How the Brain Sees: Fundamentals and Recent Progress in Modeling Vision

Stephen Grossberg

Ennio Mingolla

Department of Cognitive and Neural Systems

BOSTON UNIVERSITY

Annual Meeting of the ***Vision Sciences Society***, May 6, 2005

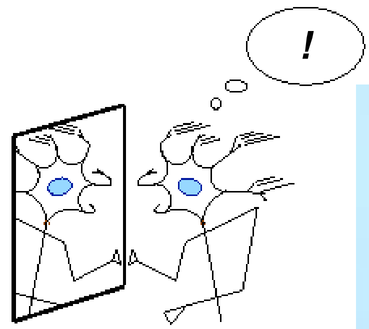
This tutorial is available for download at:

http://cns.bu.edu/techlab/

Why bother to learn about a model?

A model can

explain data by linking brain to perception,



link experiments to underlying **mechanisms** in surprising ways,

Adapted from Andres Perez-Uribe, 1993

and **suggest** exciting **new experiments.**

A possible worry

How many principles and mechanisms do we need to know?

> **"In fact, as many kinds of mathematics seem to be applied to perception as there are problems in perception. I believe this multiplicity of theories without a reduction to a common core is inherent in the nature of psychology . . . , and we should not expect the situation to change. The moral, alas, is that we need many different models to deal with the many different aspects of perception.**
>
> **Sperling, 1981**

Claim: *A few* principles and mechanisms explain a lot!

Styles of explanation

Some think:
"The brain is a bag of tricks."

Others think:
Studying statistics of the visual world suffices.

Who needs [to study] brains?!

25 years of modeling suggest . . .

A real *theory* can be had

A small number of *mechanisms*

- short-term memory
- long-term memory
- habituation
- adaptive gain control -- normalization
- local circuits with feedback -- bottom up, top down, and lateral connections

A somewhat larger number of *functional modules*

- filters of various kinds
- center-surround networks
- gated dipoles -- "nature's flip-flops"

A still larger number of *architectures*

specialized combinations of **mechanisms** and **modules** for cognition, audition, **vision**, …

LAMINART Architecture

How does the **cerebral cortex** work?

How do **cortical layers** support intelligence?

Quantitative simulations of
electrophysiologically identified cells in
anatomically supported networks
produce **laminar circuit dynamics** whose emergent properties
mimic **percepts.**

Is this just “more of the same . . .”?

New principles and new computational paradigms
generate basic questions that are easy to state.

These turn an **impenetrable mystery** into a
workable hard problem.

Today:

Use **experiments** to introduce **models.**
Use **models** to explain **data.**
Show how **models** suggest **new experiments.**

Why do we see?

surface color

Possible answer: **Seeing** **helps to recognize objects**

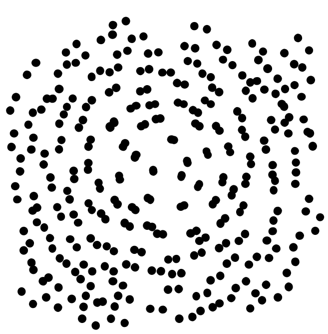
Counterexample: ***Amodal percepts***

Bregman, 1981
Kanizsa, 1979

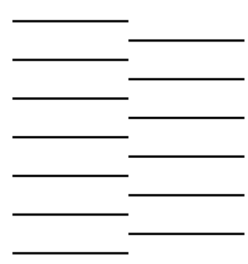
boundary completion

Seeing vs. knowing

We can **know** a form without **seeing** it.



Glass pattern



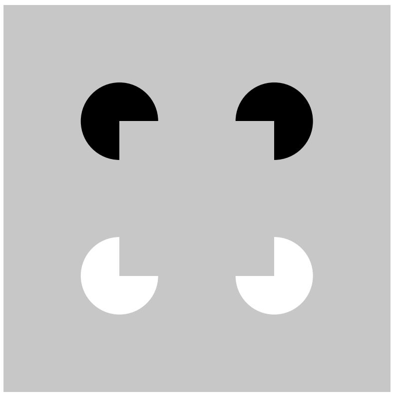
offset grating

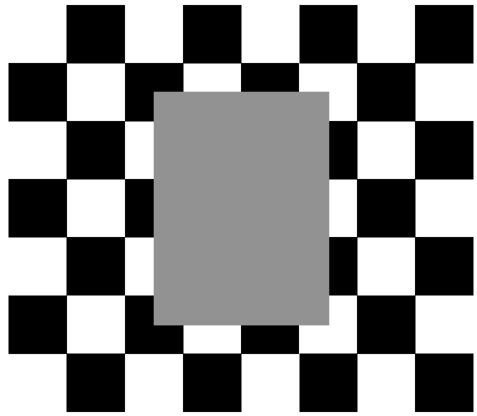
boundary completion

Why are some boundaries *amodal?*

Prediction: G & M, 1985

***All* boundaries are amodal** (within the boundary stream)





objects in textured scenes

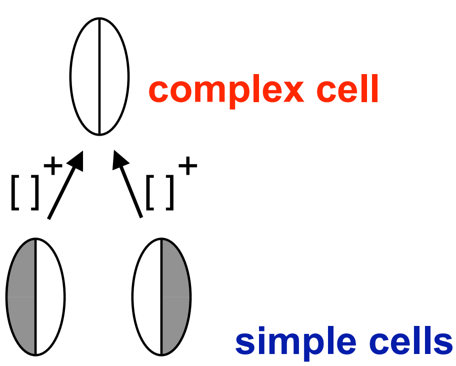
Kanizsa

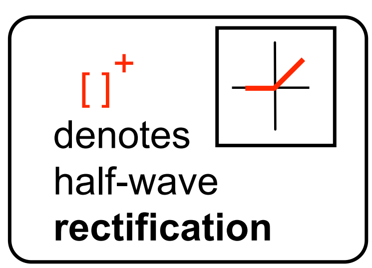
Complex cells pool opposite contrast polarities

Both achromatic and chromatic

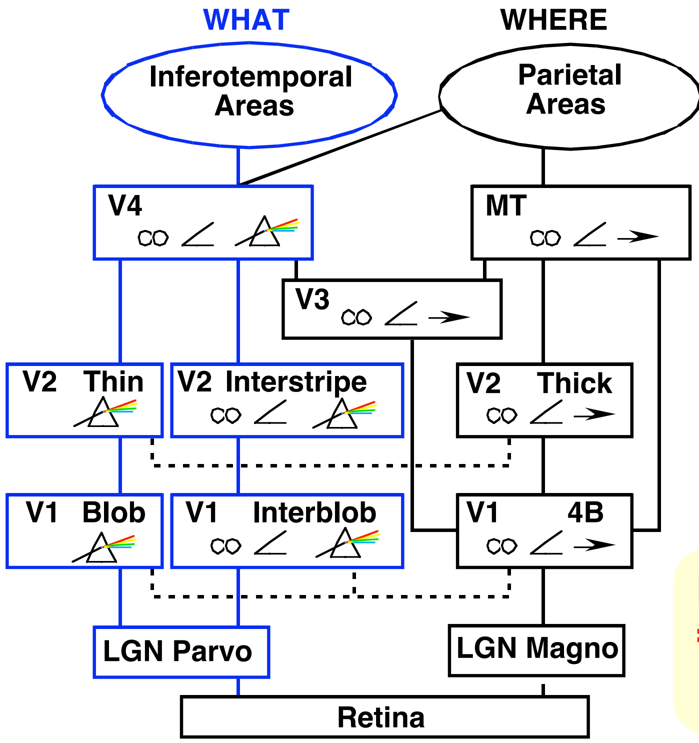
Complex cells as **amodal boundary detectors** not obvious . . .

Thorell, DeValois, and, Albrecht, 1984:
Complex cells **"must surely be considered color cells in the broadest sense"** because they pool inputs from multiple achromatic and chromatic cells.





Prediction, 1985



DeYoe and van Essen, 1988

Interacting
BOUNDARY
and
SURFACE
Streams
in “WHAT” pathway

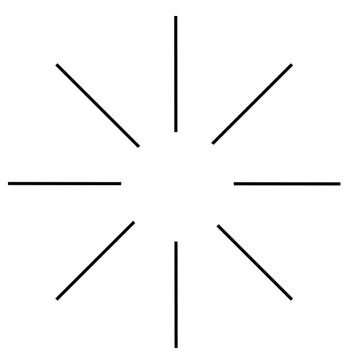
boundary & surface
≠ orientation & color

Livingstone and Hubel, 1987

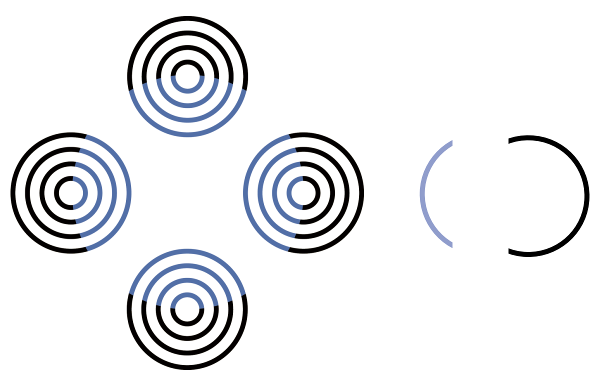
If boundaries are amodal, how do we *see?*

How can we **see** properties that are not “in the stimulus”?

When ***do*** we see?



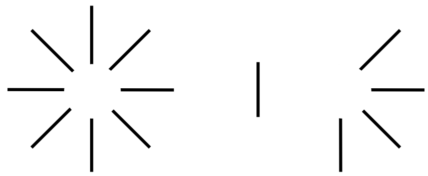
Ehrenstein, 1941



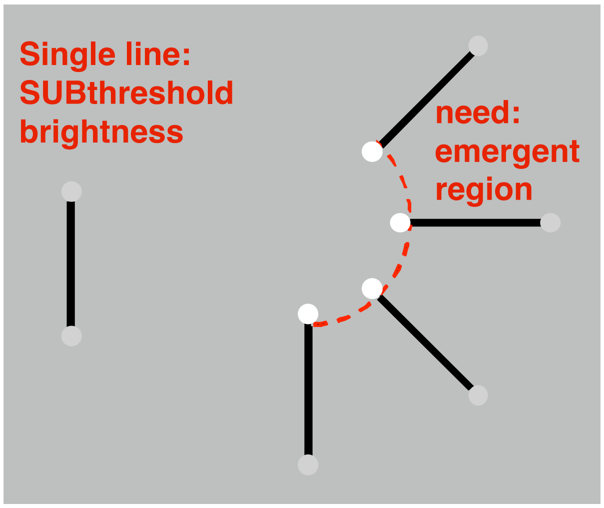
Varin, 1971

filling-in **of surface color**

Boundary and Surface Interaction



Kennedy, 1979

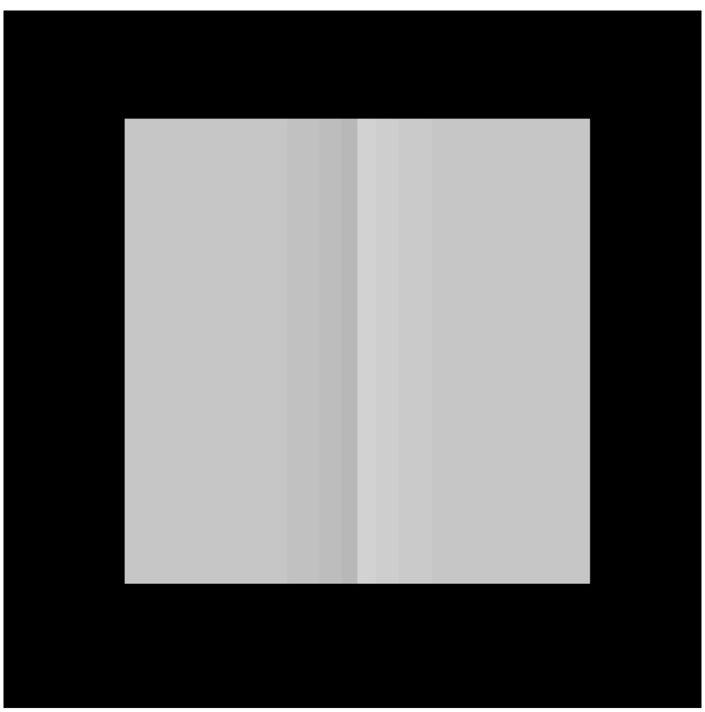


Boundary completion

Filling-in

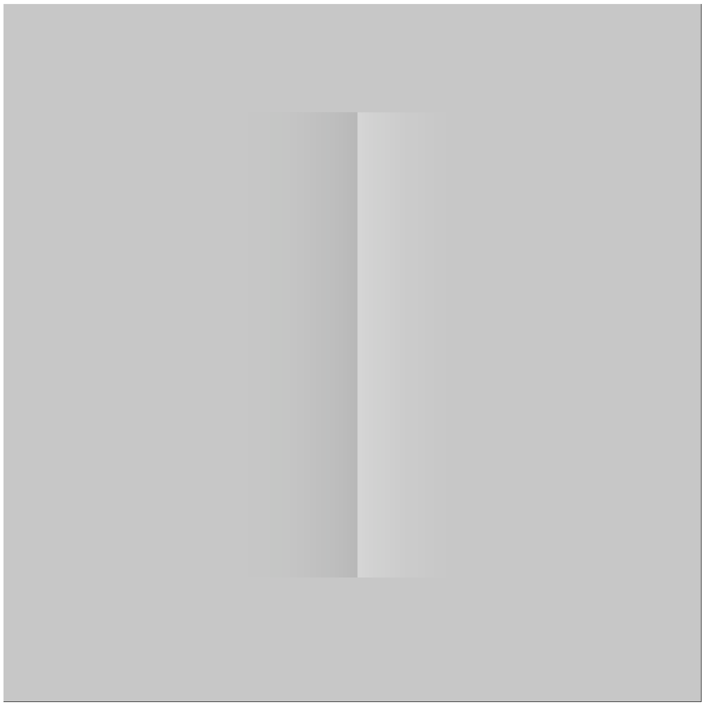
What signals are you filling in?

Craik-O'Brien-Cornsweet Effect, A



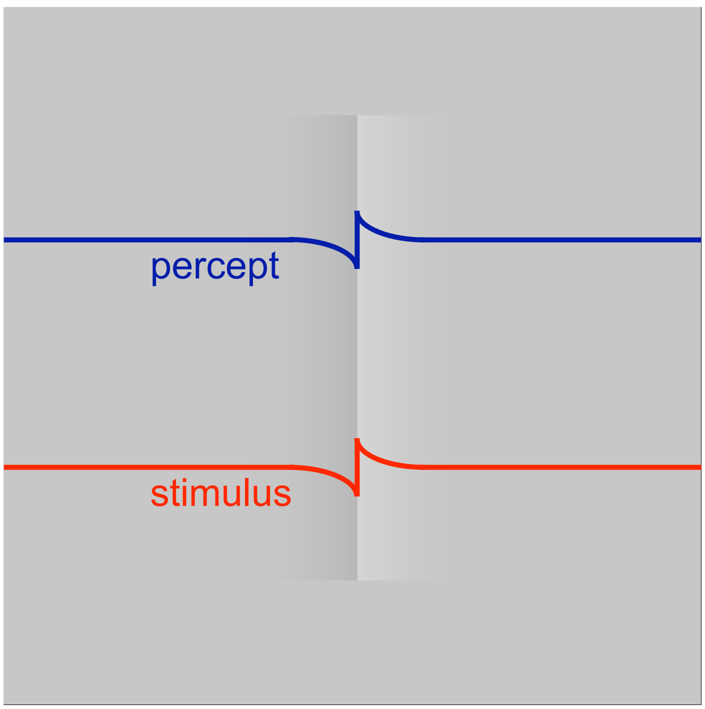
Todorović, 1987

Craik-O'Brien-Cornsweet Effect, B



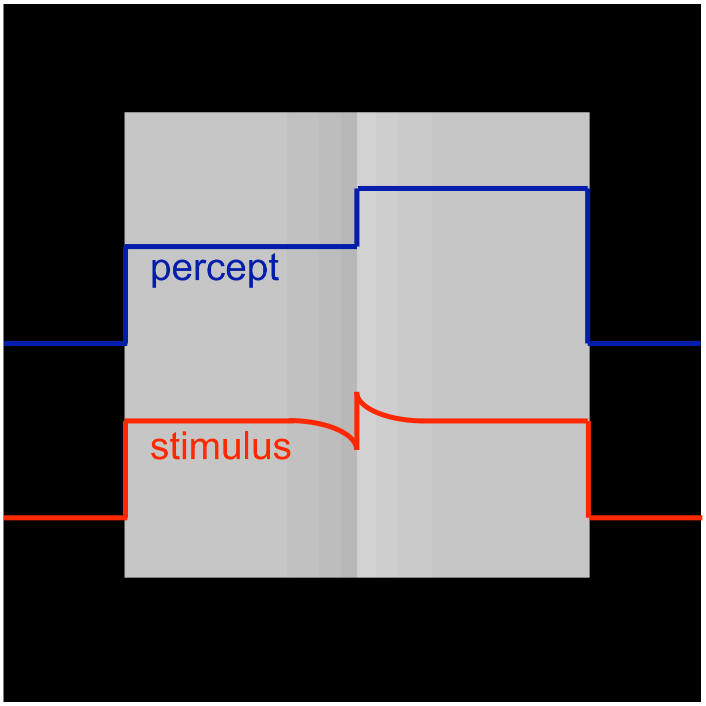
Todorović, 1987

Craik-O'Brien-Cornsweet Effect, C



Todorović, 1987

Craik-O'Brien-Cornsweet Effect, D



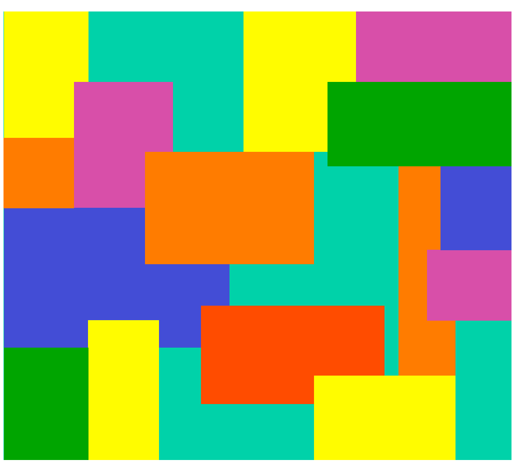
Boundary completion
defines
filling-in compartments.

Filling-in determines
what we **see**
in each compartment.

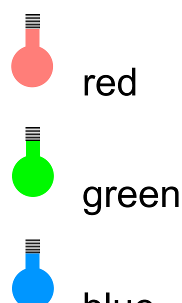
Why filling-in?

Todorović, 1987

Discounting the illuminant



Helmholtz



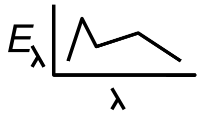
Land -- McCann
"Mondrians" 1971

FIRST EXPERIMENT:
Red illuminant intensity increases
Colors look "much the same"

We factor away the "extra" red

Land -- McCann Mondrians

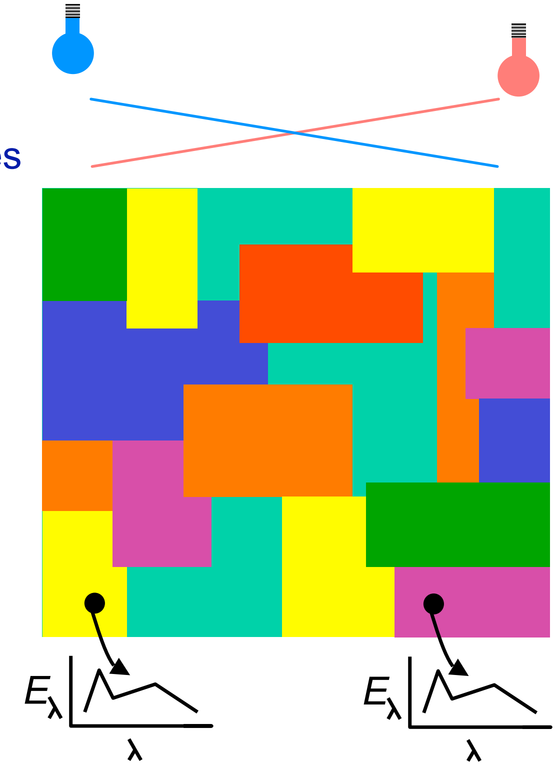
Gradients of illumination
create same spectrum patches



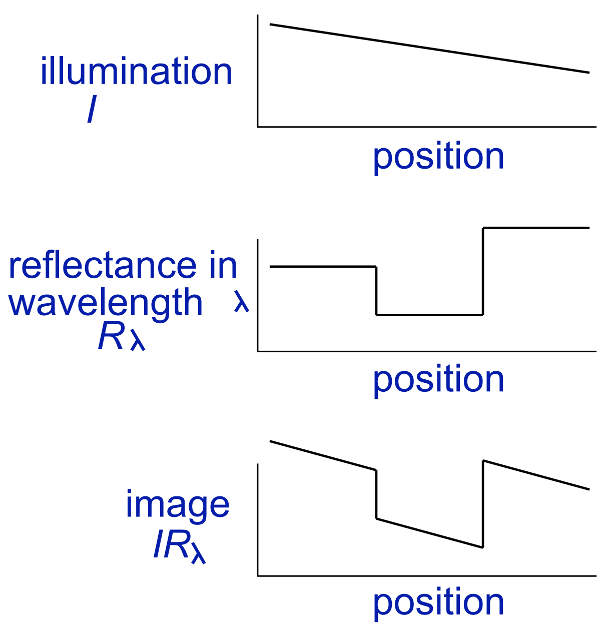
with different reflectances

Different colors seen from
the ***same*** **spectrum**
. . . similar to those
seen in white light

"discount the illuminant"

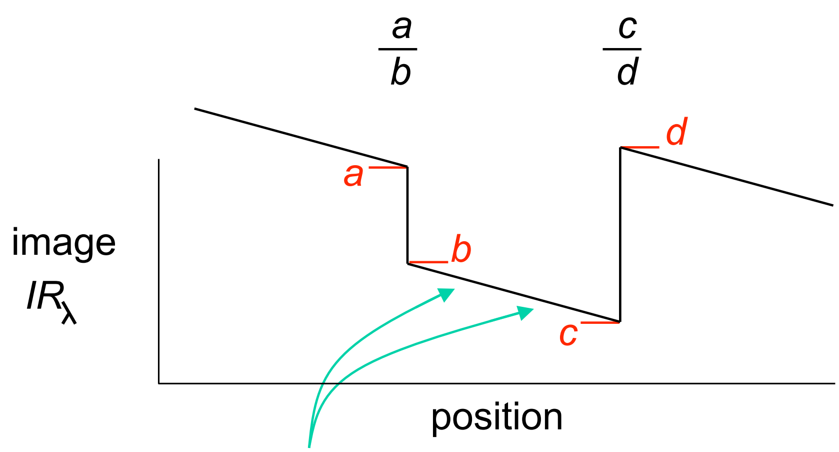


How are illuminants discounted?



"Retinex" Strategy

1. Recover relative reflectances ***(ratios)*** near image *edges.*



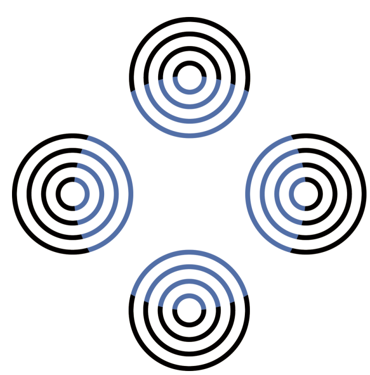
2. ***Suppress*** information from slowly varying region *interiors.*

How are *boundary* and *surface* computations related?

How are perceptual **boundaries** formed?

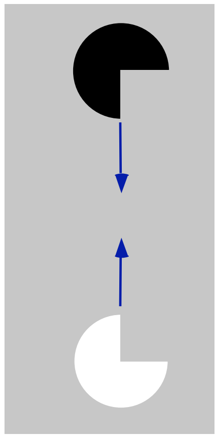
How does surface **filling-in** occur?

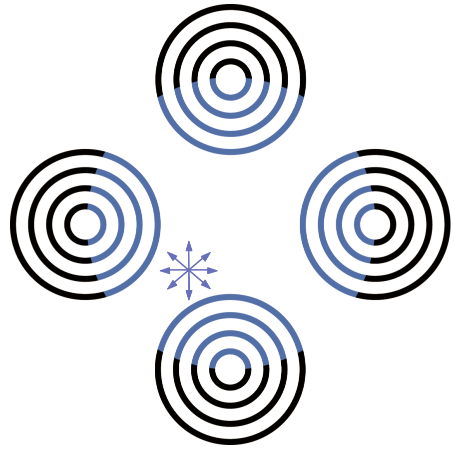
Are these **independent modules**?



No!
The answer is more interesting.

Boundary and surface computations are *COMPLEMENTARY*

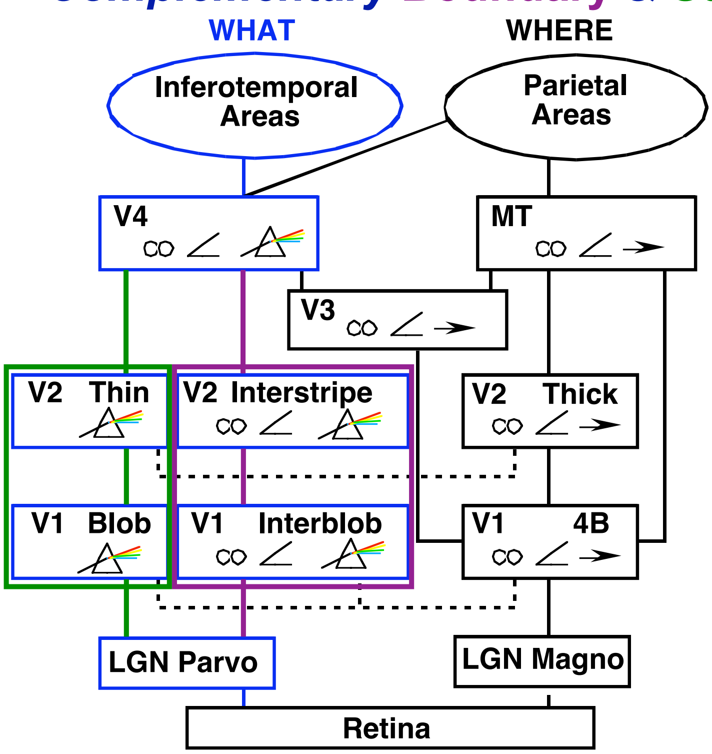




Boundary Contour System
BCS: *Completion*
- oriented
- inward
- insensitive to direction-of-contrast

Feature Contour System
FCS: *Filling-in*
- unoriented
- outward
- sensitive to direction-of-contrast

Complementary Boundary & Surface Streams

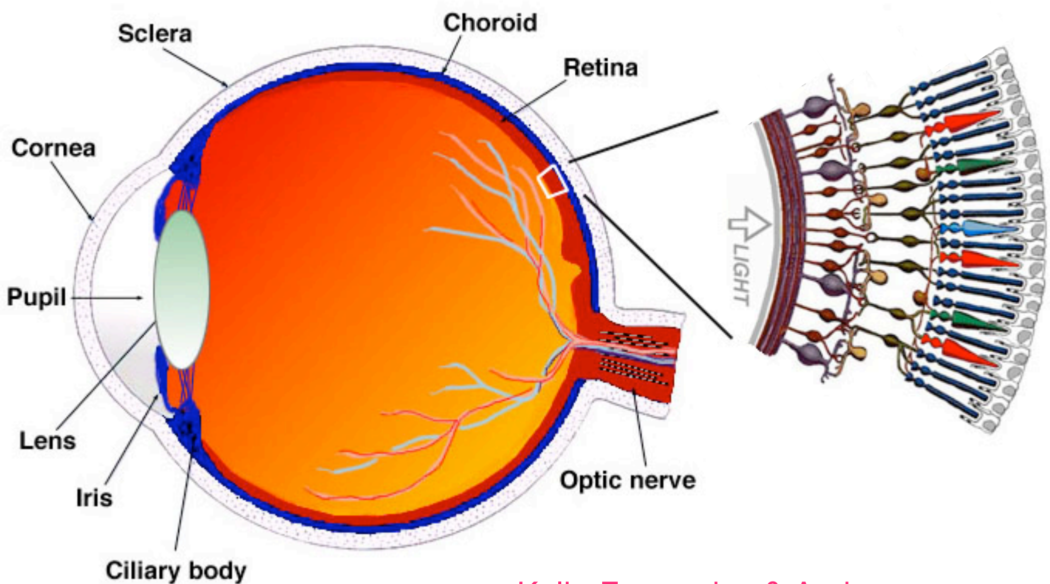


DeYoe and van Essen, 1988

Boundaries
interblob stream

Surfaces
blob stream

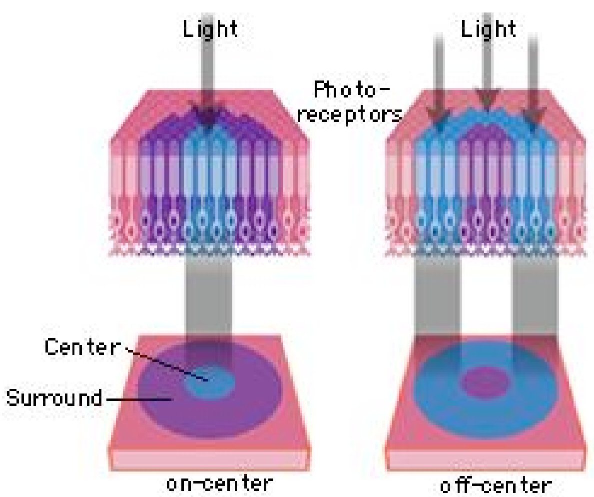
How does the brain compute *ratios* and *filling-in?*



Kolb, Fernandez & Anderson
http://retina.umh.es/Webvision/sretina.html

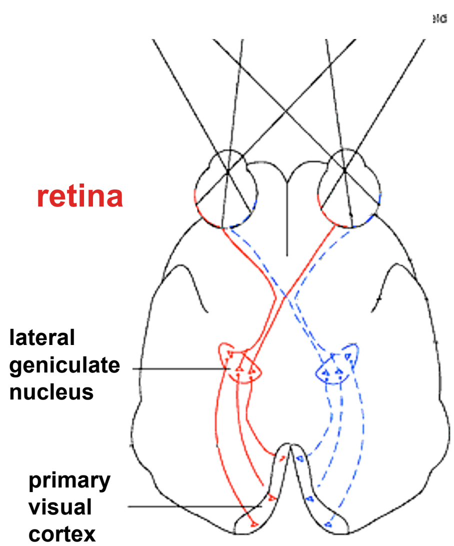
Center-Surround Receptive Fields

Happens *everywhere,*
starting in the retina



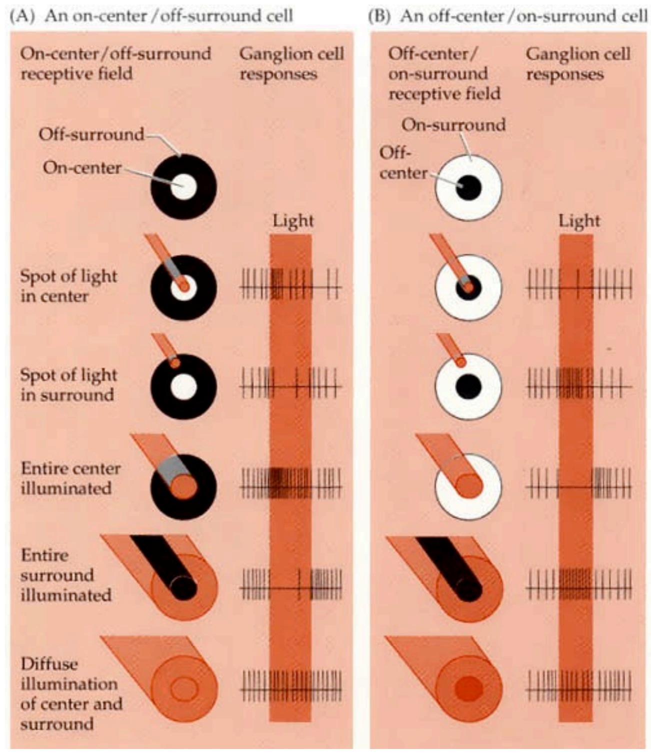
brainconnection.com

Kuffler, 1953



adapted from
blindeyemedia.com

Center-Surround Receptive Fields



How *does* the brain process ratios?

Use a **THOUGHT EXPERIMENT** to clarify **basic issues**.

Simple algebra naturally expresses these issues.

In general, ***math makes understanding simpler!***

How can *center-surround networks* compute ratios?

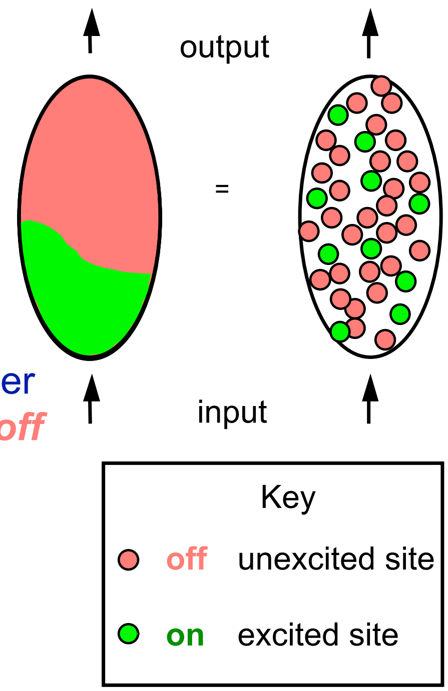
Two parts to this question:

1. **What is a neuron,**
   computationally speaking?

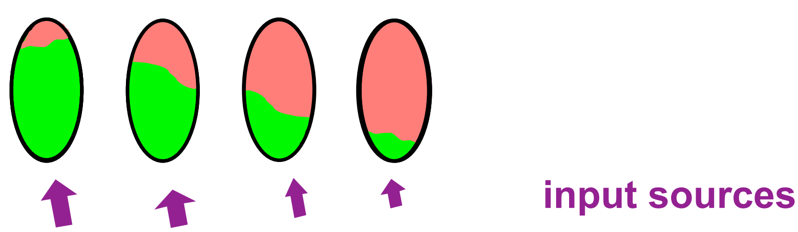
   a maximum and a minimum number of **excitable sites** that turn ***on*** or *off*

   Infinity does not exist in biology.

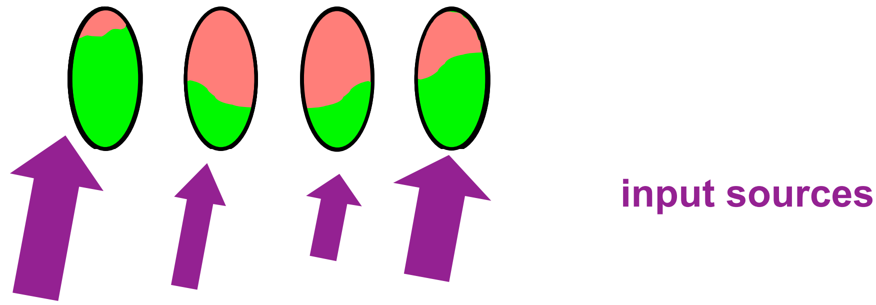
2. **Why do neurons compete?**



Pattern Processing by Cell Networks, A



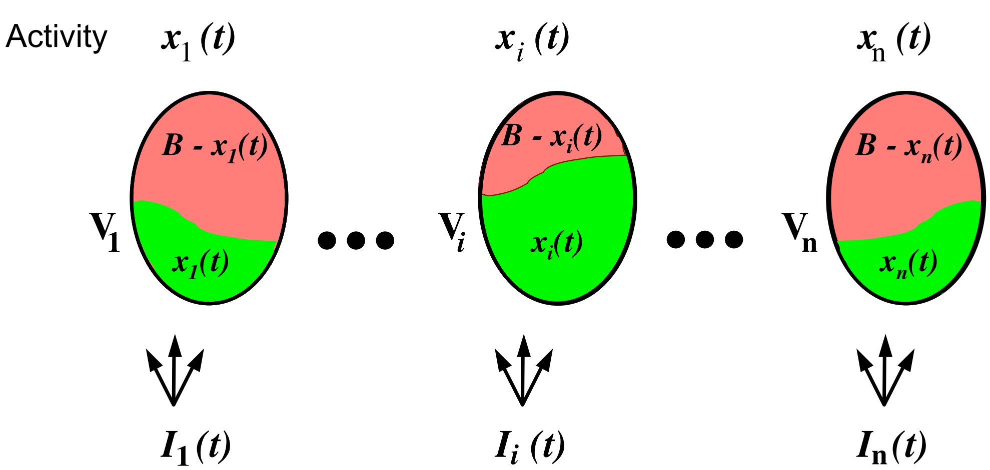
Pattern Processing by Cell Networks, B



Total **size** of inputs to each cell varies wildly through time.

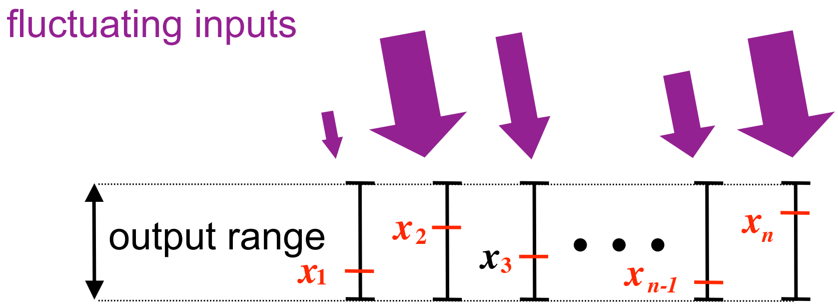
How do cells maintain **sensitivity** to varying input ***patterns***?

Computing in a Bounded Activity Domain



B	**excitable sites** (a constant)	
$x_i(t)$	**excited sites**	**Inhibitory** inputs affect **only** $x_i(t)$
$B - x_i(t)$	**unexcited sites**	**Excitatory** inputs affect **only** $B - x_i(t)$

The Noise-Saturation Dilemma



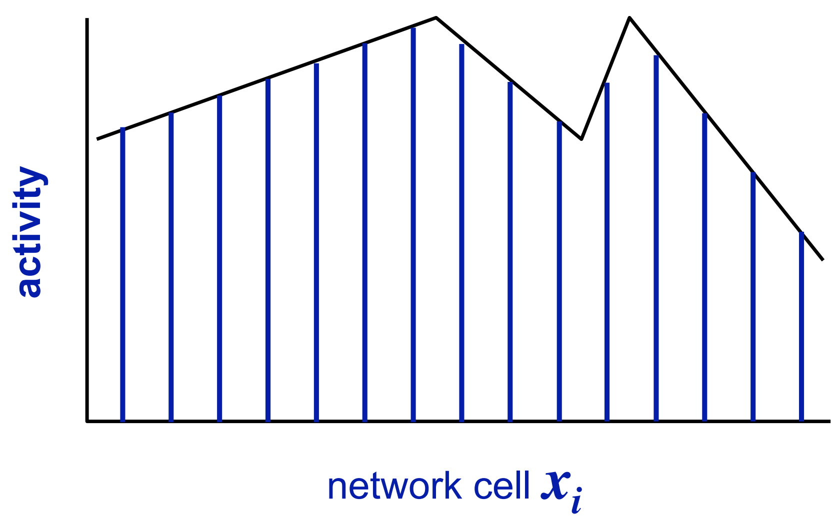
Fixed output range
Fixed output signal functions

Grossberg, 1973

If x_i's are **sensitive to small** inputs, why don't they ***saturate*** in response to large inputs?

If x_i's are **sensitive to large** inputs, why don't small inputs get lost in endogenous ***noise?***

Graphical convention



Noise-Saturation Dilemma

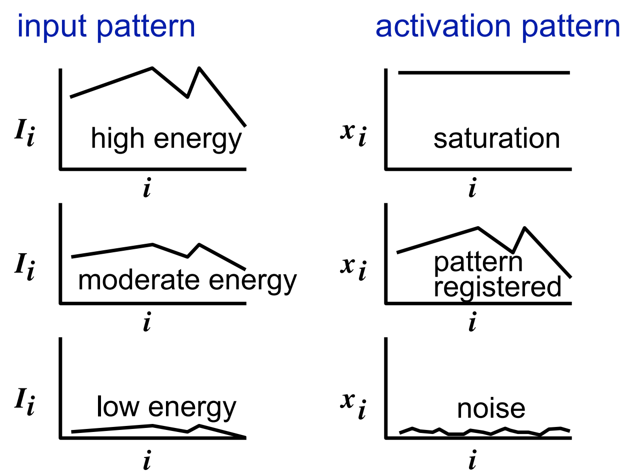
Problem:

remain **sensitive** to

input ratios $\theta_i = \dfrac{I_i}{\sum_j I_j}$

as total input

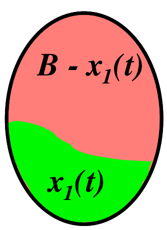
$$I = \sum_j I_j \to \infty$$

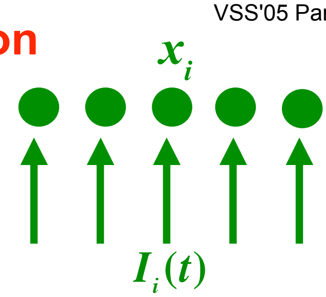


Solution:

Shunting ON-center, OFF-surround networks possess *automatic gain control* that can generate an *wide dynamic range* for effective pattern processing under variable input loads

Shunting Saturation





A, B are constants

$$\frac{d}{dt} x_i = \underbrace{-Ax_i}_{\text{(a)}} + \underbrace{(B - x_i)I_i}_{\text{(b)}}$$

no interactions

(a) Spontaneous decay of activity x_i to equilibrium

(b) Turn on unexcited sites $B - x_i$ by inputs I_i (mass action)

Inadequate response to a *spatial pattern* of inputs:

$$I_i(t) = \theta_i I(t)$$

θ_i **relative intensity (cf., reflectance)**

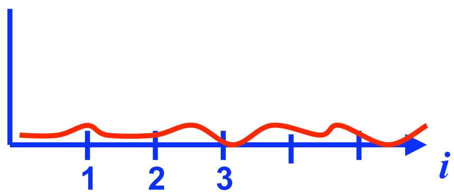
$I(t)$ **total intensity (cf., luminance)**

Shunting Saturation

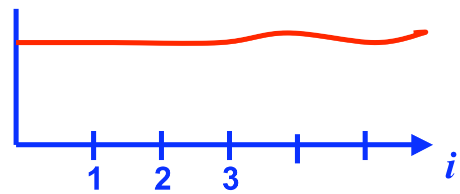
At equilibrium: $$0 = \frac{d}{dt}x_i = -Ax_i + (B - x_i)I_i$$

$$x_i = \frac{BI_i}{A + I_i} = \frac{B\theta_i I}{A + \theta_i I} \rightarrow B \quad as \quad I \rightarrow \infty$$

$$I_i = \theta_i I \qquad I = \sum_j I_j$$



I small: lost in noise



I large: saturates

Sensitivity loss to relative intensity as total intensity increases

Computing with Patterns

How to compute the pattern-sensitive variable:

$$\theta_i = \frac{I_i}{\sum_{k=1}^{n} I_k} \quad ?$$

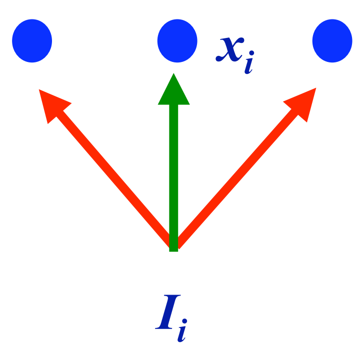
the **ratio** of **one input** to the **sum of all inputs**

Need interactions! What type?

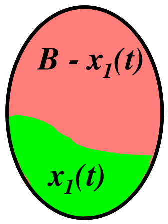
$$\theta_i = \frac{I_i}{I_i + \sum_{k \neq i} I_k}$$

$I_i \uparrow \Rightarrow \theta_i \uparrow$ **excitation**

$I_k \uparrow \Rightarrow \theta_i \downarrow$ **inhibition**



Shunting Dynamics



unexcited sites are "switched **ON**" by mass action from "their" (excitatory) inputs, and

excited sites are "switched **OFF**" by mass action from "other" (inhibitory) inputs:

$$\frac{dx_i}{dt} = \underbrace{-Ax_i + (B - x_i)I_i}_{\text{before}} \underbrace{- x_i \sum_{k \neq i} I_k}_{\text{new}}$$

Effects of *Shunting Inhibition*

At equilibrium: $$x_i = \underbrace{\theta_i}_{\text{PATTERN}} \underbrace{\frac{BI}{A+I}}_{\text{ENERGY}} \qquad x_i \to \theta_i \ \ as \ \ I \to \infty$$

"factorization"

Input to a node: I_i or $I_i(t)$ for $i = 1, \ldots n$

Total input: $I = \sum_j I_j$

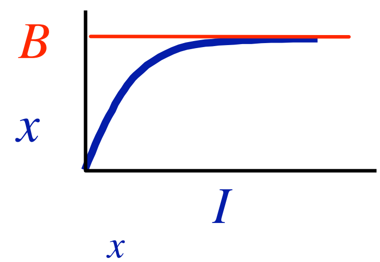
Normalized input: $\theta_i = \dfrac{I_i}{I}$

ratio sensitivity over a wide dynamic range: ***automatic gain control***

Ratios require **ON-center OFF-surround** anatomies!

And moreover . . .

$$x = \sum_k x_k = \frac{BI}{A+I} \leq B, \text{ since } \sum_k \theta_k = 1$$



Total network activity is ***bounded*** for all inputs.

Normalization! . . . limited capacity

Hodgkin/Huxley Equations and Shunting Networks

Shunting ON-center, Off-surround networks

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (x_i + C)\sum_{k \neq i} I_k$$

hyperpolarization constant

are consistent with ***membrane equations*** of physiology

$$C\frac{\partial V}{\partial t} = \underbrace{\overset{\text{excitatory}}{(V^+ - V)g^+}}_{\text{Na}^+\text{ channel}} + \underbrace{\overset{\text{inhibitory}}{(V^- - V)g^-}}_{\text{K}^+\text{ channel}} + \underbrace{\overset{\text{passive}}{(V^P - V)g^P}}_{\text{Cl}^-\text{ channel}}$$

$V^+ \to B$

$C \to 1$

$V^- \to C$

$V^P \to 0$

$g^P \to A$

$g^+ \to I_i$

$g^- \to \sum_{k \neq i} I_k$

a link between dynamics and anatomy

Hodgkin and Huxley, 1952
Grossberg, 1968
Carpenter, 1981

Mudpuppy Retina Neurophysiology

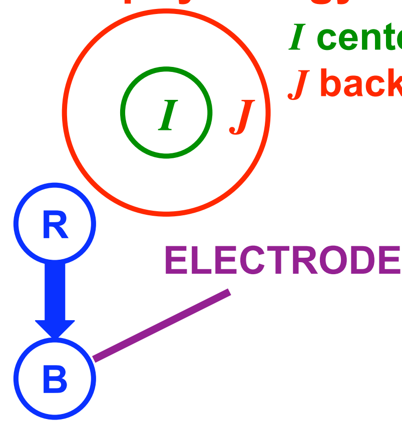
a) Relative figure-to-ground θ_i

b) Weber-Fechner $\dfrac{I}{A+J}$

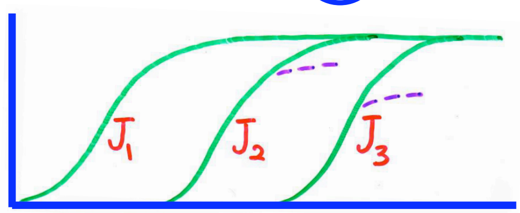
c) No hyperpolarization
SHUNT: Silent inhibition

d) Shift property:

I center
J background



$x_i(K,J)$



$K = ln(I)$

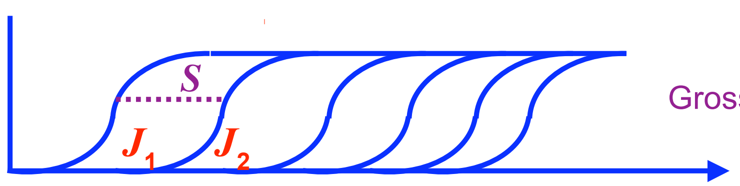
Werblin, 1970

ADAPTATION: sensitivity *SHIFTS for different backgrounds*

NO COMPRESSION

Weber Law, Adaptation, and Shift Property

$x_i(K,J)$



Grossberg, 1981

$K = ln\left(I_i\right)$

Convert to logarithmic coordinates:

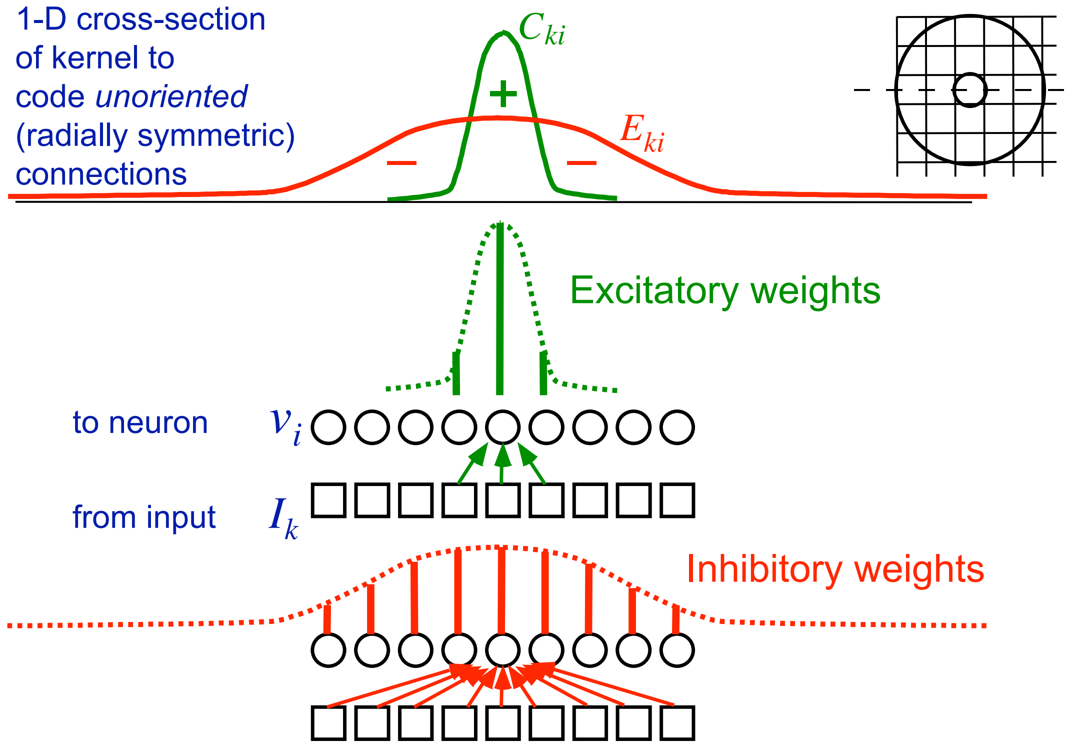
$$K = ln\left(I_i\right), \quad I_i = e^K, \quad J = \sum_{k \neq i} I_k$$

$$x_i(K,J) = \frac{Be^K}{A + e^K + J}$$

$$x(K+S, J_1) \equiv x(K, J_2), \quad S = ln\left(\frac{A+J_1}{A+J_2}\right)$$

size of SHIFT

Generalize to *Multiple Spatial Channels:* Distance-Dependent Kernels



Shunting Network with Distance-Dependent Terms

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)\sum_{k=1}^{n} I_k C_{ki} - (x_i + D_i)\sum_{k=1}^{n} I_k E_{ki}$$

$$C_{ki} = C\exp\left[-\mu(k-i)^2\right]$$

$$E_{ki} = E\exp\left[-\nu(k-i)^2\right]$$

Note: ***both*** subtractive and shunting terms.

Equilibrium of Distance-Dependent Network

Set $\frac{dx_i}{dt} = 0$ and recall that $I_k = I\theta_k$

$$x_i = \frac{I\sum_{k=1}^{n}\theta_k\left(BC_{ki} - DE_{ki}\right)}{A + I\sum_{k=1}^{n}\theta_k\left(C_{ki} + E_{ki}\right)}$$

Numerator: **DoG**

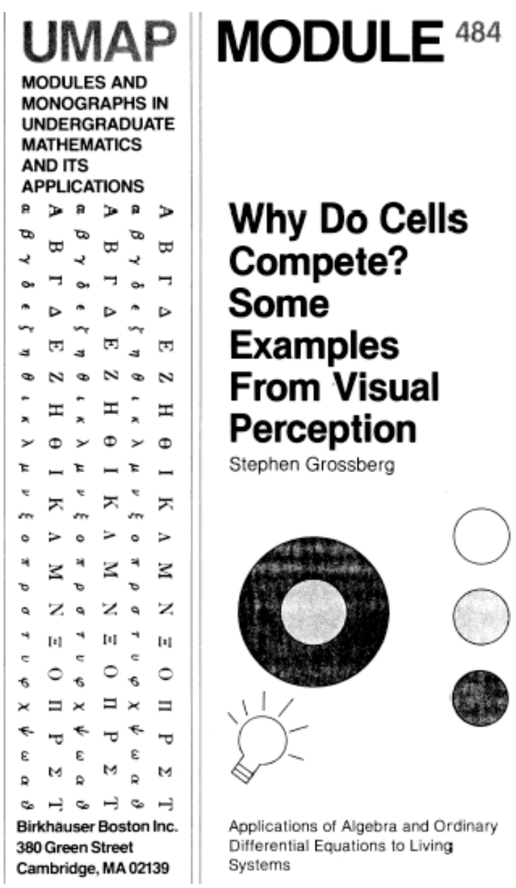
difference of Gaussians

Denominator: **SoG**

sum of Gaussians

scaled against ***A***

Grossberg, 1973
Heeger, 1992
Douglas et al. 1995



102

Intermodular Description Sheet: UMAP Unit 484

Title: WHY DO CELLS COMPETE?
SOME EXAMPLES FROM VISUAL PERCEPTION

Author: Stephen Grossberg
Department of Mathematics
Boston University
Boston, MA 02215

Review Stage/Date: III 10/30/81

Math Fields: Algebra and ordinary differential equations.

Application Field: Living systems.

Target Audience: Students in beginning courses in differential equations or mathematical modeling.

Abstract: This unit introduces current data about brightness constancy and contrast in visual perception, and shows that data of this kind are a consequence of a principle that holds in all cellular systems. Variations on this principle imply other visual properties, such as edge detection, spatial frequency detection, and pattern-matching properties.

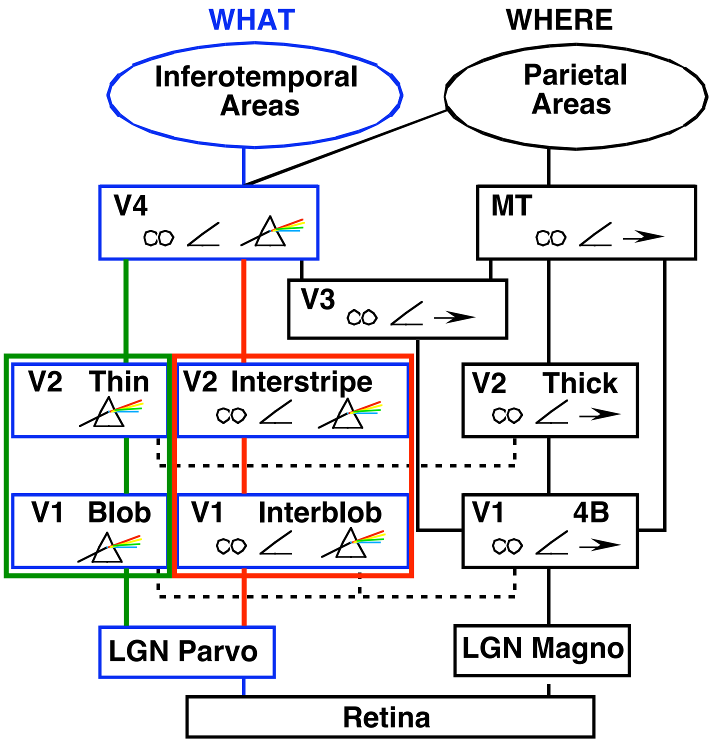
Corequisites: Elementary first-order differential equations.

Related Unit: A Strange Result in Visual Perception (Unit 251)

The UMAP Journal, Vol. III, No. 1, 1982

Downloadable from
http://cns.bu.edu/techlab/

Next: Boundary & Surface Streams

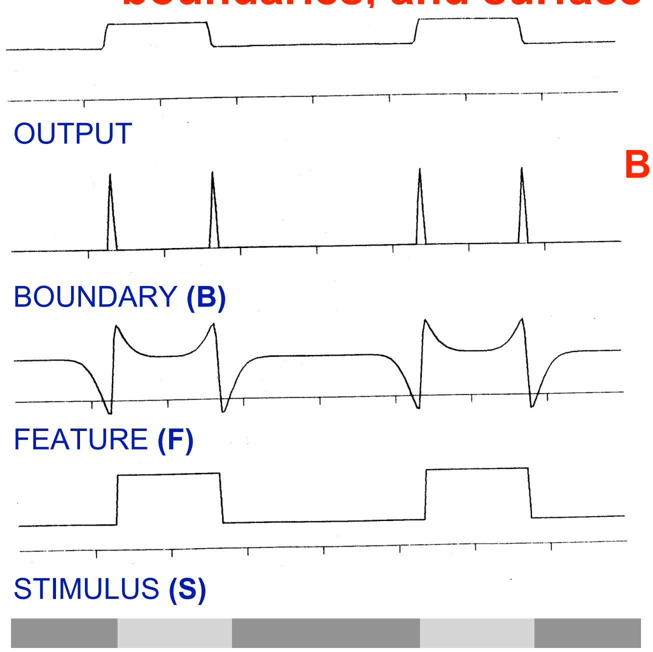


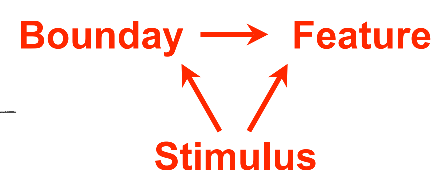
DeYoe and van Essen, 1988

Boundaries
interblob stream

Surfaces
blob stream

Combine shunting network, boundaries, and surface filling-in, A



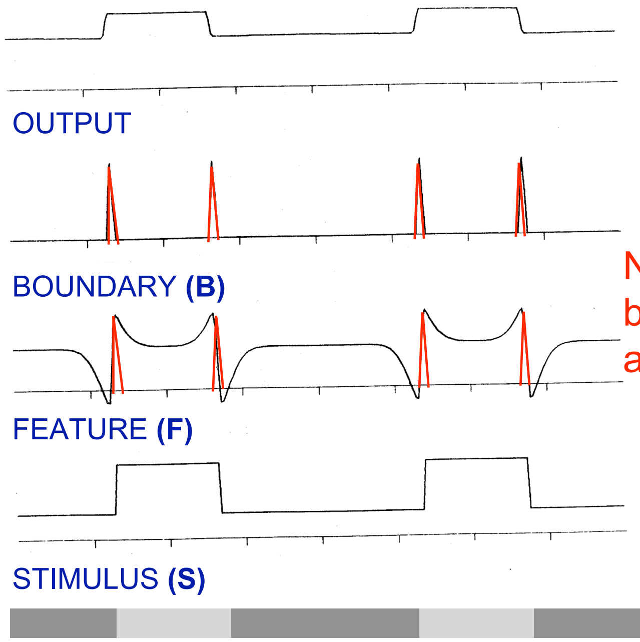


Veridical!

Grossberg and Todorović, 1988

Boundary peaks are spatially *narrower* than featural peaks

Combine shunting network, boundaries, and surface filling-in, B

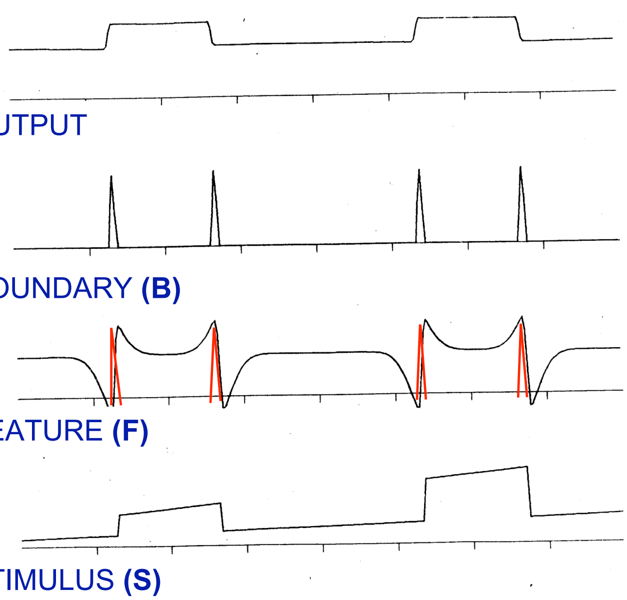


Note spatial registration of boundary (red highlight) and feature signals

Grossberg and Todorović, 1988

Boundary peaks are spatially *narrower* than featural peaks

Brightness Constancy

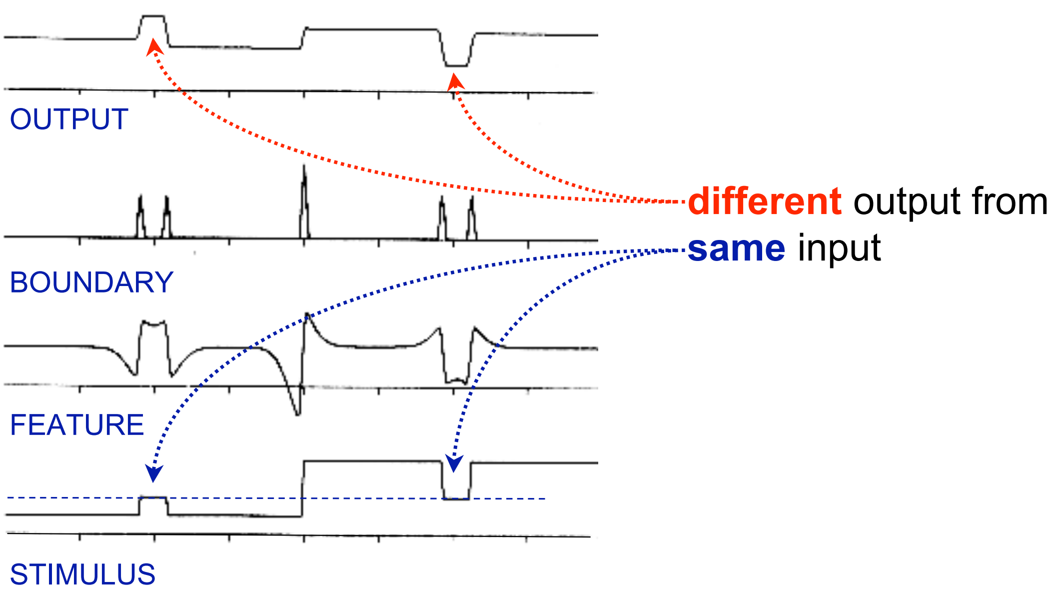


B → F
S

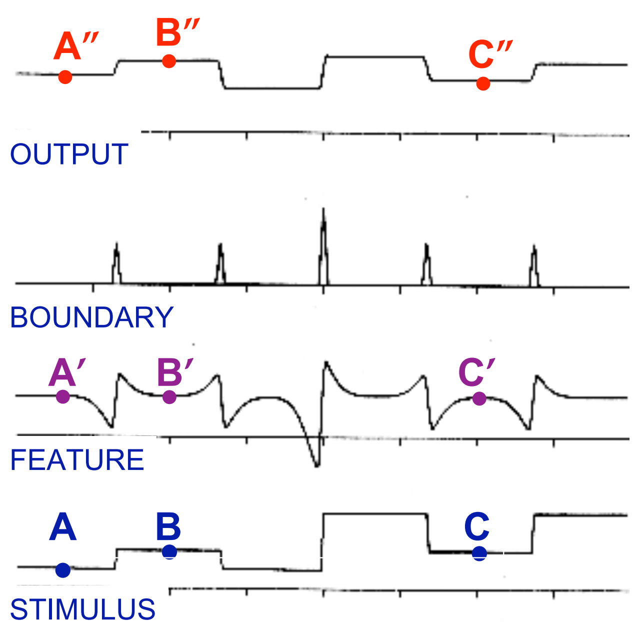
Not veridical, but useful!

ratio-sensitive edges **in** ***FCS***

Brightness Contrast: Small regions



Brightness Contrast, Large Targets



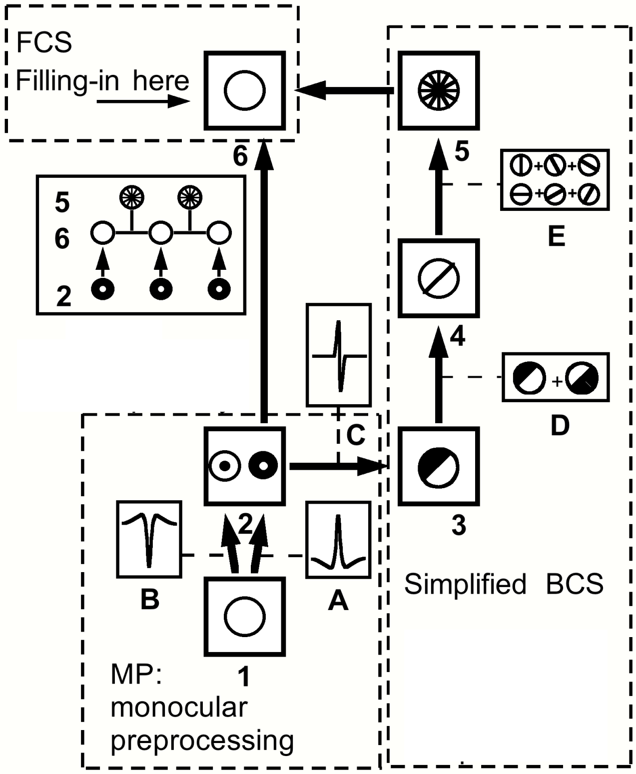
different output at
A″, **B″**, and **C″** from

same *FCS signal*
at **A′**, **B′**, and **C′**

. . . and
same input at **B** and **C**

Requires ***filling-in*** to understand

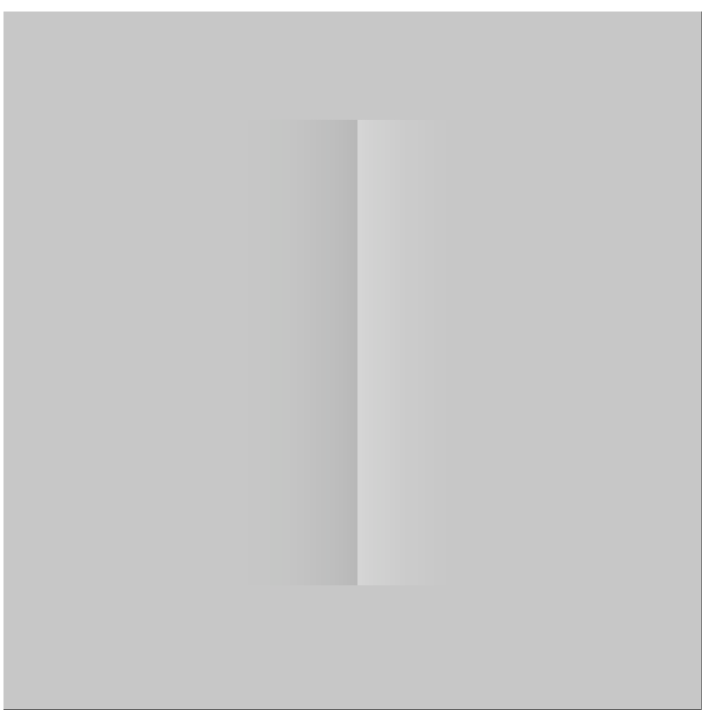
Grossberg and Todorović 1988 Macrocircuit



Pooling over **orientations** and **contrast polarity**

oriented filtering for boundary detection

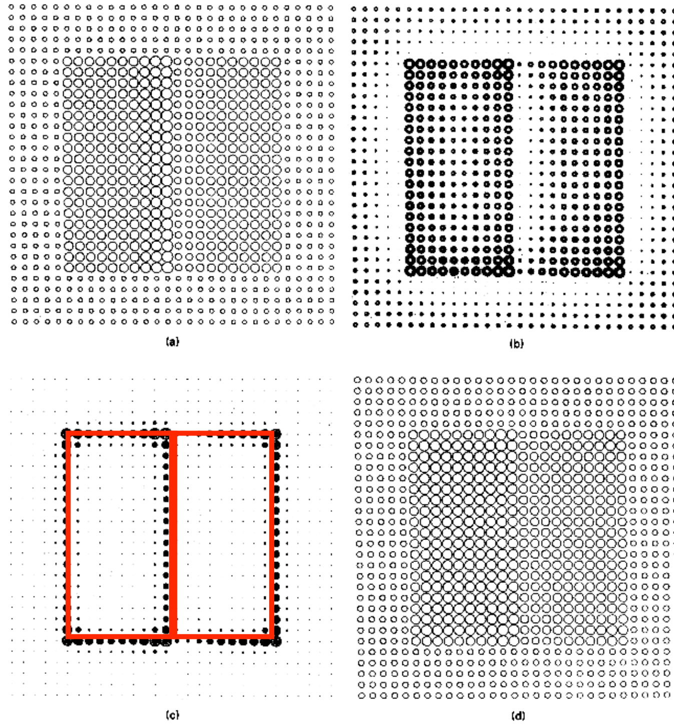
Craik-O'Brien-Cornsweet Effect



Boundary completion defines **filling-in** compartments.

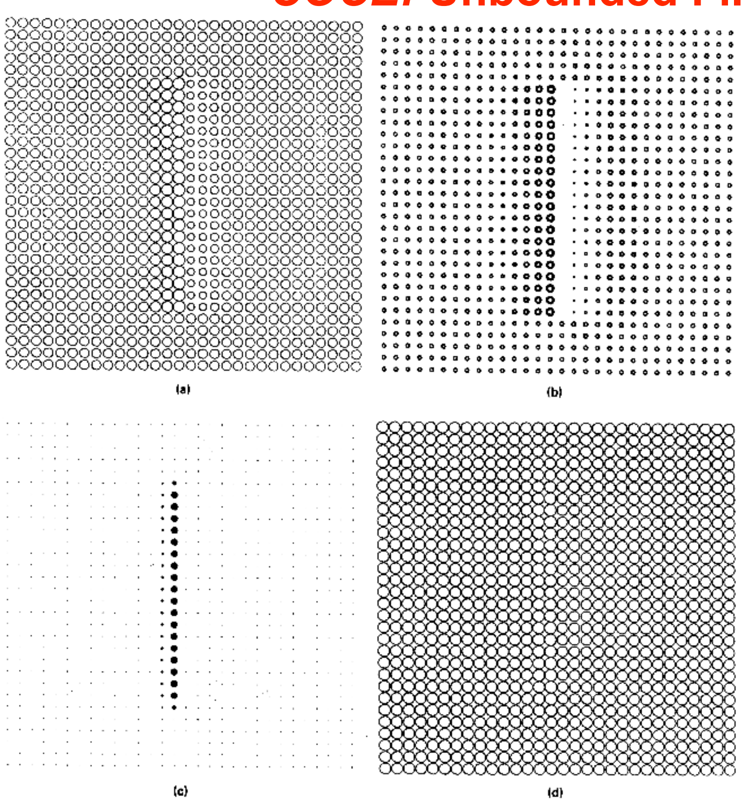
Filling-in determines what we **see** in each compartment.

COCE: Closure and Filling-in



Note the crucial role of
closed compartments

COCE: Unbounded Filling-in



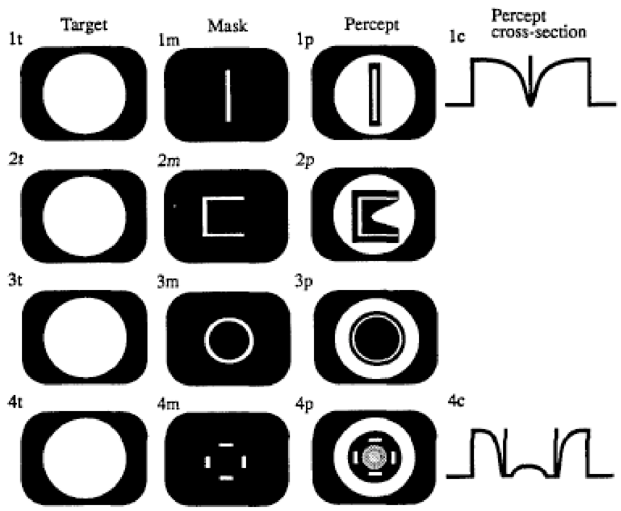
No outer boundaries
➔ no *illusion.*

Not just
“attenuation of low
spatial frequencies”

MANY experiments on filling-in!

Paradiso and Nakayama, 1991

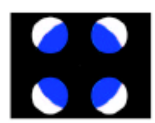
Catching filling-in "in the act!"



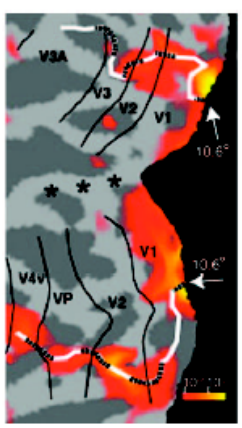
Cortical Loci of Boundary Completion and Filling-in

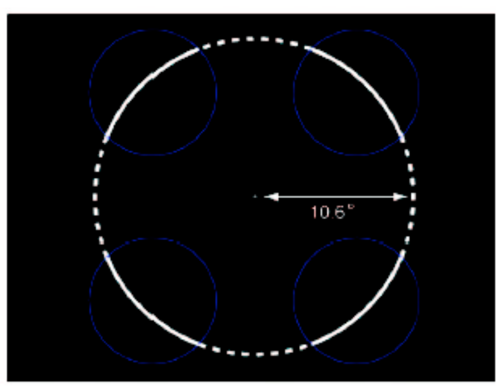
Sasaki and Watanabe, 2004

Boundary: V1, V2, V3/VP, V4v



Neon filling-in: V1 only





Oriented filtering is not enough

Need:

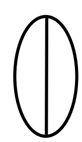
grouping
boundary completion
3-D figure/ground [Part 2 today]

to get the right perceptual compartments for ***filling-in:***

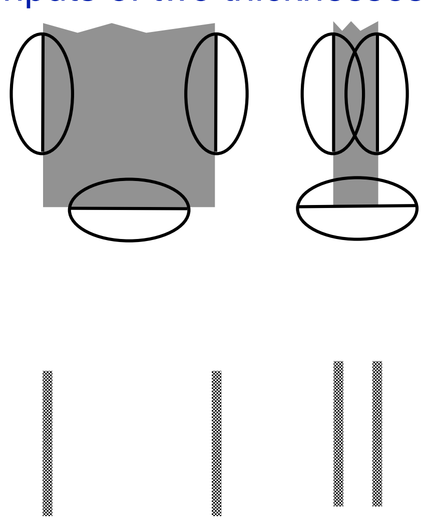
Gillam, 1987

How Thin Is "Thin"?

For a given receptive field size:



Inputs of two thicknesses:



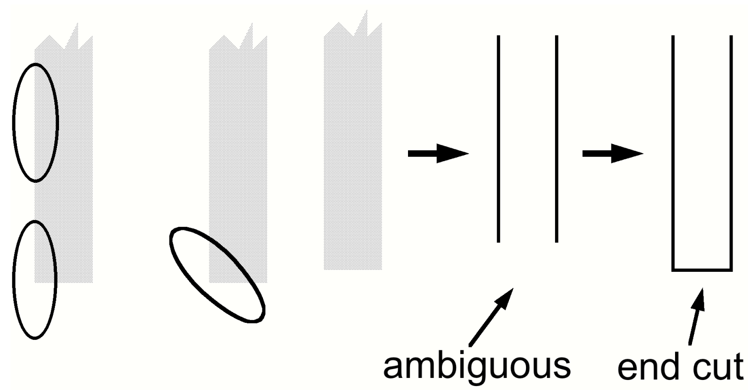
For a *thin* line

no detector perpendicular to line end can respond "enough"

. . . based on bottom-up input alone.

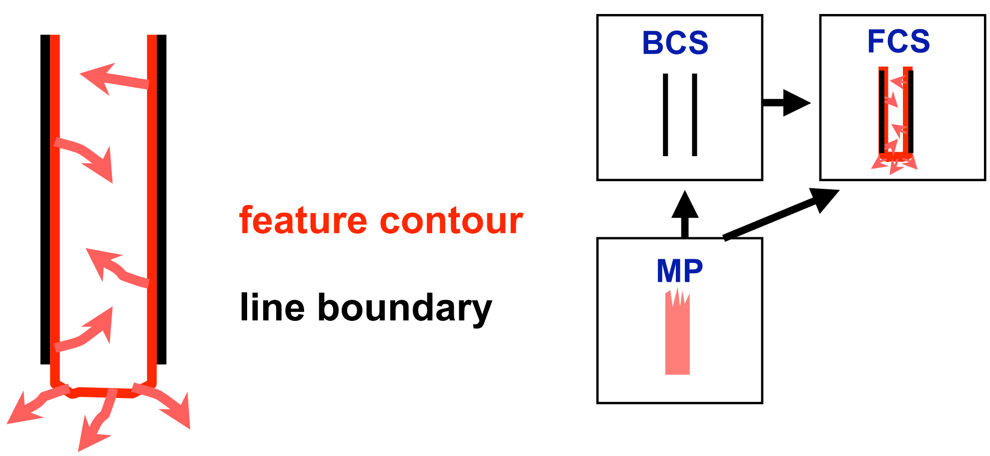
End Cuts

Visual system
must *synthesize*
a line end.



If No End Cuts . . .

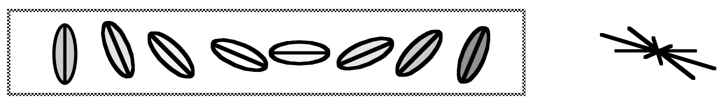
A PERCEPTUAL DISASTER
in the Feature Contour System



Color flows from line end!

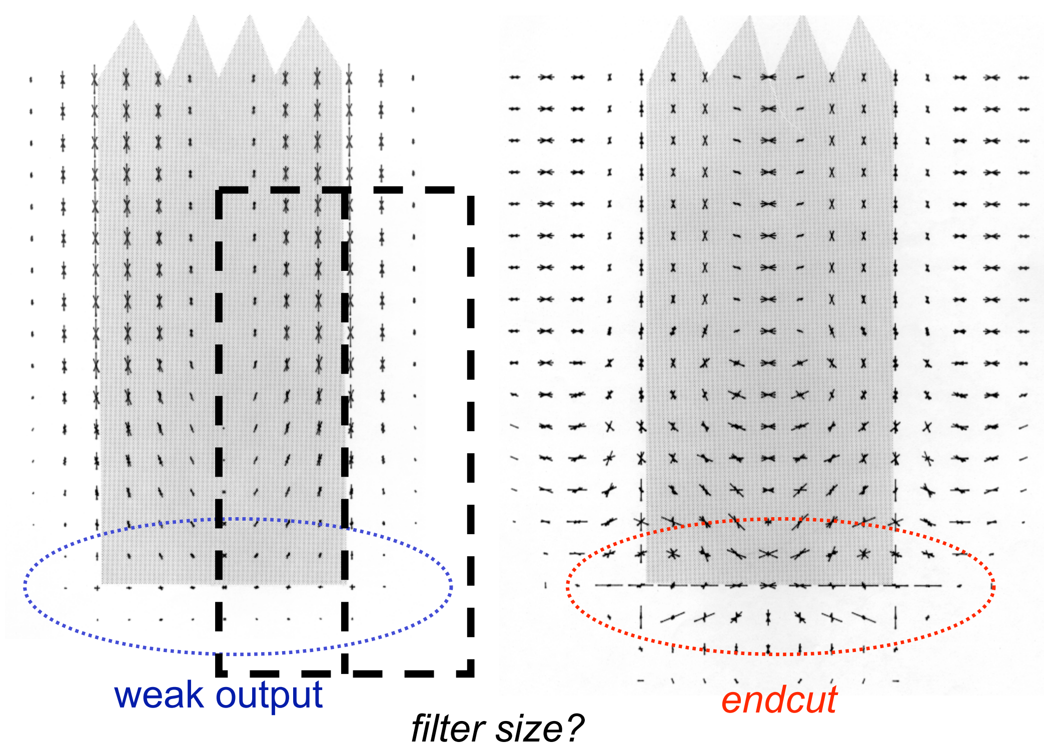
Graphical Notation

Orientation hypercolumn



More active cells have lighter shading

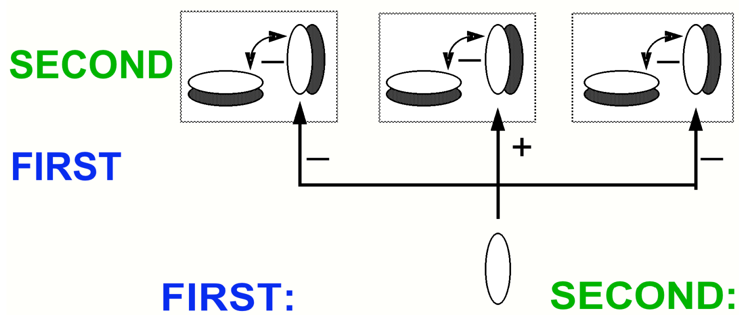
Endcut simulation



BCS: Short-Range Competition

End cuts (via 1985 mechanism)

Two stages of *short-range competition*



FIRST:

across location
same orientation

SECOND:

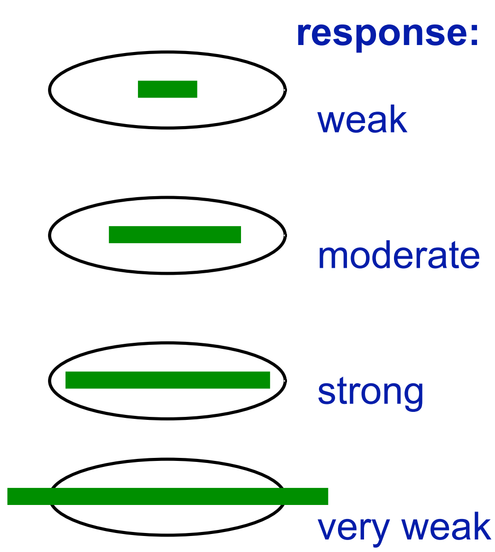
same location
across orientation

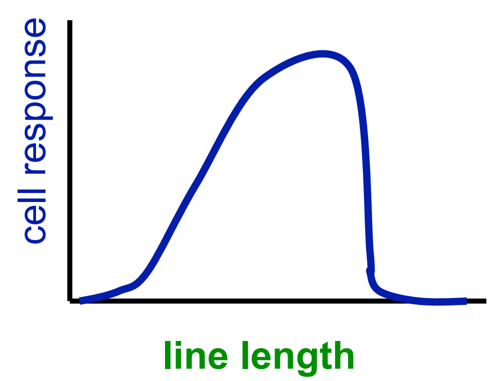
---

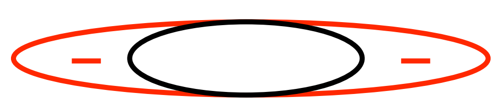
* Not just between perpendiculars

Endcut = endstopped plus . . .

Complex and (even) *"simple"* cells may be **endstopped.**
How can you tell?





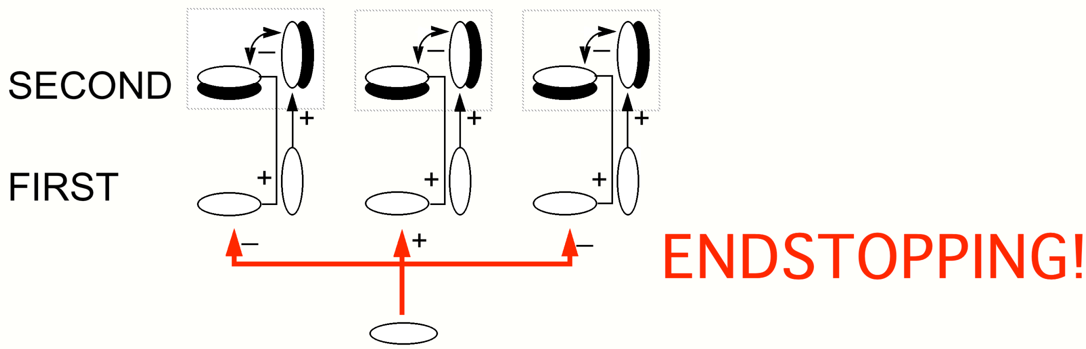


inhibition

Endstopping: The First Competitive Stage

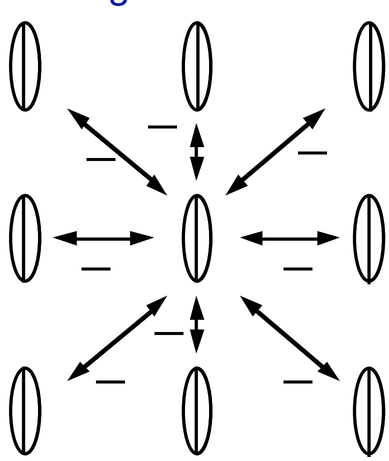
Mechanism for generating end cuts:

Two stages of short-range competition



In other words . . .

"Lateral inhibition" among neighboring cells with similarly oriented receptive fields can generate **endstopping.**



(Overlapping: ellipses are 10 times illustrated size.)

Variations on Shunting Network Equations

Shunting competition:
within orientations, k
across positions, pq to ij

$$\frac{d}{dt} w_{ijk} = -w_{ijk} + I + f\left(J_{ijk}\right) - w_{ijk} \sum_{(p,q)} J_{pqk} A_{pqij}$$

Just a variation of "center-surround" equation,

. . . but with additional indices for
2-D position and **orientation**

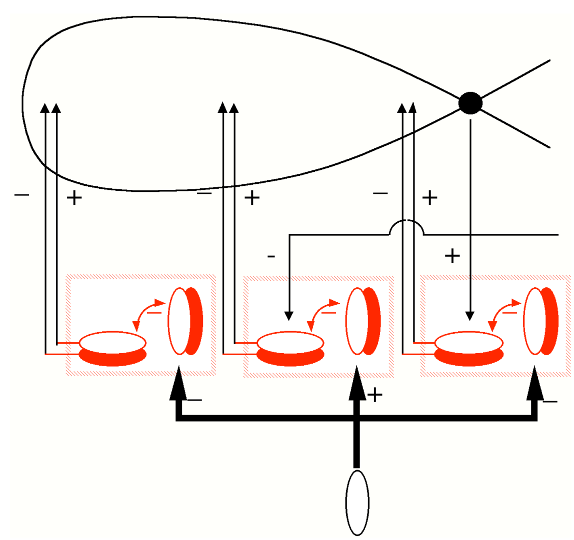
Second Competitive Stage

Begin with: **push-pull opponent process**

$$x_{ijk} = w_{ijk} - w_{ijK}$$

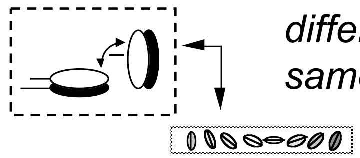
where orientation k is **perpendicular** to orientation K

followed by . . .



Normalization in cross-orientation inhibition

normalization across orientations
at each position (dashed boxes)



different icons; same structure

$$O_{ijk} = O(x_{ijk}) = C\left[w_{ijk} - w_{ijk}\right]^{+}$$

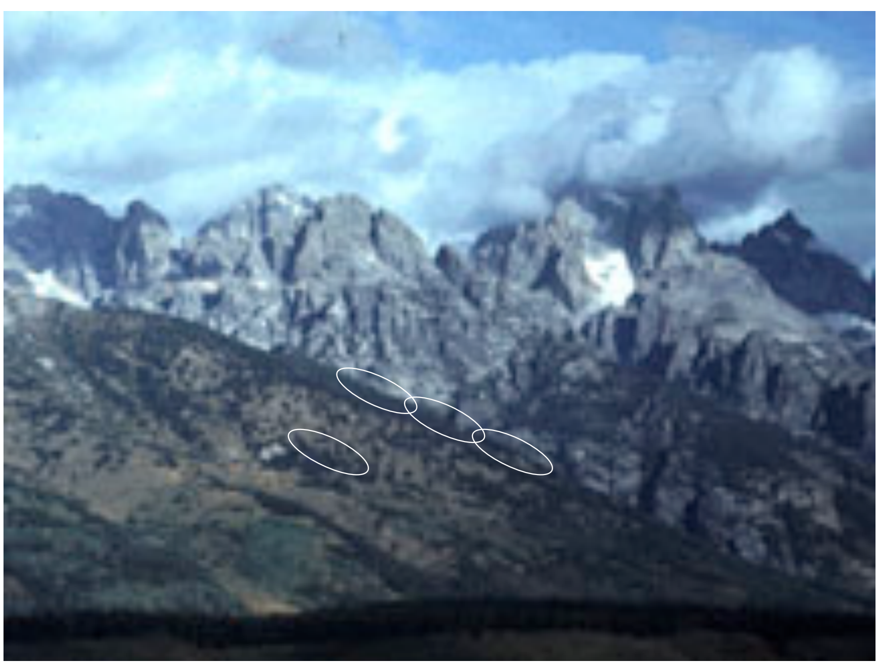
$$\frac{d}{dt}y_{ijk} = -Dy_{ijk} + \left(E - y_{ijk}\right)O_{ijk} - y_{ijk}\sum_{m\neq k} O_{ijm}$$

At equilibrium: $y_{ijk} = \dfrac{EO_{ijk}}{D + O_{ij}}$

$$\text{where} \quad O_{ij} = \sum_{m} O_{ijm}$$

Grossberg, 1973
Heeger, 1993

Boundary completion in the real world?

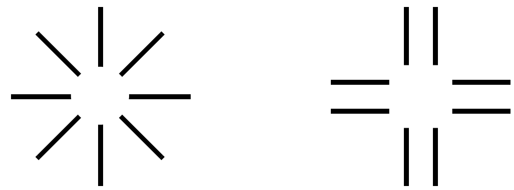


Need: long-range oriented cooperation -- feedback!

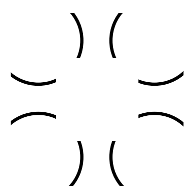
Cooperative-Competitive Nonlinear Feedback

1985:
Use cooperative-competitive nonlinear feedback
CC Loop
to complete and sharpen boundaries.

Recall: *Perpendicular induction* at line ends:



Long-range cooperation
can win over locally preferred orientations

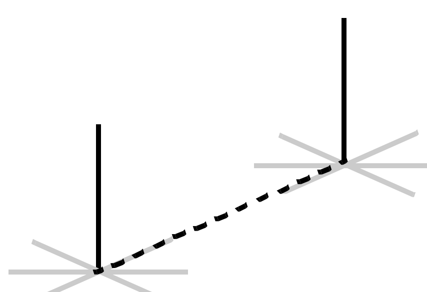


Kennedy, 1979

Boundary Grouping

Each line end induces a “fuzzy band” of “almost perpendicular” **candidate directions for grouping**

When aligned across perceptual space, **cooperative completion** of boundaries



From Fuzzy to Sharp

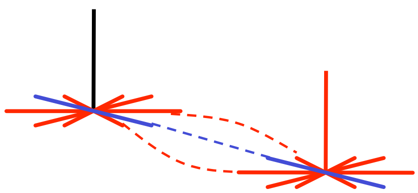
Why do we not always perceive fuzzy illusory contours?

Hierarchical resolution of uncertainty:
1) **Need fuzziness** to initiate grouping.
2) Risk loss of acuity.

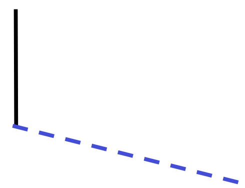
CC LOOP is a decision process.

CHOOSE: the contextually best orientation -- **cooperation!**

SUPPRESS: other local orientations -- **competition!**

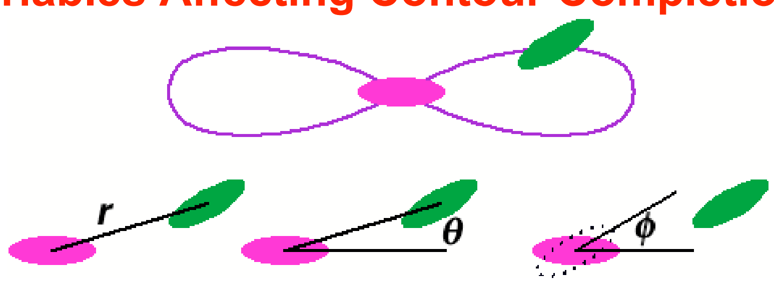


before choice
(transient)



after choice
("equilibrium")

Variables Affecting Contour Completion



proximity ***r*** of center of "inducing unit" to center of "receiving unit"

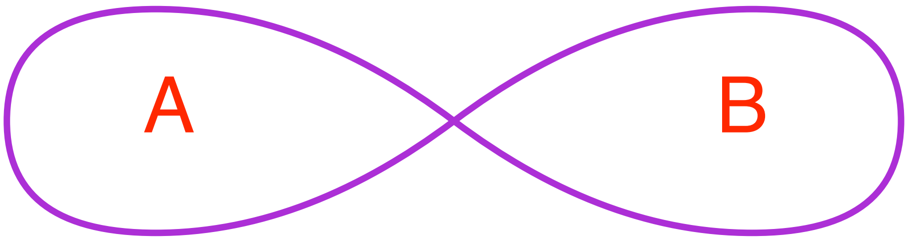
alignment θ angle formed by inducing unit's center relative to preferred axis of receiving unit

orientation ϕ difference in preferred orientation of inducing and receiving units

The Bipole Property

"Completable" perceptual gap bridged in one or two cycles

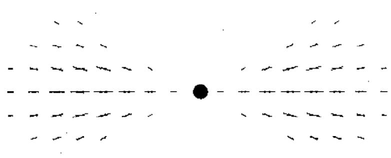
completion via long-range cooperative units



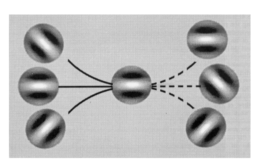
fuzzy "AND" gate

Bipoles Through the Ages

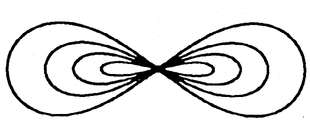
Grossberg & Mingolla, 1985



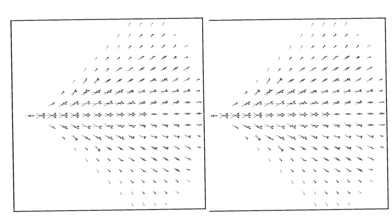
Field, Hayes, & Hess, 1993
"association field"



Heitger & von der Heydt, 1993



Williams and Jacobs, 1997

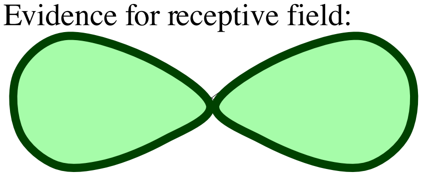


Cf. **"relatability"** -- geometric constraints on which contours get to group with which -- Kellman & Shipley, 1991
Also, Ullman, Zucker, Mumford, Guy & Medione **"tensor voting"**

Long-Range Boundary Completion

Stimulus: / Probe location:	Cells in V2 Response?
	YES
	NO
	NO
	YES
(more contrast)	NO
	YES

Evidence for receptive field:



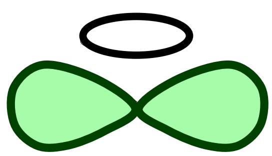
von der Heydt, Peterhans, & Baumgartner, 1984

Peterhans & von der Heydt, 1988

More Data

Horizontally tuned cells:

Probe location:



Stimulus	V1 response	V2 response
	Strong	Strong
	None	Weak, with orientationally FUZZY receptive field
	None	Stronger, with orientationally SHARPER receptive field (same cell as above)

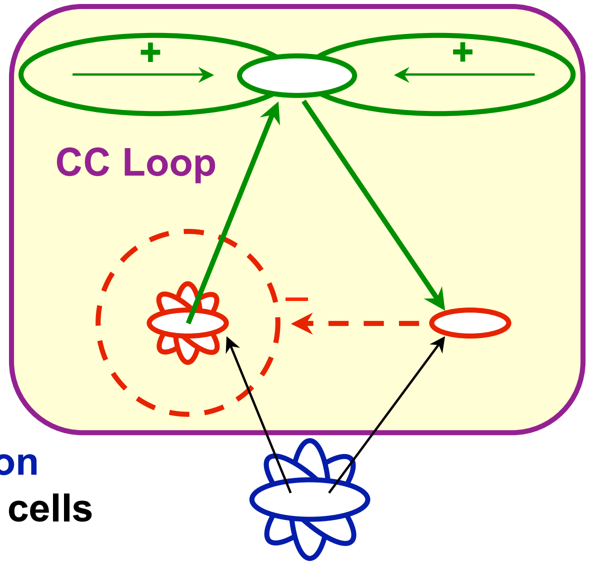
Evidence for:
1) orientationally fuzzy **end cuts**
2) oriented, long-range **cooperation.**

von der Heydt, Peterhans, & Baumgartner, 1984

Peterhans & von der Heydt, 1988

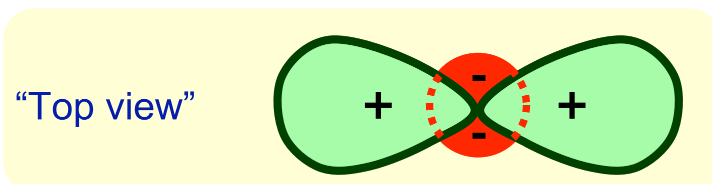
Cortical BCS Stages

Long-range cooperation



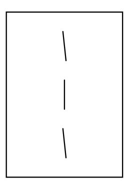
Short-range competition
across position
across orientation

Oriented boundary detection
simple and complex cells

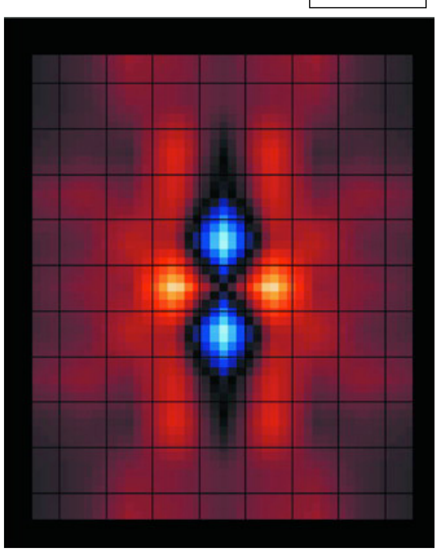


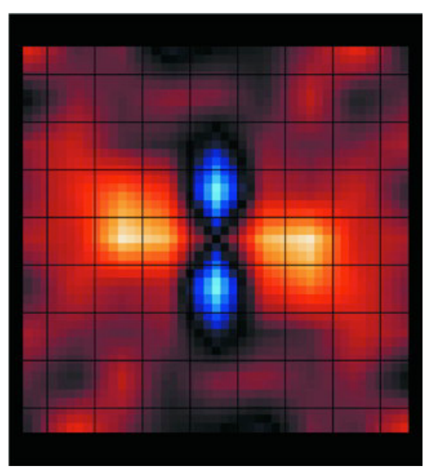
Parallel Studies

Psychophysics



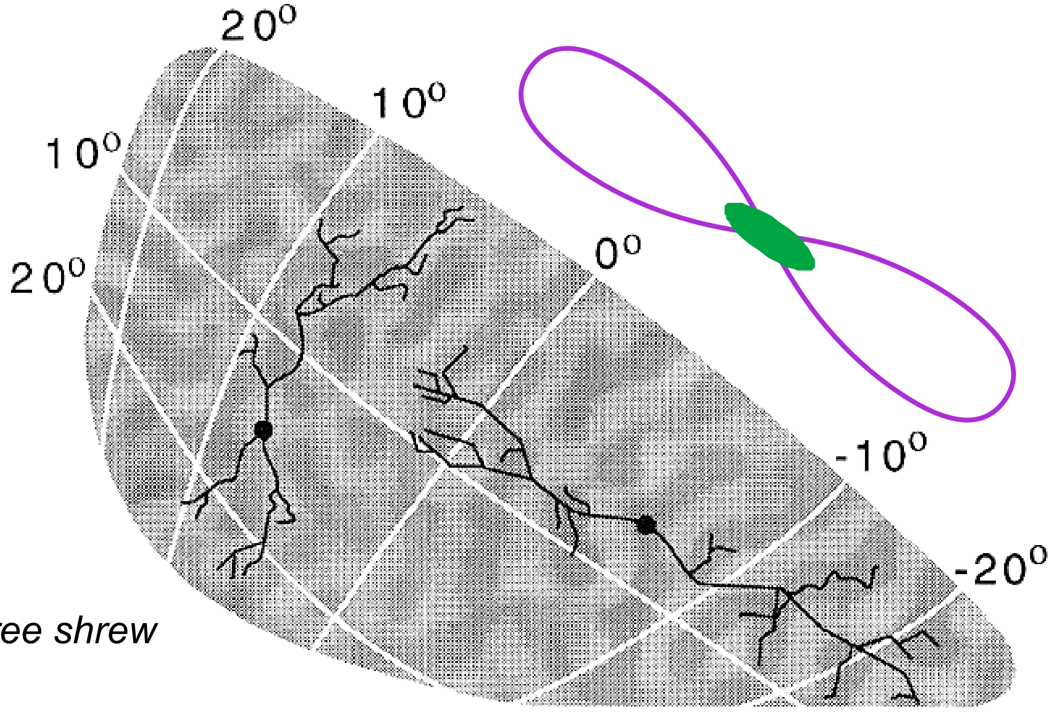
Physiology





Kapadia, Ito, Gilbert, and Westheimer 1995

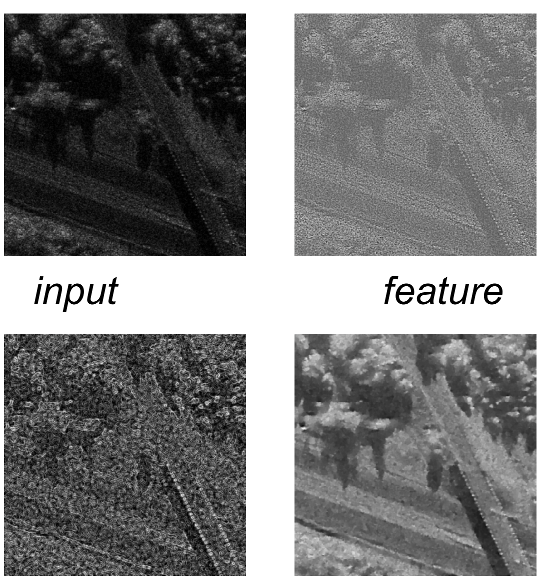
Horizontal Connections in Striate Cortex



tree shrew

Bosking, *et al.*, 1997

Do these ideas work on hard problems?



input *feature*

boundary *filling-in*

Mingolla et al., 1999

Application:
Image Enhancement

Synthetic aperture radar

signal: 5 orders of magnitude

multiplicative noise

sparse high-intensity pixels

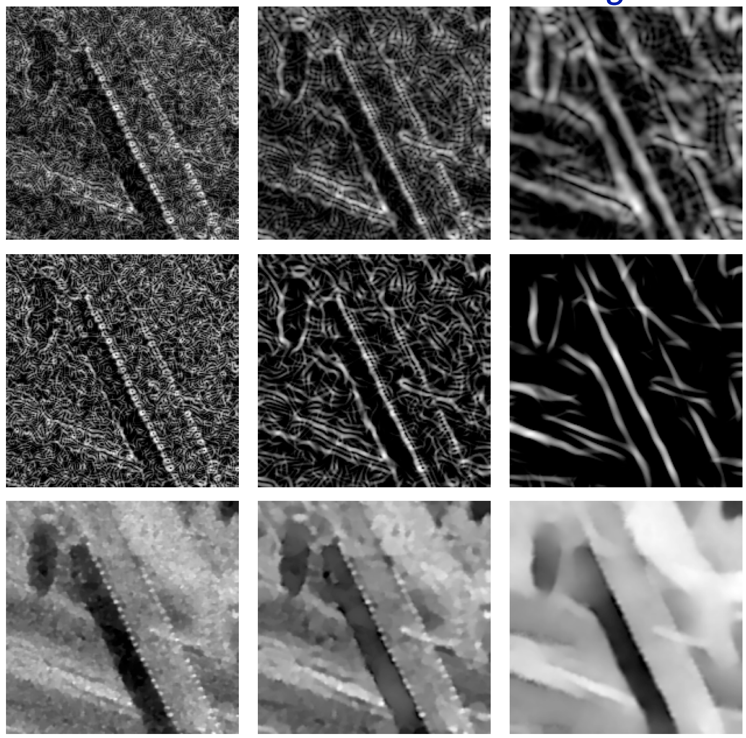
Details of Image Enhancement

Scale: *small* *medium* *large*

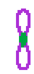
boundaries before completion

boundaries after completion

filling-in



large scale bipole:



Design Themes

Theorems: A foundation for designing more realistic networks

Role of **nonlinear signal functions** in choosing **strongest groupings.**

Role of **competition** in **self-normalizing networking activity**

Role of **short-term memory** in **storing winning grouping** and providing **coherence**

-- same issues in cognitive information processing

Recurrent Shunting Networks in Vision

To join **grouping** with **coherent binding**, we need:

spatial and orientational kernels (e.g. bipole)

multiple nested **layers with feedback loops**

Earlier analysis of **feedforward shunting ON-center, OFF-surround network** is not enough!

Grouping: Combining Cooperation and Competition

Bottom up: The **competition** influences the **cooperation.**

But the strongest **cooperation** also biases the **competition:**

COOPERATION

COMPETITION

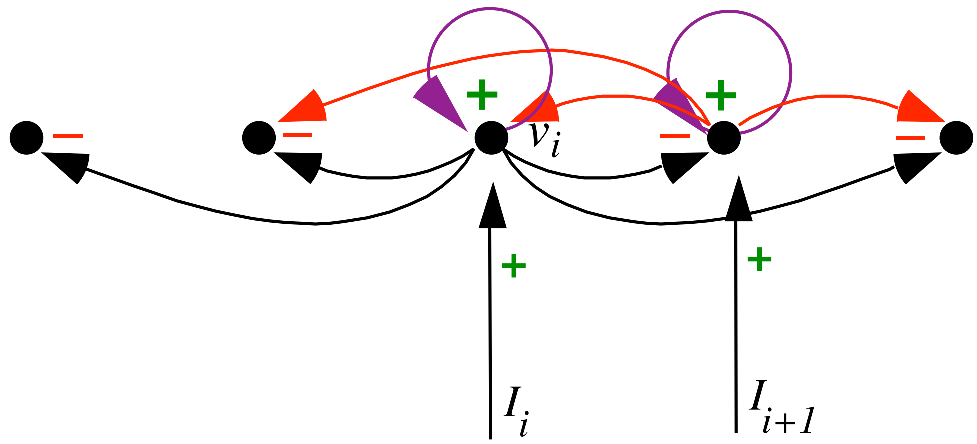
Need: FEEDBACK NETWORKS

Classify: information processing and storage abilities

New property: Coherent binding
The grouping is the emergent unit. **Grossberg, 1973+**

Noise-Saturation Dilemma -- Again!

Need: ON-center, OFF-surround with FEEDBACK



More complicated situation

Greater need for mathematical analysis to clarify . . .

Feedback Shunting Networks

Given a network's anatomy, it signal functions, parameter restrictions, and initial conditions, ask:

STABILITY: Is there **storage** of a pattern (**short-term memory**)?

PATTERN TRANSFORMATION:

What happens to initial activity pattern?

Is it preserved, destroyed, smoothed, contrast-enhanced, …?

Properties of Recurrent Competitive Networks

Grossberg, 1973:

What happens to x *(total* network activity) as (time) $t \rightarrow \infty$?

Possibilities:

- $x \rightarrow 0$ "collapse" of all activity
- $x \rightarrow \infty$ network "blows up"
- $x \rightarrow$ constant (stability) **storage!**
- $x \rightarrow$ one of finitely many values
- $x \rightarrow$ one of infinitely many (finite) values
- x oscillates
- x is chaotic (not in 1973!)

Key result:
Network ***anatomies*** *(patterns of connections) and* ***signal functions*** *constrain outcomes.*

Recurrent Network Analysis

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)\Big[\underset{\uparrow}{f(x_i)} + I_i^+\Big] - x_i\Big[\sum_{k \neq i} \underset{\uparrow}{f(x_k)} + I_i^-\Big]$$

feedback **feedback**

Let inputs I_i^+, I_i^- be "on" (i.e., positive in value) during some time interval, [-T,0].

This generates an ***initial pattern*** **of activities**, $x_i(0)$, i =1, 2, … n.

Study "reverberations," $\lim_{t \to \infty} x_i$ with inputs shut off.

Factorize Pattern and Total Activity

Method of proof: Change variables to:

pattern: $X_i = \dfrac{x_i}{x}$ **total activity:** $x = \sum_{k=1}^{n} x_k$

feedback signal: $f(w)$ $g(w) = \dfrac{f(w)}{w}$

Why $g(w)$**?**

"How nonlinear *IS* it?"

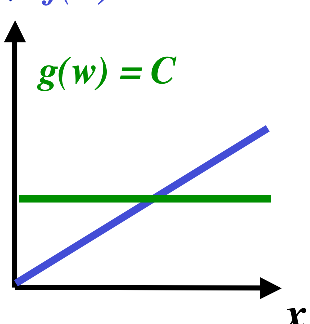
Shape of Nonlinear Feedback

feedback signal: $f(w)$ $g(w) = \dfrac{f(w)}{w}$

linear

e.g., $f(w) = Cw$

$g(w) = C$

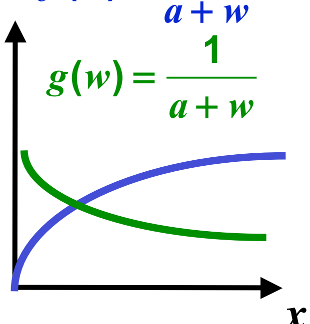


no advantage across size of x

slower than linear

$f(w) = \dfrac{w}{a+w}$

$g(w) = \dfrac{1}{a+w}$

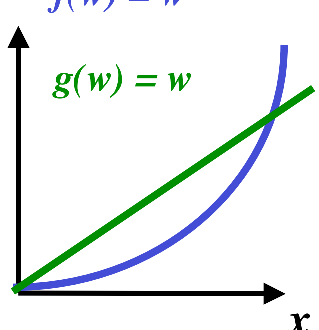


relatively **stronger** for **small** x

faster than linear

$f(w) = w^2$

$g(w) = w$

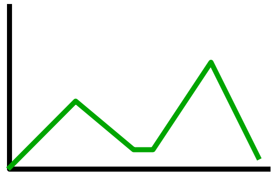


relatively **stronger** for **large** x

A Series of Global Theorems

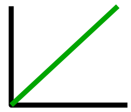
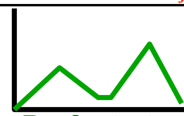
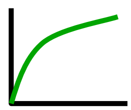
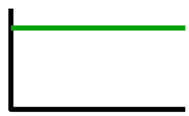
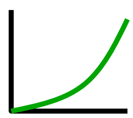
Grossberg, 1973, Studies in Applied Math

$x_i(0)$



Initial pattern i

Nonlinear

f	$X_i(\infty) = \frac{x_i(\infty)}{\Sigma_j x_j(\infty)}$	$x(\infty) = \Sigma_j x_j(\infty)$
 Linear	 ✓ Perfect storage of any pattern (∞)	Amplifies noise (or no storage) ✗
 Slower-than-linear	 ✗ Saturates	Amplifies noise ✗
 Faster-than-linear	 ? Chooses max Winner-take-all	Suppresses noise Normalizes total activity ✓

Factorize Pattern and Total Activity

Pattern variable equation:

$$\frac{d}{dt} X_i = B X_i \sum_{k=1}^{n} X_k \left[g(X_i x) - g(X_k x) \right]$$

Who wins the competition?

Total activity equation:

$$\frac{d}{dt} x = x \left[-A + (B - x) \sum_{k=1}^{n} X_k g(X_k x) \right]$$

Is my network stable?
How does it treat noise?

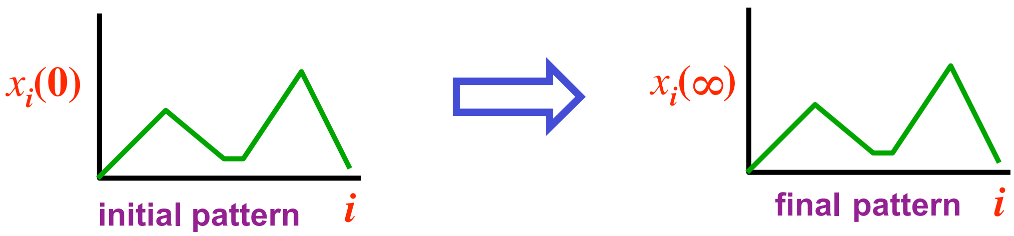
Pattern Transformation

Linear *f* perfectly stores any pattern

Pattern variable equation:

$$\frac{d}{dt}X_i = BX_i \sum_{k=1}^{n} X_k \left[g(X_i x) - g(X_k x) \right]$$

$$f(w) = Cw, \;\; g(w) = C, \;\; \frac{dX_i}{dx} = 0$$



Pattern Transformation

Faster-than-linear *f* makes a *choice*

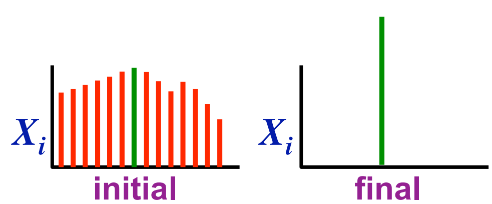
Pattern variable equation:

$$\frac{d}{dt}X_i = BX_i \sum_{k=1}^{n} X_k \left[g(X_i x) - g(X_k x) \right]$$

$$e.g.,\; f(w) = w^2, \;\; g(w) = w$$

$$X_i(0) > X_k(0),\; k \neq i \Rightarrow$$

$$\frac{dX_i}{dt}(t) > 0, \frac{dX_k}{dt}(t) < 0,\; k \neq i$$



largest GROWS; the rest CRASH

First network with
WINNER-TAKE-ALL!

Moral of the story: Keep track of **signs of derivatives!**

When is activity stored in short-term memory?

What happens to total activity x through time?

$x \Rightarrow 0$ no storage

$x \Rightarrow$ finite constant -- *storage*

$$\frac{d}{dt}x = x\left[-A + (B-x)\sum_{k=1}^{n} X_k g(X_k x)\right]$$

$$\frac{d}{dt}x = x(B-x)\left(G - \frac{A}{B-x}\right)$$

where

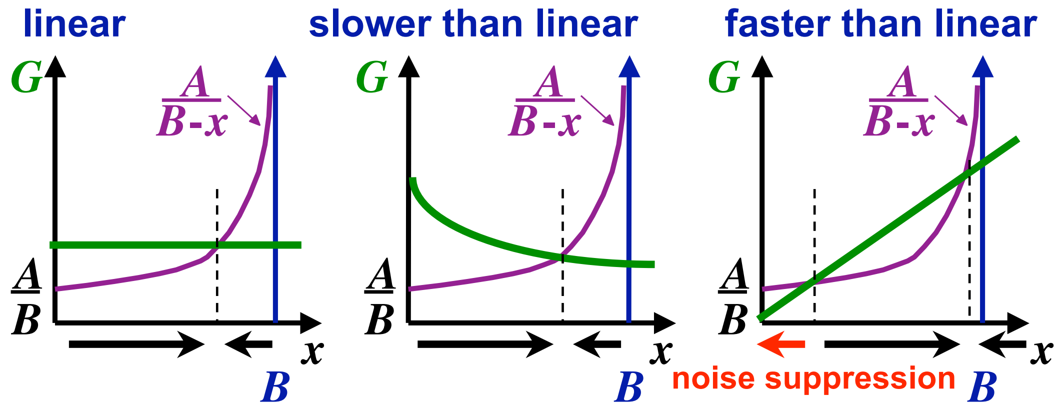
$$G = \sum_{k=1}^{n} X_k g(X_k x)$$

weighted average of $g(X_k x)$'s

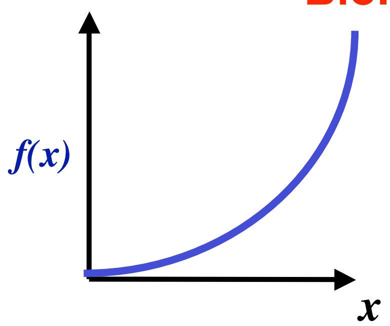
Short Term Memory : Noise Suppression or Quantized Storage

$$\frac{d}{dt}x = x(B-x)\left(G - \frac{A}{B-x}\right)$$

Sign of $\frac{d}{dt}x$ is the **sign of:** $G - \frac{A}{B-x}$



Biological Realism

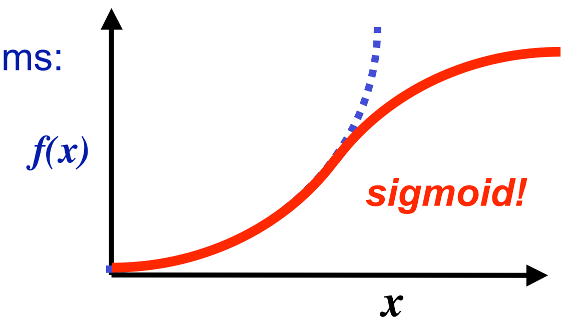


Faster-than-linear feedback signal function supports **noise suppression**

But, as $x \to \infty$, $f(x) \to \infty$
. . . not realistic

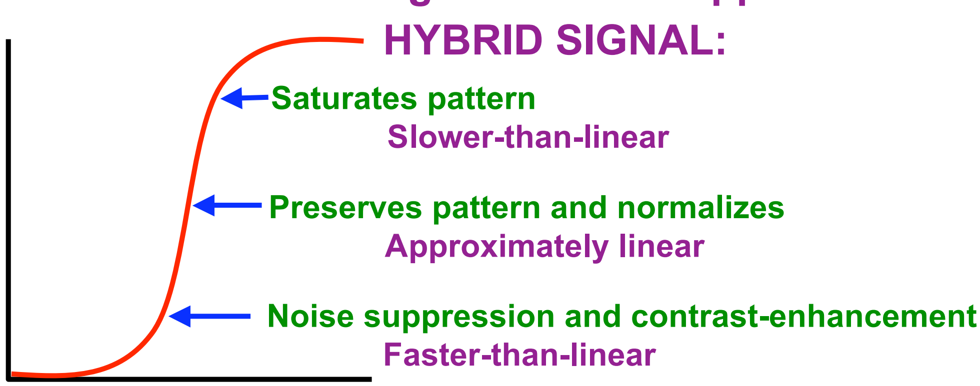
Winner-take-all **noise suppression is too severe**
Network **only stores one feature**

One change solves both problems:



Sigmoid Signal Function

Distributed Processing and Noise Suppression



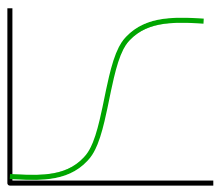
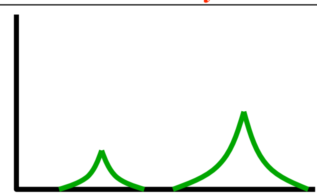
The faster-than-linear part suppresses noise and starts to contrast-enhance the pattern

As total activity normalizes, the approximately linear range is reached and tends to store the partially contrast-enhanced pattern

Sigmoid Signal Function

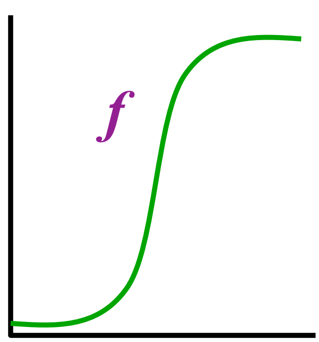
Distributed processing and noise suppression

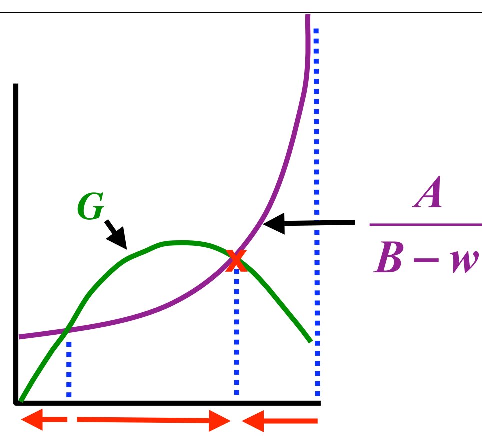
Quenching Threshold (QT)

f	$X_i(\infty)$	$x(\infty)$
 Sigmoid	 Tunable filter √	Suppresses noise √

The QT can be dynamically tuned; e.g., pay attention better after unexpected event; choose max…

Sigmoid Signal Function





One stable equilibrium point

Total activity normalization

Cf. "bubbles" in self-organizing feature maps -- Kohonen, 1984

CC Loop of *BCS* Built on Preceding Theorems

Feedback exists ***between cortical streams***

boundary grouping, completion, and ***filling-in***

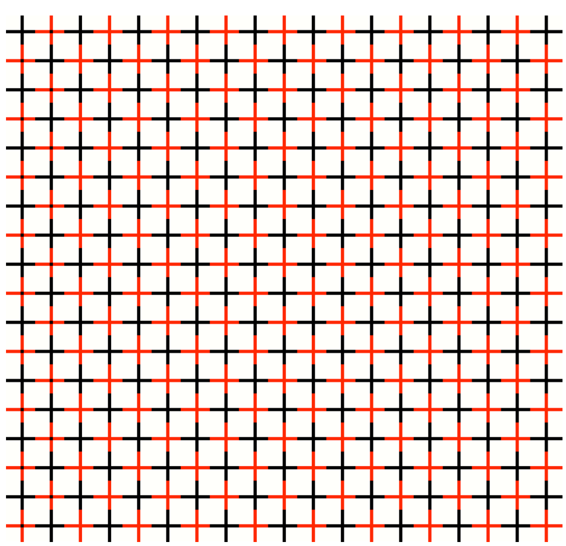
Visual processing is not conducted by:
- independent modules
- intrinsic images
- feature maps

Boundary strength is not the same as **lightness** or **color**

Next: Early model analysis of such issues

Neon Grid

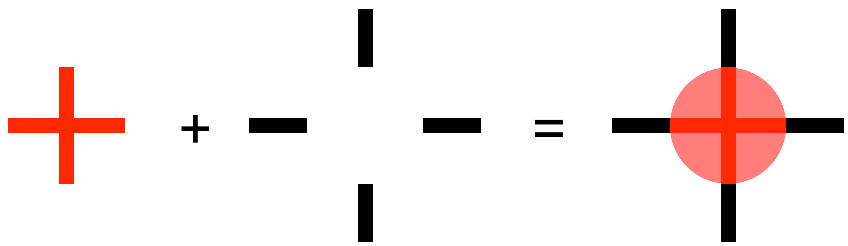
Visible evidence for how groupings form and contain color filling-in



Redies & Spillmann, 1981

Reality *vs* Illusion

BCS/FCS theory explains how:
a red cross placed inside an Ehrenstein figure



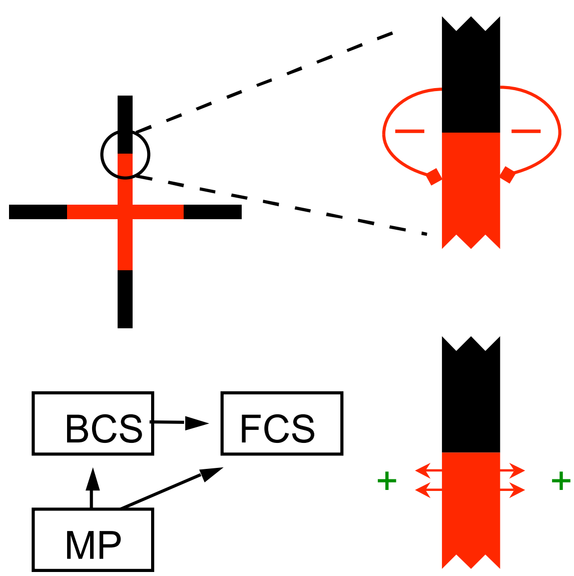
produces color spreading

Redies and Spillmann, 1981

"Real" contours of small cross cannot enclose red featural quality;

"Illusory" contours of Ehrenstein figure ***do!***

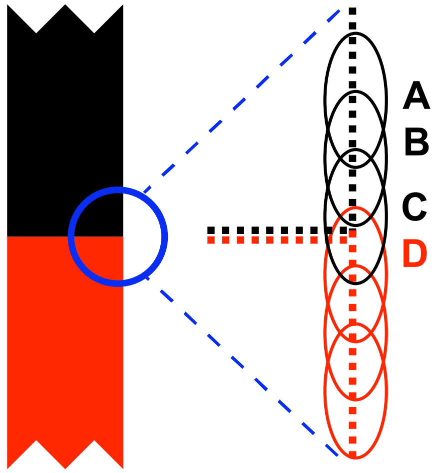
Why Does Color Spread?



BCS: inhibition
lower-contrast
boundary signals
are weakened

FCS: *no* inhibition
feature signals
survive and disperse

Relative Contrast with Background



If boundary of black line inhibits the boundary of the red, why doesn't the black boundary self-annihilate?

BCS's **First Competitive Stage**: ***shunting inhibition***

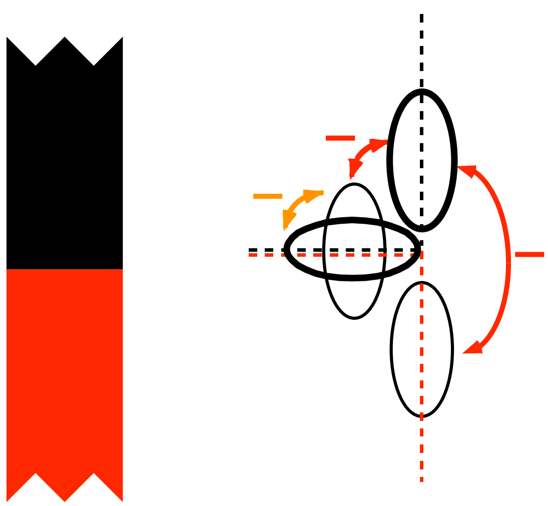
Divisive inhibition at **A** and **B** is **balanced.**

C inhibits D more due to **higher contrast with background.**

Strength of* neon effect varies with *amount of contrast.

van Tuijl & de Weert, 1979; Redies & Spillmann, 1981

Trapping the Escaping Color



1st and 2nd competitive stages

same orientation,
***across position* inhibition**

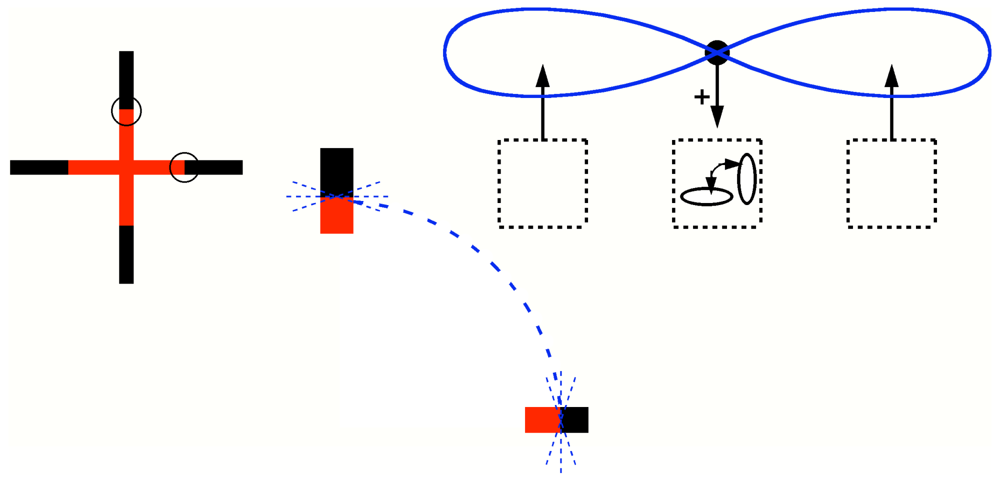
then

across orientation,
***same position* inhibition**

to generate ***end cuts***

enhanced horizontal boundary

Emergent Boundary Formation



The cooperative-competitive loop (CC Loop)

long-range cooperation and **short-range inhibition**
choose coherent boundaries and suppress alternatives

Transition to 3-D figure/ground

BCS/FCS theory was good for its time, but . . .

Neon color spreading and related phenomena raise issues of

transparency

3-D surface organization

figure/ground perception

and more . . .

THREE THEMES

How is grouping organized in the visual cortex?

A larger issue: How do the LAMINAR CIRCUITS
of visual cortex enable us to see?

How does the visual cortex carry out 3D vision?

stereopsis
planar 3D surface perception
curved and slanted 3D surface perception
bistable percepts and binocular rivalry
anchoring of surface lightness and color

How does the visual cortex separate figure from ground?

completion and recognition of partially occluded objects

transparency	Benary cross
3D neon color spreading	Kanizsa stratification
White's effect	Bregman-Kanizsa f-g separation

HOW IS GROUPING ORGANIZED IN THE VISUAL CORTEX?

Grouping is not a separate process

It interacts with several other processes in the brain's architecture for seeing

Study it as part of a larger issue:

HOW DOES THE CEREBRAL CORTEX WORK?

HOW DOES THE CEREBRAL CORTEX WORK?

It supports the highest levels of biological intelligence in all modalities

VISION, SPEECH, COGNITION, ACTION

Why does the cortex have LAYERS?

How does LAMINAR COMPUTING give rise to biological intelligence?

1. How does visual cortex stably DEVELOP and LEARN to optimize its structure to process different environments?

2. How does visual cortex GROUP distributed information?

3. How does top-down ATTENTION bias visual processing?

A recent breakthrough shows how 1 implies 2 and 3!

LAMINAR COMPUTING

A New Paradigm

Proposes how the cerebral cortex achieves:

Stable development

Stable learning throughout life

ANALOG COHERENCE
Coherently group distributed information without a loss of analog sensitivity (binding problem)
Hybid of digital and analog computing

Pay attention to important events

A synthesis of: **Bottom-up adaptive filtering**
Horizontal associative grouping
Top-down hypothesis testing and attention
in ALL of its processing stages

LAMINAR COMPUTING

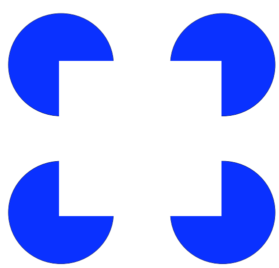
How does it compare with earlier BCS?

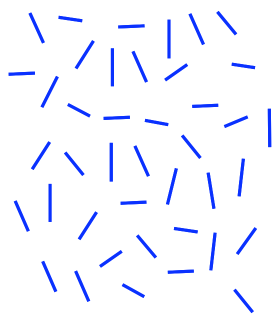
Uses similar combination of mechanisms:
properties and problems of old BCS forced
discovery of laminar model

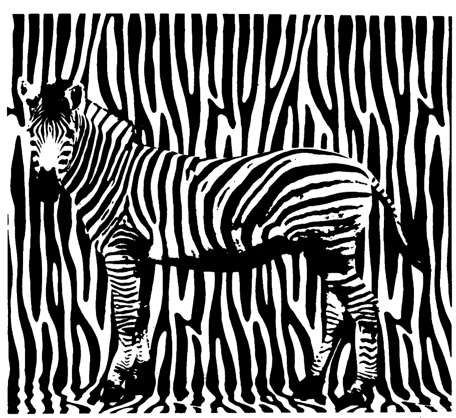
A much more ingenious, parsimonious, and beautiful
circuit

Can explain a MUCH larger data base!
unifies development
learning
grouping
attention
figure-ground perception…

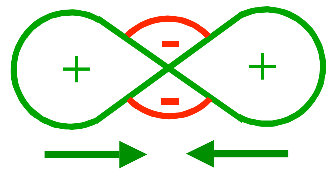
PERCEPTUAL GROUPING

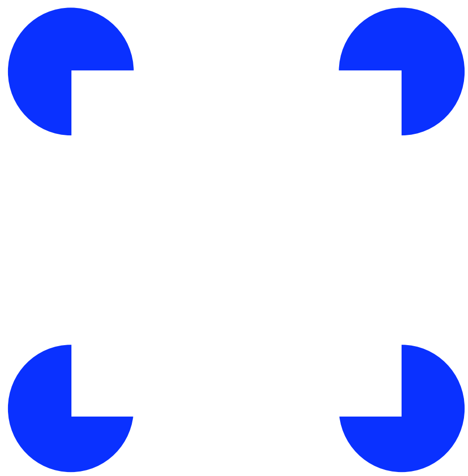






BIPOLE PROPERTY





Problems with old bipole:

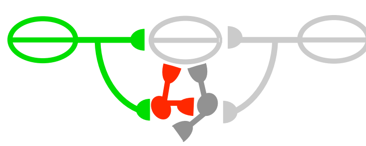
1. Inward selectivity of bipole

vs. outward horizontal signals in (e.g.) layer 2/3:

2. Hard to get groupings with analog sensitivity

Grossberg & Mingolla, 1985

LAMINAR BIPOLE PROPERTY



Long-range horizontal excitatory connections

Shorter-range disynaptic inhibitory connections

Input on just one side

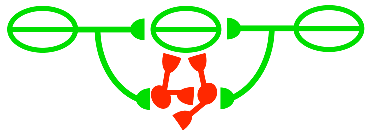
ONE-AGAINST-ONE:

Balanced Excitation and Inhibition

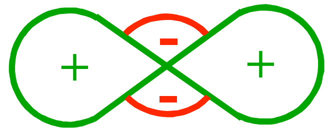
Cell is not excited

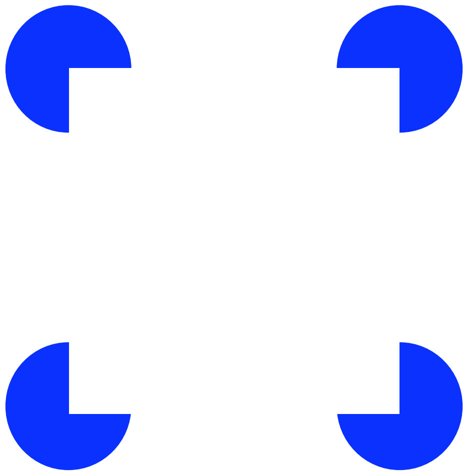
Grossberg, Mingolla & Ross, 1997

LAMINAR BIPOLE PROPERTY



vs.





Collinear input on both sides

Excitatory inputs summate

Inhibitory inputs normalize

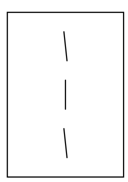
Shunting inhibition!

TWO-AGAINST-ONE

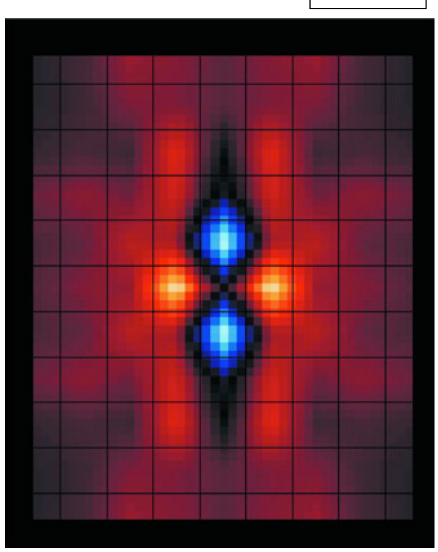
Cell is excited

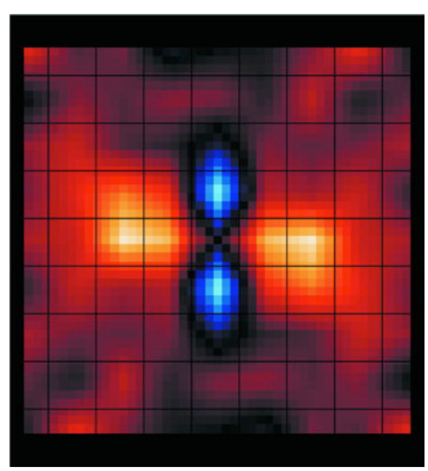
KAPADIA, ITO, GILBERT & WESTHEIMER (1995)

Psychophysics



Neurophysiology

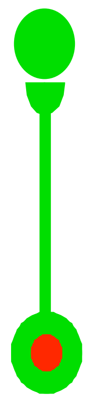




HOW ARE BIPOLE CELLS ACTIVATED?

DIRECT BOTTOM-UP ACTIVATION OF LAYER 4

V1
layer 4



LGN

**Strong bottom-up
LGN input to layer 4**

Stratford et al. (1996)
Chung & Ferster (1998)

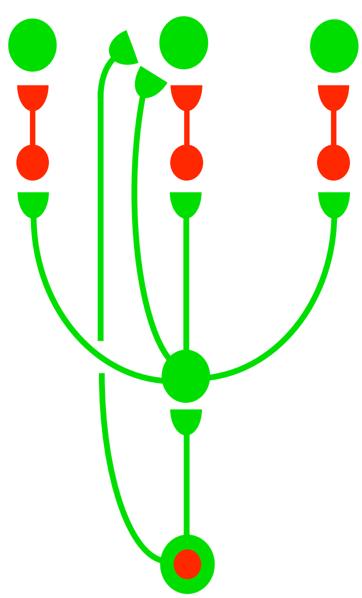
ANOTHER BOTTOM-UP INPUT TO LAYER 4: WHY?

LAYER 6-TO-4 ON-CENTER OFF-SURROUND

4

6

LGN



LGN projects to layers 6 and 4

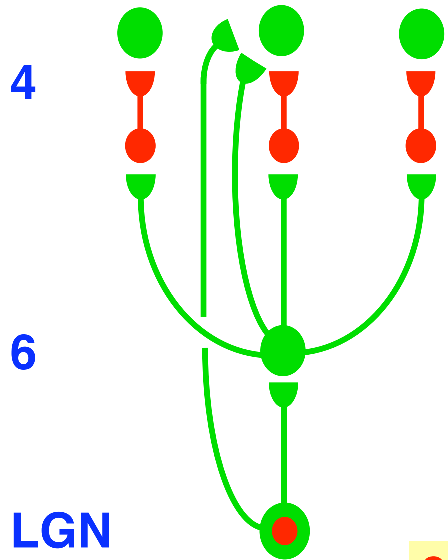
**Layer 6 excites spiny stellates
in column above it**

**Medium-range connections
onto inhibitory interneurons**

**6-to-4 path acts as
on-center off-surround**

Grieve & Sillito, 1991, 1995
Ahmed et al., 1994, 1997

BOTTOM-UP CONTRAST NORMALIZATION



Together, direct LGN-to-4 path and 6-to-4 on-center off-surround provide contrast normalization

Grossberg, 1973
Heeger, 1992
Douglas et al., 1995

SHUNTING
on-center off-surround

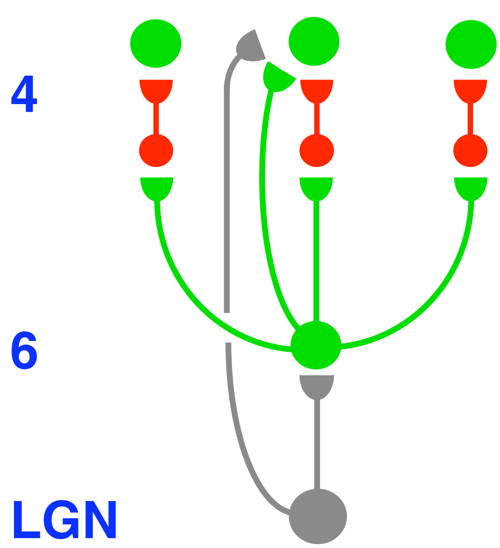
Spatial competition: cf. old BCS

Do not discuss oriented RFs; discuss new circuit ideas

MODULATION OR PRIMING BY 6-TO-4 ON-CENTER

On-center 6-to-4 excitation is inhibited down to being modulatory (priming, subthreshold)

Stratford et. al, 1996
Callaway, 1998



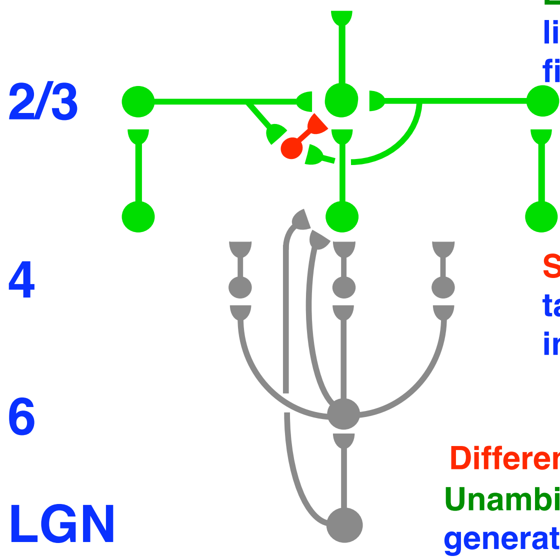
On-center 6-to-4 excitation cannot activate layer 4 on its own

Plays key role in stable development and learning

Need direct LGN-to-4 path to drive cortical activation

GROUPING STARTS IN LAYER 2/3

Bipole Property!



Long-range horizontal excitation links collinear, coaxial receptive fields

Gilbert & Wiesel, 1989
Bosking et al., 1997
Schmidt et al, 1997

Short-range disynaptic inhibition of target pyramidal via pool of interneurons

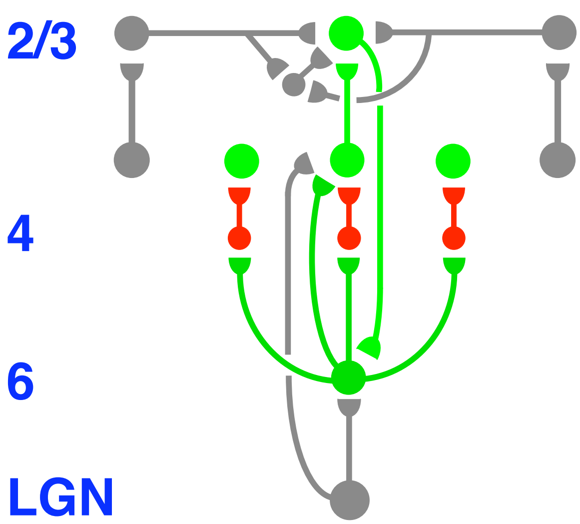
Hirsch & Gilbert, 1991

Difference with old BCS:
Unambiguous groupings can form and generate feedforward outputs quickly

Thorpe et al, 1996

Orientational competition: cf. old BCS

HOW IS THE FINAL GROUPING SELECTED?
FOLDED FEEDBACK



Groupings in layer 2/3 feed back into 6-to-4 on-center off-surround

Can also go via layer 5

Blasdel et al., 1985
Kisvarday et al., 1989

Strongest grouping enhanced by its on-center

Inputs to weaker groupings suppressed by off-surround

Interlaminar feedback creates functional columns

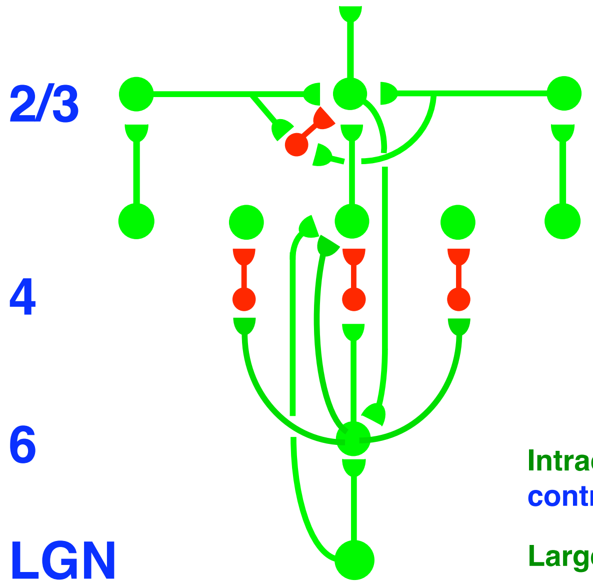
An application of theorems about recurrent shunting on-center off-surround networks!

A BRAIN WITHOUT BAYES

Real-time Decision Making under Uncertainty

A Hybrid of Feedforward and Feedback Processing

A Self-Organizing System that Trades Certainty Against Speed



Rapid feedforward processing when data are unambiguous

Activities of conflicting groupings are reduced by self-normalizing inhibition: Ambiguous processing slows down

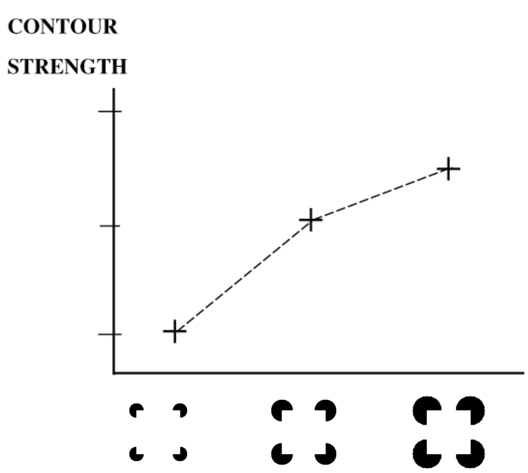
Self-normalizing inhibition creates real-time normalized activity distributions ("probabilities") that reflect system uncertainty

Intracortical feedback selects and contrast-enhances a winning grouping

Large activity speeds up processing of unambigous winning grouping

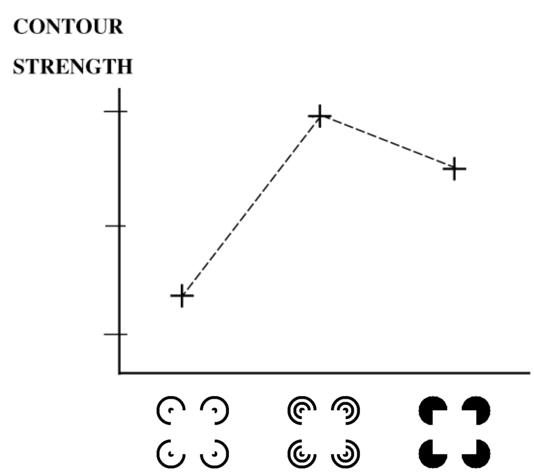
When can correct answer catch up to ambigous one? cf. speed/accuracy tradeoff

ANALOG-SENSITIVE BOUNDARY COMPLETION



Increases with *"support ratio"*

Shipley and Kellman, 1992

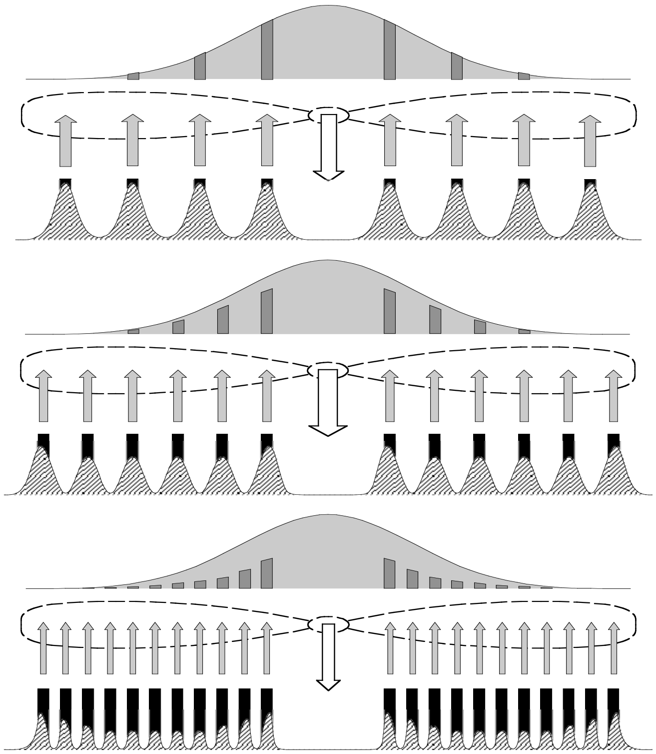


Inverted-U

Lesher and Mingolla, 1993

cf. Soriano, Spillmann and Bach, 1994 (shifted gratings)

COOPERATION *AND COMPETITION*

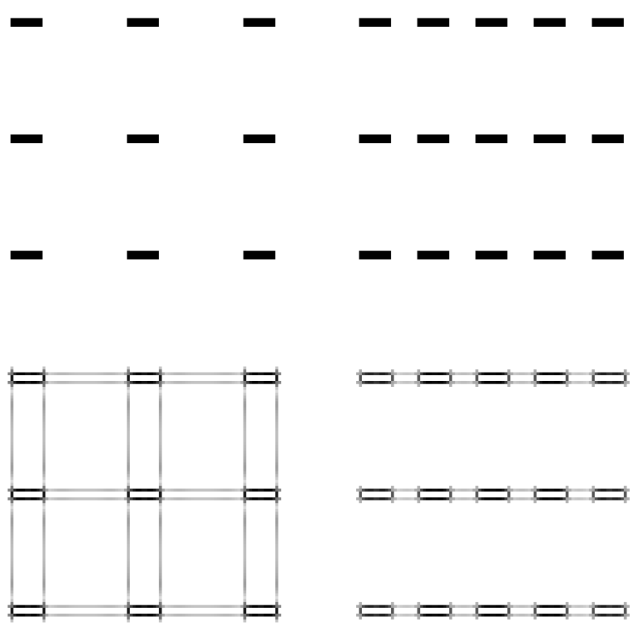


few lines,
wide spacing

more lines
overcome slight
inhibition from
neighbors

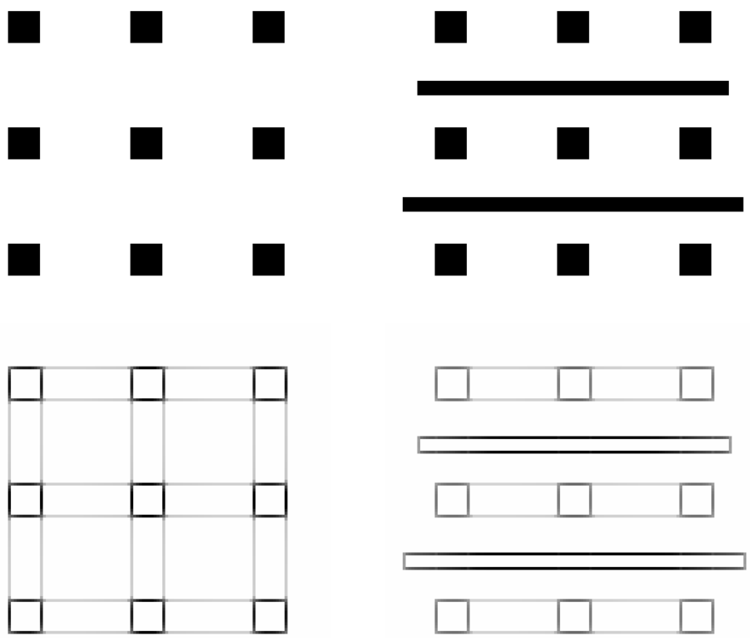
crowding lowers
overall effective
input to cooperation

GESTALT GROUPING SIMULATION



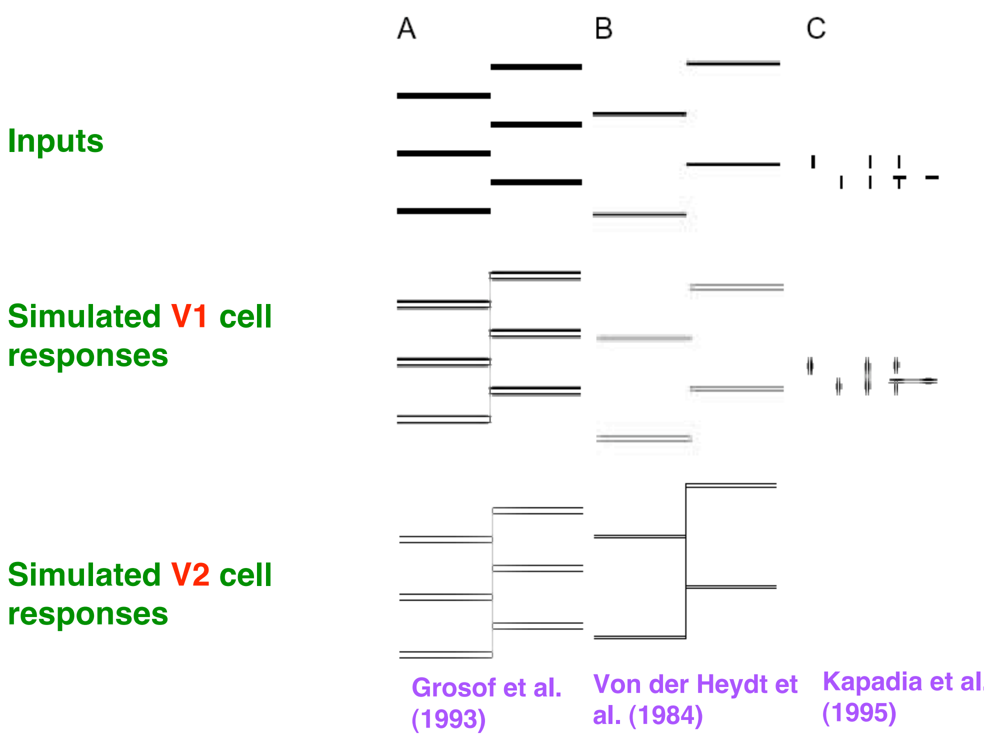
Proximity: cooperation strengthens horizontal grouping
competition breaks vertical grouping

GESTALT GROUPING SIMULATION

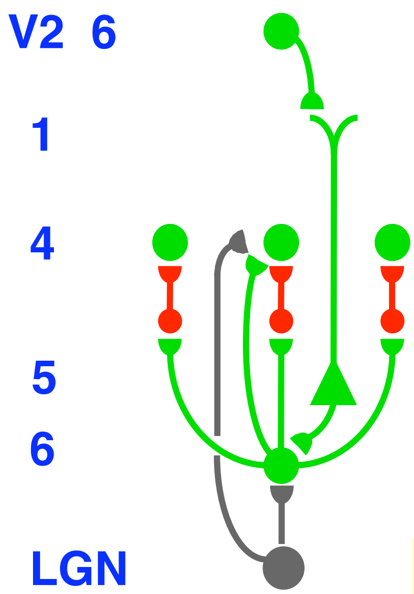


Good Continuation: competition breaks vertical groupings

GROUPING SIMULATIONS: V1 AND V2



HOW DOES TOP-DOWN ATTENTION FIT IN?
FOLDED FEEDBACK AGAIN



Attentional signals also feed back into 6-to-4 on-center off-surround

1-to-5-to-6 feedback path
- Macaque: Lund & Boothe, 1975
- Cat: Gilbert & Wiesel, 1979

DATA: V2-to-V1 feedback is on-center off-surround and affects layer 6 of V1 the most
- Bullier et al., 1996
- Sandell & Schiller, 1982

Attended stimuli enhanced
Ignored stimuli suppressed

Attention acts via a TOP-DOWN MODULATORY ON-CENTER OFF-SURROUND NETWORK

WHY IS THE MODEL CALLED LAMINART?

LAMINART = LAMINAR ART

ART = ADAPTIVE RESONANCE THEORY

Grossberg (1976, 1980), Carpenter and Grossberg (1987),…

ART is a perceptual and cognitive theory that proposes how stable development and learning occur throughout life using top-down attention

ART predicted in the 1980's that attention is realized by a top-down modulatory on-center off-surround network!

Such a network helps to dynamically stabilize learning

SUPPORT FOR ART PREDICTIONS

ATTENTION HAS AN ON-CENTER OFF-SURROUND

Bullier, Jupe, James, and Girard, 1996
Caputo and Guerra, 1998
Downing, 1988
Mounts, 2000
Reynolds, Chelazzi, and Desimone, 1999
Smith, Singh, and Greenlee, 2000
Somers, Dale, Seiffert, and Tootell, 1999
Sillito, Jones, Gerstein, and West, 1994
Steinman, Steinman, and Lehmkuhne, 1995
Vanduffell, Tootell, and Orban, 2000

"BIASED COMPETITION"

Desimone, 1998
Kastner and Ungerleider, 2001

SUPPORT FOR ART PREDICTIONS

ATTENTION CAN FACILITATE MATCHED BOTTOM-UP SIGNALS

Hupe, James, Girard, and Bullier, 1997
Luck, Chellazi, Hillyard, and Desimone, 1997
Roelfsema, Lamme, and Spekreijse, 1998
Sillito, Jones, Gerstein, and West, 1994
and many more…

INCONSISTENT WITH MODELS WHERE TOP-DOWN MATCH IS SUPPRESSIVE

Mumford, 1992
Rao and Ballard, 1999: Bayesian Explaining Away

SUPPORT FOR ART PREDICTIONS

LINK BETWEEN ATTENTION AND LEARNING

VISUAL PERCEPTUAL LEARNING
Ahissar and Hochstein, 1993

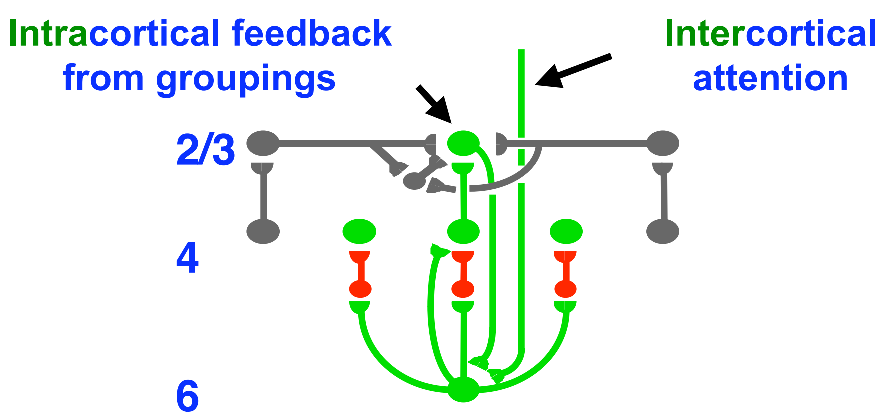
AUDITORY LEARNING
Gao and Suga, 1998

SOMATOSENSORY LEARNING
Krupa, Ghazanfar, and Nicolelis, 1999
Parker and Dostrovsky, 1999

Also clarifies Watanabe et al (2002+) data on perceptual learning without attention (use intracortical feedback)

GROUPING AND ATTENTION SHARE DECISION CIRCUIT

The preattentive grouping is its own "attentional" prime!

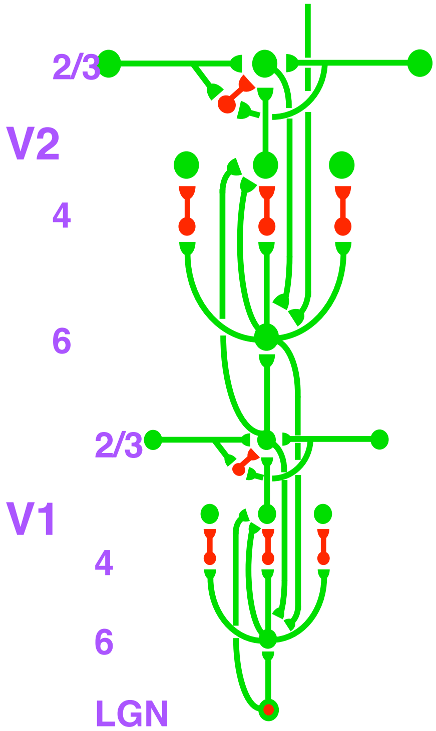


Attention acts via a
TOP-DOWN
MODULATORY ON-CENTER
OFF-SURROUND NETWORK

Why so many debates about pre-attentive and attentive processing?

They share a decision circuit!

V2 REPEATS V1 CIRCUITRY AT LARGER SPATIAL SCALE



V2 layer 2/3 horizontal axons longer-range than in V1

Amir et al. (1993)

Therefore, longer-range groupings can form in V2

Von der Heydt et al. (1984)

WHAT IS THE RELATIONSHIP BETWEEN GROUPING AND ATTENTION?

Attention and perceptual grouping coexist in the same cortical areas

Both processes have many shared properties

But they obey seemingly contradictory constraints

SHARED PROPERTIES OF ATTENTION AND GROUPING

ENHANCEMENT of weak, near-threshold stimuli

Attention: Reynolds et al., 1996; Hupe et al., 1998
Grouping: Kapadia et al., 1995; Polat et al., 1998

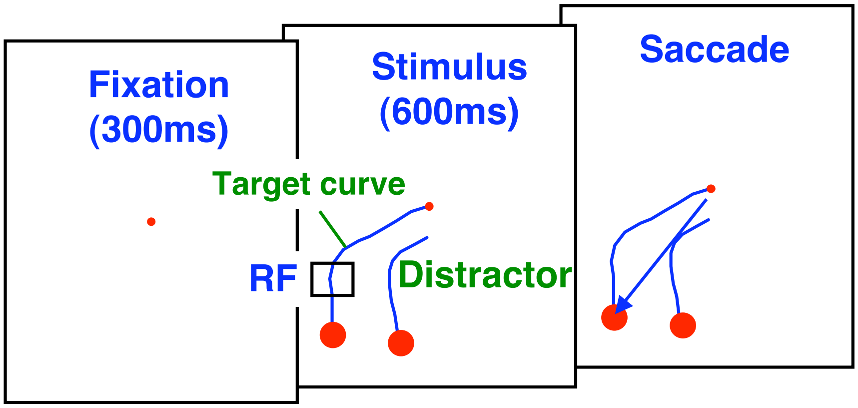
SUPPRESSION of competing stimuli / rival groupings

Attention: Luck et al., 1994; Caputo & Guerra, 1998
Grouping: van Lier et al., 1997; Kubovy et al., 1998

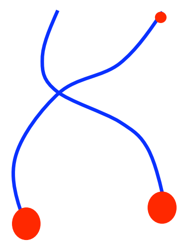
HOW CAN ATTENTION SELECT A WHOLE OBJECT?

Attention and grouping share a decision circuit!

ATTENTION FLOWS ALONG CURVES: ROELFSEMA ET AL. (1998): MACAQUE V1

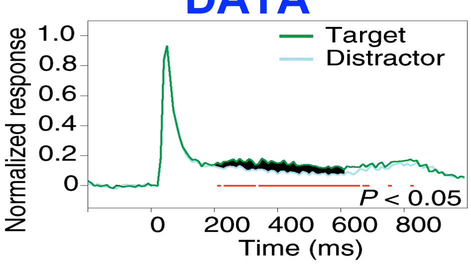


Crossed-curve condition:
Attention flows across junction between smoothly connected curve segments (Good Continuation)

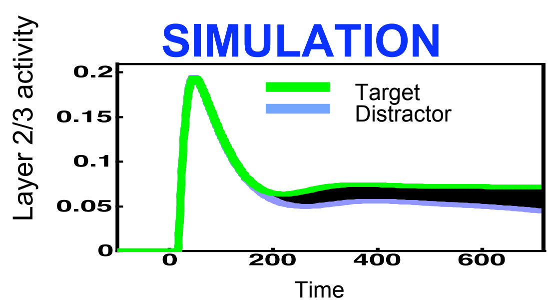


SIMULATION OF ROELFSEMA ET AL. (1998)

DATA



SIMULATION

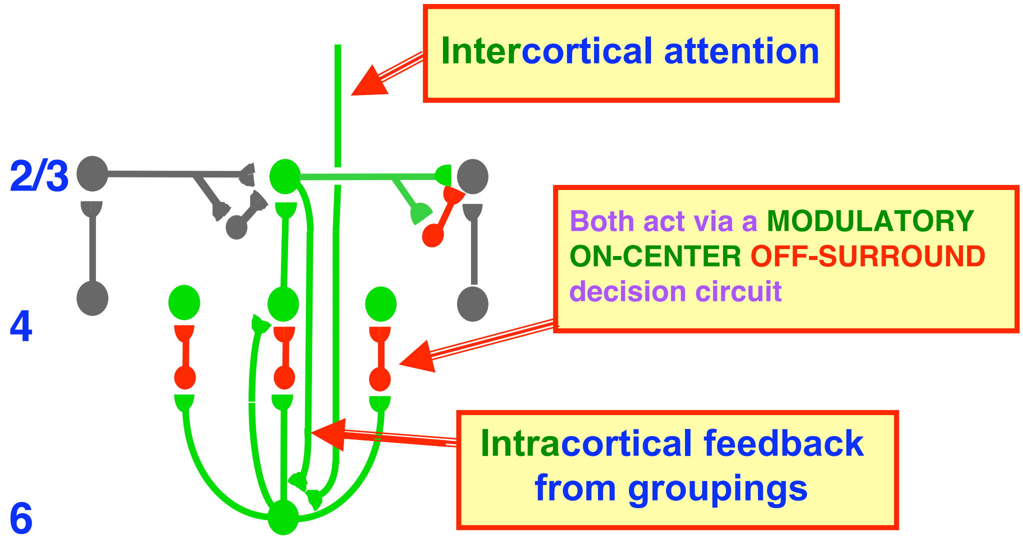


Attention directed only to far end of curve

Propagates along active layer 2/3 grouping to distal neurons

Grossberg and Raizada (2000, Vision Research)

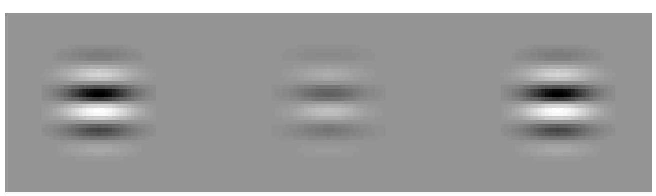
EXPLANATION: GROUPING AND ATTENTION SHARE THE SAME MODULATORY DECISION CIRCUIT



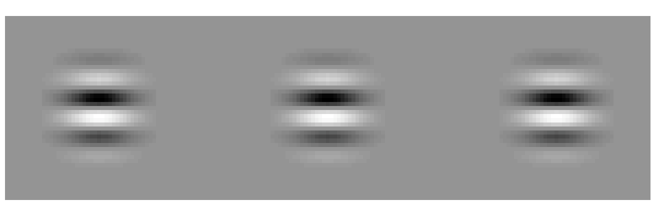
POLAT ET AL. (1998): CAT AREA 17 (V1) CONTRAST-SENSITIVE GROUPING

TARGET: Variable-contrast Gabor in neuron's Classical RF
FLANKERS: Constant-contrast collinear Gabors outside RF

Collinear flankers ENHANCE response to near-threshold target



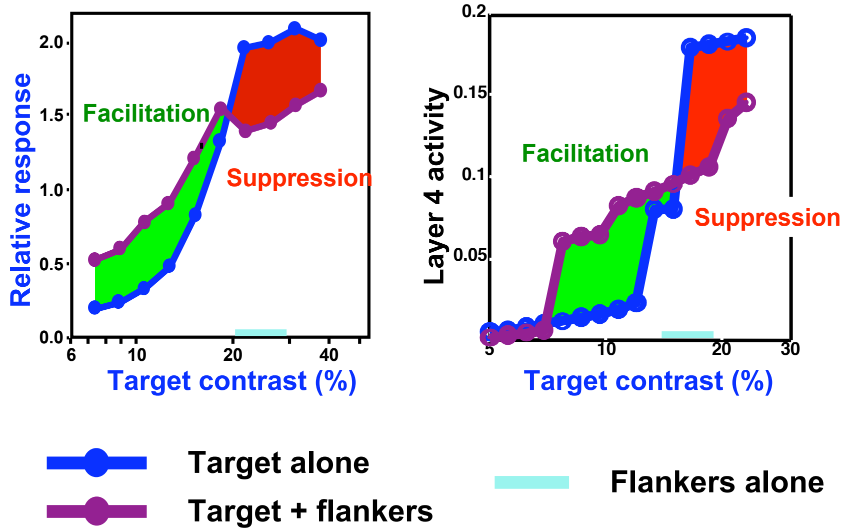
Flankers SUPPRESS response to high contrast target



SIMULATION OF POLAT ET AL. (1998)

Depends on Shunting Inhibition of Layer 6

DATA **SIMULATION**



SEEMINGLY CONTRADICTORY CONSTRAINTS ON ATTENTION AND GROUPING RESOLVED

Attention cannot produce above-threshold activity where there is no bottom-up visual input

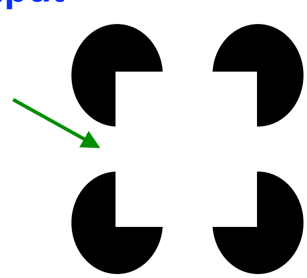
Prime to see a yellow ball

Do not hallucinate seeing a yellow ball

Modulatory on-center

Grouping can produce above-threshold activity where there is no bottom-up visual input

Illusory contour seen here, but no bottom-up contrastive input



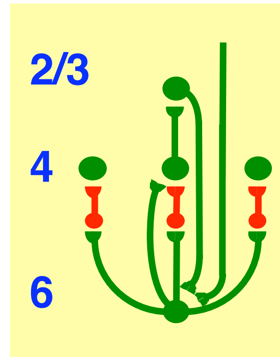
Groupings can form in layer 2/3

Needs the layers; not in old BCS!

WHAT DOES LAMINAR COMPUTING ACHIEVE?

1. SELF-STABILIZING DEVELOPMENT AND LEARNING

2. Seamless fusion of
   PRE-ATTENTIVE AUTOMATIC
   BOTTOM-UP PROCESSING
   and
   ATTENTIVE TASK-SELECTIVE
   TOP-DOWN PROCESSING



3. ANALOG COHERENCE: Solution of the BINDING PROBLEM without a loss of analog sensitivity

Even the earliest cortical stages carry out active adaptive information processing:
LEARNING, GROUPING, ATTENTION

LAMINAR COMPUTING: A NEW WAY TO COMPUTE

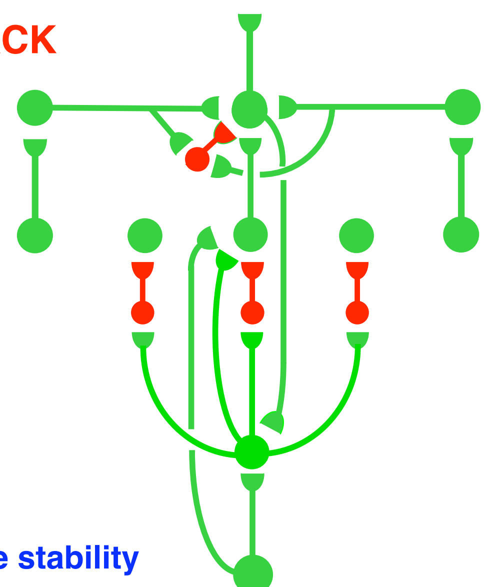
1. FEEDFORWARD AND FEEDBACK

Rapid feedforward processing when data are unambiguous

Feedback chooses among ambiguous alternatives:
self-normalizing competition

A self-organizing system that trades certainty against speed

cf., Bayesian models



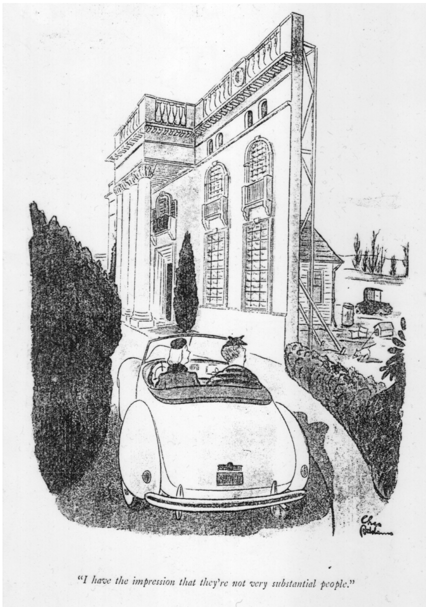
2. ANALOG AND DIGITAL

ANALOG COHERENCE combines the stability of digital with the sensitivity of analog

3. PRE-ATTENTIVE AND ATTENTIVE LEARNING

A pre-attentive grouping is its own “attentional” prime!

3D VISION AND FIGURE-GROUND PERCEPTION



How are
3D BOUNDARIES
and
3D SURFACES
formed?

Form
And
Color
And
DEpth theory

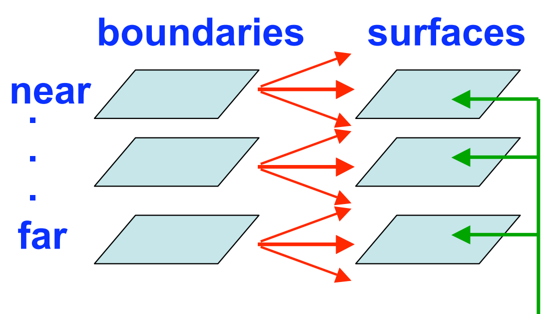
How the world looks so real without assuming naïve realism

Grossberg (1987, 1994, 1997)

Prediction: Visible figure-ground-separated Form-And-Color-And-DEpth are represented in cortical area V4

3D SURFACE FILLING-IN

From filling-in of surface
LIGHTNESS and COLOR
to filling-in of surface
DEPTH



Prediction: Depth-selective boundary-gated filling-in defines the 3D surfaces that we see

Prediction: A single process fills-in lightness, color, and depth

Can a change in brightness cause a change in depth? YES!
e.g., proximity-luminance covariance
Egusa (1983), Schwartz & Sperling (1983)

Why is depth not more unstable when lighting changes?
Prediction: Discounting the illuminant limits variability

STEREOGRAM SIMULATION:
SURFACE LIGHTNESSES ARE SEGREGATED IN DEPTH

Fang & Grossberg (2004, 2005; see poster #577 on Saturday)

Left input

Right input

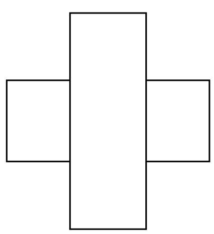
Far plane

Fixation plane

Near plane

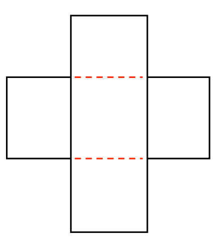
Cf. algorithms that just compute disparity matches and let computer code build the surface; e.g., Marr & Poggio (1974) et al

FIGURE-GROUND SEPARATION
AND AMODAL COMPLETION



Why are 2D pictures often perceived as 3D representations of occluding and occluded surfaces?

Why is completion of the horizontal boundary amodal?



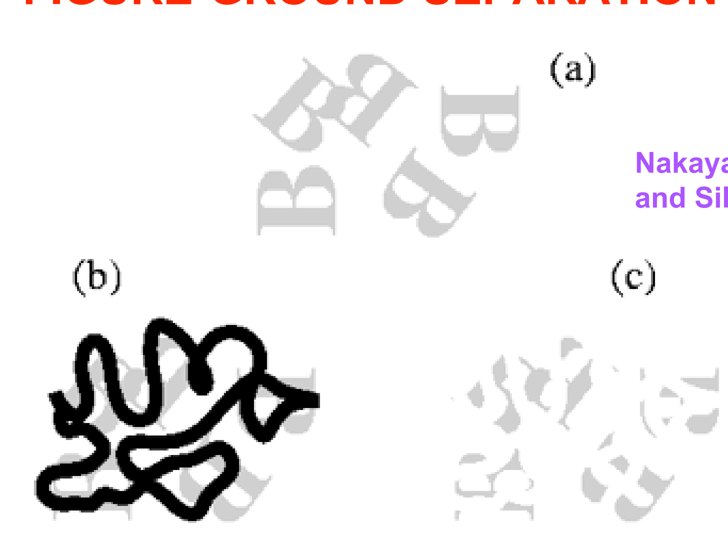
Easy! ALL boundaries are invisible!

Amodal boundary completion helps to recognize partially occluded objects

Hard: Why we see only unoccluded parts of partially occluded opaque surfaces

Hard because this is not always true: cf., transparent surfaces

BREGMAN-KANIZSA FIGURE-GROUND SEPARATION



Nakayama, Shimojo, and Silverman (1988)

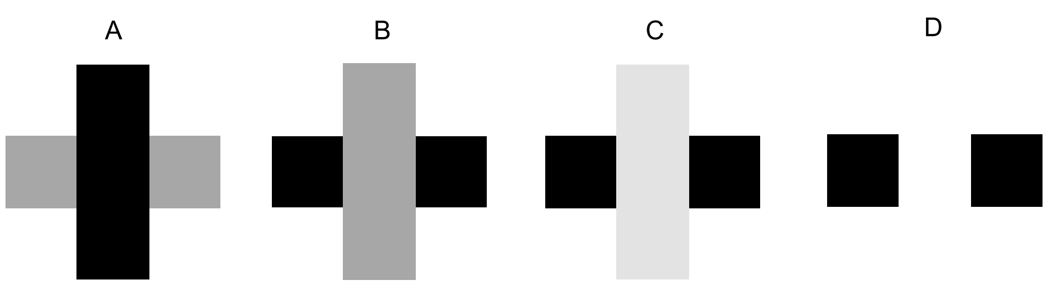
Black occluder helps to recognize gray B's because shared black/gray boundaries "belong" to black occluder:

Extrinsic vs. intrinsic boundaries

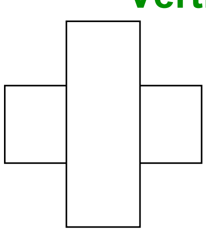
INTERACTION OF GEOMETRY AND CONTRAST

Opaque Surfaces

Depth perception can depend on contrast



Vertical near **Horizontal near**



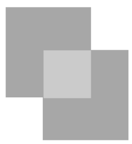
The same geometry in all cases

INTERACTION OF GEOMETRY AND CONTRAST

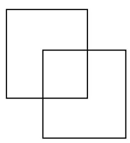
Transparent Surfaces

Unique transparency

Bistable transparency



No transparency



The same geometry in all cases

HOW SMART IS BRAIN EVOLUTION?

How can evolution discover a process as subtle as figure-ground perception of occluding and occluded objects? …of opaque vs. transparent objects?

Prediction:
Solution of simpler problems imply figure-ground properties

CONSISTENCY IMPLIES FIGURE-GROUND SEPARATION!

I. BOUNDARY-SURFACE COMPLEMENTARITY
versus
BOUNDARY-SURFACE CONSISTENCY

We SEE one unified percept!

II. FIGURE-GROUND RECOGNITION
versus
VISIBLE SURFACE PERCEPTION

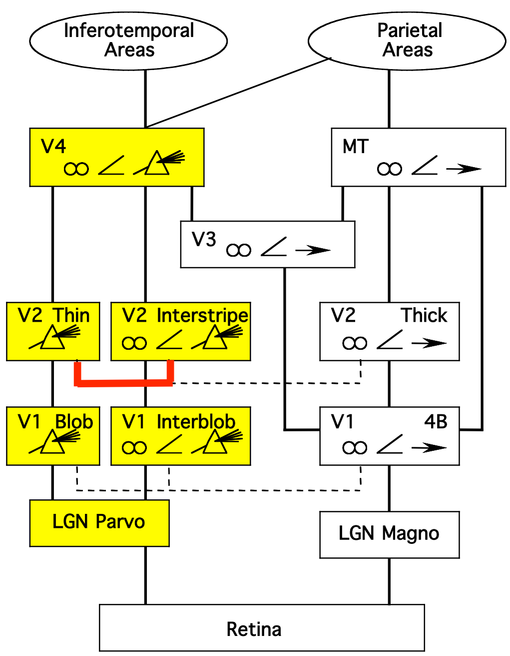
How do we RECOGNIZE a partially OCCLUDED object?

Why do we NOT SEE partially OCCLUDED object parts when the occluder is OPAQUE?

Why do not all OCCLUDING objects look TRANSPARENT?

The same process handles both I and II!

INTERSTREAM FEEDBACK ENSURES CONSISTENCY



Prediction:
Feedback between V2 boundary and surface streams ensures consistency and initiates figure-ground separation

What sort of feedback?!

DeYoe and Van Essen, 1988,
Trends in Neurosciences, 11, 219-226

HOW DOES THE CORTEX DO BINOCULAR VISION?

Most models consider only V1 stereopsis
e.g., disparity energy model

Most models do not explain 3D SURFACE PERCEPTS

Most models do not include CORTICAL LAYERS

Can the LAMINART model be self-consistently extended?

YES!

3D LAMINART MODEL

Grossberg and Howe (2003); Grossberg and Swaminathan (2004);
Cao and Grossberg (2005): Grossberg and Yazdanbakhsh (2005)

Unifies and further develops

LAMINART model of development, learning, grouping, and attention
Grossberg, Mingolla, Raizada, Ross, Sietz, Williamson

FACADE model of 3D vision and figure-ground perception
Grossberg, Grunewald, Kelly, McLoughlin, Pessoa

It shows how interactions between V1, V2, and V4 can explain many data about 3D vision

3D LAMINART SIMULATIONS

Contrast variations of dichoptic masking (McKee et al., 1994)

Correspondence Problem (Smallman & Mckee, 1995)

Panum's limiting case (Gillam et al., 1995; McKee et al., 1995)

Venetian blind illusion (Howard & Rogers, 1995)

Stereopsis with polarity-reversed stereograms (Nakayama & Shimojo, 1990)

Venetian blind illusion (Howard & Rogers, 1995)

Da Vinci stereopsis (Nakayama & Shimojo, 1990; Gillam et al., 1999)

Craik-O'Brian-Cornsweet lightness illusion (Todorovic, 1987)

The effect of interocular contrast differences on stereothresholds (Schor & Heckman, 1989)

Closure relationships and variations of Da Vinci stereopsis (Cao & Grossberg, 2004, 2005)

Simulate properties of:

3D perception of slanted and curved surfaces and bistable Necker cube (Grossberg & Swaminathan, 2004)

3D surface percepts of dense and sparse stereograms (Fang & Grossberg, 2005; VSS poster #577 on Saturday at 2-7 PM)

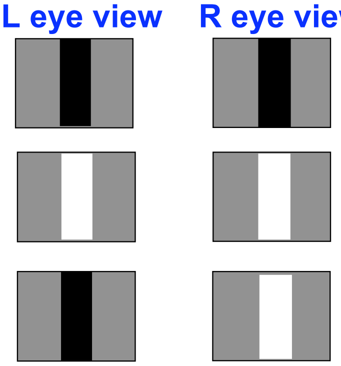
3D transparency, neon color spreading, and stratification (Grossberg & Yazdanbakhsh, 2005)

Binocular rivalry (Yazdanbakhsh & Grossberg, 2005; VSS talk on Wednesday at 8:30 AM)

HOW TO UNIFY CONTRAST-SPECIFIC BINOCULAR FUSION WITH CONTRAST-INVARIANT BOUNDARY PERCEPTION?

Contrast-specific binocular fusion

L eye view R eye view

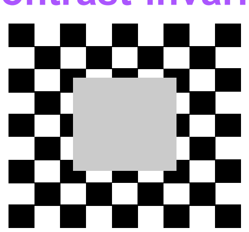


Binocular fusion

Binocular fusion

No binocular fusion

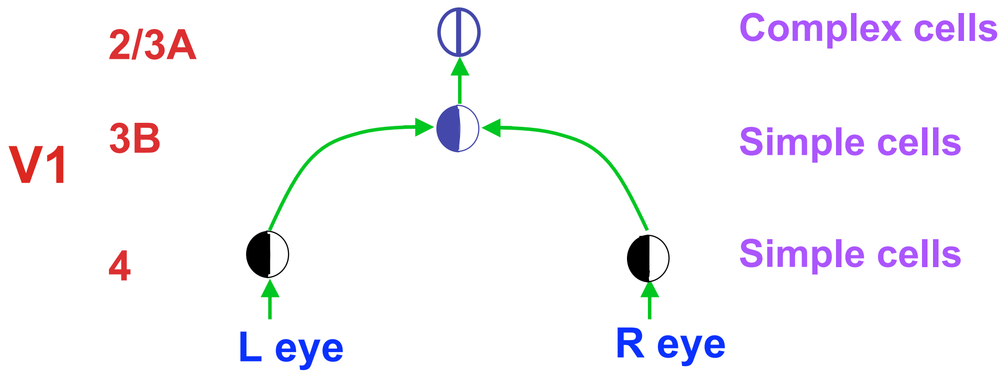
Contrast-invariant boundary perception



Contrast polarity along the gray square edge reverses

Opposite polarities are pooled to form object boundary

MODEL UNIFIES CONTRAST-SPECIFIC BINOCULAR FUSION WITH CONTRAST-INVARIANT BOUNDARY PERCEPTION



Contrast-specific stereoscopic fusion by disparity-selective simple cells

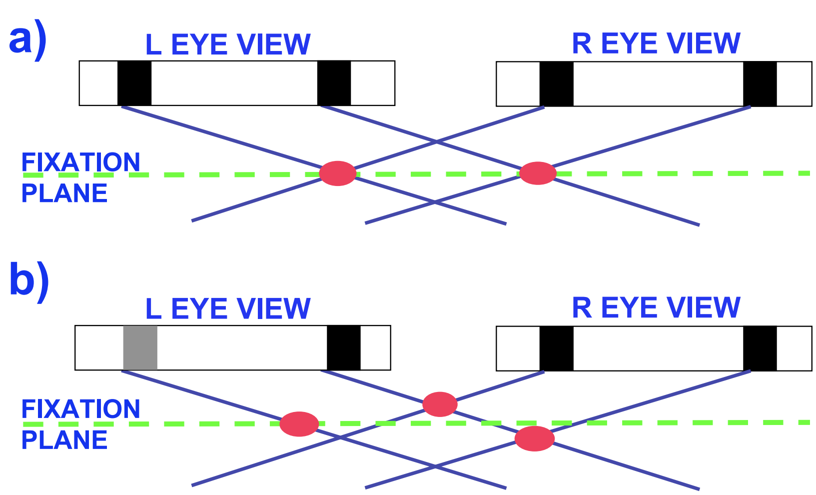
Contrast-invariant boundaries by pooling opposite polarity binocular simple cells at complex cells in layer 2/3A

Ohzawa et al,. 1990; Grossberg & McLoughlin, 1997

CONTRAST CONSTRAINT ON BINOCULAR FUSION

Left and right input from same object has similar contrast

Percept changes when one contrast is different:

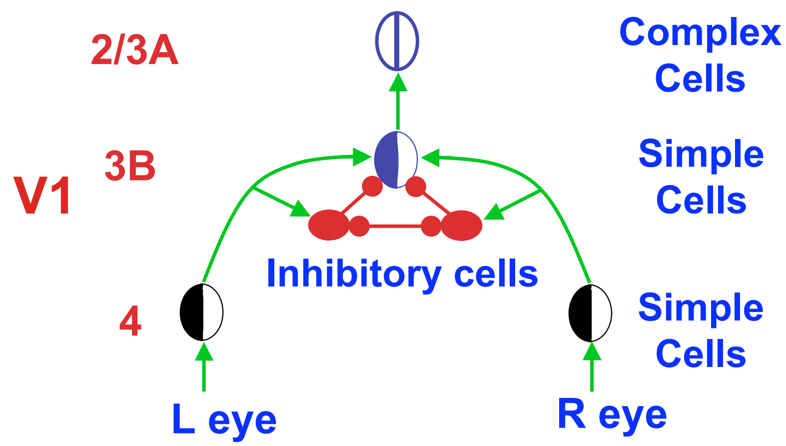


Fusion only occurs between bars of similar contrast

McKee et al., 1994

MODEL IMPLEMENTS CONTRAST CONSTRAINT ON BINOCULAR FUSION

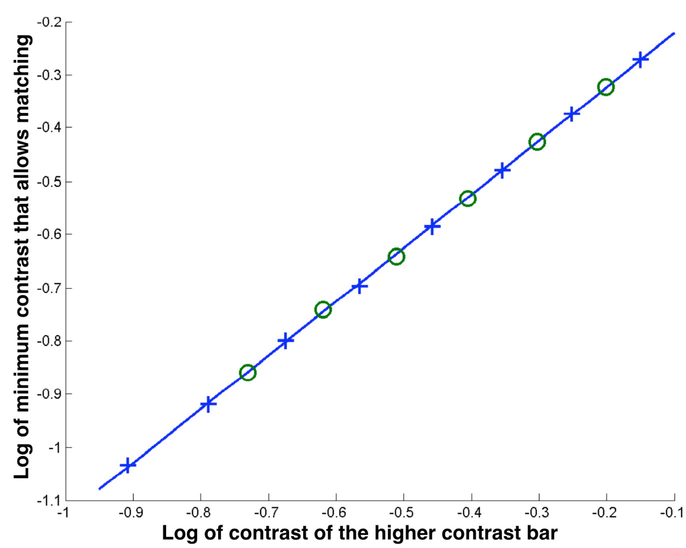
An Ecological Constraint on Cortical Development



Inhibitory cells (red) ensure that fusion occurs when contrasts in left and right eye are approximately equal (cf. "obligate" cells Poggio, 1991).

RATIO CONSTRAINT ON BINOCULAR FUSION

Smallman and McKee (1995)



Data: line of best fit has a slope of 1

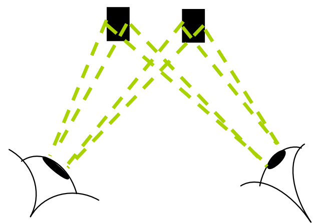
Simulation: + and o are model simulations

HOW TO SOLVE THE CORRESPONDENCE PROBLEM?

How does the brain inhibit false matches?

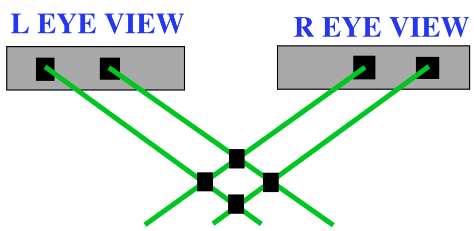
Contrast constraint is not enough

a)



Stimulus

b)

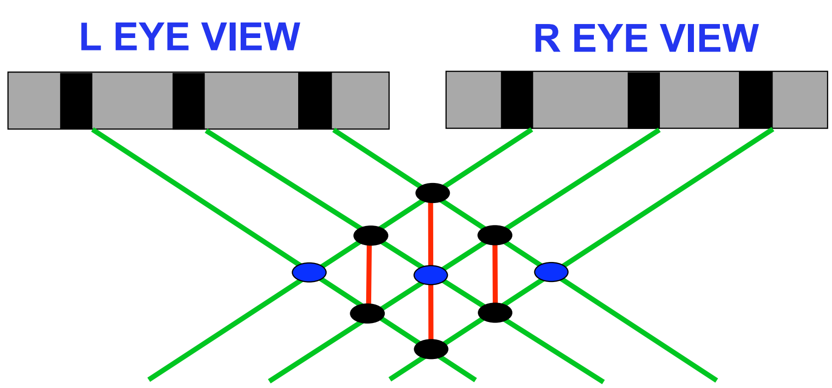


Multiple possible binocular matches

Which squares in the two retinal images must be fused to form the correct percept?

MODEL V2 DISPARITY FILTER SOLVES THE CORRESPONDENCE PROBLEM

An Ecological Constraint on Cortical Development

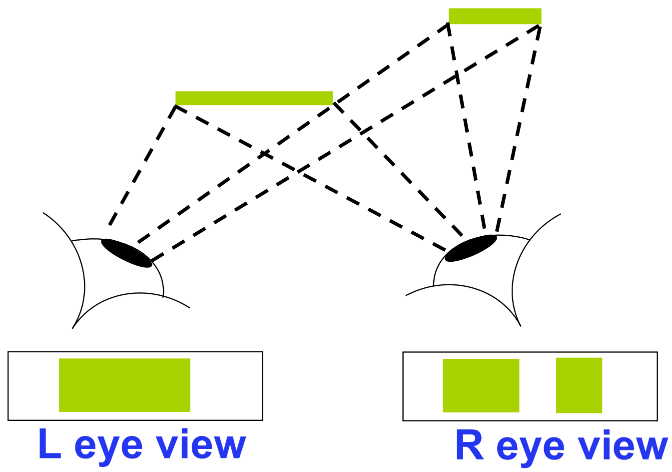


False matches (black) suppressed by line-of-sight inhibition (green lines) and cyclopean inhibition (red lines)

"Cells that fire together wire together"

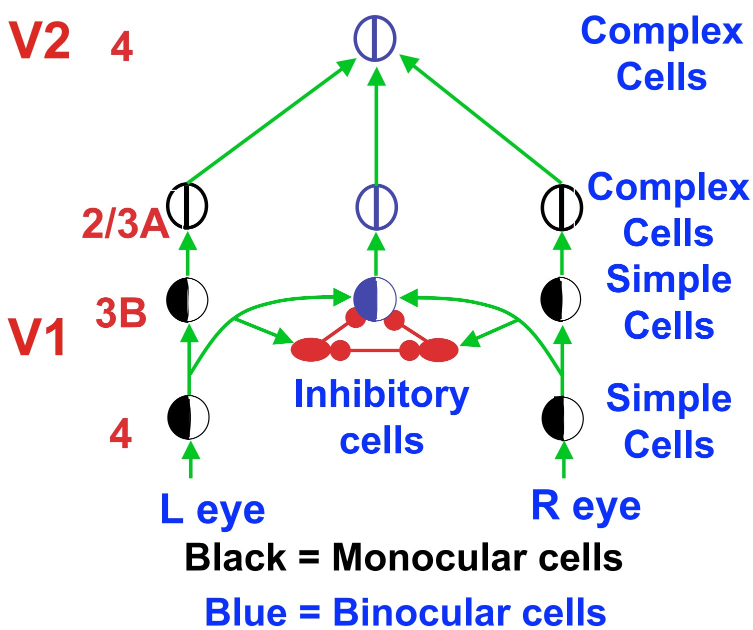
HOW DOES MONOCULAR INFORMATION CONTRIBUTE TO DEPTH PERCEPTION?

DaVinci Stereopis



Only by utilizing monocular information can visual system create correct depth percept (Gillam et al,. 1999)

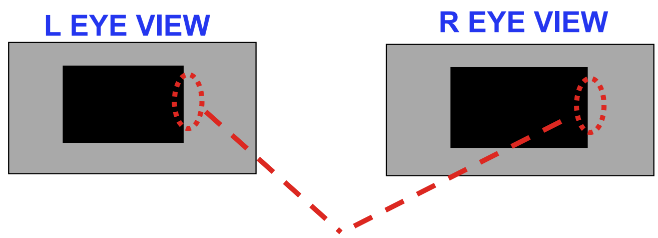
MODEL UTILIZES MONOCULAR INFORMATION



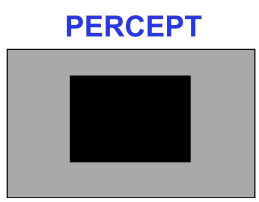
In V2, monocular inputs add to binocular inputs and contribute to depth perception

HOW TO FORM SURFACE PERCEPTS?

a) Neurons accomplish disparity sensitivity by matching edges
e.g. Cumming & DeAngelis, 2001

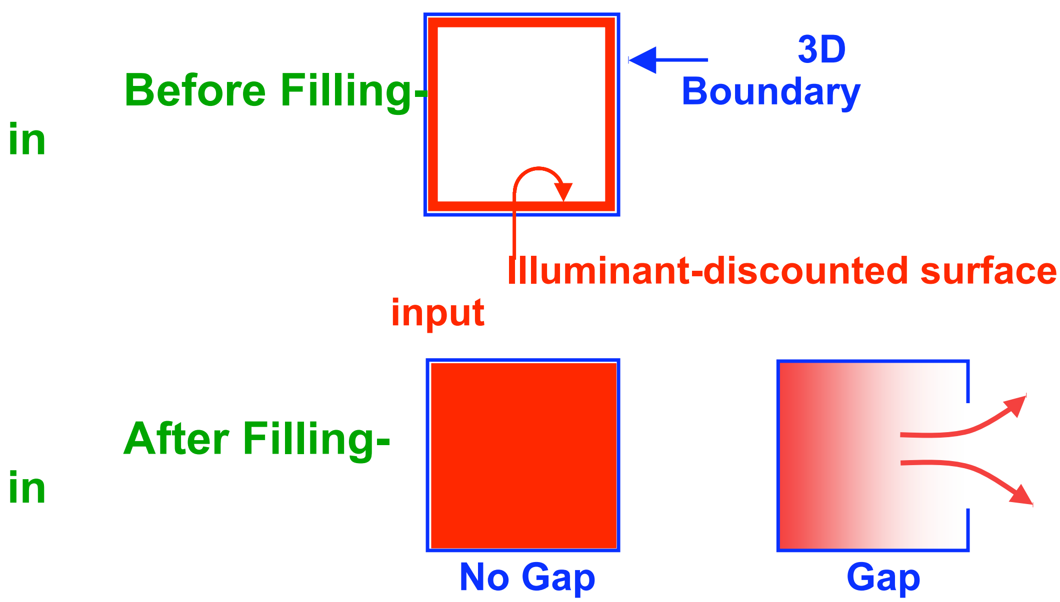


b) Why then do we see entire surfaces, not just edges?



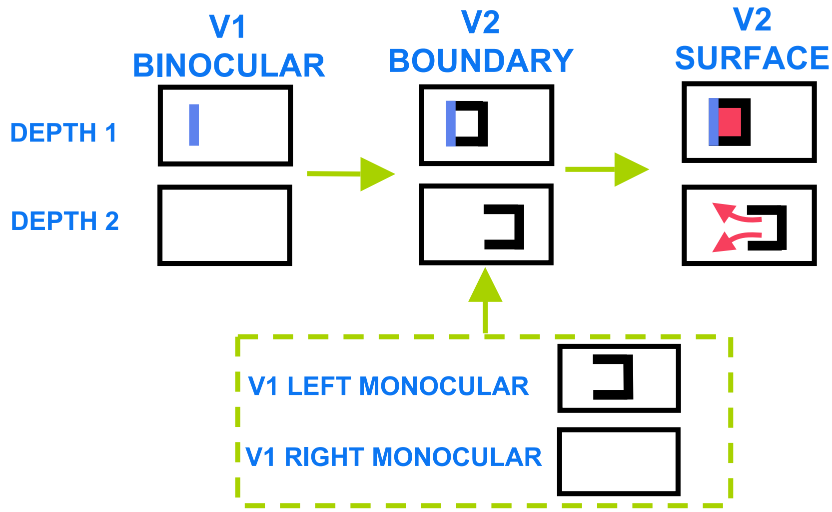
3D boundary-gated surface filling-in

CLOSED BOUNDARIES SURROUND VISIBLE SURFACE REGIONS



Cf. role of closed 2D boundaries in explaining COCE
Grossberg & Todorovic (1988)

3D BOUNDARY-GATED SURFACE FILLING-IN



Prediction: Monocular boundaries are added to ALL binocular boundaries

Regions that are surrounded by a CLOSED boundary can depth-selectively contain filling-in of lightness and color signals

CONNECTED VS BROKEN BOUNDARIES

Helps to explain lots of data

Stereopsis and 3D surface perception

3D figure-ground separation

Transparency

3D neon color spreading

Experimental test of this prediction:

e.g., Yazdanbakhsh and Watanabe, 2004

Confirmed asymmetric interaction of horizontal boundaries and depth-selective vertical boundaries

GROSSBERG & HOWE (2003) 3D LAMINART MODEL

Fill-in visible 3D surface within connected boundaries

Inhibit false matches

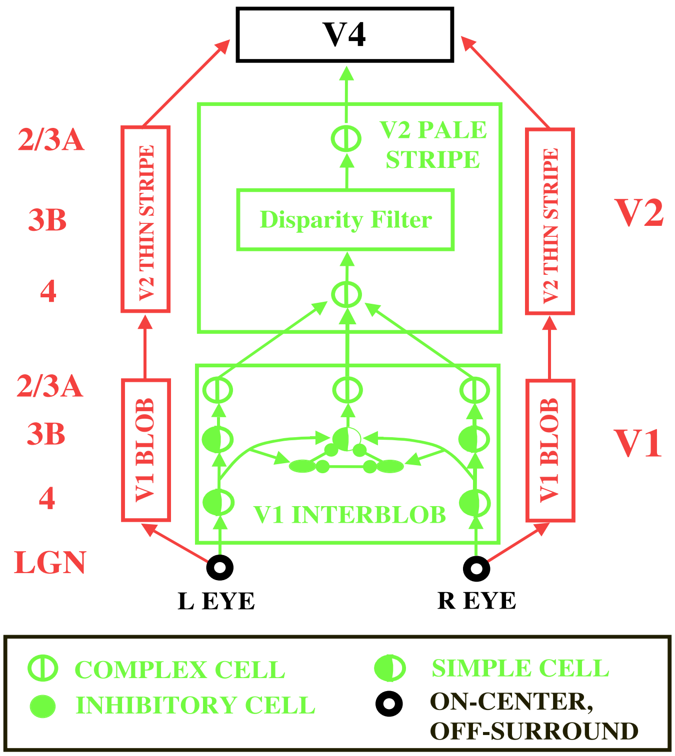
Pool binocular and monocular cells

Polarity-pooling complex cells: Monocular, binocular

Polarity-sensitive simple cells: Monocular, binocular

Polarity-sensitive monocular simple cells

Discount illuminant



SUPPORTING ANATOMICAL AND PHYSIOLOGICAL DATA

LGN: Has circularly symmetric receptive fields (Kandel et al, 2000), parvocellular, but not magnocellular component, critical for fine stereopsis (Shiller et al 1990a,b)

V1 in general: V1 interblob regions more concerned with orientation (i.e. form) information whereas V1 blob regions more concerned with color (Livingstone & Hubel, 1984). V1 contains "obligate" cells that respond to binocular, but not to monocular, simulation (Poggio 1991)

V1 Layer 4: Major recipient of the LGN parvocellular input, mainly monocular, outputs to layer 3B, but not to layer 2/3A (Callaway, 1998), contains simple cells (Hubel & Wiesel, 1968; Schiller et al., 1976)

V1 Layer 3B: Contains simple cells (Dow, 1974), monocular and binocular cells (Hubel & Wiesel, 1968; Poggio, 1972), inputs independent of ocular dominance (Katz et al., 1989), projects to 2/3A (Callaway, 1998)

V1 Layer 2/3A: Contains monocular and binocular cells (Poggio, 1972), many complex cells (Hubel & Wiesel, 1968; Poggio, 1972)

SUPPORTING ANATOMICAL AND PHYSIOLOGICAL DATA

V2 in general: Binocular (Hubel & Livingstone, 1987; Mausell & Newsome, 1987; Roe & Ts'o, 1997), disparity-sensitive (Poggio and Fischer, 1977; von der Heydt et al., 2000), fewer false matches in V2 than in V1 (Bakin et al, 2000)

V2 Pale stripes: Receives projections from V1 interblob but few from V1 blob regions (Livingstone & Hubel, 1984; Roe & Ts'o, 1997), particularly into layer 4 (Rockland & Virga, 1990), orientation selective (Peterhans, 1997; Roe & Ts'o, 1997), contains complex cells (Hubel & Livingstone, 1987), layer 2/3A projects to V4 (Xiao et al., 1999), contains a complete map of visual space (Roe & Ts'o, 1995), highly sensitive to orientation information (Peterhans, 1997)

V2 Thin stripes: Receives input from V1 blob but little from V1 interblob regions (Livingstone & Hubel, 1984; Roe & Ts'o, 1997), highly sensitive to color information (Peterhans, 1997), contains a complete map of visual space (Roe & Ts'o, 1995)

V4: Receives input from V2 pale stripes (Xiao et al., 1999) and V2 thin stripes (Mausell & Newsome, 1987; Xiao et al., 1999), and is disparity selective (Ghose & Ts'o, 1997)

22 SIMULATIONS WITH ONE SET OF PARAMETERS

Grossberg and Howe (2003)

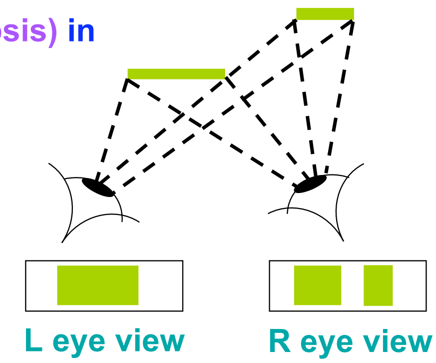
- Contrast variations of dichoptic masking (McKee et al., 1994)
- Correspondence Problem (Smallman & Mckee, 1995)
- Panum's limiting case (Gillam et al., 1995; McKee et al., 1995)
- Venetian blind illusion (Howard & Rogers, 1995)
- Stereopsis with polarity-reversed stereograms (Nakayama & Shimojo, 1990)
- Venetian blind illusion (Howard & Rogers, 1995)
- Da Vinci stereopsis (Nakayama & Shimojo, 1990; Gillam et al., 1999)
- Craik-O'Brian-Cornsweet lightness illusion (Todorovic, 1987)
- Effect of interocular contrast differences on stereothresholds (Schor & Heckman, 1989)

Illustrate model by explaining some DaVinci stereopsis percepts

ECOLOGICAL OPTICS HYPOTHESIS

Nakayama & Shimojo (1990)

N & S CLAIM: Visual systems interpret unpaired image points (DaVinci stereopsis) in terms of previous experiences with OCCLUSION RELATIONSHIPS



Cf. claim that visual STATISTICS influence what we see; e.g., Bayesian approaches to vision

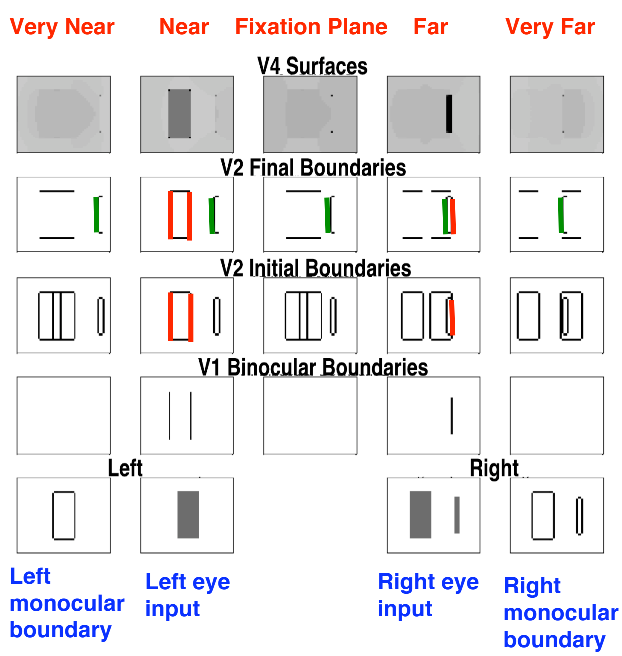
Image statistics clearly influence development of cortical maps and RFs; e.g., Wiesel and Hubel et al.

ECOLOGICAL OPTICS COUNTEREXAMPLES:
Simulate key DaVinci stereopsis percepts without explicit knowledge of occlusion relationships. However, line-of-sight inhibition and disparity-tuned complex cells develop with guidance from visual statistics

DA VINCI STEREOPSIS

Nakayama and Shimojo (1990)

An emergent property of the previous simple mechanisms working together



3D surface percept Not just disparity match!

Line-of-sight inhibition kills weaker vertical boundaries

Add monocular boundaries along lines-of-sight

Binocular match: boundaries of thick bar

Filling-in contained by connected boundaries

Vertical boundaries from monocular left edge of thin bar survive

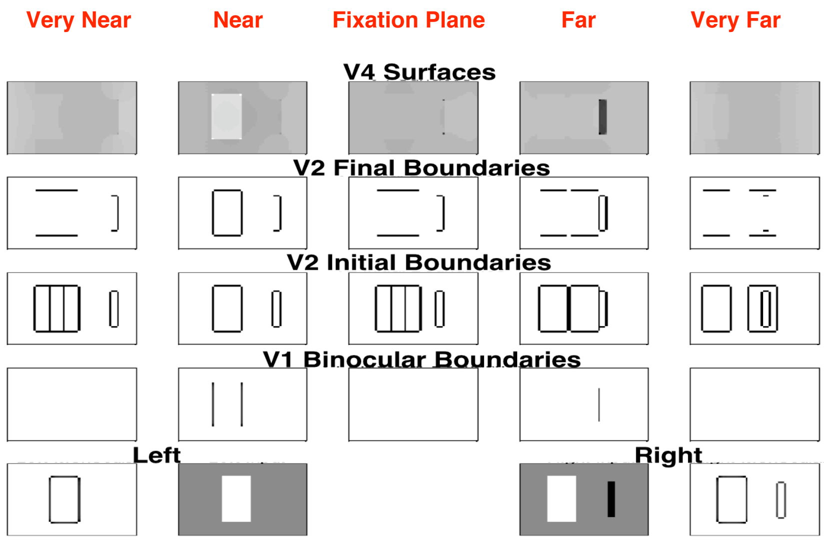
Strongest boundaries: binocular and monocular boundaries add

Binocular match: Right edge of thin and thick bars

POLARITY-REVERSED DA VINCI STEREOPSIS

Nakayama and Shimojo (1990)

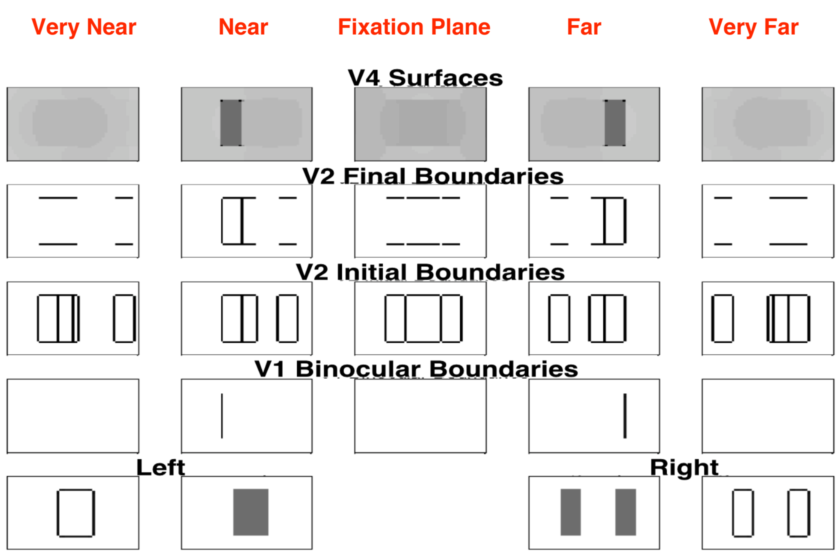
Same Explanation!



DA VINCI STEREOPSIS

Gillam, Blackburn, and Nakayama (1999)

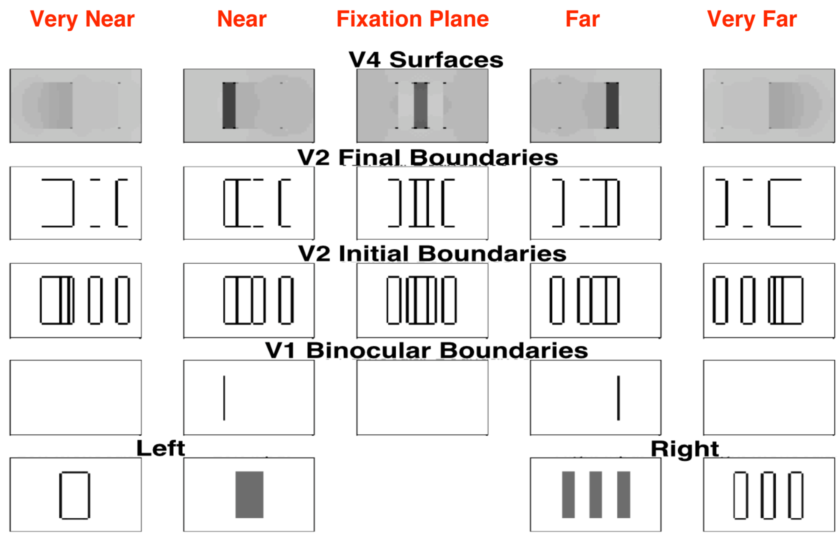
Same Explanation!



DA VINCI STEREOPSIS

Gillam, Blackburn, and Nakayama (1999)

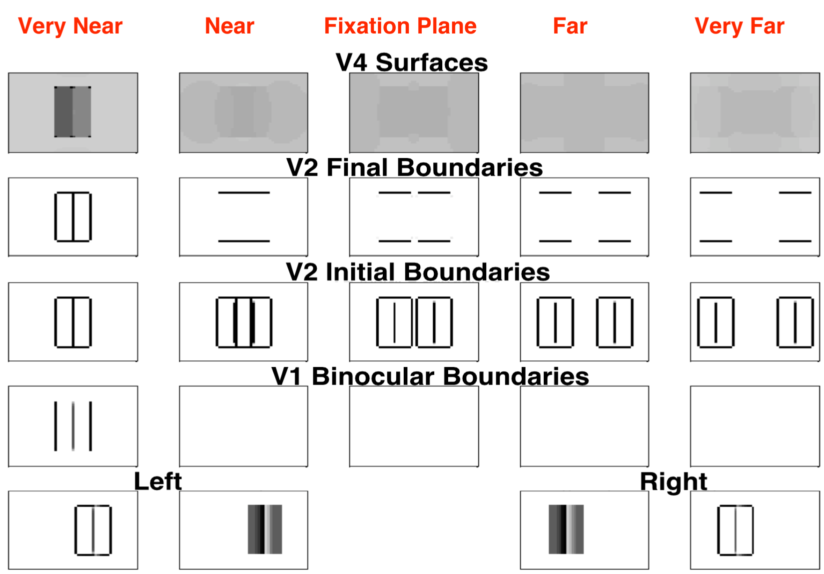
Same Explanation!



CRAIK-0'BRIAN-CORNSWEET EFFECT

Can the model simulate other surface percepts? e.g., surface brightness

Same Explanation!

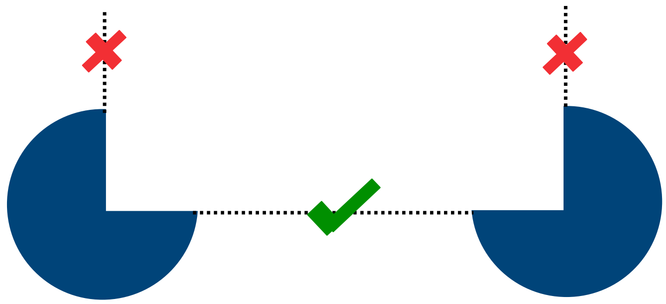


The 2D surface with the image on it is viewed at a very near depth

Adapts Grossberg & Todorovic (1988) to 3D

ROLE OF PERCEPTUAL GROUPING IN 3D PERCEPTS

How to generalize bipole grouping to 3D vision?



**How to group
3D planar, textured, slanted, and curved boundaries?**

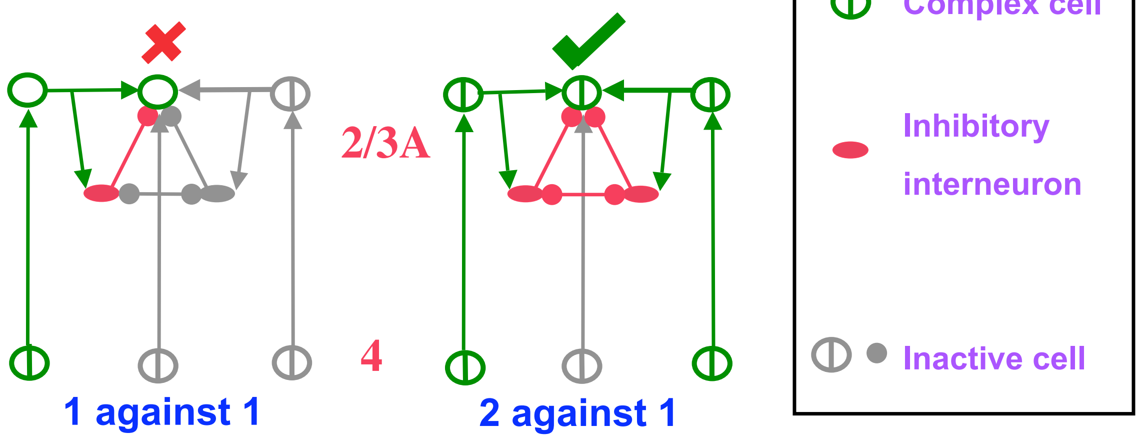
Grossberg & Swaminathan (2004); Cao and Grossberg (2004, 2005);
Fang & Grossberg (2004, 2005;)

ROLE OF PERCEPTUAL GROUPING IN 3D PERCEPTS

How to generalize bipole grouping to 3D vision?

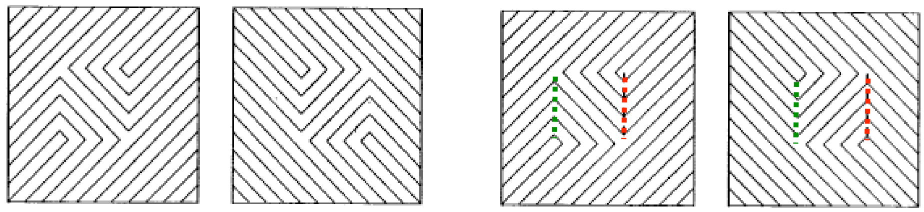
In stages: stereopsis, 3D figure-ground, slanted and curved surfaces

Bottom-up input from only one side | **Bottom-up inputs from both two sides**



3D GROUPINGS DETERMINE PERCEIVED DEPTH

Kaufman stereogram (1974)



L R

Vertical illusory contours are at different disparities than those of bounding squares

Illusory square is seen in depth

Vertical illusory contours are binocularly fused and determine the perceived depth of the square

Thin oblique lines, being perpendicular, are rivalrous:
simultaneous fusion and rivalry

3D GROUPINGS DETERMINE PERCEIVED DEPTH

Wilde (1950); Tausch (1953);

Ramachandran and Nelson (1976). Global grouping overrides point-to-point disparities. Perception, 5, 125-128

How do 3D groupings win over local disparities?

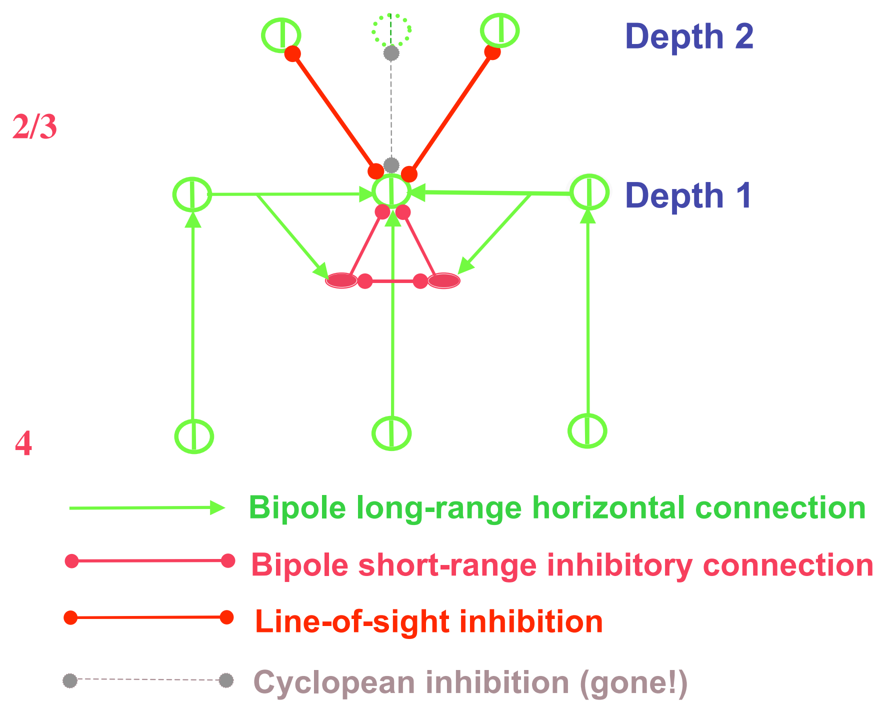
Model Hypothesis:

Disparity filter for eliminating "false matches" and 3D grouping process for eliminating "weak and incorrect groupings" are unified in V2 layer 2/3A

Eliminate all "false matches" through the 3D grouping process

Cao & Grossberg (2004, 2005)

GROUPING AND DISPARITY FILTER BOTH IN V2 LAYER 2/3



SURFACE-TO-BOUNDARY FEEDBACK

Feedback Between V2 Thin and Pale stripes

Boundaries and surfaces obey complementary rules

Surface-to-boundary feedback assures a consistent percept

It also initiates figure-ground separation!

Eliminates "extra boundaries" that hurt object recognition

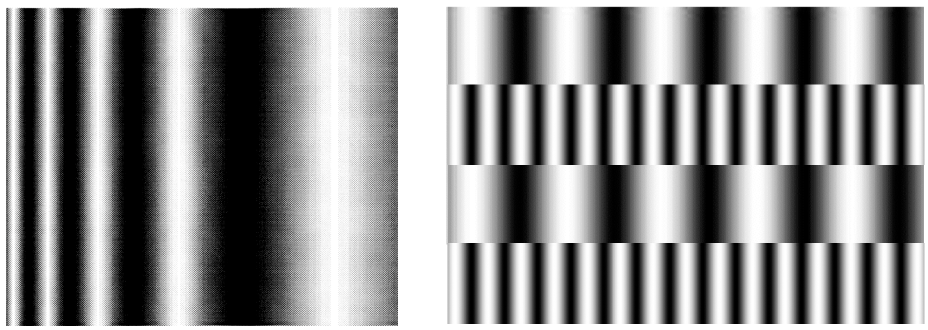
Why are there "extra boundaries"?

MULTIPLE-SCALE DEPTH-SELECTIVE GROUPINGS DETERMINE PERCEIVED DEPTH

As an object approaches, it gets bigger on the retina

Does a big scale (RF) always signal NEAR? NO!

Reversible!



Far Near Far Near

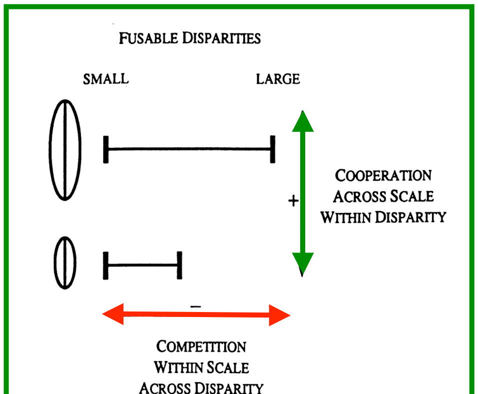
Far Near

Brown & Weisstein (1988)

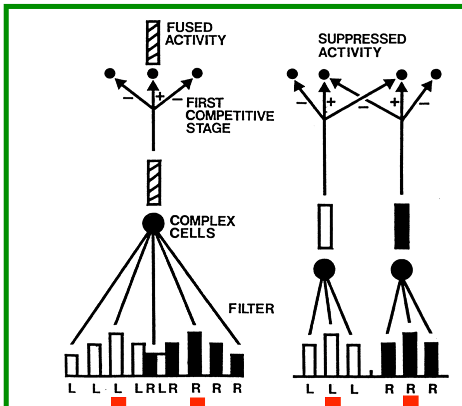
The same scale can signal either near or far

Some scales fuse more than one disparity

MULTIPLE-SCALE GROUPING AND SIZE-DISPARITY CORRELATION



Larger scales fuse more depths

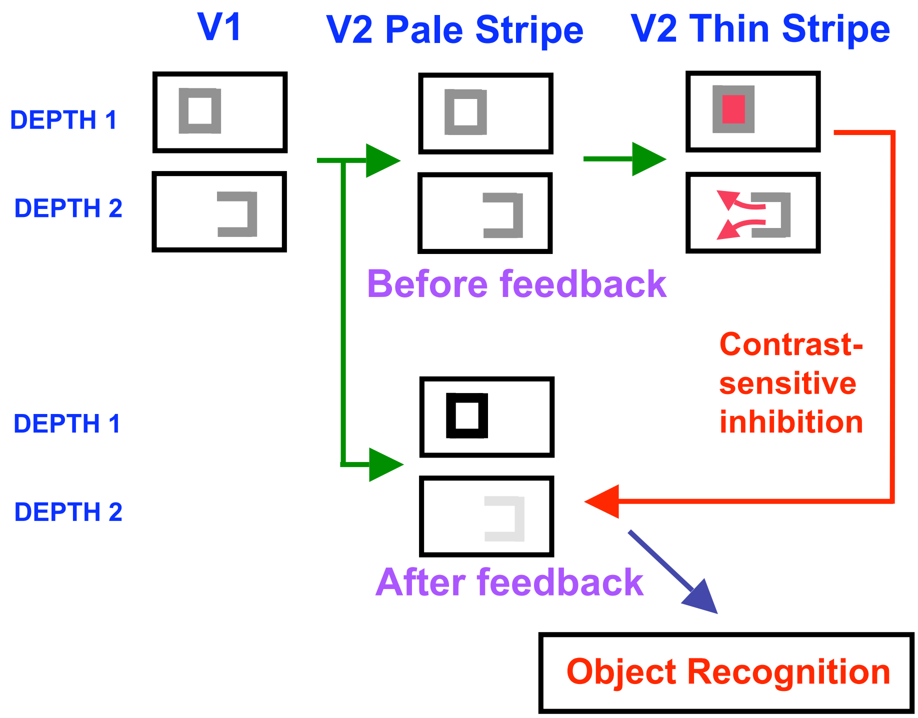


Simultaneous fusion and rivalry

Depth-selective cooperation and competition among multiple scales determines perceived depth

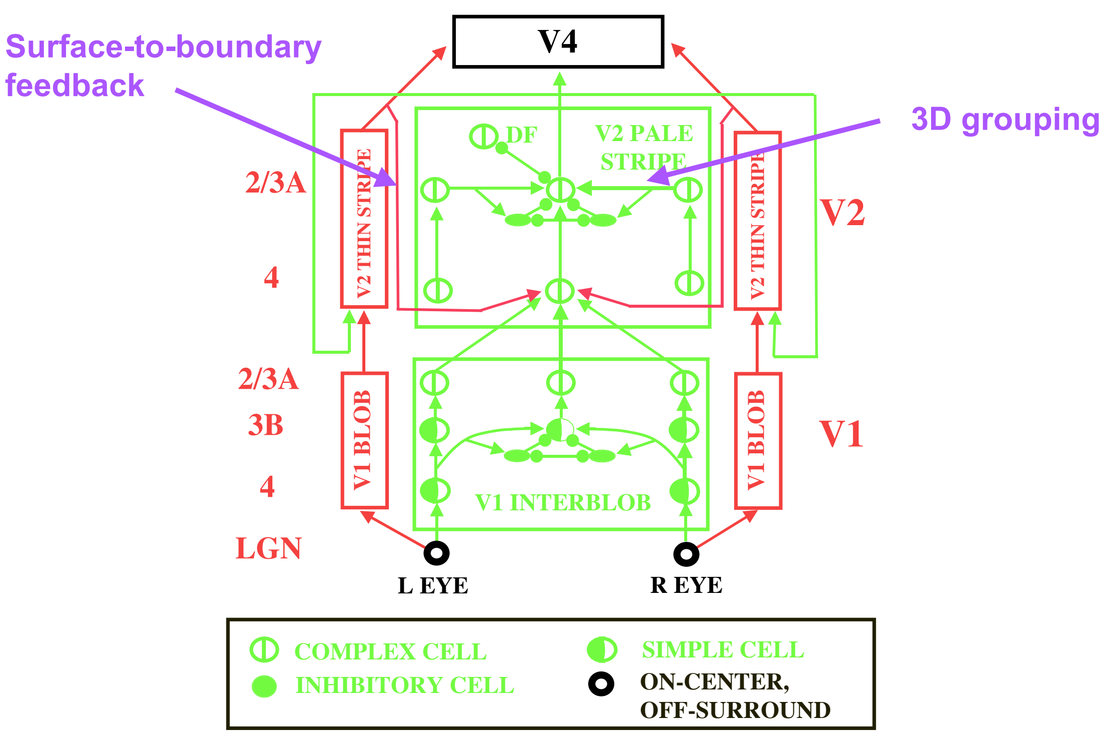
BOUNDARY PRUNING: Surface-to-boundary feedback from the nearest surface that is surrounded by a connected boundary eliminates redundant boundaries at the same position and further depths

SURFACE TO BOUNDARY FEEDBACK SUPPRESSES REDUNDANT V2 BOUNDARIES



3D LAMINART MODEL

Cao & Grossberg (2004, 2005)



27 SIMULATIONS WITH ONE SET OF PARAMETERS

This 3D LAMINART model is an extension of Grossberg and Howe (2003)

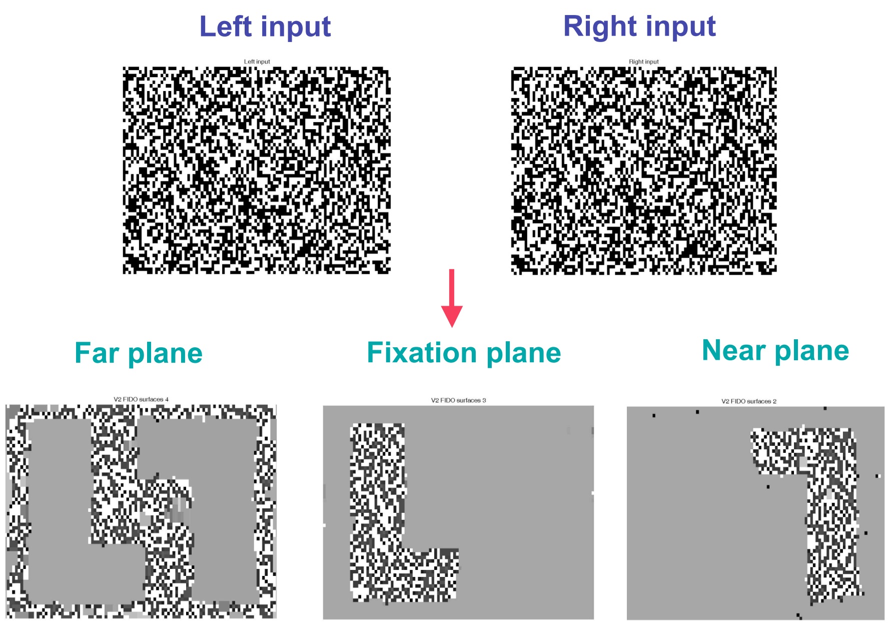
- Contrast variations of dichoptic masking (McKee et al., 1994)
- Correspondence Problem (Smallman & Mckee, 1995)
- Panum's limiting case (Gillam et al., 1995; McKee et al., 1995)
- Venetian blind illusion (Howard & Rogers, 1995)
- Stereopsis with polarity-reversed stereograms (Nakayama & Shimojo, 1990)
- Venetian blind illusion (Howard & Rogers, 1995)
- Da Vinci stereopsis (Nakayama & Shimojo, 1990; Gillam et al., 1999)
- Craik-O'Brian-Cornsweet lightness illusion (Todorovic, 1987)
- Effect of interocular contrast differences on stereothresholds (Schor & Heckman, 1989)
- Closure relationships and variations of daVinci stereopsis (Cao & Grossberg)

Other data that have been simulated using variants of this model:

- 3D slanted and curved surfaces (Grossberg & Swaminathan, 2004)
- Bistable Necker cube (Grossberg & Swaminathan, 2004)
- 3D transparency, neon color spreading, stratification (Grossberg & Yazdanbakhsh, 2005)
- Dense and sparse stereograms (Fang & Grossberg, 2005)
- Binocular rivalry (Yazdanbakhsh & Grossberg, 2005). Hear his talk at 8:30 AM on Friday
- Bregman-Kanizsa figure-ground separation, Kanizsa stratification, Muncker-White illusion, Benary cross, checkerboard percepts (Kelly & Grossberg, 2000)

STEREOGRAM SIMULATION: SURFACE LIGHTNESSES ARE SEGREGATED IN DEPTH

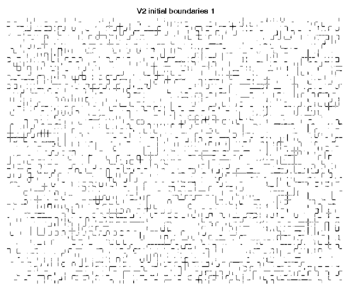
Fang & Grossberg (2004, 2005; see poster #577 on Saturday)



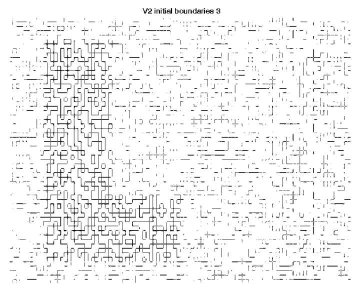
Cf. algorithms that just compute disparity matches and let computer code build the surface; e.g., Marr & Poggio (1974) et al

STEREOGRAM SIMULATION
(V2 INITIAL BOUNDARIES BEFORE S-TO-B FEEDBACK)

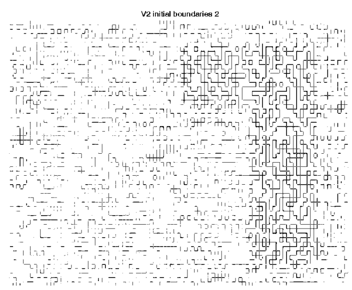
Far plane



Fixation plane

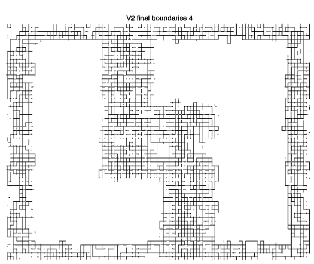


Near plane

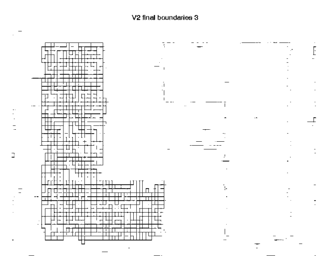


STEREOGRAM SIMULATION
(V2 OUTPUT BOUNDARIES AFTER S-TO-B FEEDBACK)

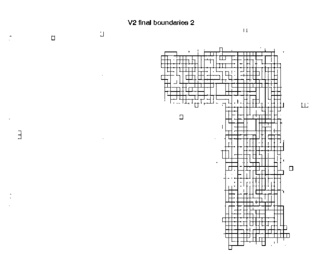
Far



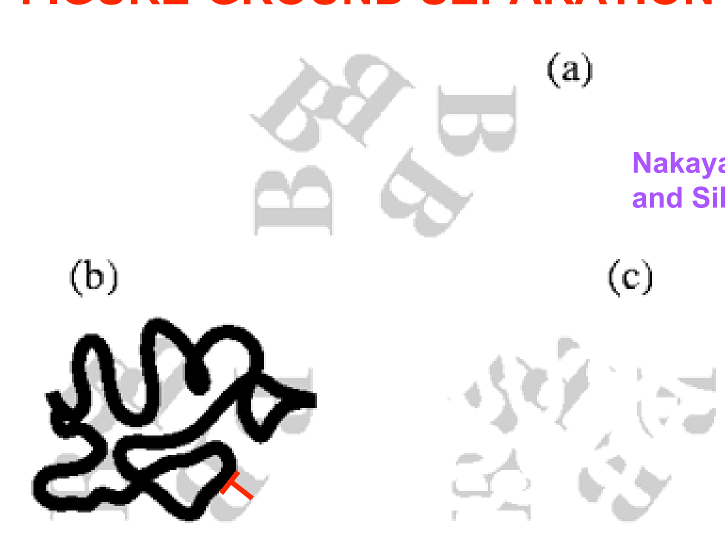
Fixation



Near plane



BREGMAN-KANIZSA
FIGURE-GROUND SEPARATION



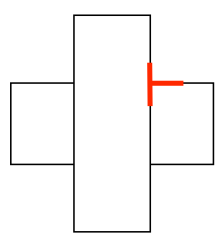
Nakayama, Shimojo, and Silverman (1988)

Black occluder helps to recognize gray B's because

shared black/gray boundaries "belong" to black occluder:

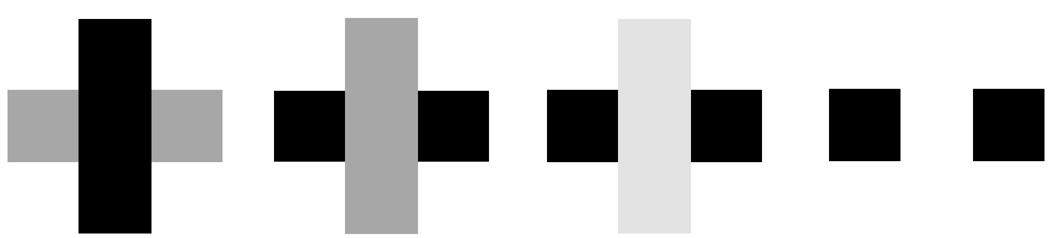
Extrinsic vs. intrinsic boundaries

DOES THE BRAIN USE T-JUNCTION OPERATORS IN FIGURE-GROUND SEPARATION?



What about the interaction of geometry and contrast?

A contrast change can reverse the answer without changing the T-junction geometry!



BIPOLE CELLS IN FIGURE-GROUND SEPARATION!

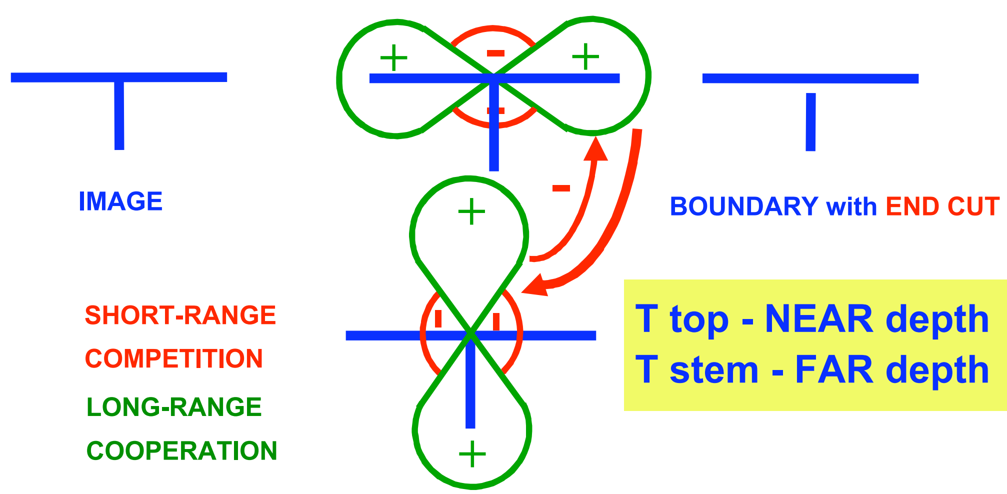
Prediction:
The bipole grouping property plays a key role in figure-ground separation

The bipole property is sensitive to both geometry and contrast

Figure-ground separation as a property of 3D boundary and surface formation

BIPOLE CELLS INITIATE FIGURE-GROUND SEPARATION

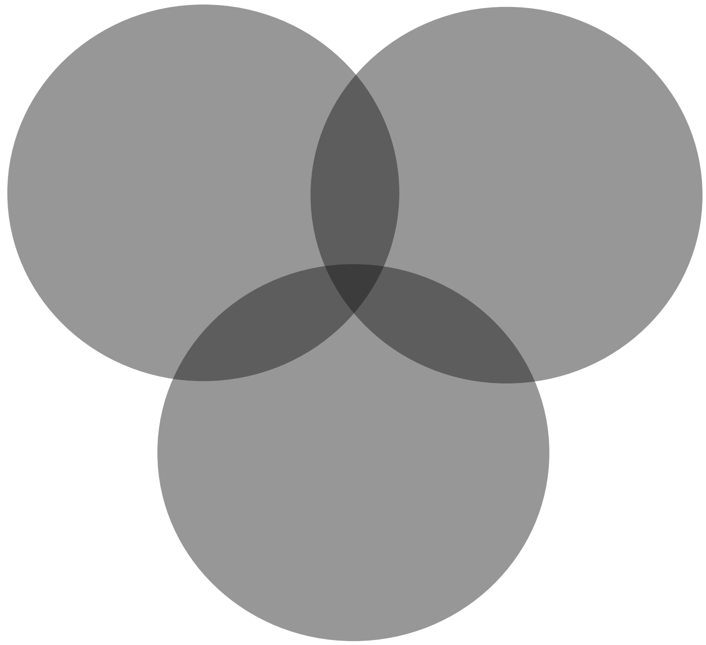
T-junction sensitivity without T-junction detectors



Prediction: 3D boundary end cuts influence depth perception

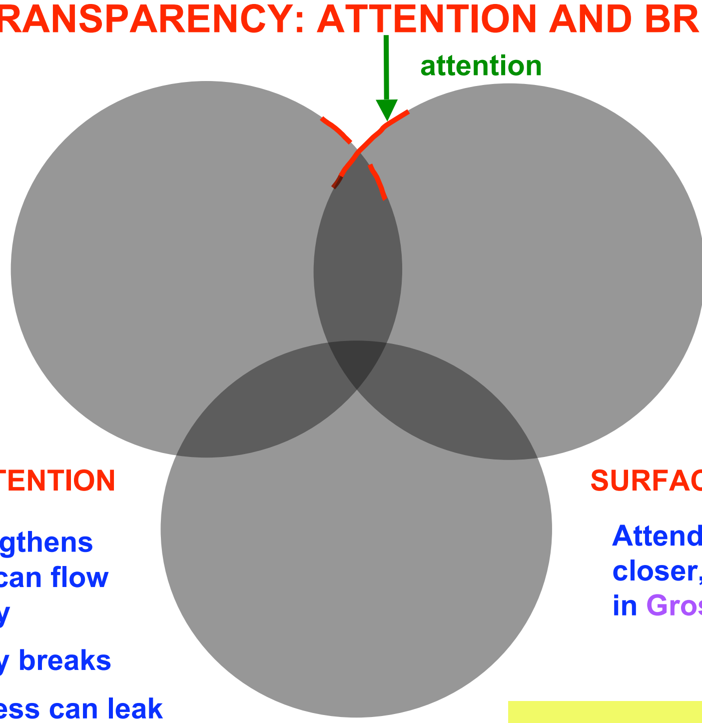
A weird idea! Do they exist? How to test?

BISTABLE TRANSPARENCY: ATTENTION AND BRIGHTNESS



Tse, P. (2005) Attention modulates the brightness of of overlapping transparent surfaces. Vision Research, in press

BISTABLE TRANSPARENCY: ATTENTION AND BRIGHTNESS



BOUNDARY ATTENTION

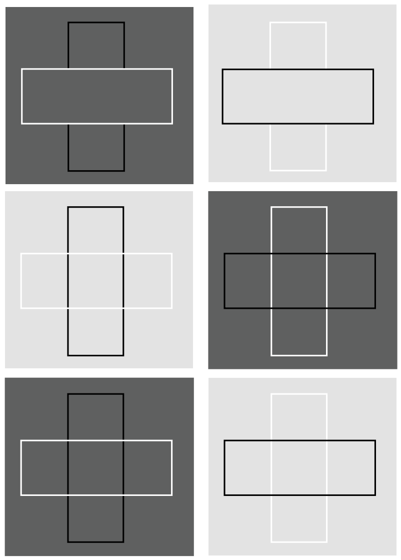
- Attention strengthens boundary and can flow along boundary
- Other boundary breaks
- Darker brightness can leak through boundary end gap
- Filling-in across gap changes brightness

SURFACE ATTENTION

Attended surface is closer, as predicted in Grossberg (1994)

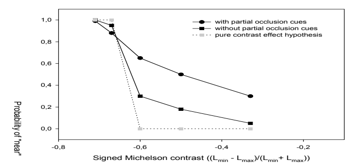
If boundary breaks, why do we not SEE the break?

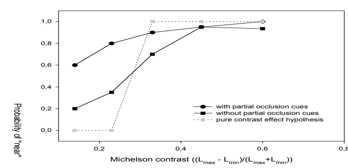
PSYCHOPHYSICAL TEST OF BIPOLES IN FIGURE-GROUND SEPARATION: GEOMETRY VS. CONTRAST



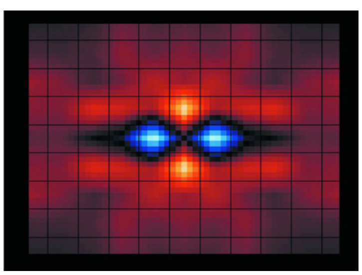
Dresp, Durand & Grossberg (2002, Spatial Vision, 15, 255-276)

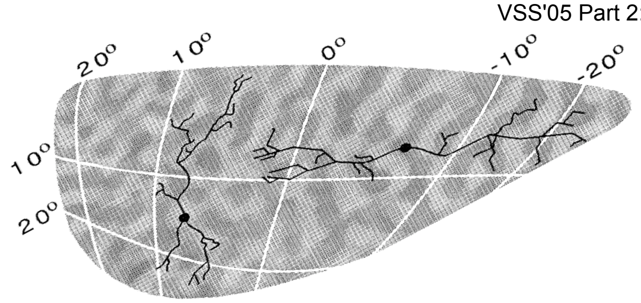
Results consistent with geometrical advantage of horizontal bipoles at occlusion T-junction, and of balanced geometrical competition at X-junctions, with increasing contrast offsetting the balance



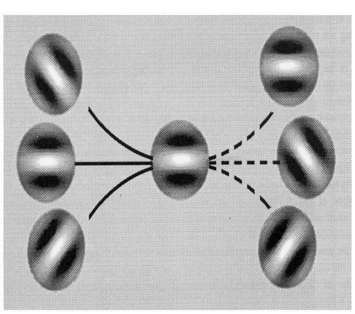


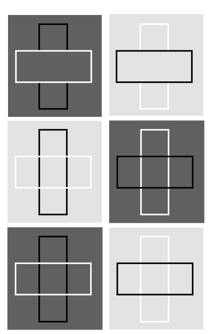
Judge if H or V looks closer as function of Michelson contrast

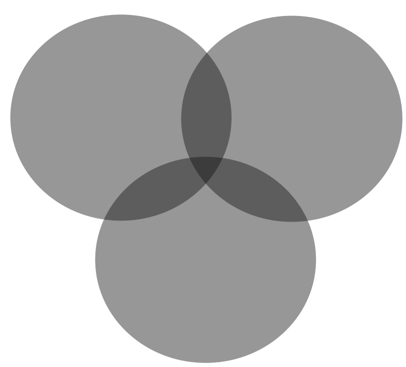




BIPOLES RULE!







CONSISTENCY IMPLIES FIGURE-GROUND SEPARATION!

I. BOUNDARY-SURFACE COMPLEMENTARITY

versus

BOUNDARY-SURFACE CONSISTENCY

We SEE one unified percept!

II. FIGURE-GROUND RECOGNITION

versus

VISIBLE SURFACE PERCEPTION

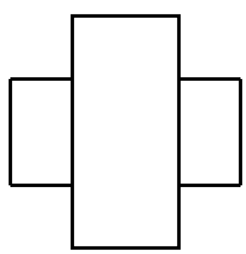
How do we RECOGNIZE a partially OCCLUDED object?

Why do we NOT SEE partially OCCLUDED object parts when the occluder is OPAQUE?

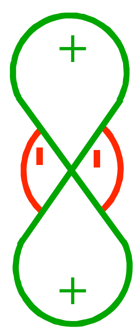
Why do not all OCCLUDING objects look TRANSPARENT?

The same process handles both I and II!

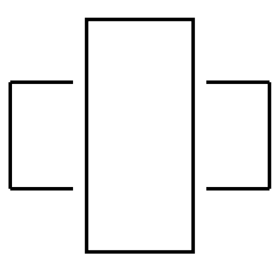
FIGURE-GROUND SEPARATION



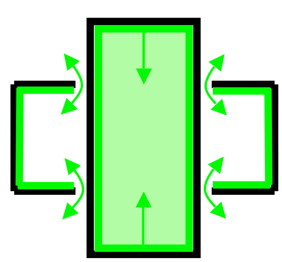
Boundary Attachment



Bipole
Cooperation
and
Competition



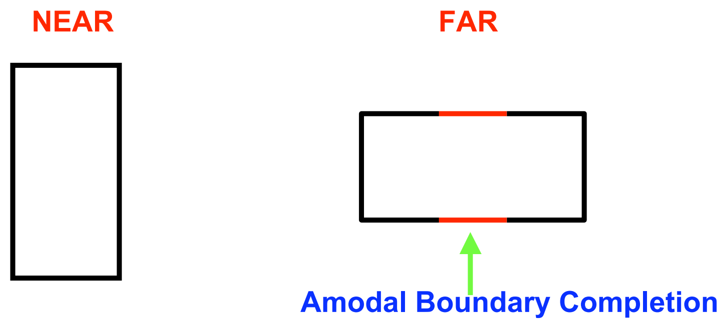
End gaps



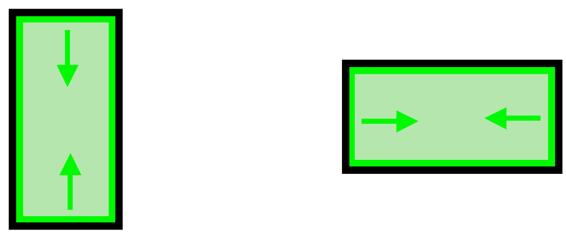
Filling-In; cf., neon color spreading

Claim: This step initiates figure-ground separation

SEPARATED V2 BOUNDARIES



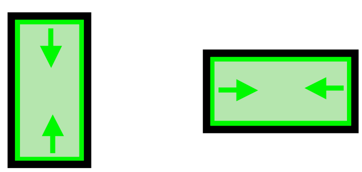
SEPARATED V2 SURFACES



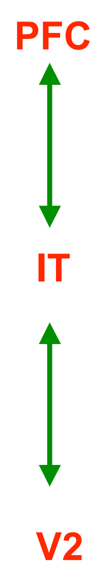
Amodal Filling-In

AMODAL COMPLETION AND RECOGNITION OF PARTIALLY OCCLUDED OBJECTS

Enables RECOGNITION of partially occluded objects:



If filling-in at this stage was modal, or visible, all occluding objects would look transparent!

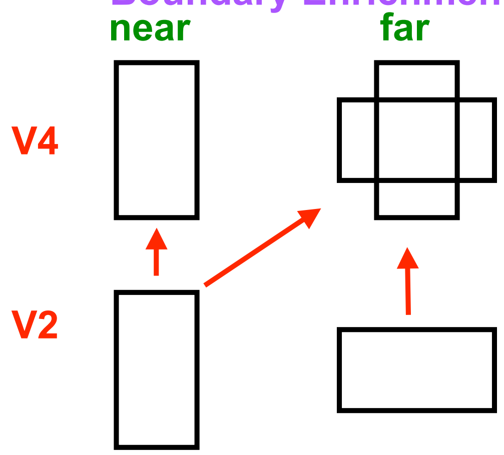


Prediction:
Direct recognition pathway for recognizing amodal boundaries and surfaces without seeing them

VISIBLE SURFACE PERCEPTION

Boundary Enrichment

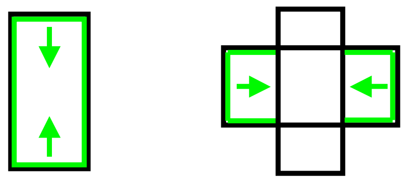
Asymmetry between near and far
cf., 3D neon color spreading



Cannot use these boundaries for occluded object recognition

Use these boundaries for occluded object recognition

Visible Surface Filling-In



Visible percept of unoccluded surface

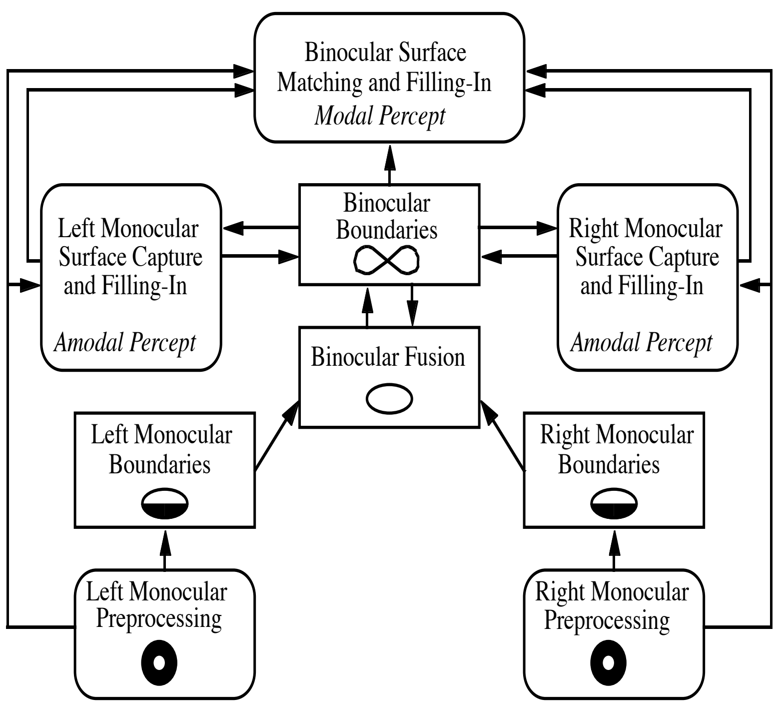
FACADE MACROCIRCUIT

V4
See unoccluded surface parts

See transparent surfaces

V2
Amodal completion

Recognition of occluded objects



BREGMAN-KANIZSA SIMULATION

INPUT STIMULUS

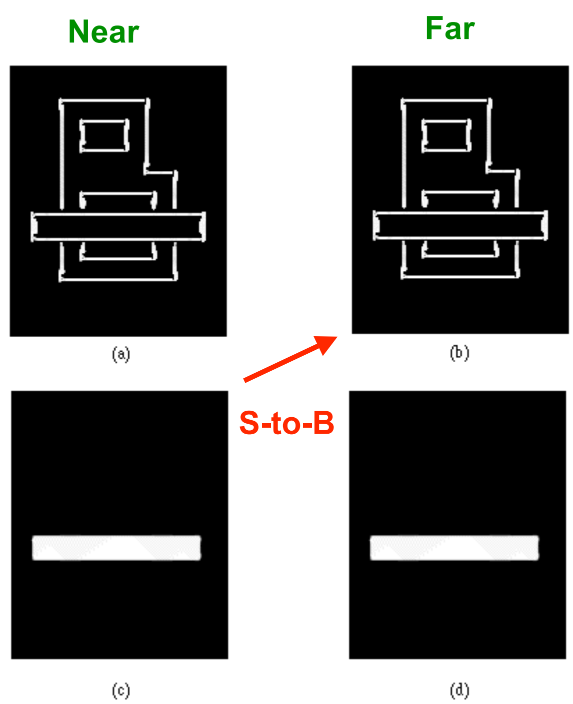
BREGMAN-KANIZSA SIMULATION

Kelly & Grossberg (2000, Perception & Psychophysics, 62, 1596-1619)

BEFORE surface-to-boundary feedback

BOUNDARIES
with end gaps at multiple depths due to size-disparity correlation

SURFACES
fill-in selectively within connected boundaries

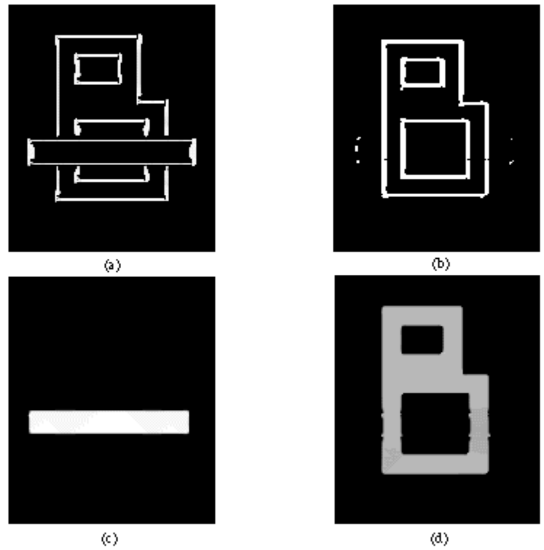


BREGMAN-KANIZSA SIMULATION

AFTER surface-to-boundary feedback

BOUNDARIES
complete occluded boundary

SURFACES
amodally fill-in occluding and occluded surfaces



Why amodal? Otherwise all occluders would look transparent!

There must be another stage where unoccluded surfaces are visible!

BREGMAN-KANIZSA SIMULATION

How to prevent all occluders from looking transparent?

Prediction:

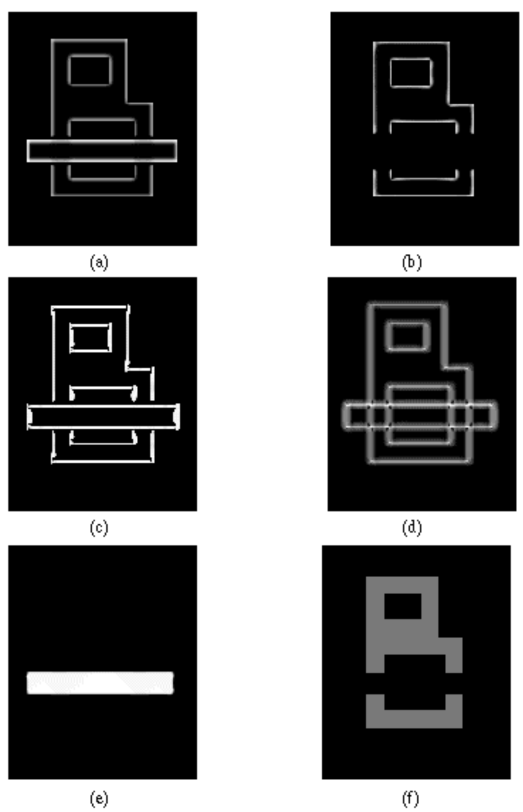
V4 boundary enrichment and modal filling-in:

Add near boundaries to far boundaries

V4 surface pruning:

Inhibit redundant surface inputs from farther depths

ASYMMETRY BETWEEN NEAR AND FAR



3-D PARSING OF OCCLUDED SURFACES

How does the laminar circuitry in areas V1 and V2 generate 3-D percepts of

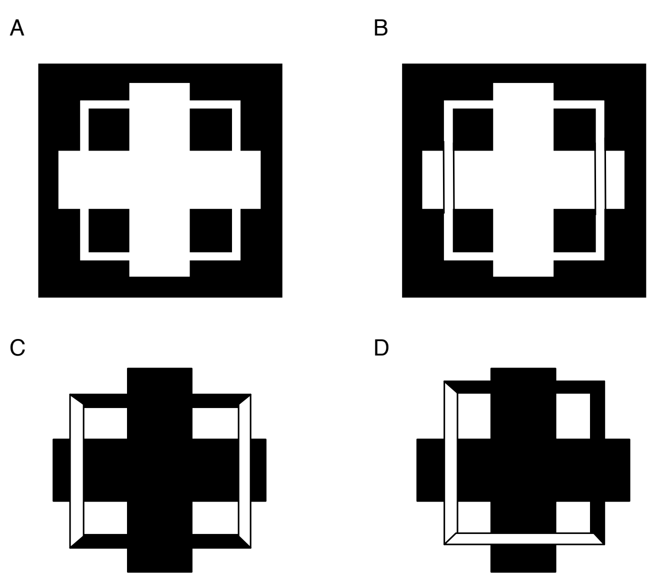
STRATIFICATION

TRANSPARENCY

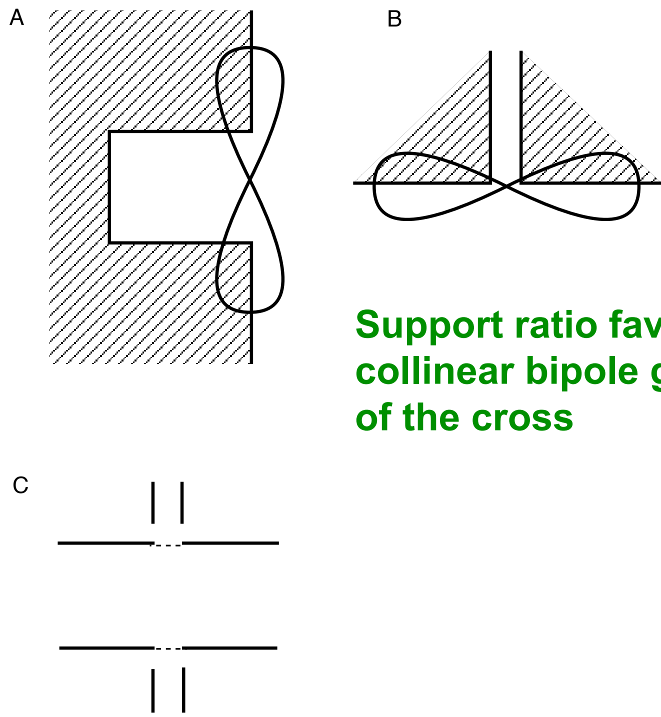
NEON COLOR SPREADING

In response to 2-D pictures and 3-D scenes?

KANIZSA STRATIFICATION



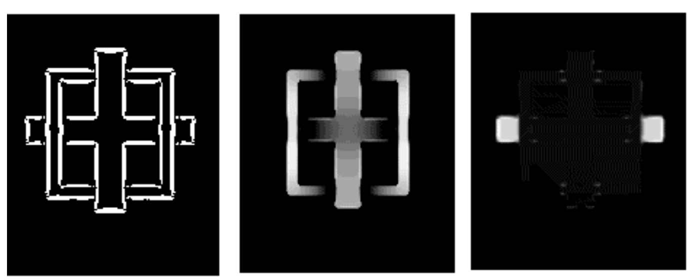
STRATIFICATION BOUNDARIES



Support ratio favors collinear bipole grouping of the cross

Petter, 1956

STRATIFICATION SIMULATION



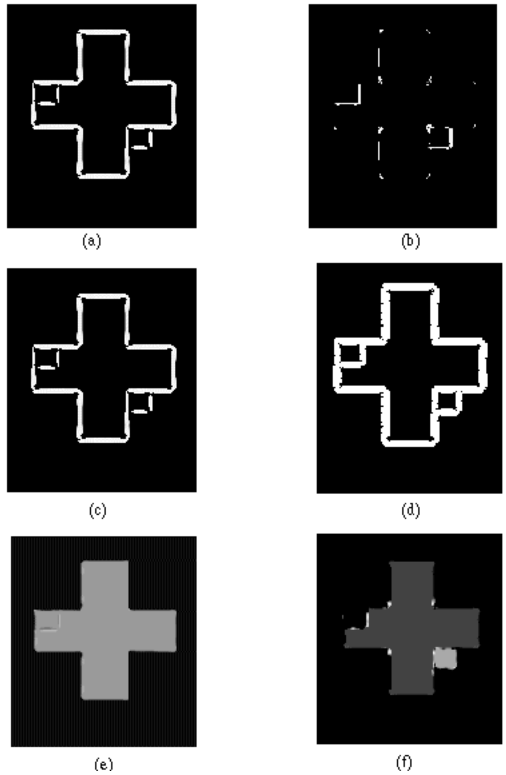
NEAR FAR

Endgaps and filling-in

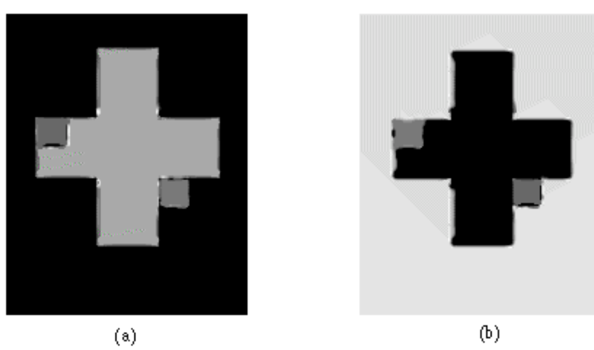
at near and far depths

Kelly & Grossberg, 2000,
Perception and Psychophysics, 62, 1596-1619

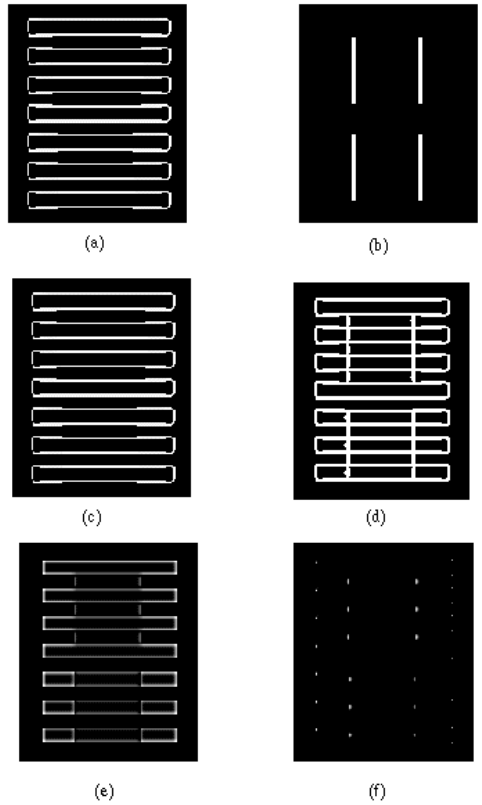
BENARY CROSS SIMULATION



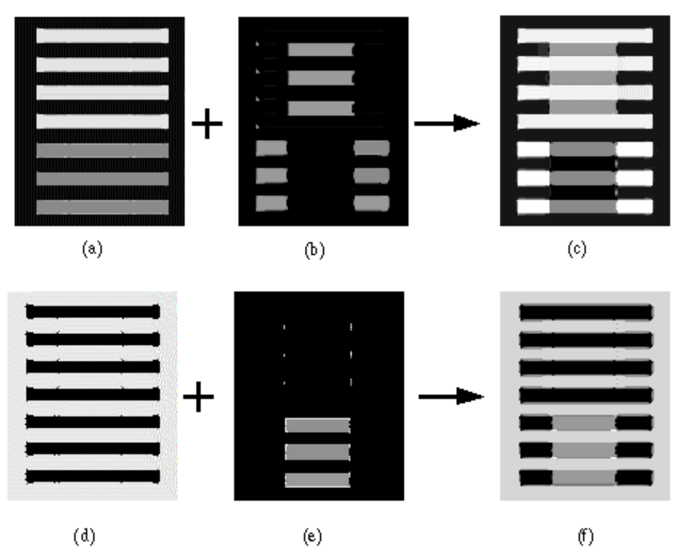
BENARY CROSS SIMULATION



WHITE'S EFFECT SIMULATION



WHITE'S EFFECT SIMULATION



HOW TO EXPLAIN TRANSPARENCY AND 3D NEON COLOR SPREADING?

Grossberg and Yazdanbakhsh
Vision Research, 45, 1725-1743

Explanation already implicit in the model if we include cortical development work of
Grossberg and Williamson (2001)

But we did not realize this!

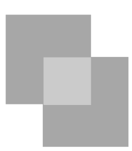
As in any real theory, hard data start falling out of the wash

The theory starts to get smarter than its creators…

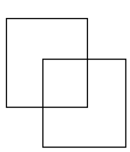
CONTRAST RELATIONSHIP IN TRANSPARENCY

Unique transparency

Bistable transparency



No transparency



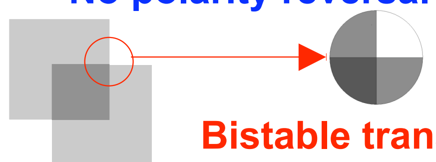
The same geometry

CONTRAST RELATION IN TRANSPARENCY

Single polarity reversal

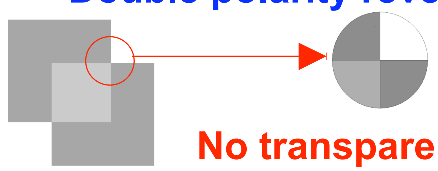
Unique transparency

No polarity reversal



Bistable transparency

Double polarity reversal

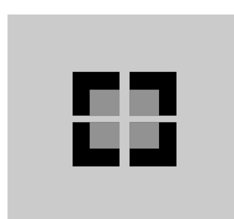


No transparency

How does polarity alignment influence transparency

Adelson, 2000; Anderson, 1997; Beck, 1984;
Metelli 1974; Watanabe and Cavanagh, 1992, 1993

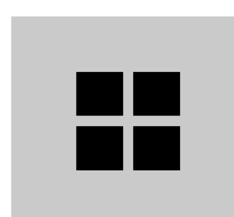
CONTRAST RELATIONS CAN INDUCE NEON SPREADING



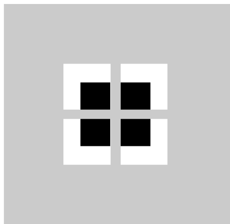
This contrast relation supports neon spreading

Percept

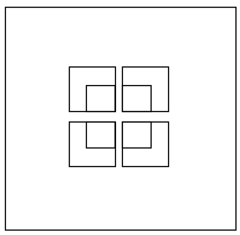
Over



CONTRAST RELATIONS CAN BLOCK NEON SPREADING



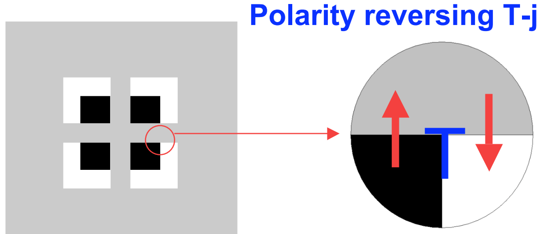
No neon spreading



Geometry is the same as the neon case

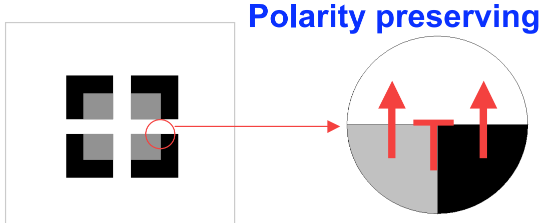
LOCAL CUES

Polarity reversing T-junction



Non-Neon

Polarity preserving T-junction



Neon

The laminar architecture should treat contrast relations in a way to let it overcome the absolute values of contrast

POLARITY ALIGNMENT INFLUENCES TRANSPARENCY AND NEON SPREADING

How early does this polarity sensitivity occur?

Claim: It occurs at layer 4 in V1

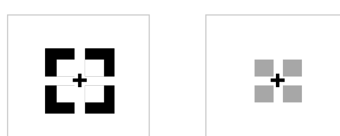
Why?

SAME OCULARITY OF CONTRAST

Same ocularity of contrast can induce neon

L R

Different ocularity of contrast can block neon

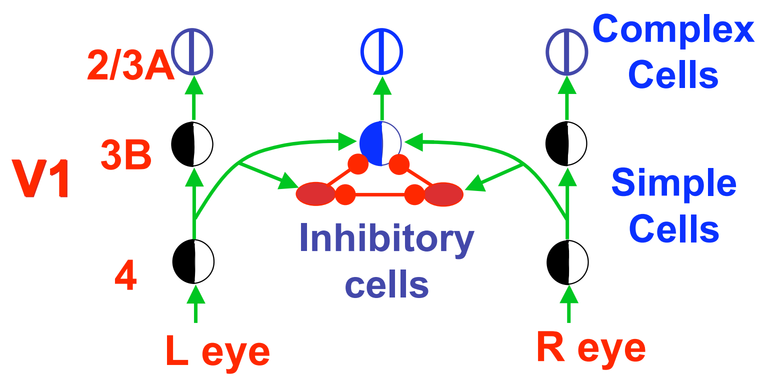


L R

Takeichi, Shimojo and Watanabe, 1992

The contrast polarity constraint is MONOCULAR

LAMINART CIRCUIT



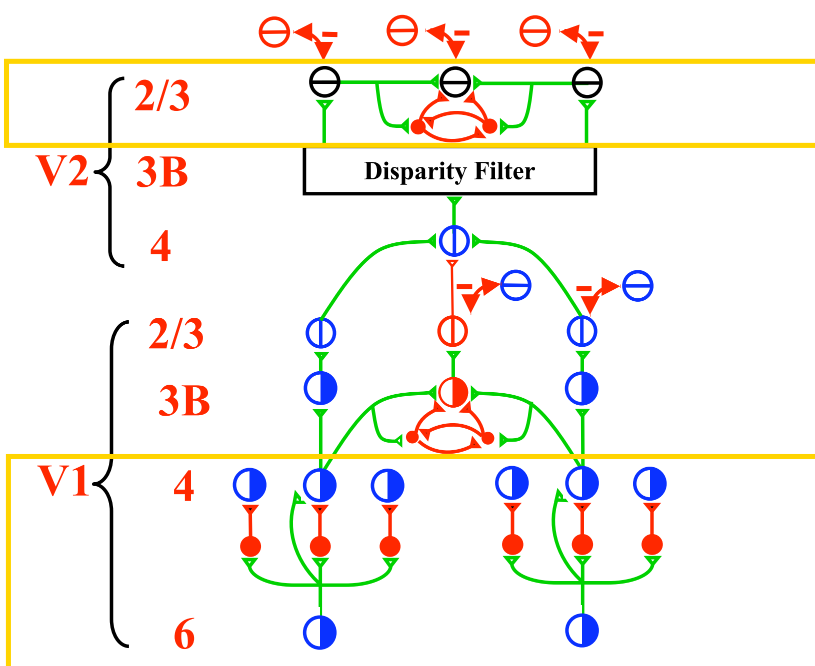
Grossberg and Howe (2003, 43, 801-829)

Binocular fusion occurs in layer 3B of V1

Prediction 1: The polarity-specific monocular process is in layer 4 of V1

Prediction 2: This process is monocular polarity-specific competition

3-D LAMINART CIRCUIT



V2 grouping pools opposite polarities

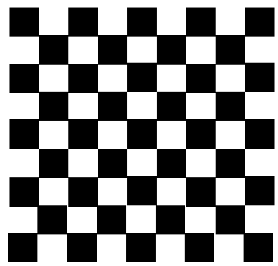
SUGGESTS NEW EXPERIMENTS

Preference for like-polarity inhibition in layer 4 of V1 is proposed to develop from normal visual statistics

Grossberg and Williamson (2001, Cerebral Cortex, 37-58)

What happens to this preference when animals are raised in abnormal visual environments?

e.g., opposite polarity textures?



SELF-NORMALIZING INHIBITION FROM V1 6-TO-4

Multiple predicted roles:

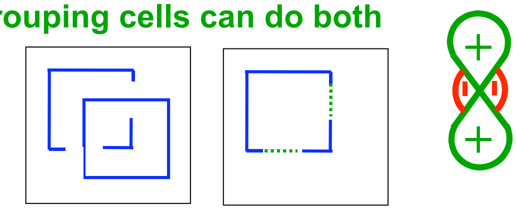
- Contrast gain control of BU inputs from LGN
- Selection and analog coherence of groupings
- Target of top-down attention
- Influences transparency percepts

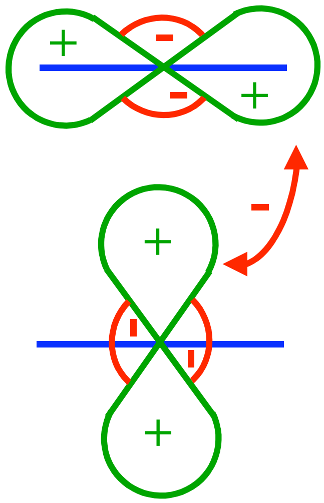
Suggests totally new kinds of experiments

Who will run with this opportunity?!

HOW ARE BOUNDARY GAPS CREATED AND COMPLETED?

Bipole grouping cells can do both

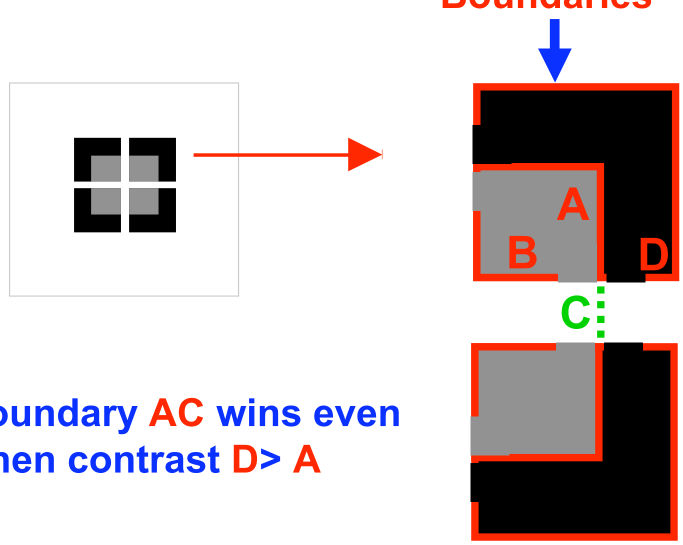




Collinear cooperation and orientational competition

SAME PROBLEM IN NEON SPREADING

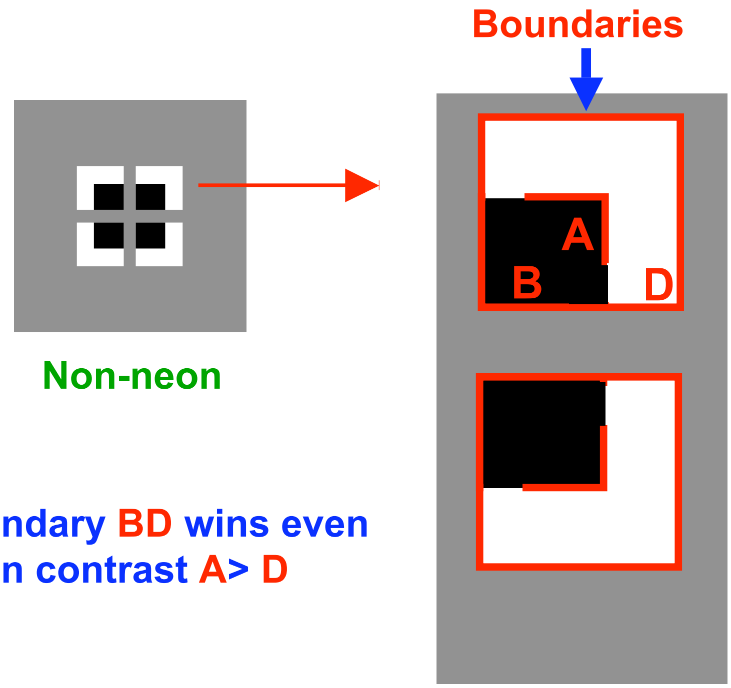
Boundaries



Boundary AC wins even when contrast D> A

Like-polarity competition between B and D allows boundary AC to win.

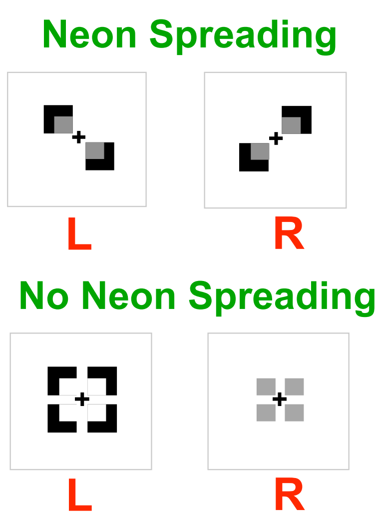
SAME PROBLEM IN BLOCKED NEON



Boundary BD wins even when contrast A> D

Opposite-polarity B and D contrasts do NOT compete.

SAME OCCULARITY OF CONTRAST CAN INDUCE NEON

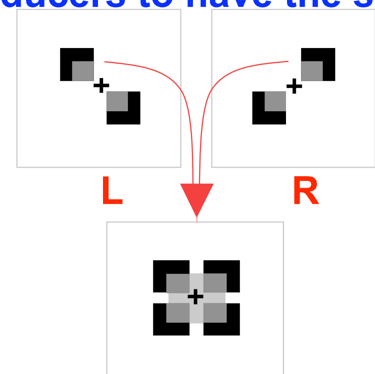


Takeichi, Shimojo and Watanabe, 1992

Explanation: In the No Neon case, different ocularity inputs bypass the monocular polarity-specific competition in V1

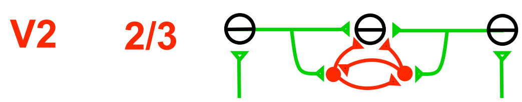
ILLUSORY CONTOUR FORMATION IS BINOCULARLY DRIVEN

Formation of illusory contours does not need inducers to have the same ocularity

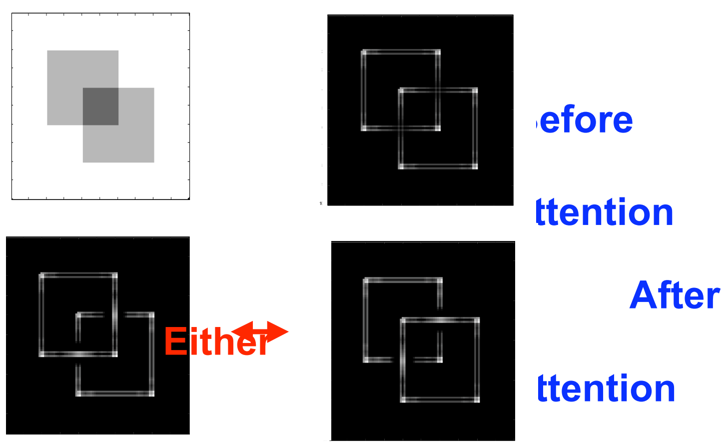


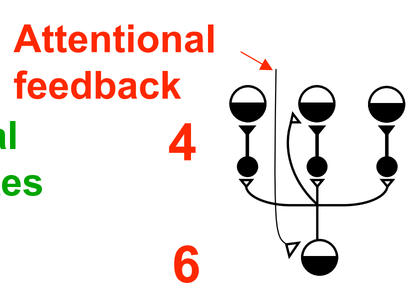
Takeichi, Shimojo and Watanabe, 1992

Layer 2/3 bipole grouping cells in V2 are binocular



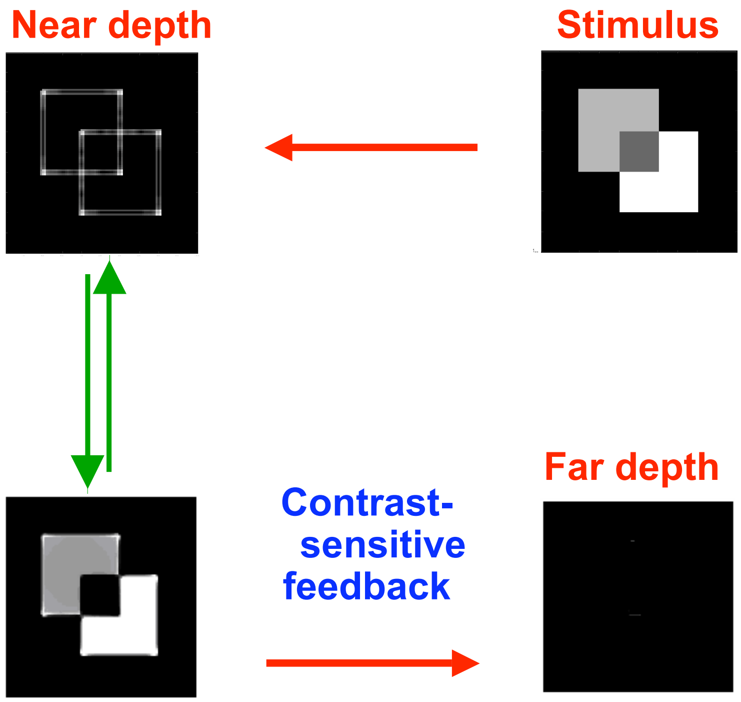
ATTENTION MAKES EITHER BOUNDARY STRONGER



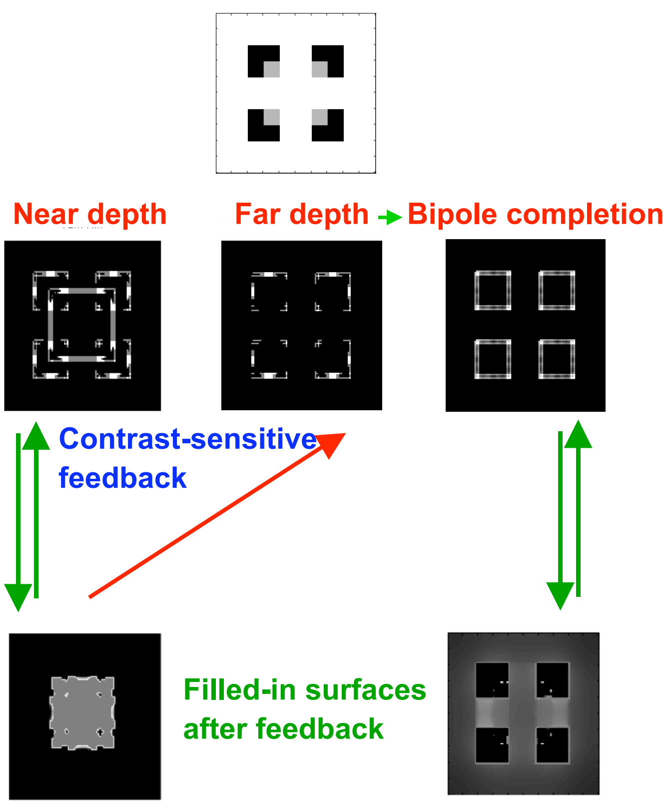


Attentional feedback activates layer 6 of V1

NON-TRANSPARENT SIMULATION

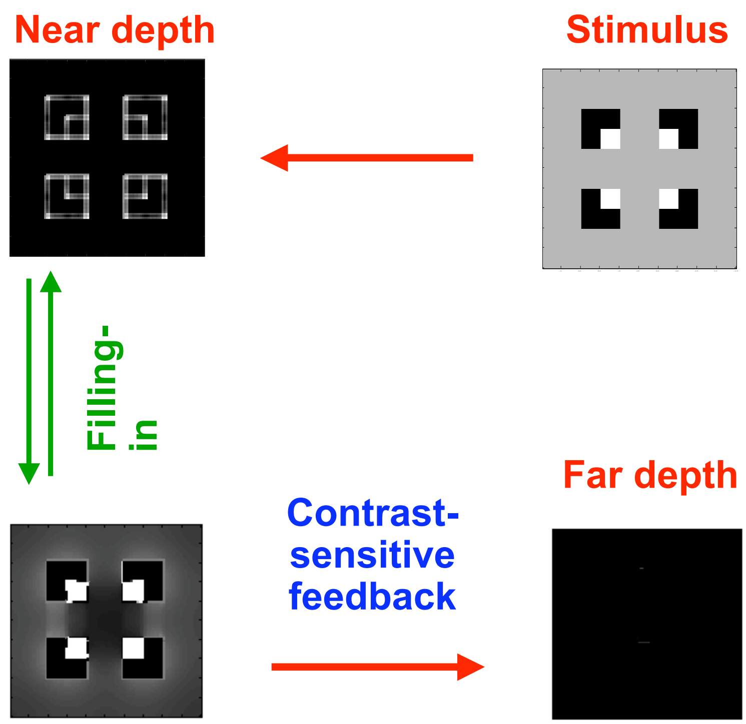


NEON SIMULATION



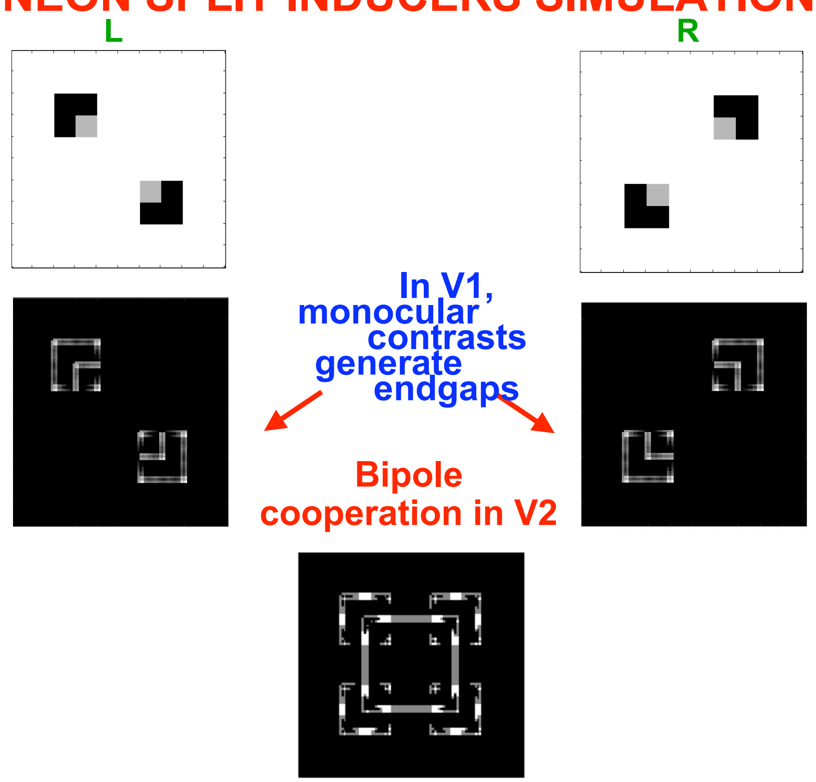
Grossberg/Mingolla
VSS'05 Part 2: 137

NON-NEON SIMULATION



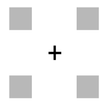
Grossberg/Mingolla
VSS'05 Part 2: 138

NEON SPLIT INDUCERS SIMULATION

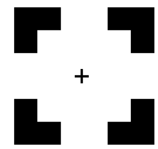


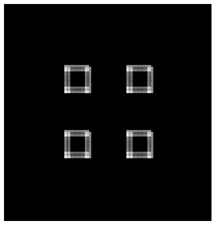
NON-NEON CASE, SPLIT CONTRAST SIMULATION

L

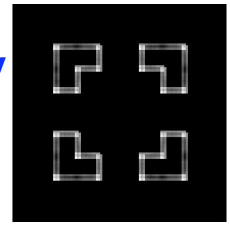


R

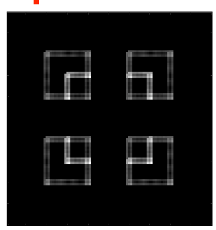




Different ocularity bypasses gap formation in V1



Long range bipole cooperation blocked by orientation competition in layer 2/3 of V2



CONCLUSIONS

Transparency and neon color spreading data uncover some constraints on depth stratification

Monocular same-polarity competition explains the contrast relation role in depth stratification

This same-polarity competition is implemented in layer 6-to-4 connections of V1, where cells are mostly monocularly driven

Implementation of monocular same-polarity competition unifies
- STRATIFICATION
- TRANSPARENCY
- NEON COLOR SPREADING

phenomena

CONCLUSIONS

Monocular same-polarity competition is consistent with model of inhibitory layer 4 development by Grossberg and Williamson (2001, Cerebral Cortex)

Question: What happens to layer 4 inhibition if animals are reared in opposite polarity textures?

HOW DOES THE CORTEX HANDLE SLANTED AND CURVED 3D SURFACES?

Previous model only handles PLANAR 3D surfaces

Can the model be self-consistently extended?

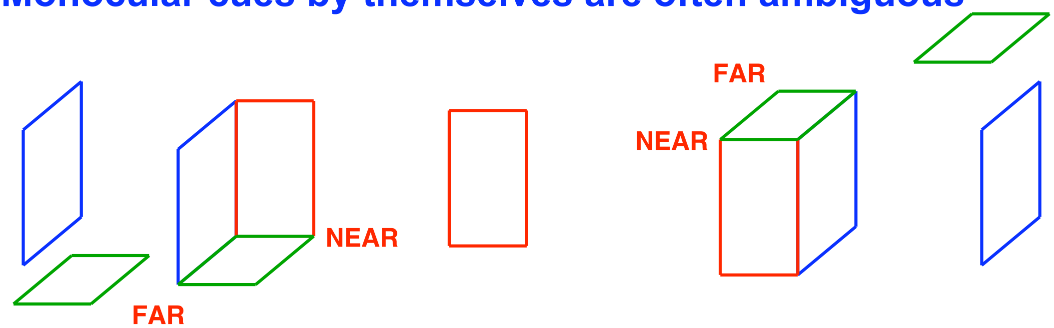
YES!

Grossberg and Swaminathan (2004, Vision Research, 44, 1147-1187)

3D REPRESENTATION OF 2D IMAGES

Monocular cues (e.g angles) can interact together to yield 3D interpretation

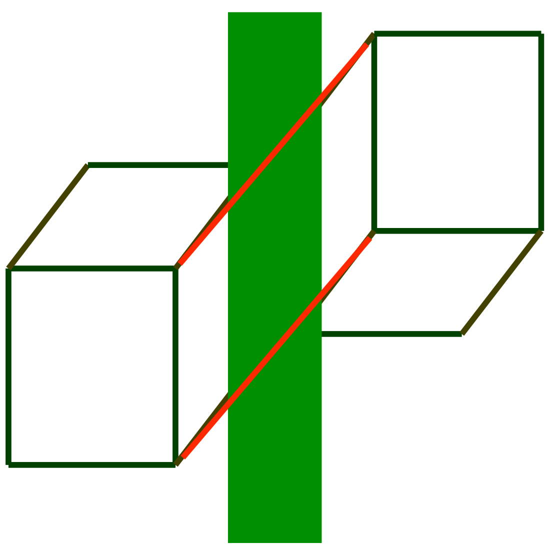
Monocular cues by themselves are often ambiguous



SAME ANGLES AND SHAPES, DIFFERENT SURFACE SLANTS

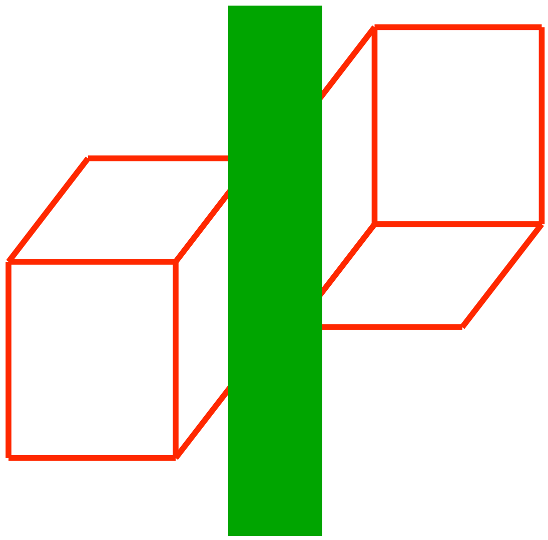
How do these ambiguous cues contextually define a 3D representation?

3D PERCEPTUAL GROUPING



Tse (1999)

3D GROUPING



Tse (1999)

3D GROUPING

A straight edge can represent a

FLAT

NEAR-TO-FAR

FAR-TO-NEAR

object contour in different figures

Spatial combinations of

ANGLES and EDGES

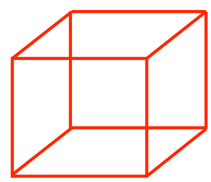
can disambiguate depth direction

3D GROUPING LAWS AND NECKER CUBE

The LAMINART model clarifies how horizontal connections can grow during development to create the
BIPOLE GROUPING property

The SAME MECHANISMS can explain development of ANGLE cells and DISPARITY GRADIENT cells which contextually represent SLANTED 3D SURFACES

Simulates BISTABLE 3D NECKER CUBE percepts!

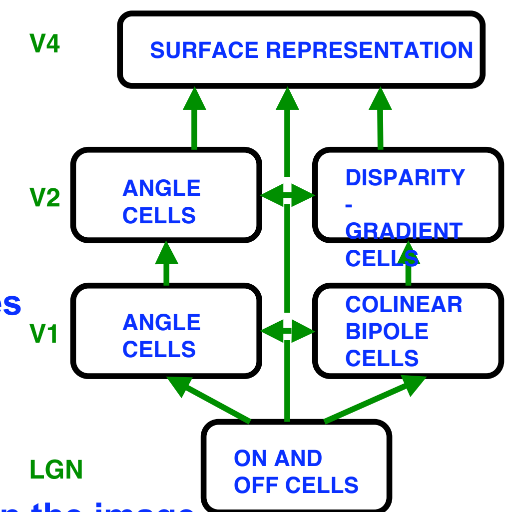


3D LAMINART MODEL

Four key additions:

Angle cells
(non-colinear bipole cells)
cells tuned to various angles

Disparity-gradient cells

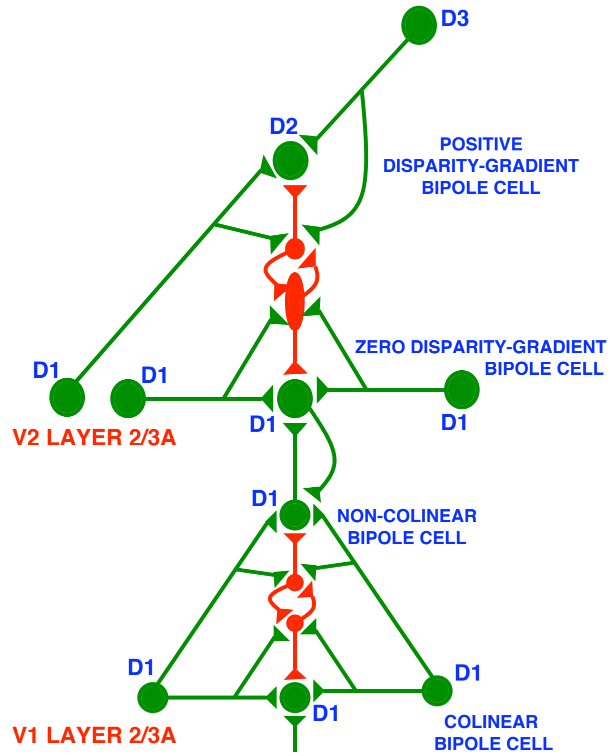


cells tuned to disparity-gradients in the image

Weights from angle cells to disparity-gradient cells
learned while viewing 3D image

Boundary grouping between disparity-gradient cells
disambiguates ambiguous groupings

3D GROUPING CIRCUIT



WHERE TO FIND MODELING ARTICLES WITH FURTHER DETAILS?

http://www.cns.bu.edu/Profiles/Grossberg