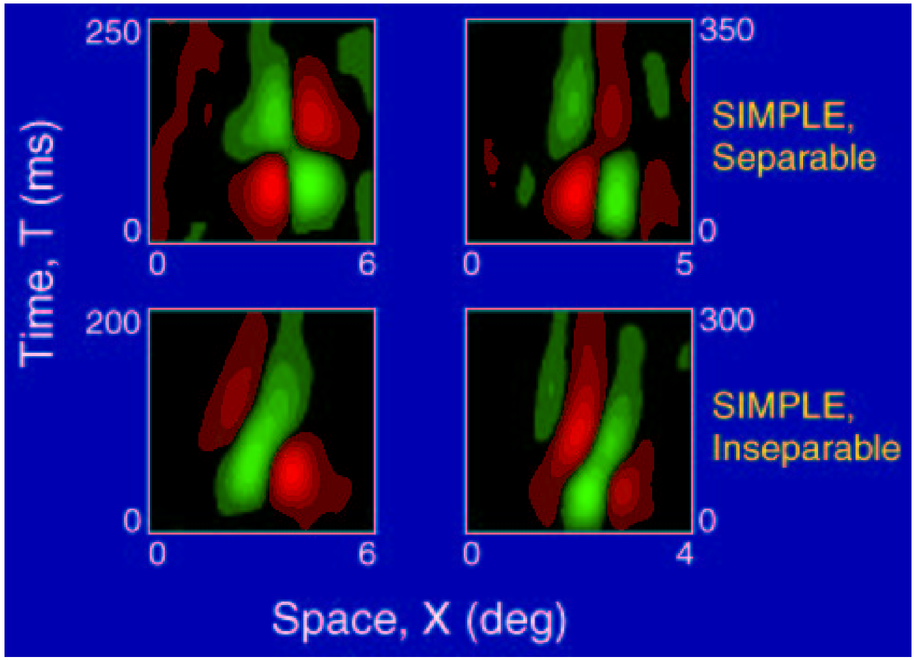
Frequency domain

Spatial domain

SPATIO-TEMPORAL *RFs* IN V1

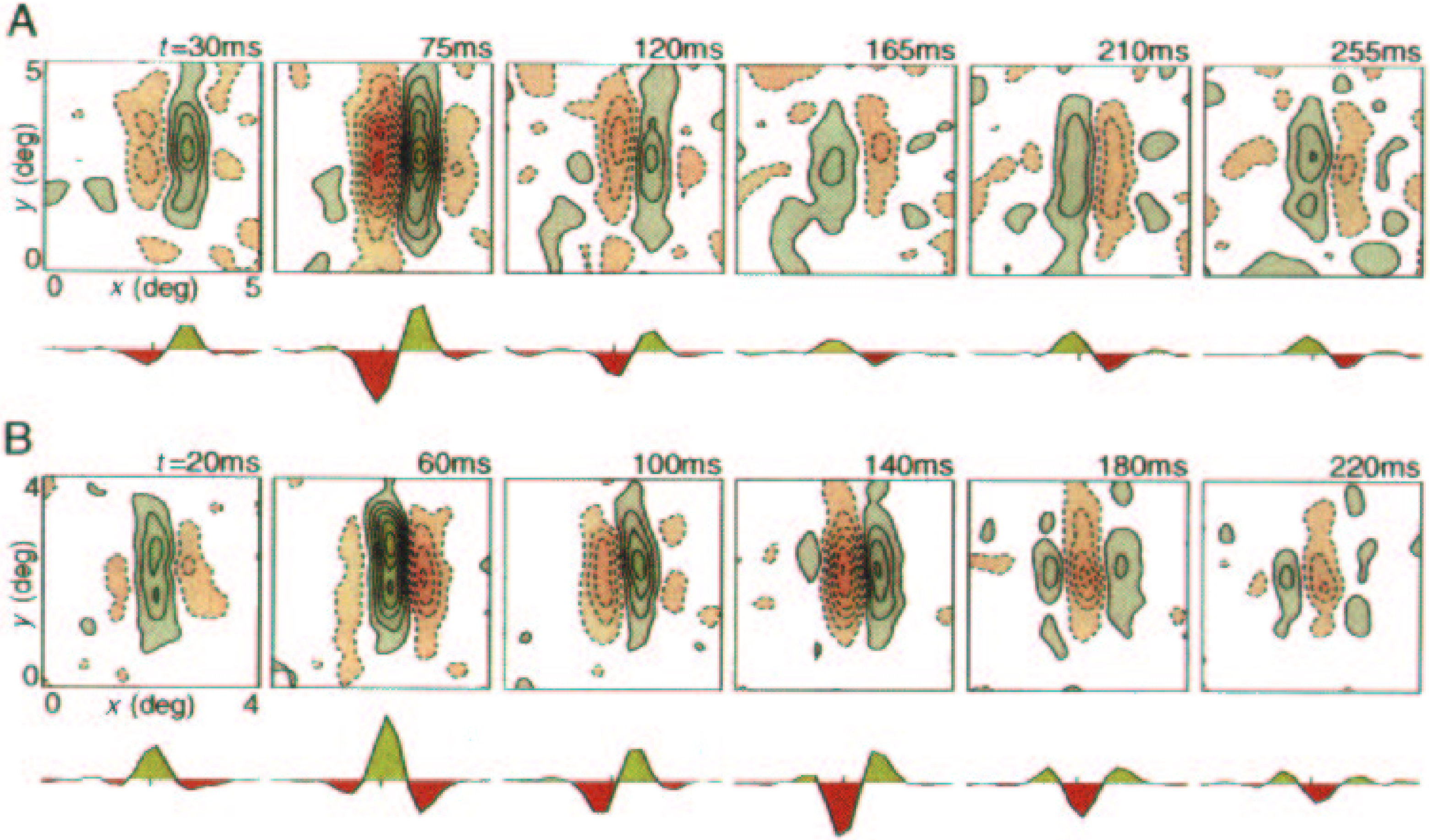


Ohzawa *et al.*, 1996



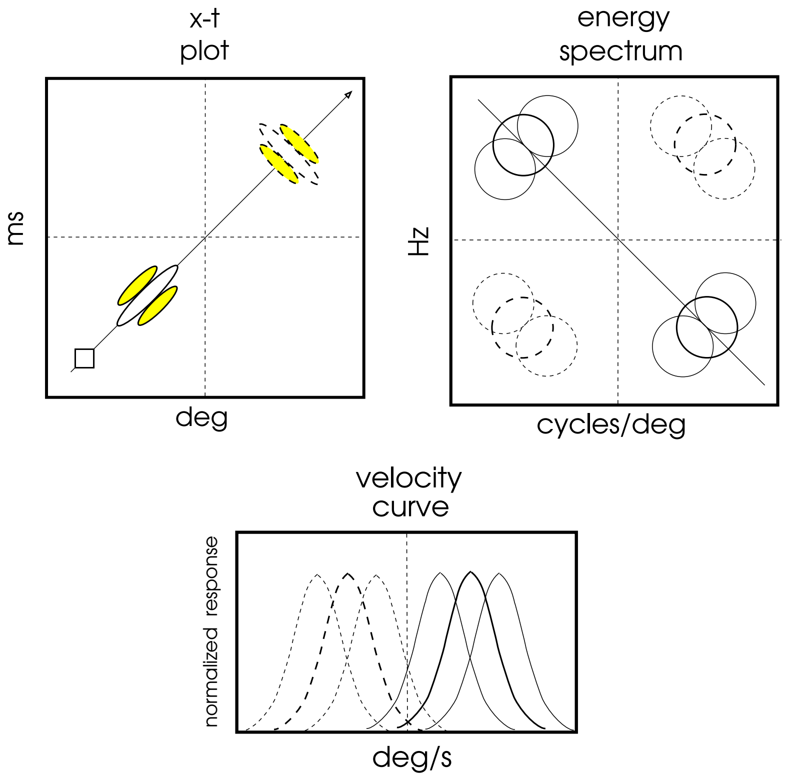
DeAngelis *et al.*, 1995

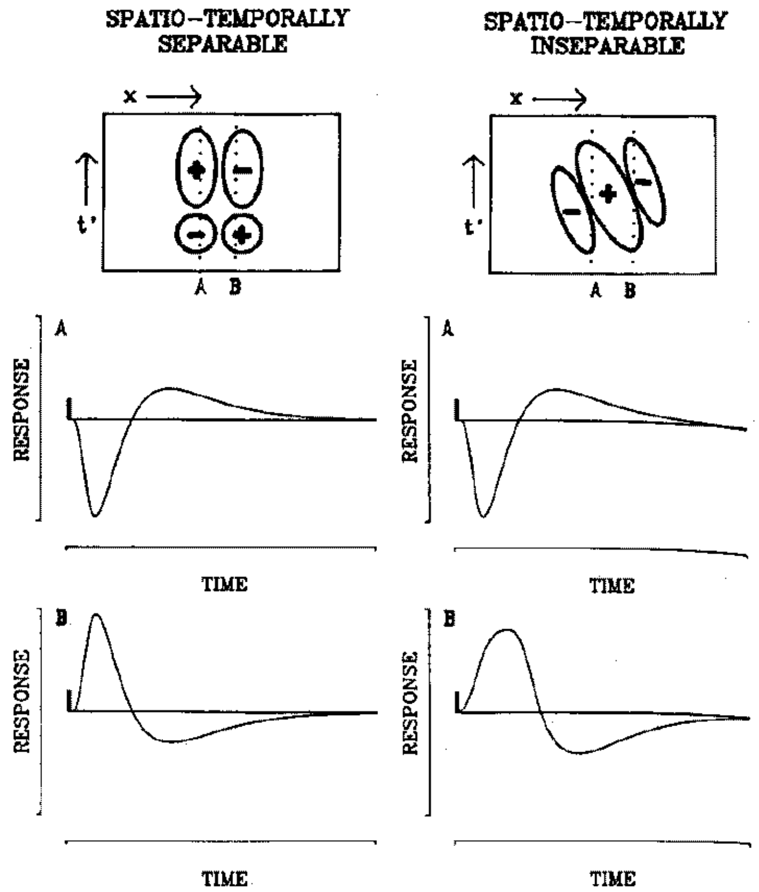
DYNAMICS OF THE *RF* STRUCTURE OF SIMPLE CELLS



DeAngelis *et al.*, 1995

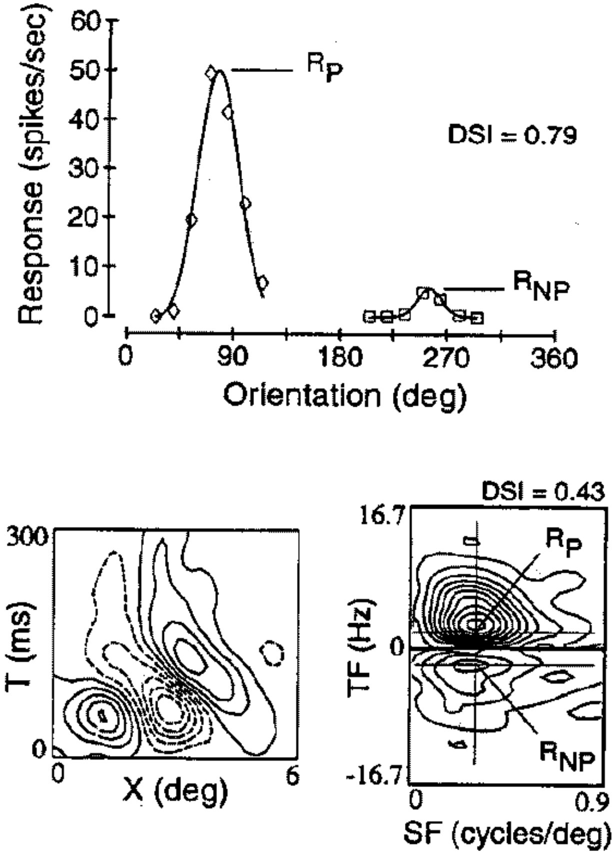
SPATIO-TEMPORAL FUNCTIONALITY





Reid *et al.*, 1991

DIRECTION SELECTIVITY



DeAngelis *et al.*, 1993

Direction Selectivity Index

$$\mathrm{DSI} = \frac{R_P - R_{NP}}{R_P}$$

Structural

- Presence of a surprising regularity
  - Subregions wax and wane smoothly in every direction
  - Spatial phase is continuosly variable (not just odd and even simmetry)
- Highly specific joint space-time behaviour

Functional

- Frequency selectivity
- Orientation tuning
- Direction selectivity
- End-stopping
- ...

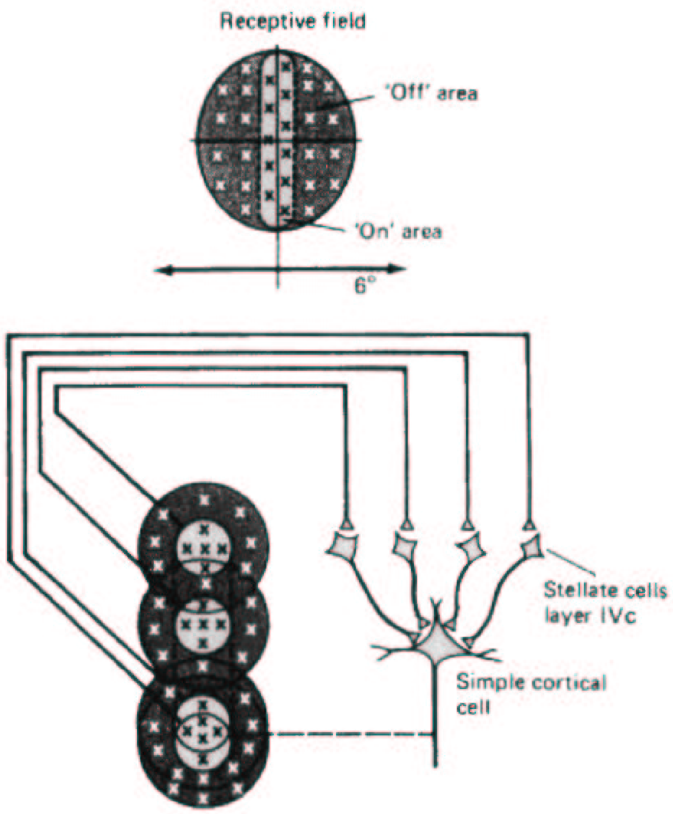
but **how do they acquire their properties?**

MODELS OF SIMPLE CELL *RF* SPECIFICITY

- Filter Models (Marcelja, 1980; Daugman, 1984; Adelson and Bergen, 1985)
- Developmental Models (von der Malsburg, 1973; Linsker, 1986; Miller, 1994; Wimbauer *et al.*, 1997)
- Architectural Models
  - Biophysically Realistic Models (Wörgötter and Koch, 1991; Somers, 1997; Suarez *et al.*, 1995)
  - Functional/Structural Models (Mitchison, 1985; Matsubara *et al.*, 1985)
  - Neural Field Models (Krone *et al.*, 1986; Mallot, 1996; Amari, 1977)

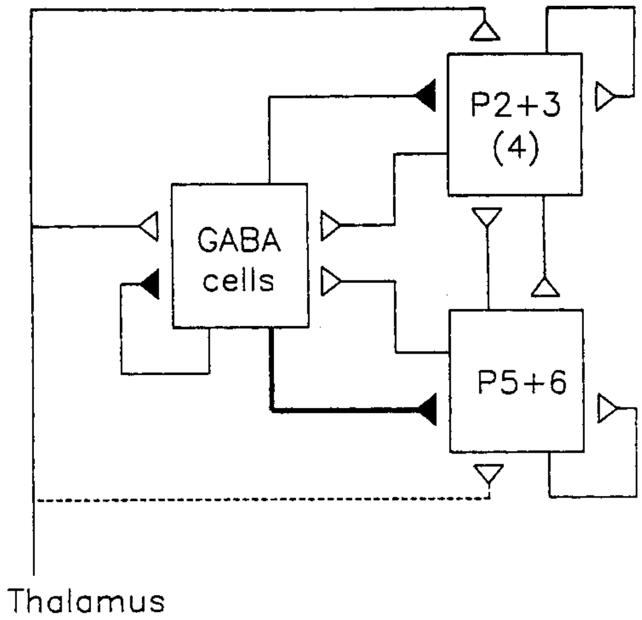
FEED-FORWARD vs FEED-BACK MODELS

Geniculo-cortical convergence



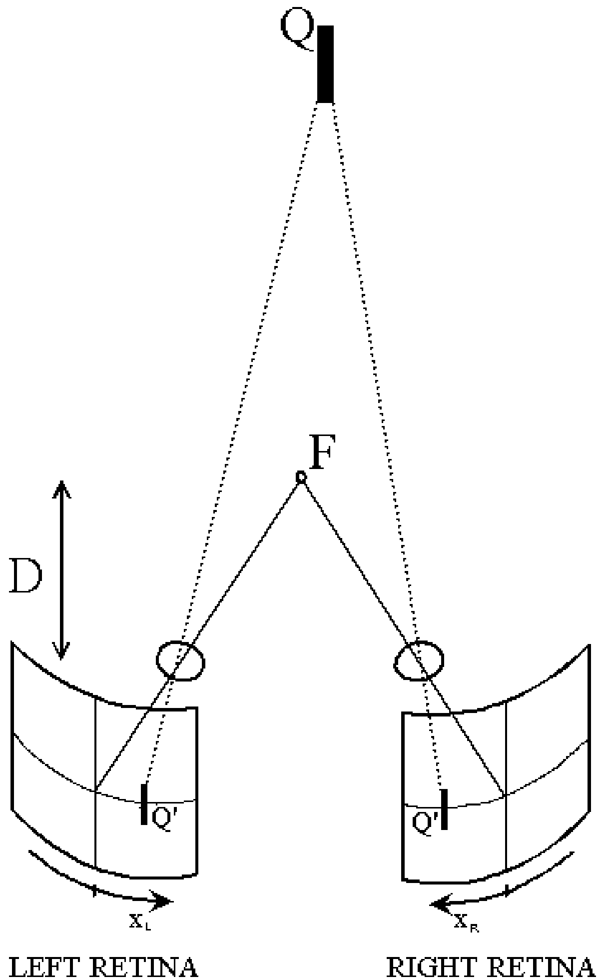
Hubel and Wiesel, 1962

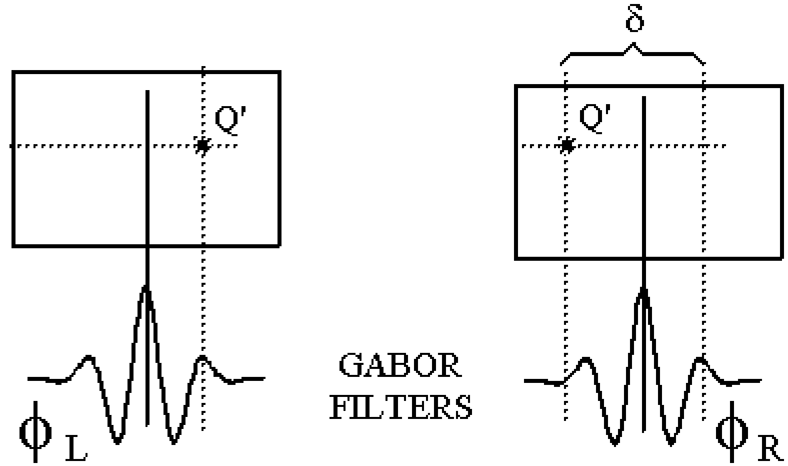
Canonical cortical microcircuit



Douglas *et al.*, 1989

THE STEREO PROBLEM





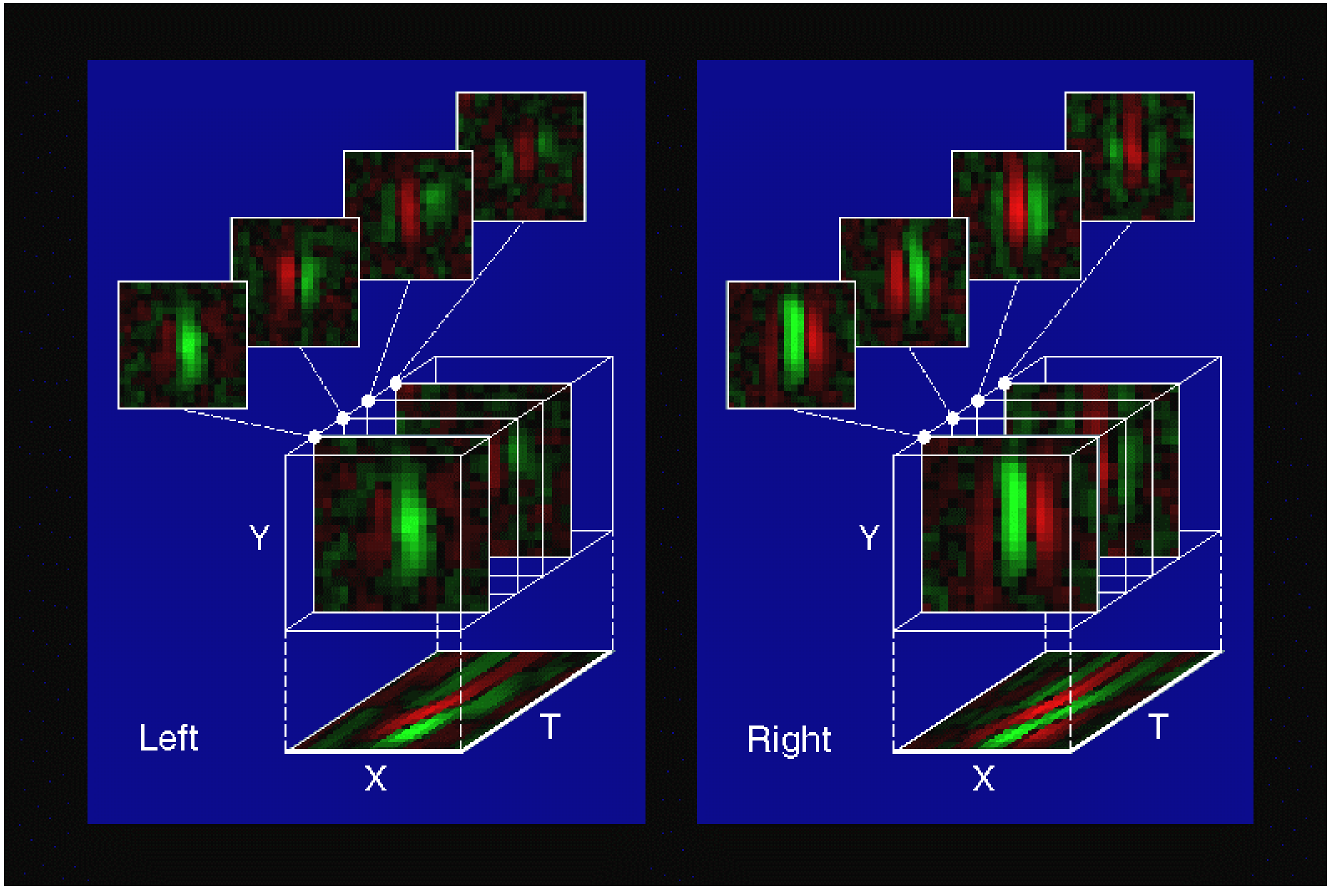
$$\delta \propto \Delta\phi = \phi_R - \phi_L$$

Fourier shift theorem

$$I^L(n) = I^R\left(n + \delta(n)\right)$$

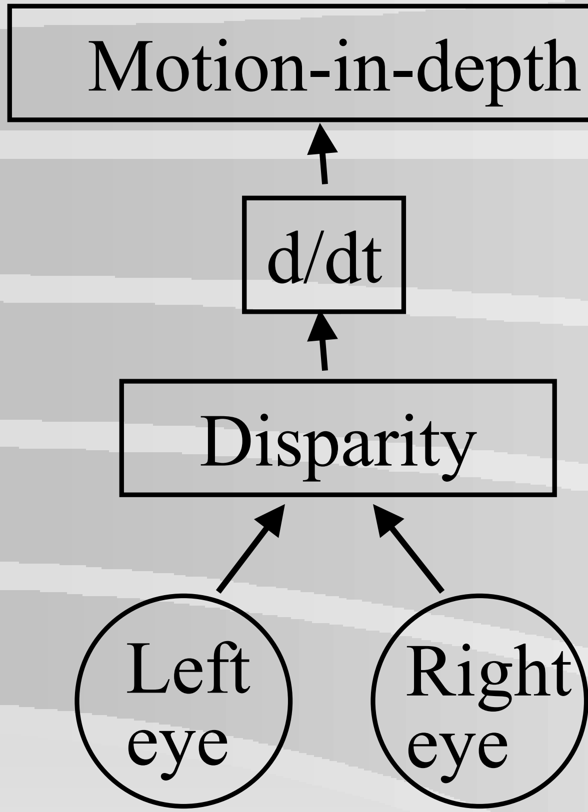
THE STEREO PROBLEM

The natural solution

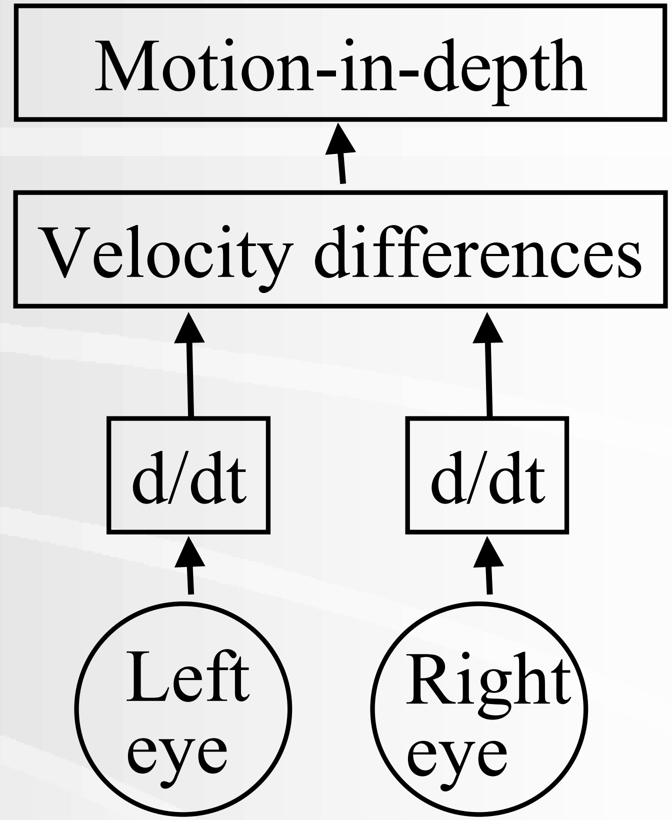


Ohzawa I. et al., The Neural Coding of Stereoscopic Depth.NeuroReport, 8 (no. 3): 3-12, 1997

BINOCULAR PERCEPTION OF MID

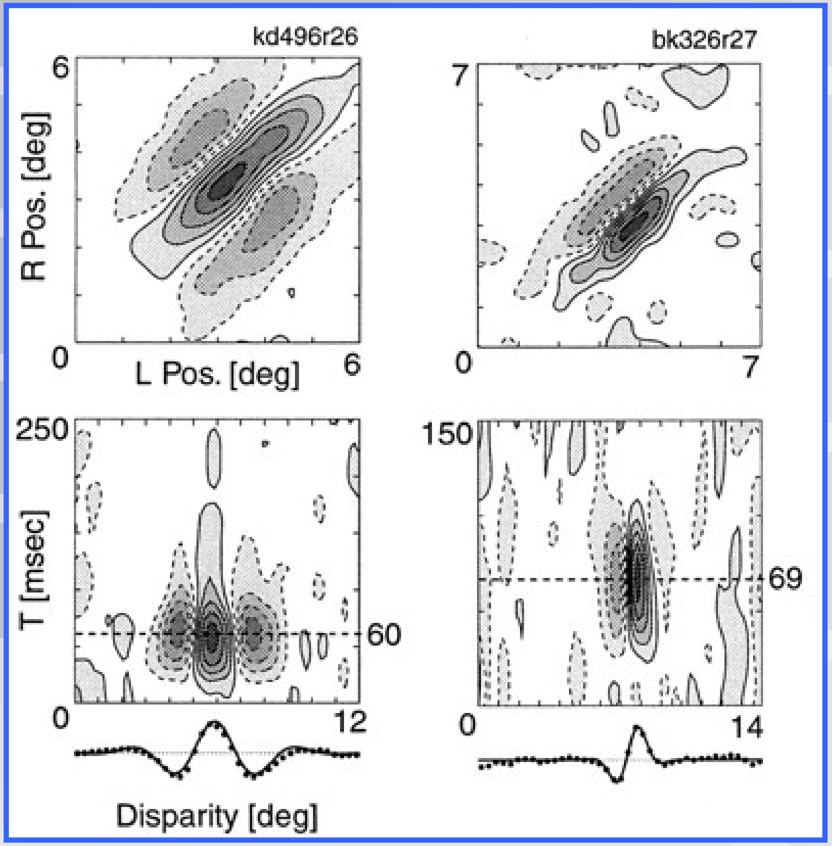


Disparity rate of change

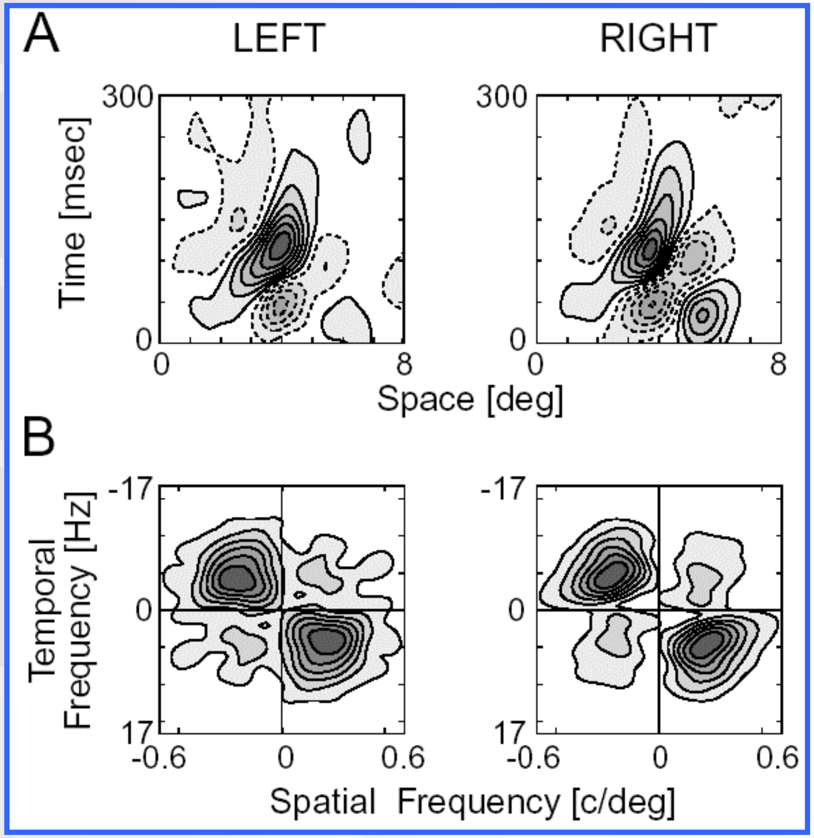


Interocular velocity difference

ENCODING OF DYNAMIC 3-D VISUAL INFORMATION IN THE VISUAL CORTEX



[Ohzawa *et al.* 1997]



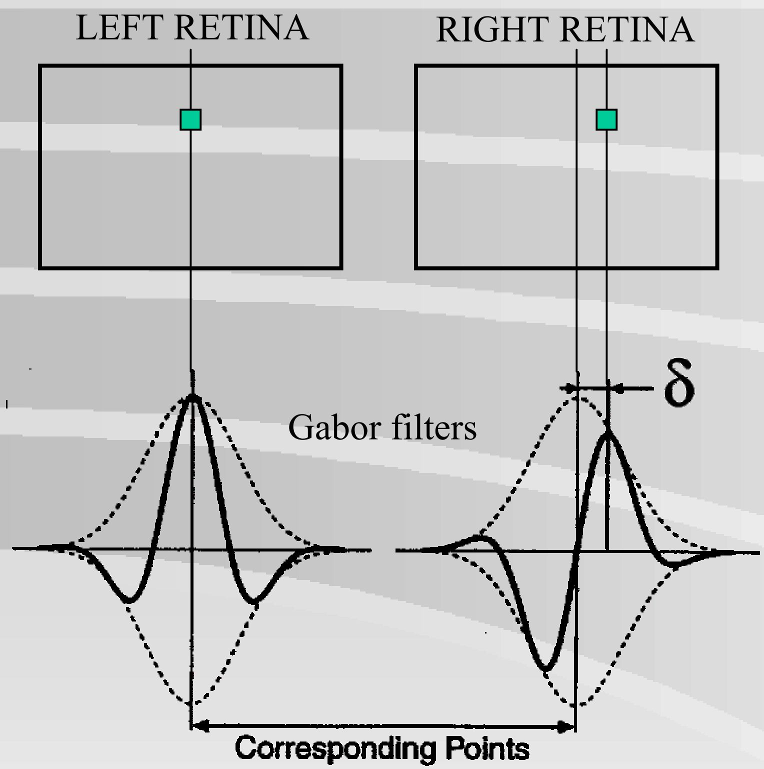
[Ohzawa *et al.* 1996]

PSPC
The Physical Structure of Perception and Computation

Dibe

PHASE-BASED DYNAMIC STEREOPSIS

Disparity as phase difference



$$I^R(x) = I^L[x + \delta(x)]$$

$$I^L[x + \delta(x)] \xleftrightarrow{\mathscr{F}} \tilde{I}(k)e^{jk\delta}$$

$$\delta(x) = \frac{\lfloor \Delta\phi(x) \rfloor_{2\pi}}{k_0} = \frac{\arg[Q^L] - \arg[Q^R]}{k_0}$$

where

$$Q(x) = \int I(\xi)h(x - \xi; k_0)d\xi = $$
$$= C(x) + jS(x)$$

with $h(x; k_0) = e^{-x^2/\sigma^2} e^{jk_0 x}$

PHASE-BASED DYNAMIC STEREOPSIS

Motion-in-depth

The total rate of variation of disparity can be written as

$$\frac{d\delta}{dt} = \frac{\partial\delta}{\partial t} + \frac{v^L}{k_0}(\phi_x^L - \phi_x^R)$$

Considering the conservation properties of local phase

$$\phi_x^L = -\frac{\phi_t^L}{v^L} \qquad \text{and} \qquad \phi_x^R = -\frac{\phi_t^R}{v^R}$$

The Physical Structure of Perception and Computation

PHASE-BASED DYNAMIC STEREOPSIS

Motion-in-depth

The total rate of variation of disparity can be written as

$$\frac{d\delta}{dt} = \frac{\partial\delta}{\partial t} + \frac{v^L}{k_0}(\phi_x^L - \phi_x^R)$$

Considering the conservation properties of local phase

$$\phi_x^L = -\frac{\phi_t^L}{v^L} \quad \text{and} \quad \phi_x^R = -\frac{\phi_t^R}{v^R}$$

Thus

$$\boxed{\frac{d\delta}{dt} = \frac{\phi_x^R}{k_0}(v^R - v^L) \simeq \frac{\partial\delta}{\partial t} = \frac{\phi_t^L - \phi_t^R}{k_0} \simeq v^R - v^L}$$

PSPC The Physical Structure of Perception and Computation

Dibe

PHASE-BASED DYNAMIC STEREOPSIS

Motion-in-depth

The total rate of variation of disparity can be written as

$$\frac{d\delta}{dt} = \frac{\partial \delta}{\partial t} + \frac{v^L}{k_0}(\phi_x^L - \phi_x^R)$$

Considering the conservation properties of local phase

$$\phi_x^L = -\frac{\phi_t^L}{v^L} \quad \text{and} \quad \phi_x^R = -\frac{\phi_t^R}{v^R}$$

Information hold in the interocular velocity difference is the same of that derived from the derivative of binocular disparity, if a phase-based disparity encoding scheme is assumed

PHASE-BASED DYNAMIC STEREOPSIS

Spatio-temporal operators

Partial derivative of disparity can be directly computed by convolutions (S, C) of stereo image pairs and by their temporal derivatives (S_t, C_t)

$$\frac{\partial \delta}{\partial t} = \left[\frac{S_t^L C^L - S^L C_t^L}{(S^L)^2 + (C^L)^2} - \frac{S_t^R C^R - S^R C_t^R}{(S^R)^2 + (C^R)^2} \right] \frac{1}{k_0}$$

where

$$4SC_t = (S + C_t)^2 - (C_t - S)^2, \qquad 4S_tC = (S_t + C)^2 - (S_t - C)^2$$

PHASE-BASED DYNAMIC STEREOPSIS

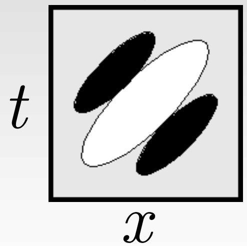
Spatio-temporal operators

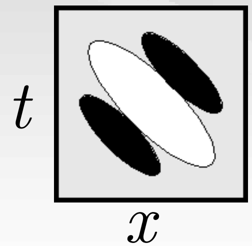
Partial derivative of disparity can be directly computed by convolutions (S, C) of stereo image pairs and by their temporal derivatives (S_t, C_t)

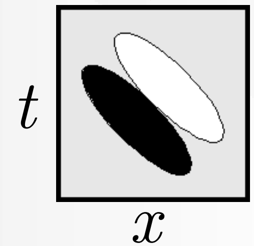
$$\frac{\partial \delta}{\partial t} = \left[\frac{S_t^L C^L - S^L C_t^L}{(S^L)^2 + (C^L)^2} - \frac{S_t^R C^R - S^R C_t^R}{(S^R)^2 + (C^R)^2} \right] \frac{1}{k_0}$$

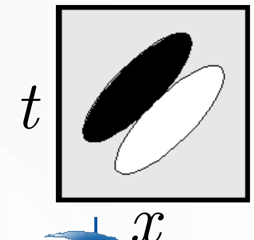
where

$$4SC_t = (S + C_t)^2 - (C_t - S)^2, \qquad 4S_tC = (S_t + C)^2 - (S_t - C)^2$$









PHASE-BASED DYNAMIC STEREOPSIS

Spatio-temporal operators

Partial derivative of disparity can be directly computed by convolutions (S, C) of stereo image pairs and by their temporal derivatives (S_t, C_t)

$$\frac{\partial\delta}{\partial t}=\left[\frac{S_t^L C^L - S^L C_t^L}{(S^L)^2+(C^L)^2}-\frac{S_t^R C^R - S^R C_t^R}{(S^R)^2+(C^R)^2}\right]\frac{1}{k_0}$$

where

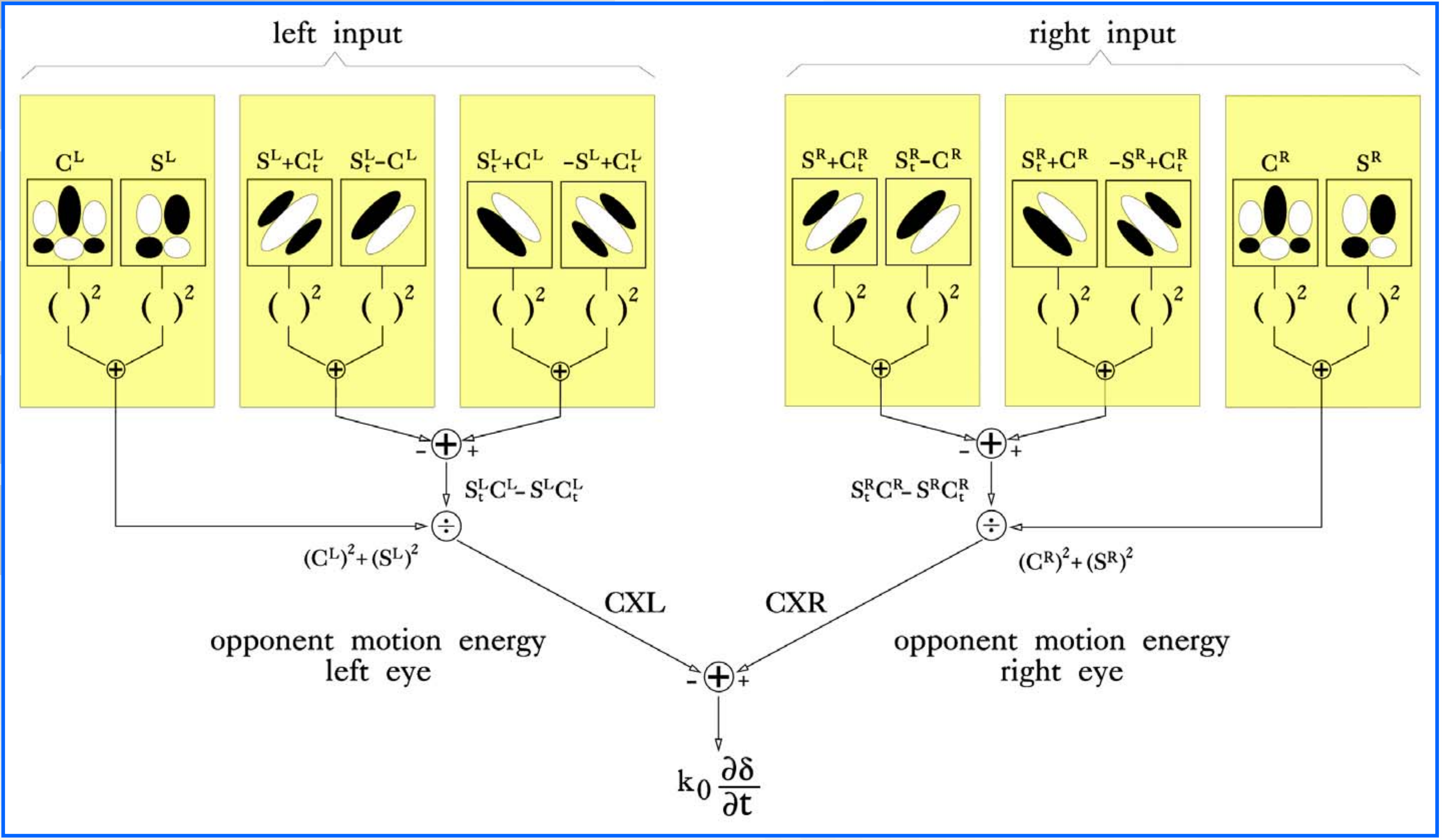
$$4SC_t=(S+C_t)^2-(C_t-S)^2, \qquad 4S_tC=(S_t+C)^2-(S_t-C)^2$$

$\partial\delta / \partial t$ is expressed through convolutions with a set of spatio-temporal filters whose shapes resemble simple cell RFs

PSPC The Physical Structure of Perception and Computation

Dibe

CORTICAL MODEL (I)



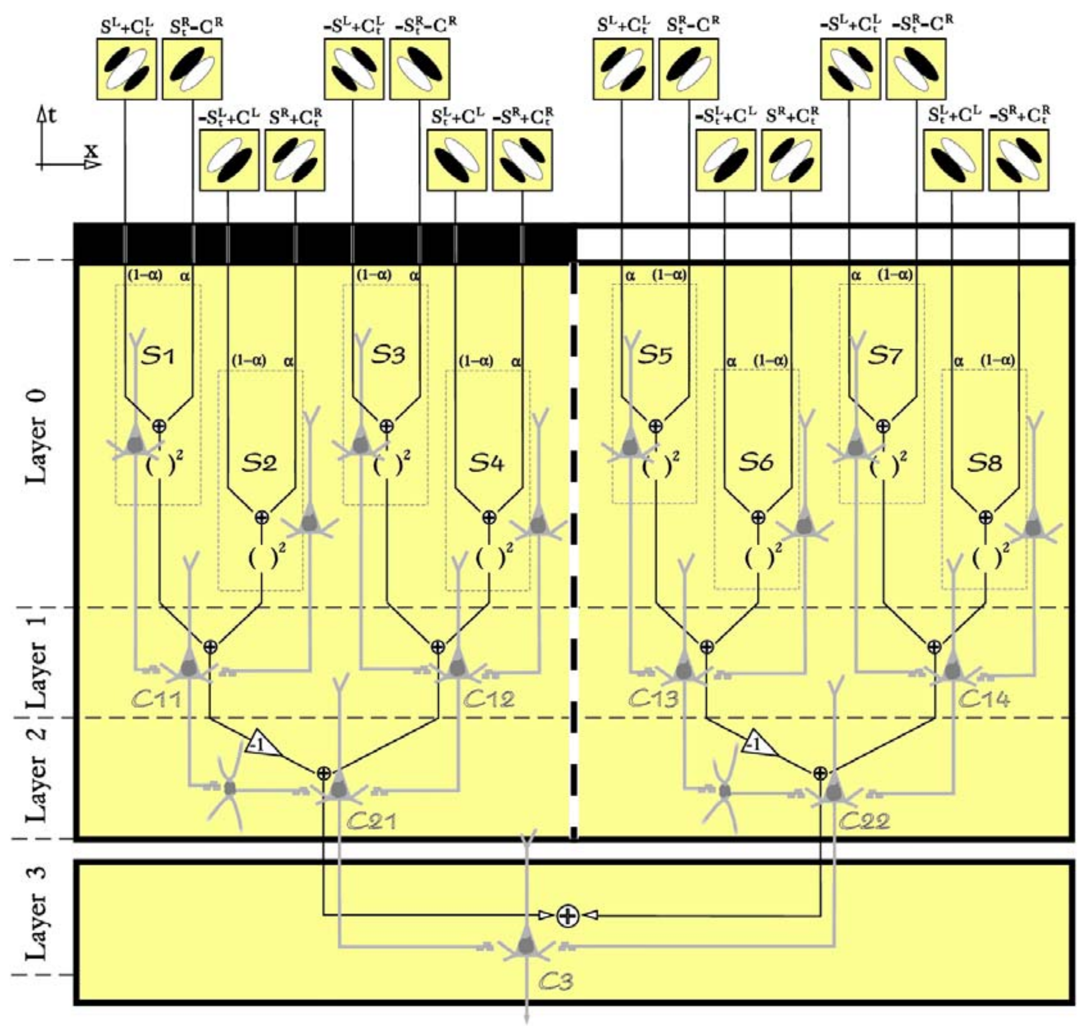
CORTICAL MODEL (II)

$$S_1 = (1-\alpha)(C_t^L + S^L) - \alpha(C^R - S_t^R)$$
$$S_2 = (1-\alpha)(C^L + S_t^L) + \alpha(C_t^R + S^R)$$
$$S_3 = (1-\alpha)(C_t^L - S^L) - \alpha(C^R + S_t^R)$$
$$S_4 = (1-\alpha)(C^L + S_t^L) + \alpha(C_t^R - S^R)$$
$$S_5 = \alpha(C_t^L + S^L) - (1-\alpha)(C^R - S_t^R)$$
$$S_6 = \alpha(C^L - S_t^L) + (1-\alpha)(C_t^R + S^R)$$
$$S_7 = \alpha(C_t^L - S^L) - (1-\alpha)(C^R + S_t^R)$$
$$S_8 = \alpha(C^L + S_t^L) + (1-\alpha)(C_t^R - S^R)$$

$$C_{11} = S_1^2 + S_2^2 \qquad C_{12} = S_3^2 + S_4^2$$
$$C_{13} = S_5^2 + S_6^2 \qquad C_{14} = S_7^2 + S_8^2$$

$$C_{21} = C_{12} - C_{11}$$
$$C_{22} = C_{13} - C_{14}$$

$$C_3 = C_{12} + C_{22}$$

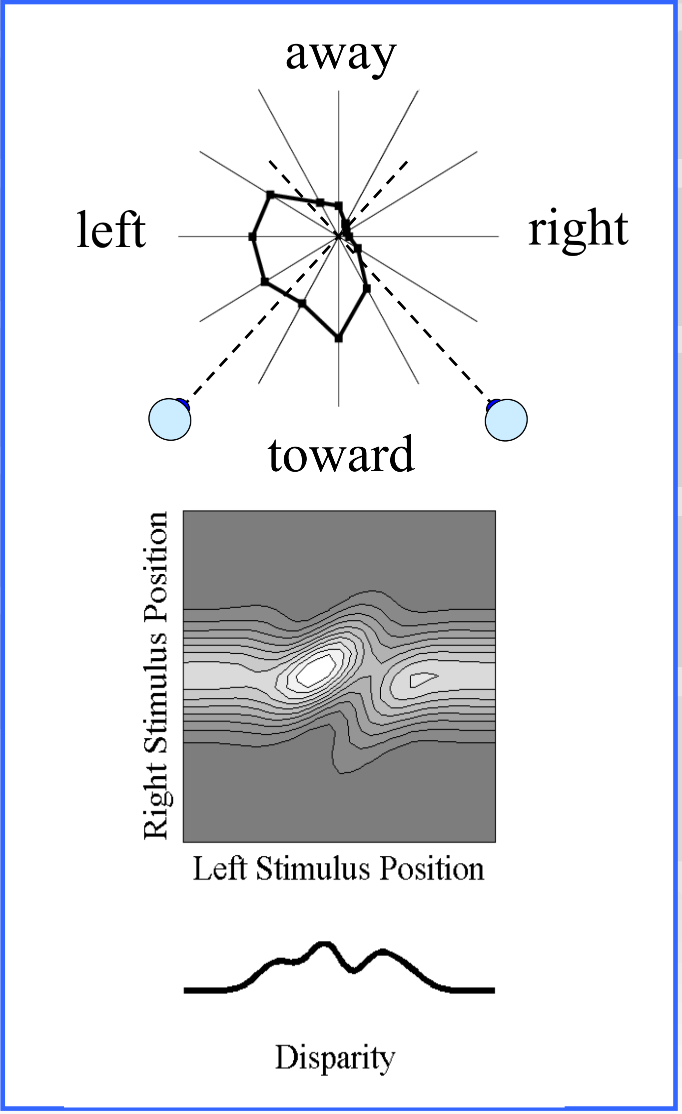


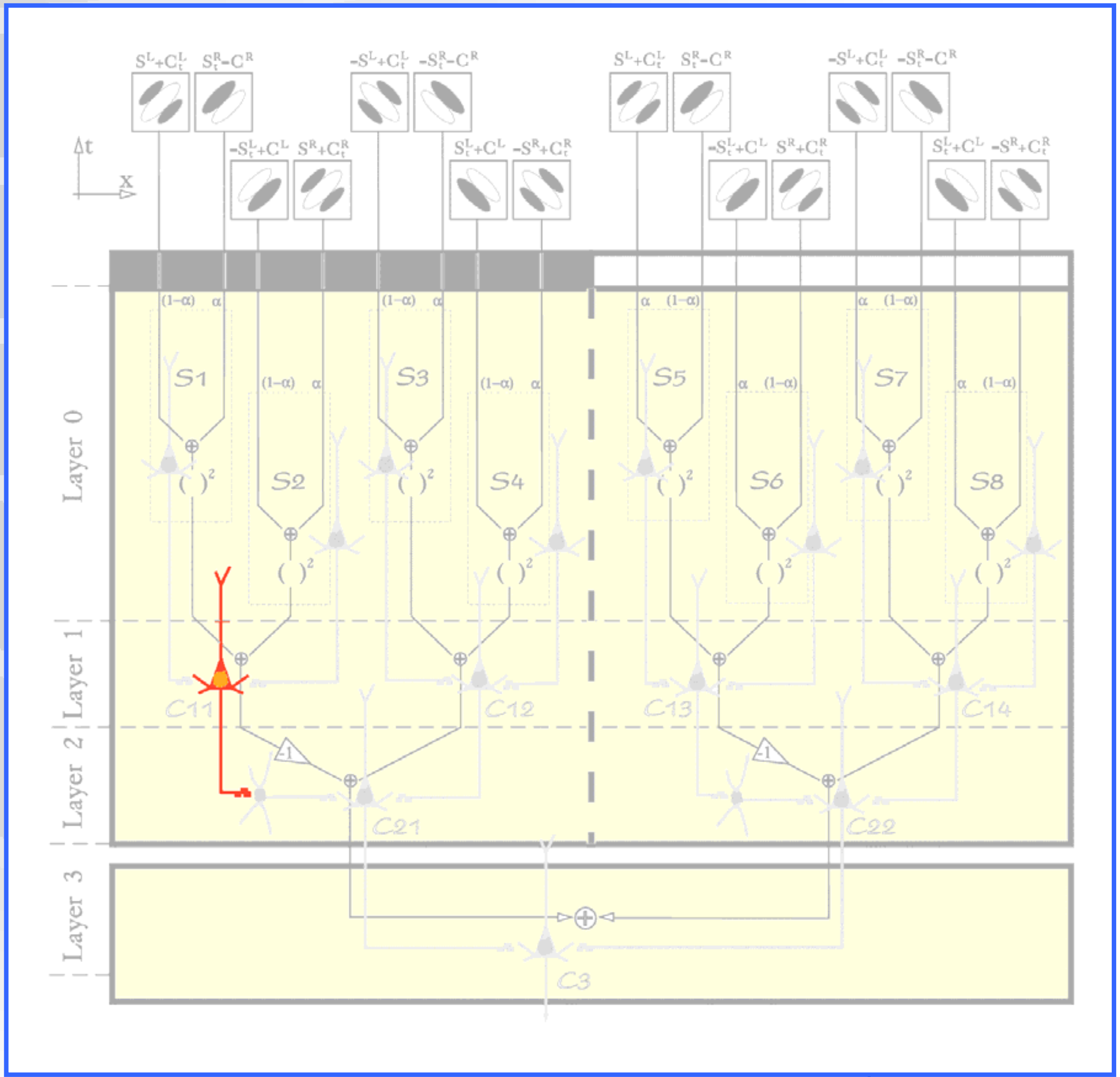
$$(1-2\alpha)(S_t^L C^L - S^L C_t^L - S_t^R C^R + S^R C_t^R)$$

PSPC The Physical Structure of Perception and Computation

Dibe

CORTICAL MODEL (II) – cont'd

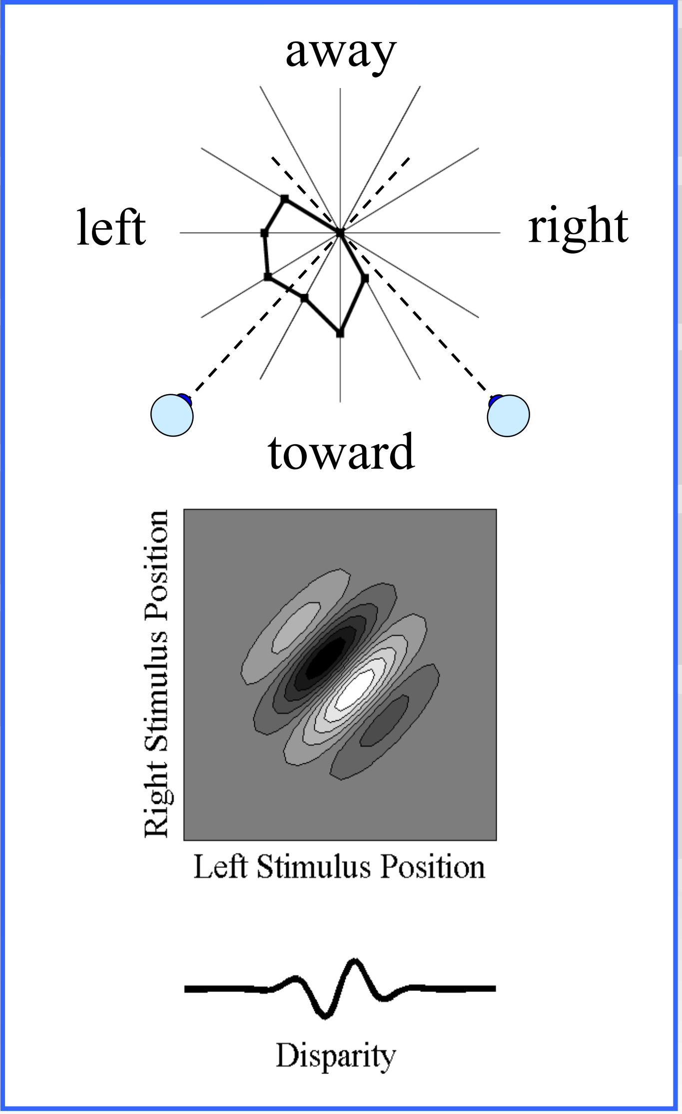


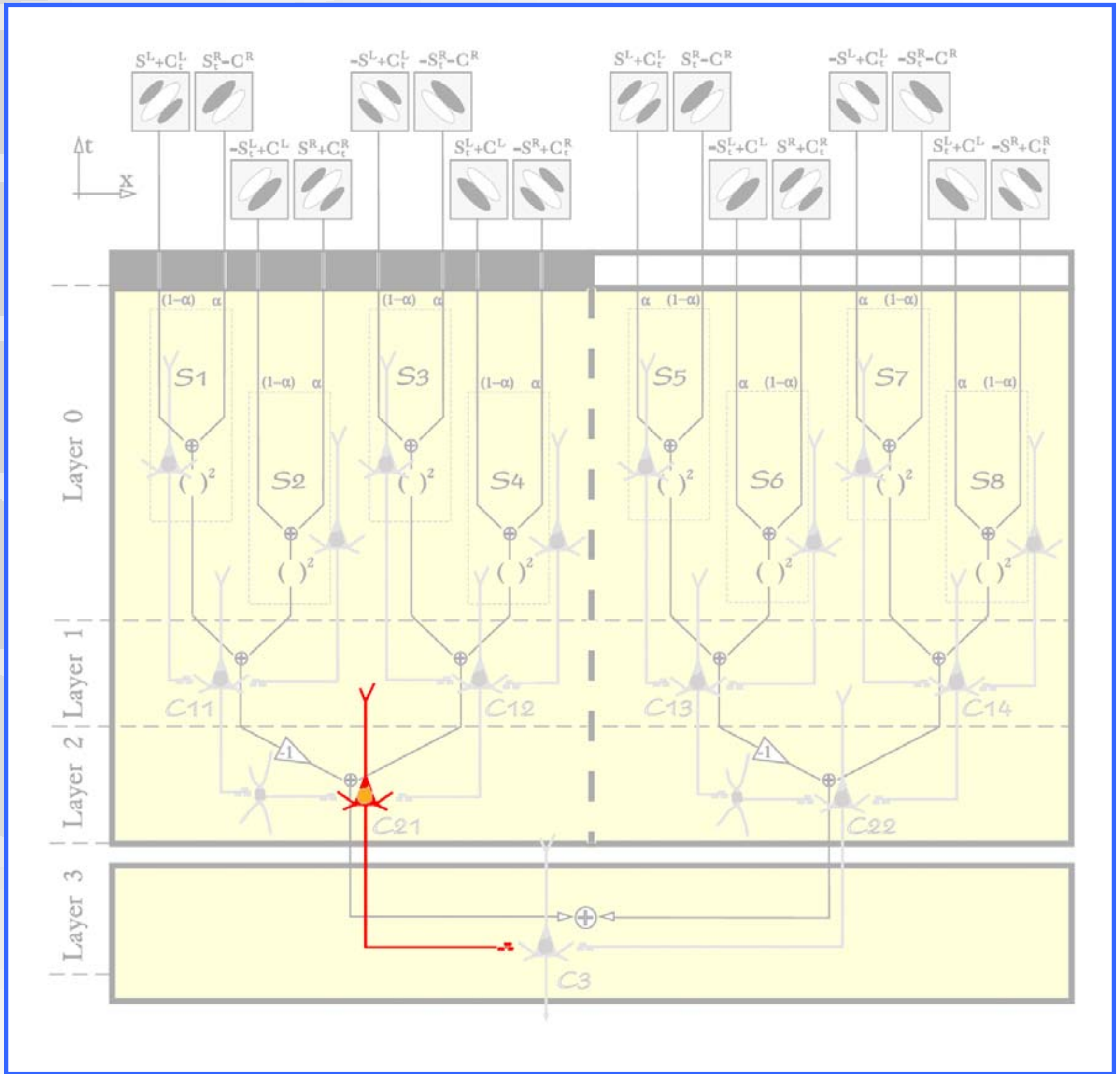


PSPC The Physical Structure of Perception and Computation

Dibe

CORTICAL MODEL (II) – cont'd

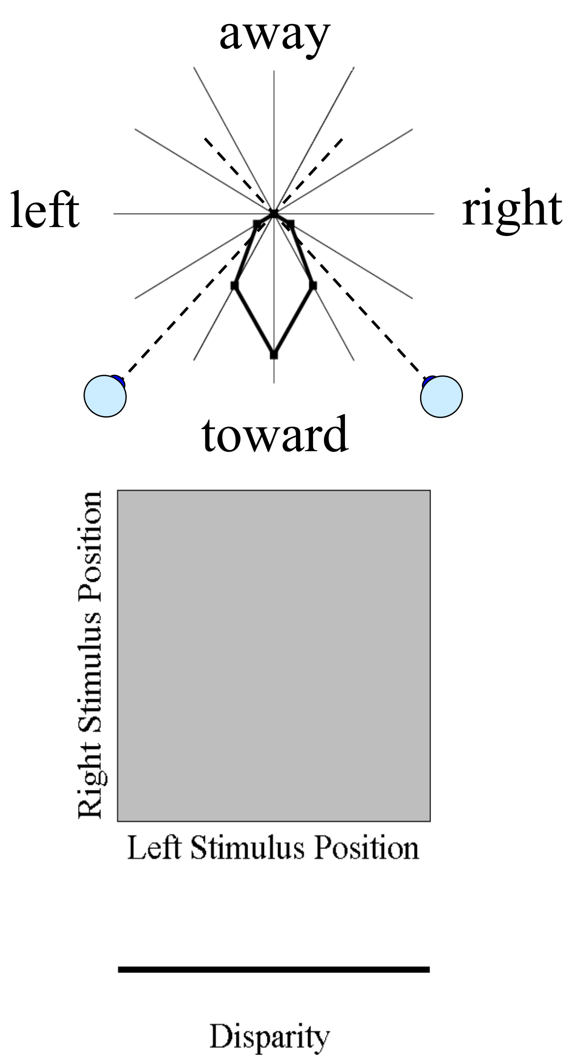


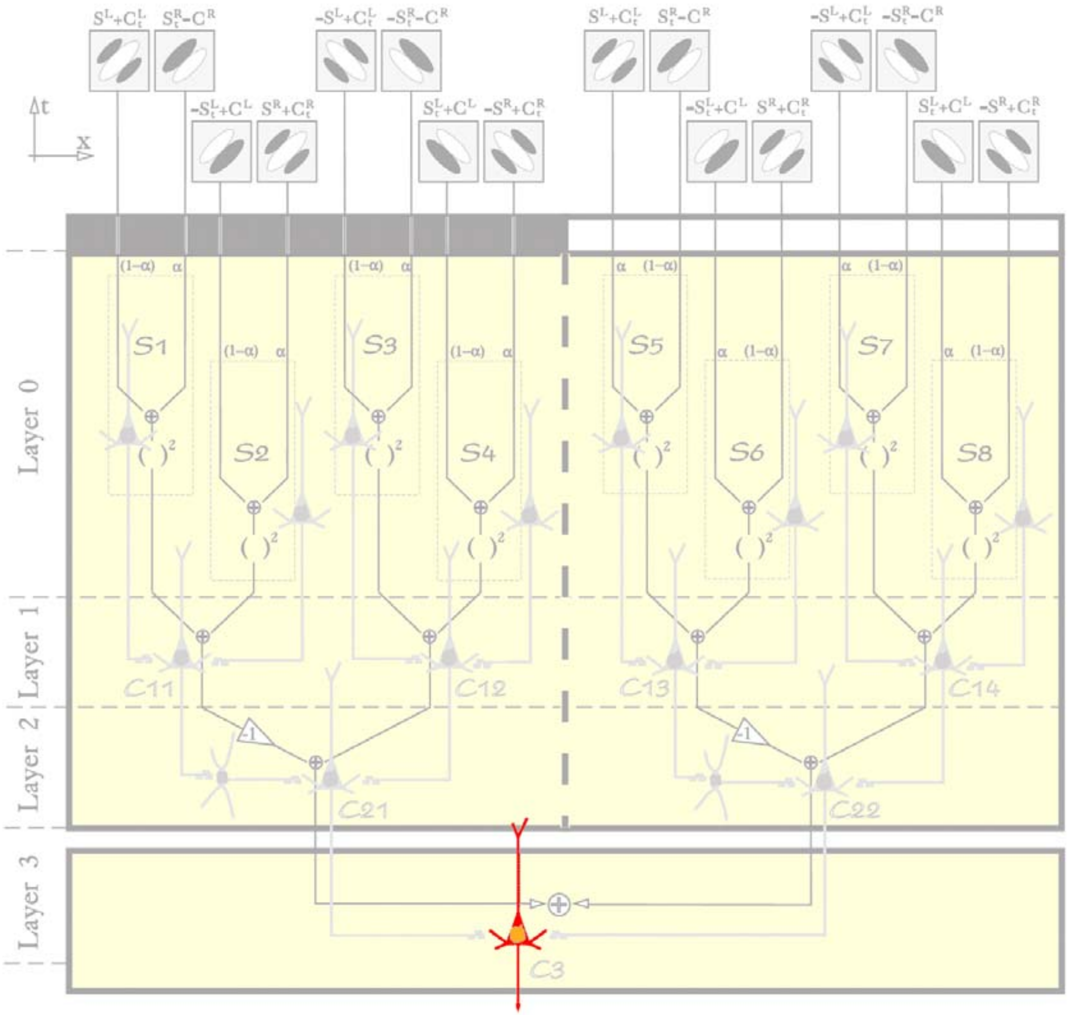


PSPC The Physical Structure of Perception and Computation

Dibe

CORTICAL MODEL (II) – cont'd





PSPC The Physical Structure of Perception and Computation

Dibe