

## WEEK 1: FUNDAMENTAL PROBLEMS OF VISION

- 1) Unit formation and grouping
- 2) Seeing and recognizing -- form/color interactions
- 3) Retinal veins and blind spot
- 4) Perceiving surface color: Constancy, contrast, and discounting the illuminant
- 5) Stabilized images: Boundaries and featural color and brightness
- 6) Complementary processing:  
Unoriented and oriented filters
- 7) The noise--saturation dilemma
- 8) Reflectances and ratios; shunting and mass action

## PREFACE

The following lecture notes for Week 1 of CN 530 contain many terms that are likely to be unfamiliar to you. In addition, many ordinary-seeming words are used in a manner that can only be fully understood in the context of material that you will not encounter until later in the course.

The purpose of this lecture is to (dis?)orient you as quickly as possible to the scope of issues addressed and methods of inquiry adopted in the course.

Please do not worry if parts, or even most, of this lecture make no sense on first presentation. If you are looking at these notes before attending the first lecture, please be aware that they may be difficult to interpret without a spoken "sound track." Lecture notes for later weeks are more straightforward.

Please use this lecture as a "site map" for locating the contents of subsequent lectures with respect to the goals of the entire course.

In general, panels that appear as plain text without headers in these notes are treated as "hidden slides" during the lecture and not presented. They are intended as notes for further study.

## WHY STUDY VISION?

Among all the studies of natural causes and reasons Light chiefly delights the beholder; and among the great features of Mathematics the certainty of its demonstrations is what preeminently (tends to) elevate the mind of the investigator. Perspective, therefore must be preferred to all the discourses and systems of human learning. In this branch [of science] the beam of light is explained on those methods of demonstration which form the glory not so much of Mathematics as of Physics and are graced with the flowers of both. But its axioms being laid down at great length, I shall abridge them to a conclusive brevity, arranging them on the method both of their natural order and of mathematical demonstration; sometimes by deduction of the effects from the cause, and sometimes arguing the causes from the effects; adding also to my own conclusions some which, though not included in them, may nevertheless be inferred from them.

Leonardo da Vinci  
General Introduction to the Book on Painting, Section 13.  
**The Notebooks of Leonardo da Vinci, Volume I.**  
Edited by Jean Paul Richter, New York: Dover.

## THE UNITS OF VISION

The objects of perception and the space in which they seem to lie are not abstracted by a rigid metric but a far looser one than any philosopher ever proposed or any psychologist dreamed.

Lettvin, 1981

**GEOMETRY OF VISION**

	<i>Euclidean</i>	<i>Quotidian</i>
unoriented	• point	dot
oriented	/ line	bar (line)
	▀ plane	surface (region)
	infinitesimal	extended

In what substrate do these functional units exist?  
 Are they biochemical processes?? ... algorithms??  
 ... network patterns?? ... aspects of consciousness??

**EMERGENT SEGMENTATION AND GROUPING**



Proximity → Collinearity

Gestalt "Laws"

The "Gestalt Laws" refer to perceptual grouping or unit formation include on the basis of properties such as:

- similarity
- proximity
- closure
- symmetry
- pragnanz* (literally "pregnant")
- "good continuation"
- "common fate"

See Koffka, 1935 and Köhler, 1947

See also: <http://www.ship.edu/~cgboeree/gestalt.html>

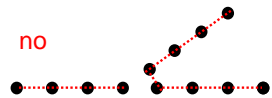
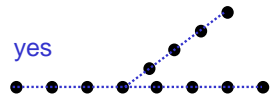
### GESTALT PROPERTIES



Bozzi, 1969

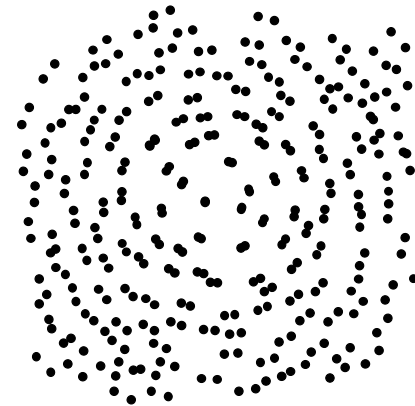
proximity < collinearity

long range "cooperation"



These emergent groupings are "invisible."

### COHERENT PATTERNS



global randomness; local correlations

Glass, 1969

Coherence: select globally consistent groupings

### Glass patterns

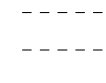
Coherence: We select the globally most consistent groupings from among all possible local groupings.

These emergent boundary groupings are *invisible*. We continue to "see" only dots, although we are aware of and "recognize" circular groupings.

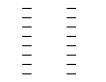
### LINE ENDINGS AND GROUPING

Possible groupings of oriented lines segments:

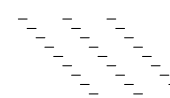
collinear



perpendicular



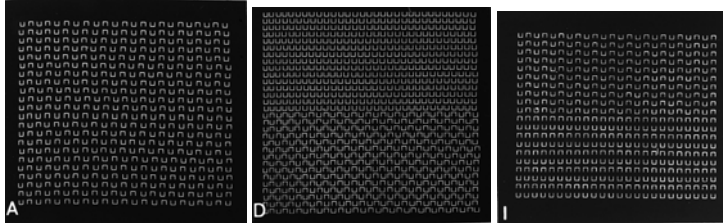
oblique



Global grouping may or may not be in the same orientation as local line segments.

**BECK TEXTURE DISPLAYS**

*Linking* forms emergent features, which support region segregation.



tachistoscopic display: 100 msec

Beck, 1983

Beck textures

“Emergent features” can form via *linking* of local features to enable us to segment one image region from another.

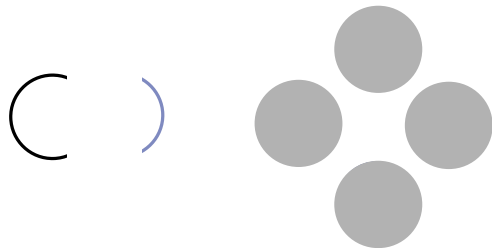
We “*see*” the same kind of image contrast for a given element, *whether or not* that element is part of some larger grouping.

Q: How are these displays constructed? Why?

The observer’s task in these displays is to say whether the top and bottom halves are “same” or “different.”

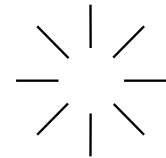
**NEON COLOR SPREADING**

Some groupings do change what we *see*.



Varin, 1971

**ILLUSORY CONTOURS**



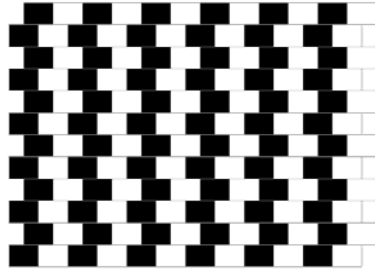
Ehrenstein 1941/1987

The form we *recognize* is not “in the image.”

We *see* the disk region as brighter than the background.

### DOMINANCE OF GLOBAL GROUPINGS

can distort local information -- even collinear groupings



See applet at:

<http://lite.bu.edu/vision/applets/Form/CafeWall/CafeWall.html>

### DISCLAIMER: COGNITIVE FACTORS



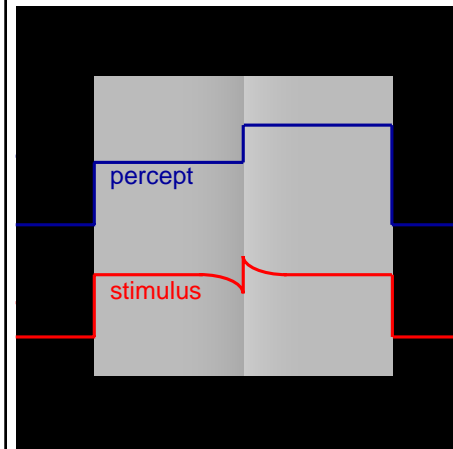
From Marr 1982 -- R. C. James  
... by way of Sinha & Adelson, 1997

### PERCEPTION "RESISTS" COGNITIVE INTRUSIONS



Fig 18, Kanizsa & Vicario, 1968

### CRAIK-O'BRIEN-CORNSWEET EFFECT



long-range interactions  
in what we **see**  
in a *region*.

After Todorović, 1987

The COCE is the first of many examples that we will encounter that suggests that . . .

a perceptual **surface** is not simply a union of local patches.

“Hold this thought” to compare with later evidence.

The kind of taxonomy of visual phenomena that we have been pursuing so far is not sufficiently constrained to help us model visual processing.

How, then, should we proceed?  
computational constraints? physiology?

## FROM PHENOMENOLOGY TO GEOMETRY

**Bar (line):** not a union of dots

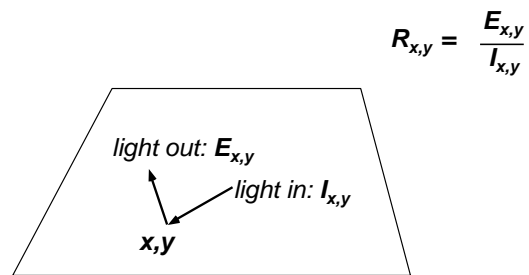
**Surface:** not a union of local patches

For these units:    **extension**  
                          **interiors**  
                          **contours** separating them from surround

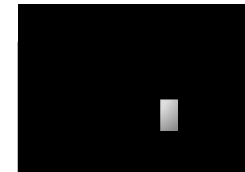
Not Euclidean geometry or calculus of Newton/Leibniz

**“Fact:** *The notion of “surface area” is devoid of any meaning unless you specify the resolution at which it is to be assessed. The same goes ipso facto for the notion of “arc length.”*  
Koenderink **Solid Shape** 1990

## ILLUMINATION, REFLECTANCE, AND VISION



## AMBIGUITY OF LOCAL INFORMATION



$$E_{x,y} = f(I_{x,y} R_{x,y})$$

How could you tell from purely local information whether a surface patch is planar under nonuniform illumination or smoothly curved in depth?

If you cannot make this distinction at a certain point in visual processing, what *can* you base subsequent processing on?

### Illumination vs. Reflectance effects

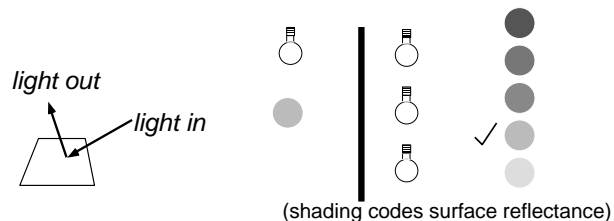
Claim: We must be able to distinguish the visual effects of surface shape (orientation) and reflectance from the effects of variations in illumination\* *at a fairly early stage* of visual processing, because virtually every other important visual competence depends on being able to make this distinction.

**Note:** How animals do this is still *unknown* for general illumination conditions! [That's a major understatement; put another way: we're still far away from having a general computer vision system.]

NOTE: We perceive a world of objects and events, **not** of light. (Gibson, 1950, 1966, 1979)

\* E.g. cast shadows; variable distances from source, . . .

### BRIGHTNESS CONSTANCY



**REFLECTANCE: A surface property**

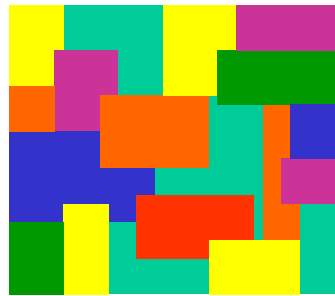
We seem to estimate  
the *ratio* of reflectances across *boundaries*.

Physical reflectance is a surface property that is constant over time, specifically w/r/t variable illumination.

It would be nice if our perceptions of surface color were similarly constant.

Note: We will confront two definitions of "reflectance;" besides **physical reflectance** (radiance as a ratio of incident illumination), we have **Grossbergian reflectance**, the *ratio of the magnitude of input to one node in a network relative to total input to the network*.

**LAND'S EXPERIMENTS IN COLOR CONSTANCY**



(Shading within boxes stands for colored pigments.)

red (Land, 1971)

green

blue

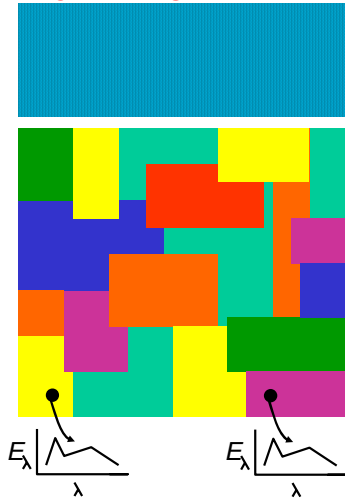
**Land -- McCann "Mondrians," I**

FIRST EXPERIMENT: If the intensity of the red (long wavelength) illuminant is doubled or tripled, the colors in the Mondrian still look "much the same."

We somehow factor away the "extra" red.  
Helmholtz: **"Discounting the illuminant"**  
(wavelength and intensity)

**LAND -- McCANN MONDRIANS**

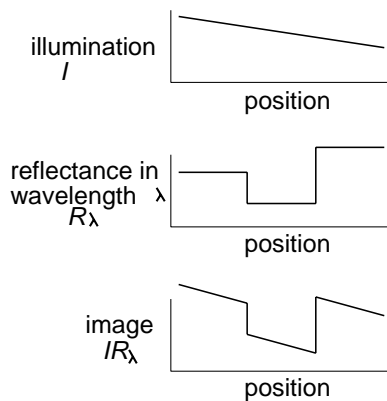
Use different illumination gradients in different wavelengths, adjusted so as to "offset" the effects of spectral reflectance of two patches.



*Different* colors seen from the *same* spectrum

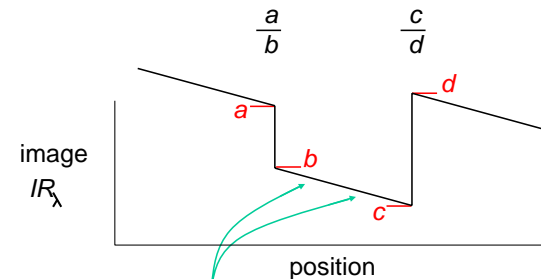
... similar to those seen in white light

**GRADIENTS OF ILLUMINATION AND REFLECTANCE**



**"RETINEX" STRATEGY**

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



2. Suppress information from slowly varying region interiors.

How to go from  $IR_\lambda$  to  $f(R_\lambda)$ ?

Ideally, some *simple* function

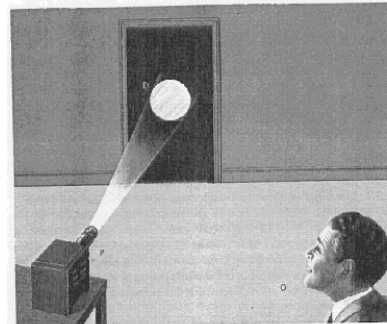
To be able to “discount” the illuminant (intensity) with a Retinex-like strategy, the *illumination gradient must be more gradual than the gradient of change in reflectance.*

This is not true in the *Gelb effect*.

CRUCIAL QUESTION: How realistic in general is the assumption that illumination effects on a visible scene vary more gradually (across space) than surface effects (e.g. pigment changes)?

Exercise: Look around you and point to the surface of an object whose appearance would be closest to that of a moon rock, if that moon rock were in the room with you now.

## THE GELB EFFECT

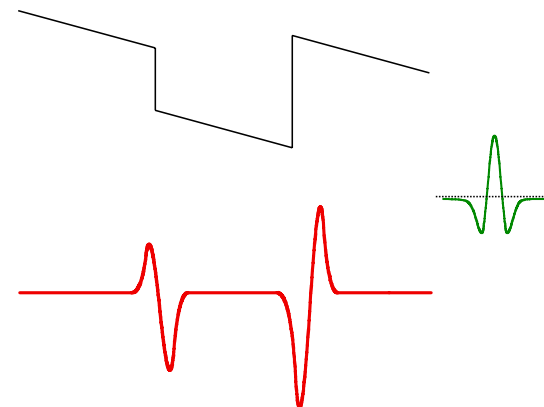


From Kaufman, 1974.

Figure 5-1. The Gelb experiment. A spinning black paper disc D is illuminated by a projector P. Stray light enters the room behind the disc so that it cannot be seen by the observer O. The disc appears luminous except when a small piece of white paper is placed in the path of the projector light.

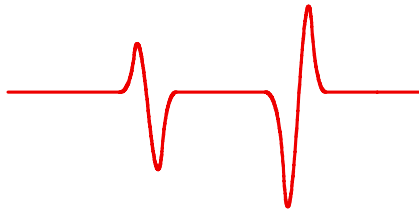
Figure 5-1. The Gelb experiment. A spinning black paper disc D is illuminated by a projector P. Stray light enters the room behind the disc so that it cannot be seen by the observer O. The disc appears luminous except when a small piece of white paper is placed in the path of the projector light.

## RATIOS FROM EARLY BIOLOGICAL PROCESSING



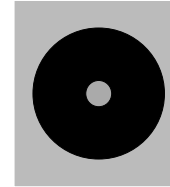
## RATIOS NOT ENOUGH

If early representation looks like this



Do we need “**filling-in**” for perceived surface interiors?  
otherwise, would we “see” a world of line drawings?

## BRIGHTNESS CONTRAST



**Normalization** *within functional spatial domains* --  
total “energy” in a representation is conserved,  
as the sum of inputs varies.

Not just “lateral inhibition” -- but “anchoring”.  
Gilchrist

## BRIGHTNESS CONTRAST

Two small disk patches of equal luminance.  
One annulus is of high luminance.  
The other is of low luminance.

### **Contrast effect:**

Perceived brightness of inner disk  
varies in direction opposite to  
luminance of annuli.

Claim: This occurs because some  
representation of the “sum of ratios” of inputs  
for the two scenes is *approximately constant* for each view.

That is, **normalization** occurs, whereby  
-- *for some functional spatial domain* --  
the quantity of total “energy” in a representation (or output  
of some network) is conserved, as the sum of inputs varies  
over some range. (Cf. Simulation Assignment 1.)

## RETINA

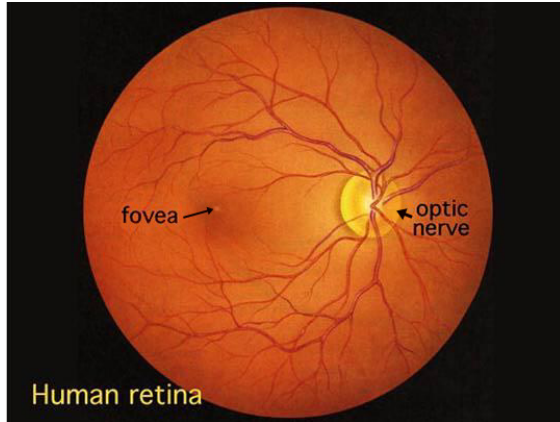
The **retina** is the “interface” between a mammal and its visual  
environment.

Starting from the interface, research on visual perception could  
be motivated by:

- 1) a detailed analysis of the visual system's mechanisms or
- 2) a detailed analysis of the visual environment.  
(cf. Gibson, 1950, 1966, 1979; Marr, 1982)

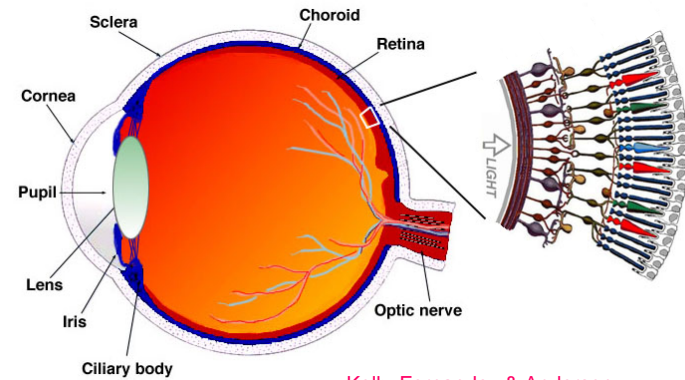
Ultimately, these two lines of inquiry would converge, because  
the visual environment has shaped the evolution of our visual  
systems.

**RETINA AS INPUT DEVICE, I**



Lines and edges are registered . . . *poorly!*

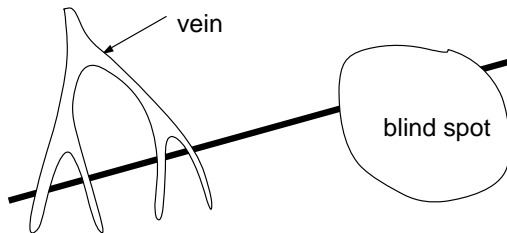
**PATH OF LIGHT**



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

**RETINA AS INPUT DEVICE, II**

*Pattern formed on retina* by a dark line



. . .is not even *connected!*

**Completion** needed for "*real*" contours

*Note: this statement is still controversial.*

**COMPLETION OVER THE BLIND SPOT**



**DEMO:** Close right eye and fixate upper cross with left eye.

Hold page at about 1 ft from the eye, and move it back and forth in depth slightly until the disk on the left disappears.

Similarly, when fixating the lower cross, the gap in the black line can be made to fall on the blind spot, and the line is seen as continuous.

Adapted from Kandel & Schwartz's (1985)  
 adaptation from Hurvich, (1981)

**EYE MICROMOVEMENTS**

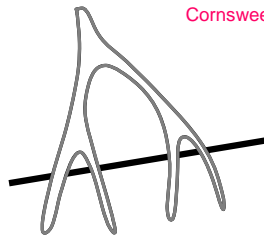
The eye jiggles constantly in its orbit.  
 (“*tremor*,” approximately 40 Hertz\*)

The shadows of retinal veins *do not move* relative to the photoreceptor mosaic.

They therefore form “stabilized images.”

But **stabilized images fade**.  
 (Time scale: seconds)

Cornsweet, 1970; Krauskopf, 1963, Ratliff, 1965; Yabus, 1967



**Note:** While the stabilization of veins accounts for our not “seeing” them, the line that we *do* see still makes an *incomplete* pattern on the retina.

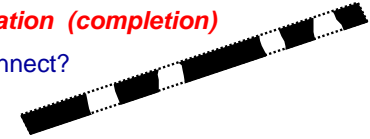
\* Local expert: Prof. Rucci

**COMPENSATION FOR RETINAL GAPS**



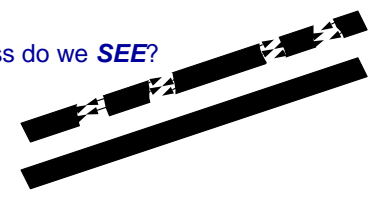
**Emergent boundary formation (completion)**

\*Which boundaries to connect?



**Featural filling-in**

What color and brightness do we **SEE?**



\* Not the same as “recognition”

Examples of emergent segmentation:

- 1) Collinear (Beck, Bozzi)
- 2) Perpendicular to line ends (Beck, Ehrenstein, Varin neon)
- 3) Diagonal, oblique (Beck, Kennedy)
- 4) “Other” (Beck’s T’s and L’s, dalmation)

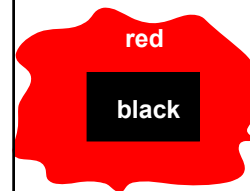
The contrasts and hues that we “see” are not the only relevant perceptual structures.

We segment and group images on the basis of “emergent boundaries,” that are *based on* image contrasts, but *not isomorphic* to the pattern of contrasts.

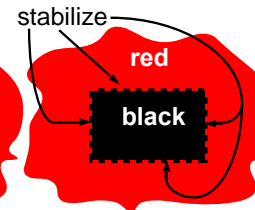
Why?

**BOUNDARIES AND FILLING-IN**

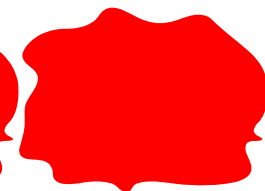
An image like this



when stabilized on the retina,



looks like this:



NOTE: time scale of **seconds**

**STABILIZED IMAGE EXPERIMENTS**

How to stabilize an image. . . the hard way!

Caps needed to be of low mass and low moment of inertia; why?



Fig 24 from **Yarbus (1967)**

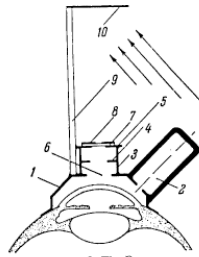


Fig. 20. The P<sub>5</sub> cap.

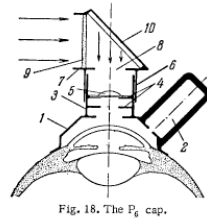
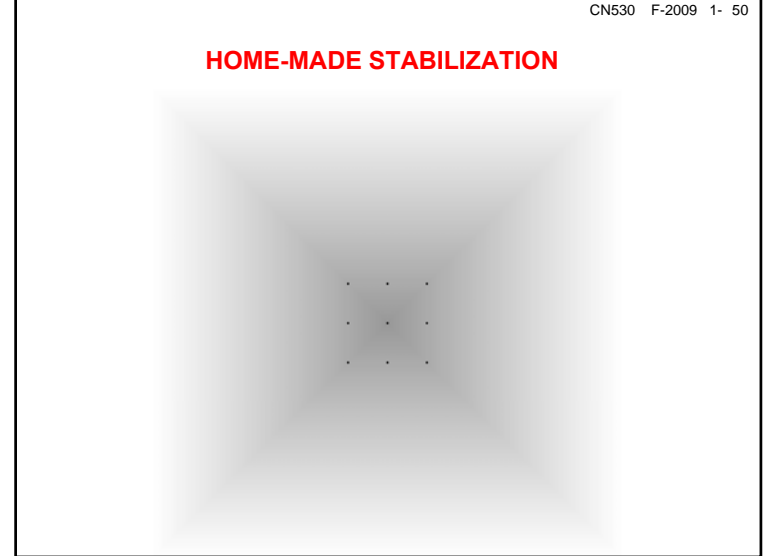


Fig. 18. The P<sub>6</sub> cap.

**HOME-MADE STABILIZATION**



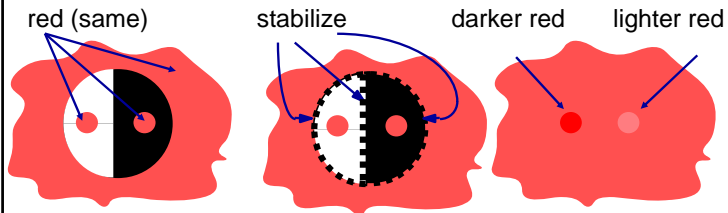
**YARBUS EXPERIMENT, 1967**

A red dot moves back and forth.

When indicated boundaries are stabilized, . . .

. . . white and black are **no longer visible**, but

. . . the effects of **contrast remain!**



**INFERENCES FROM YARBUS EXPERIMENT**

**Boundaries restrict featural filling-in** (color and brightness).  
When boundaries fade, colors flow.  
Boundary fading does not imply color fading.

**Two subsystems in early vision:**  
**boundary system**  
**feature system** (brightness, color)

Note: The above nomenclature owes to a specific theoretical approach (i.e., Boundary Contour System/Feature Contour System). This approach is not universally agreed upon. It's basic premises are, however, widely known, and many researchers adopt a similar distinction (e.g. "borders" and "surface color") without "buying into" all details of the theory.

**BOUNDARIES AND VISIBILITY**

Q: What happens if there are **no boundaries** in a visual field?

**Ganzfeld\***: Completely homogeneous visual field -- can be approximated by wearing goggles made from halves of a ping-pong ball.

A: **When viewing a ganzfeld, brightness fades.**  
 Note: the time scale of fading is **long** (minutes).

Note: Homogeneous regions of any image are *de facto* stabilized.  
 e.g. "clear blue sky"

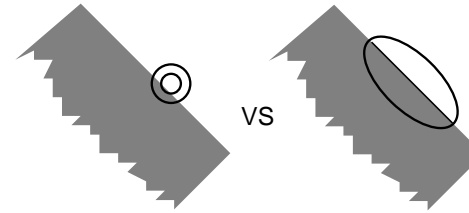
Homogeneous regions carry no information.

\*As the term is used in vision research.

**BOUNDARY PROCESSING**

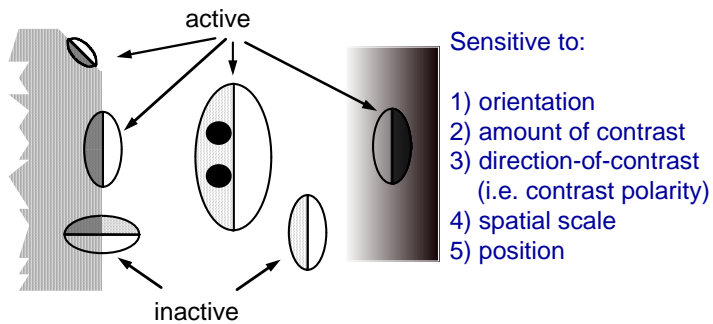
Boundaries: how to **detect**, **sharpen** (sometimes) and **complete**?

Unoriented and **oriented receptive fields** (masks, filters, kernels)



Hubel & Wiesel, 1962 . . . Nobel Prize

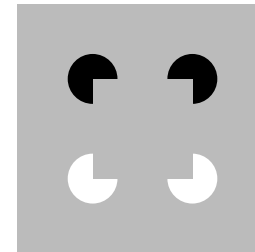
**SIMPLE CELLS: ORIENTED LOCAL CONTRAST FILTERS**



**Not** (just) "edge detectors."

**REVERSE-CONTRAST KANIZSA SQUARE 1**

Local oriented-contrast filtering alone is *not* enough:



Cf: Shapley & Gordon, 1985

Again, boundary ≠ brightness.

**REVERSE-CONTRAST KANIZSA SQUARE 2**

Evidently, while sensitive to:

- 1) orientation,
- 2) amount of contrast,
- 3) spatial scale, and
- 4) position,

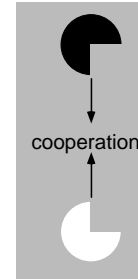
human boundary processing is



**insensitive** to *direction of contrast*  
(i.e., **contrast polarity** of “edges”),

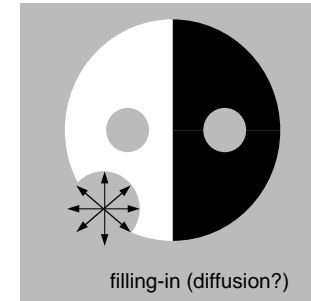
when completing over gaps  
(i.e., over a sufficiently large spatial scale.)

**COMPLEMENTARITY**



**Boundaries: Completion**

oriented  
inward  
insensitive to  
direction-of-contrast

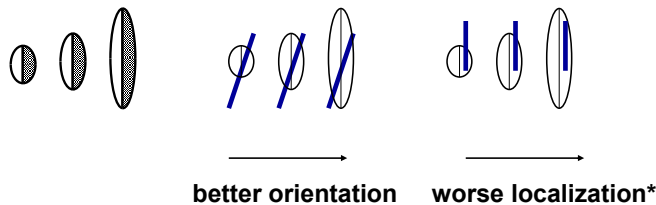


**Surfaces: Filling-in**

unoriented  
outward  
sensitive to  
direction-of-contrast

**UNITS OF VISUAL REPRESENTATIONS**

**UNCERTAINTY:** Some kinds of information are *incompatible*; they can not be represented simultaneously by a single unit.



**QUANTIZATION:** How to maximize resolution, given that vision can not “do calculus” with infinitesimals?

\*As we will see in a few weeks, specialized “endstopped” cells address this problem in part.

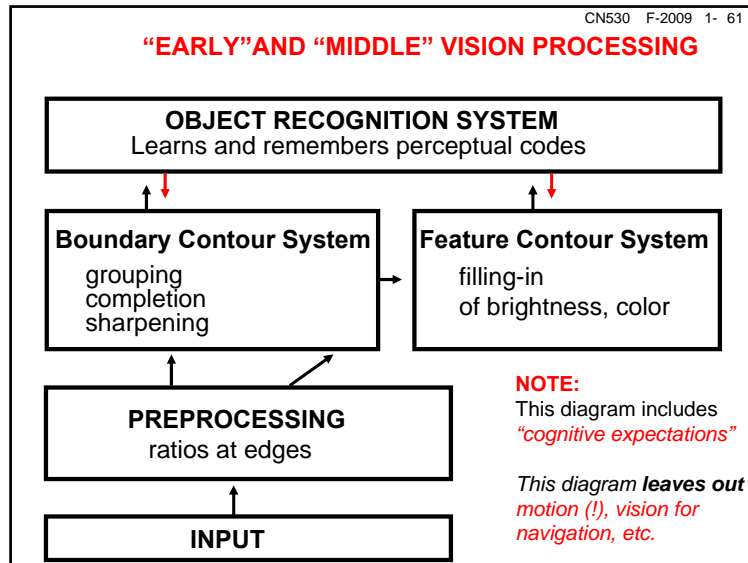
**UNCERTAINTY AND QUANTIZATION**

A thin perceptual line is, in initial processing, a fuzzy/statistical object, NOT a union of points, owing to the uncertainty inherent in using receptors (or “filters” or “detectors”) of finite size.

Even a thin perceptual line embodies aspects of *boundary* and *feature qualities*  
e.g. location (“where”) vs color (“what”)

Oriented filters are NOT (just) edge detectors

Their inherent uncertainty (w/r/t localization of “edgel”) “feeds into” mechanisms that deal with *plausible grouping directions* to yield a result that is better than can be gotten from *local information* alone.



CN530 F-2009 1- 62

### Where’s the “intelligence” in vision?

Helmholtz: “as if” unconscious inference  
 Gestaltists: field theory  
 Gibson: direct perception; “pickup” of environmental information  
 Marr: computational theory  
 Grossberg: network architecture & resonance  
 Schwartz: maps  
 You: \_\_\_\_\_

CN530 F-2009 1- 63

### WHAT ARE THE UNITS OF VISION?

This course is about the geometry of vision.  
 (Recall Lettvin quote.)

*In the visual universe, we have, at times:  
**coherence** among disconnected elements, and  
**segmentations** through homogeneous regions!  
 What kind of geometry does this?  
 What could **possibly** be its functional units?*

Consider: The “preattentive” visual system at times *makes use* of direction-of-contrast information and at other times *ignores it*, thereby rivaling the performance of the most efficient cognitive engine yet invented: *the thermos bottle!*

CN530 F-2009 1- 64

### Addendum

7) The noise--saturation dilemma  
 8) Reflectances and ratios; shunting and mass action

The preceding panels form a unit describing some basic issues in vision.

The panels that follow begin the task of modeling.

**THE PLAN (including next week)**

*Feedforward shunting competitive networks*

*Distance-dependent interactions*

Functional view of network properties:

*motivated* by physiology  
*described* abstractly.

Try to understand certain *molar* perceptual  
(or “computational”) properties  
with reference to simplified *micro*-interactions.

**How can center-surround networks compute ratios?**

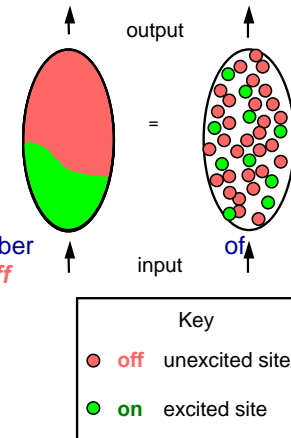
Two parts to this question:

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

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

Infinity does not exist in biology.

**2. Why do neurons compete?**



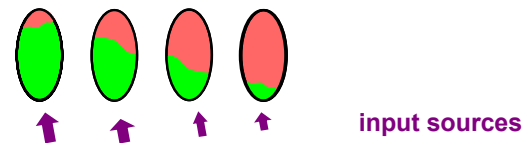
Competition is ubiquitous in the biological world.

Why competition on the *cellular* level?

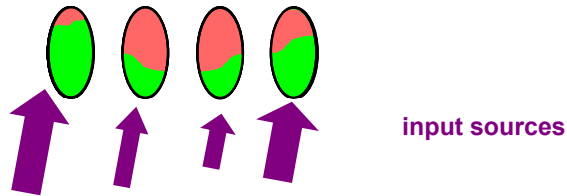
*What is a cell*, computationally speaking?

*Infinity* does not exist in biology;  
so our neural models should not assume  
that it does -- not even *implicitly*.

**Pattern Processing by Cell Networks, A**



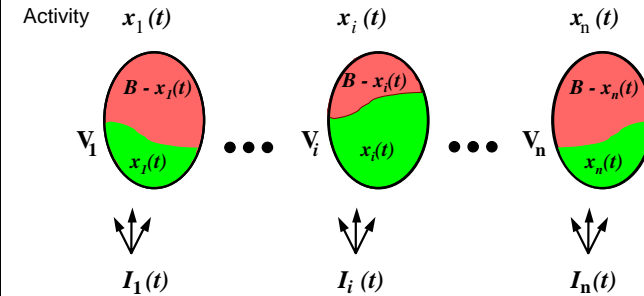
### Pattern Processing by Cell Networks, B



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**  $B - x_i(t)$

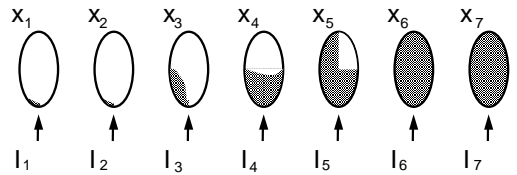
### NOISE-SATURATION DILEMMA

Pattern registration by cellular systems

Grossberg, 1973

Given the constraints of a finite number of sites with all-or-none processes, how does nature prevent the following?

Let:  
 $I_n = x^n$   
 $I$  stands for intensity, or energy of input.

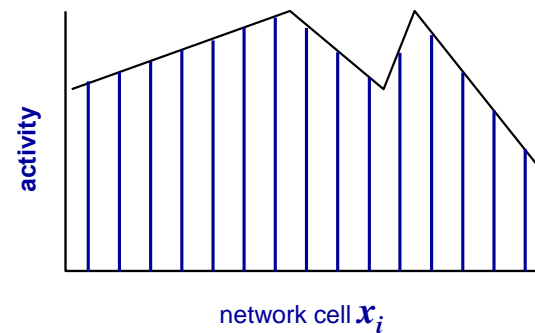


range in which input signals cannot be distinguished from endogenous noise  
 effective operating range  
 saturation range

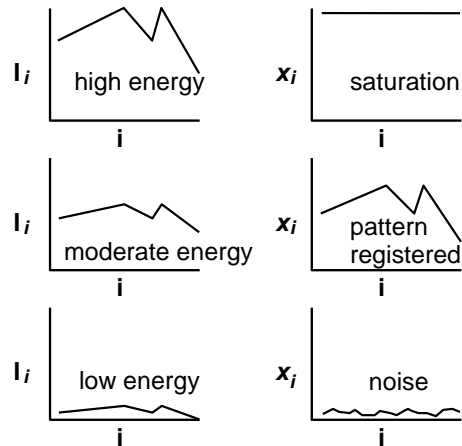
If  $x_i$ 's are **sensitive to small** inputs, why don't they **saturate** in response to large inputs?

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

### Graphical convention



**NOISE-SATURATION DILEMMA, AGAIN**

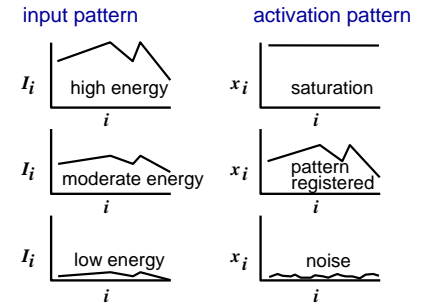


**NEED:** automatic gain control (retuning of sensitivity)

**Noise-Saturation Dilemma**

**Problem:**

remain sensitive to input ratios  $\theta_i = \frac{I_i}{\sum_j I_j}$  as total input  $I = \sum_j I_j \rightarrow \infty$

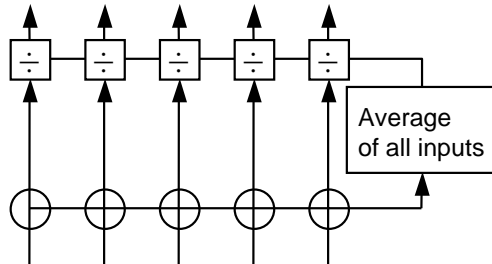


**Solution:**

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

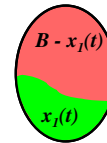
**HARDWARE AND WETWARE**

The conventional engineering solution for an automatic gain control . . .



would be physiologically "expensive."

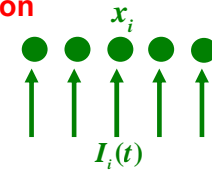
**Shunting Saturation**



A, B are constants

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

(a) (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)

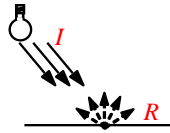
Response to a spatial pattern of inputs:

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

- $\theta_i$  relative intensity (cf., reflectance)
- $I(t)$  total intensity (cf., luminance)

**PHYSICAL REFLECTANCE**

Classical definition:  
the **ratio** of reflected (*R*)  
to incident (*I*) light for a  
Lambertian (i.e., ideal,  
perfectly diffusing) surface.



$0 < R/I < 1$  for each surface

a.k.a: *albedo*

**Jargon alert!!!**

COMPARE:

Grossberg's definition of **reflectance** of network inputs:

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

Total input:  $I = \sum_j I_j$

"Reflectance"  
of an input:  $\theta_i = \frac{I_i}{I}$

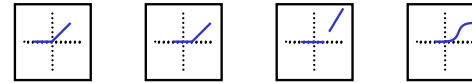
**Some Usage Conventions**

Notation:

$v_i$  *name* of the *i*'th node

$x_i$  *potential* or *activity level* of the *i*'th node

$f(x_i)$  *output* signal function of the *i*'th node  
(often involving, e. g., a rectification,  
thresholding, or sigmoidal transform)



**NOTE:** Signal functions will **NOT** "be an issue" in CN530;  
Your assignments ask you to plot the value of  $x_i$  only.

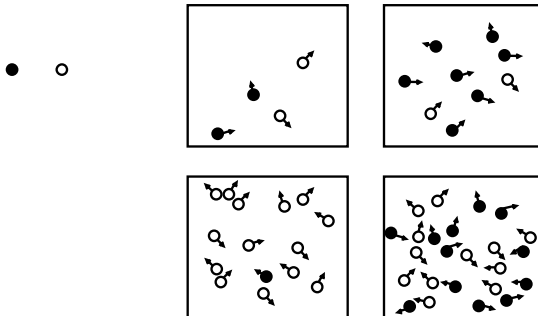
What about Grossberg's phrase "cell or population"?

**MASS ACTION**

MULTIPLICATION because of *statistical independence*

Expected value:  $E(X \text{ AND } Y) = E(X)E(Y)$

Reaction rate for particles of two types in a medium?



Analogy: Consider a chemical reaction occurring among  
particles of two types that are *sparse*ly distributed in a  
fluid container (e.g. yeast and sugar, for fermentation).

Reaction occurs *only* upon collision of particles of different types.

Random process: *Rate* of reaction is proportional to *product*  
*of the concentrations* of the two types of particles.

In network equation, *multiply* input size by measure of  
available (inactive) sites, . . .

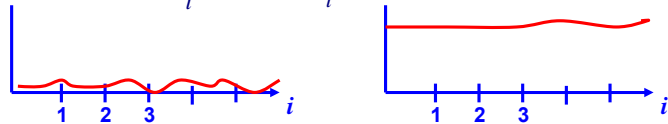
because there are many sites, and a single input is thought  
of as arriving at particular sites randomly along similarly  
numerous, "all or none" channels.

### Shunting Saturation

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

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

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



I small: lost in noise\*

I large: saturates

\*Note: There's no noise in the equation

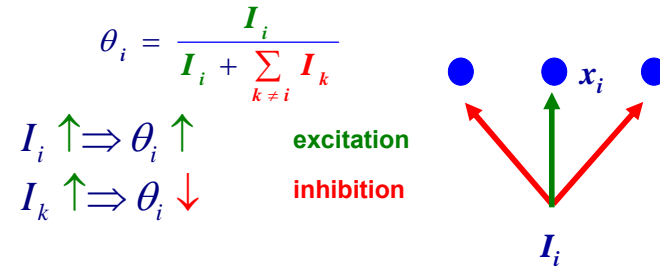
Sensitivity loss to **relative** intensity as **total** intensity increases

### Computing with Patterns

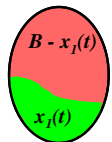
How to compute the **pattern-sensitive** variable:

$\theta_i = \frac{I_i}{\sum_{k=1}^n I_k}$  ? the **ratio** of **one** input to the **sum** of **all** inputs

Need interactions! What type?



### Shunting Dynamics



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

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

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

### Effects of Shunting Inhibition

At equilibrium:  $x_i = \theta_i \frac{BI}{A + I} \quad x_i \rightarrow \theta_i \text{ as } I \rightarrow \infty$

**PATTERN ENERGY**  
"factorization"

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

Total input:  $I = \sum_j I_j$  **ratio sensitivity** over a wide dynamic range:

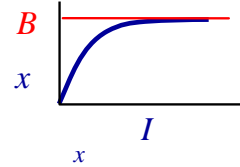
Normalized input:  $\theta_i = \frac{I_i}{I}$  **automatic gain control**

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

But NOTE: the "ON-center" here is a single node; later we will generalize to spatially "wider" centers.

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**

### THE NOISE-SATURATION DILEMMA AND VISION

CONSTANCY?	CONTRAST?	MONDRIANS?
<p><b>Reflectances:</b></p> <p>Compute ratios . . .  <b>. . . relative to a given spatial domain.</b></p> <p>e.g. this input / <b>total</b> input,  OR this input / <b>nearby</b> inputs</p>	<p><b>Normalization:</b></p> <p>Conserve a quantity of total activity. . .  <b>. . . relative to a given spatial domain.</b></p> <p><b>Tricky:</b> Is the <b>domain</b> the <i>entire scene</i>, or some spatially limited kernel (or "receptive field")</p>	<p><b>Discount the illuminant</b></p> <p>"Clip off" effect of average or total energy; ignore homogeneous regions  <b>. . . relative to ??</b></p> <p><b>MUCH WORSE!</b></p>

The previous panel's point can be put like this:

Normalization is easy; knowing when and where to normalize and when and where **to not normalize** is hard.

A current leading theory of lightness perception (Gilchrist *et al.*) is based on interactions among "frameworks" -- i.e. regions of the scene within and between which ratios of luminance are to be compared.

Good news: The formula's for computing the ratios and related quantities are pretty well worked out.

Bad news: There is nothing like an effective procedure for determining **what the frameworks are** for even modestly complex "scenes."

### What are the units?

Modeling visual *appearance* (including, e.g., brightness or lightness) requires an understanding of **perceptual unit formation**, which is one of the fundamental challenges of vision.

Unit formation is a much more subtle a process than any based on a "conventional" view of the early visual system as a **feedforward hierarchy of slabs of filters**.

**Shunting neural networks**, including those with (various kinds of) feedback, offer a more promising formalism for approaching unit formation.