CSE 332
Introduction to Visualization

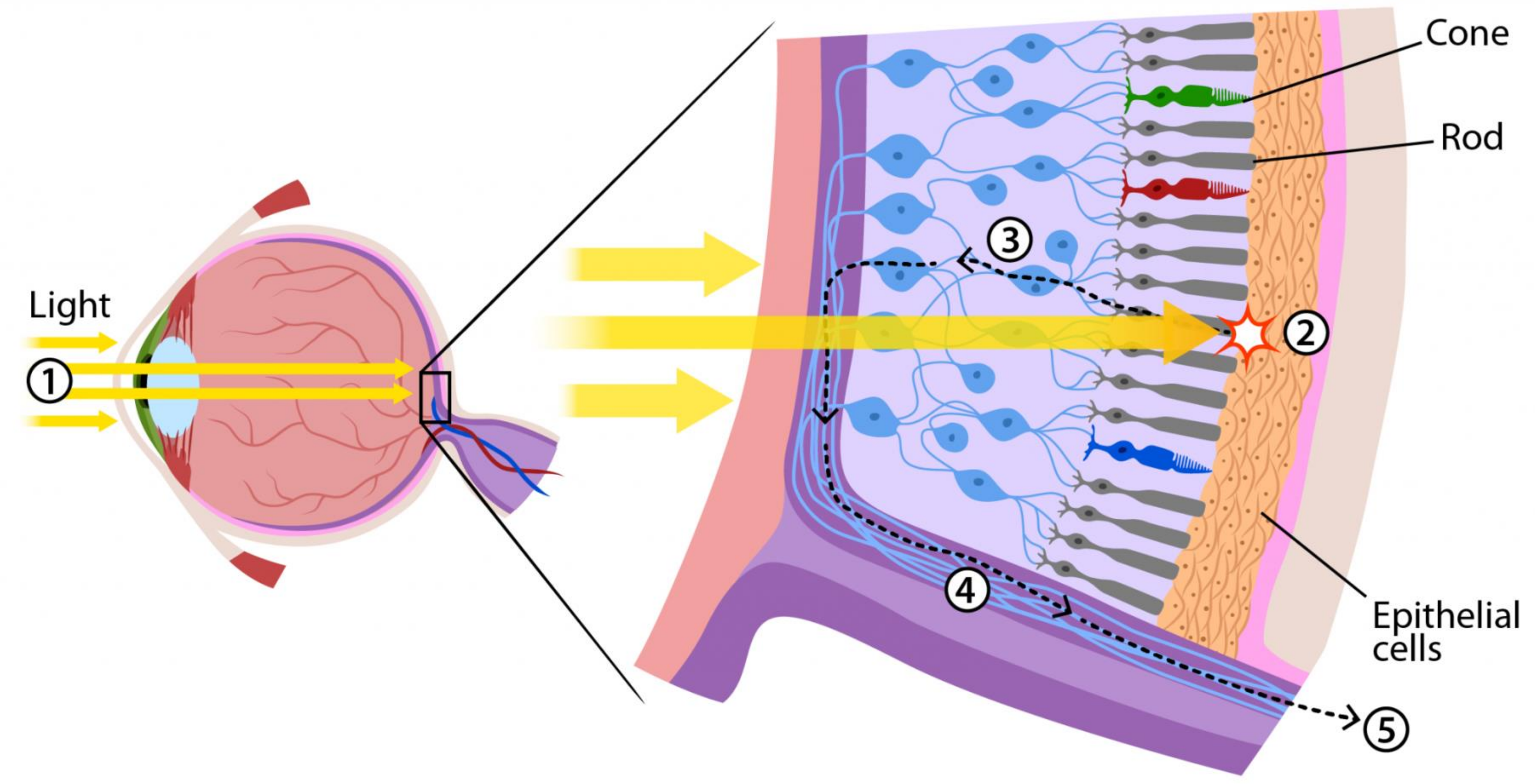
Visual Perception and Cognition

Klaus Mueller

Computer Science Department
Stony Brook University

Lecture	Topic	Projects
1	Intro, schedule, and logistics	
2	Applications of visual analytics, data, and basic tasks	
3	Data preparation and reduction	Project 1 out
4	Data preparation and reduction	
5	Data reduction and similarity metrics	
6	Dimension reduction	
7	Introduction to D3	Project 2 out
8	Bias in visualization	
9	Perception and cognition	
10	Visual design and aesthetics	
11	Cluster and pattern analysis	
12	High-Dimensional data visualization: linear methods	Project 3 out
13	High-D data vis.: non-linear methods, categorical data	
14	Computer graphics and volume rendering	
15	Techniques to visualize spatial (3D) data	
16	Scientific and medical visualization	
17	Scientific and medical visualization	
18	Non-photorealistic rendering	Project 4 out
19	Midterm	
20	Principles of interaction	
21	Visual analytics and the visual sense making process	
22	Visualization of graphs and hierarchies	
23	Visualization of text data	Project 5 out
24	Visualization of time-varying and time-series data	
25	Memorable visualizations, visual embellishments	
26	Evaluation and user studies	
27	Narrative visualization and storytelling	
28	Data journalism	

The Human Eye



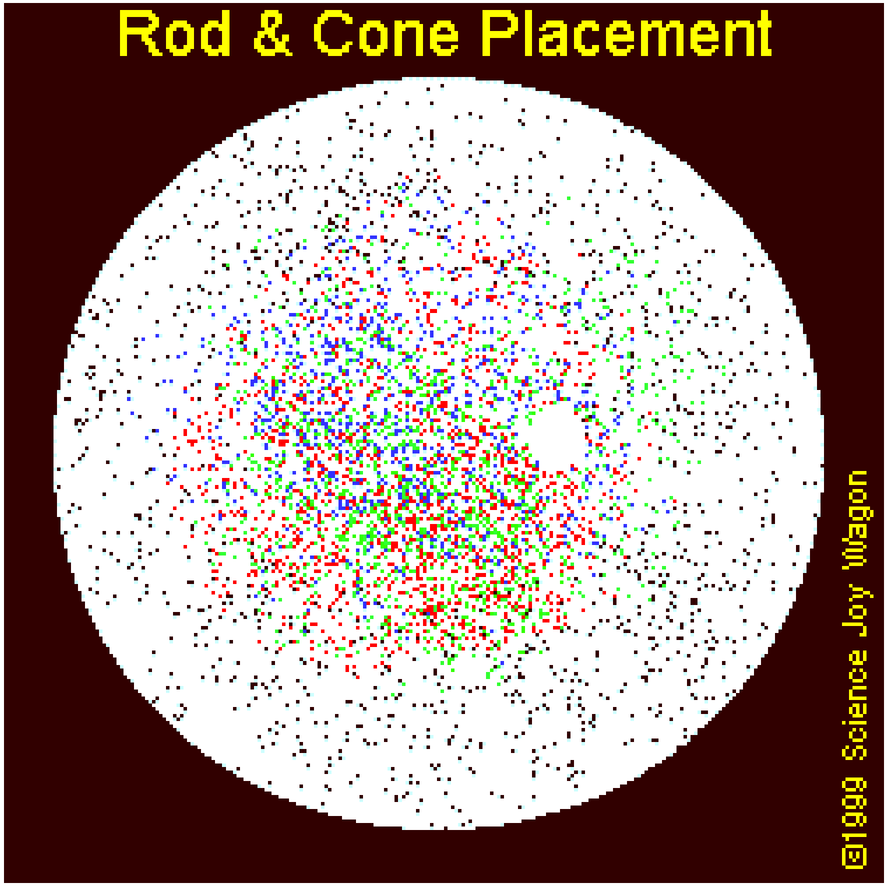
The discs of rods hold rhodopsin and the discs of cones hold photopsin. Both of these photoreceptor proteins are special molecules that change shape when activated by light. This shape change allows the proteins to activate a second special protein molecule that then starts causing other changes involved in sending a visual signal. For the signal to be sent through the cell, charged molecules called ions are let in and out of the cell in an action potential.

https://askabiologist.asu.edu/rods-and-cones

Retinal Distribution of Rods and Cones

What can you observe here?

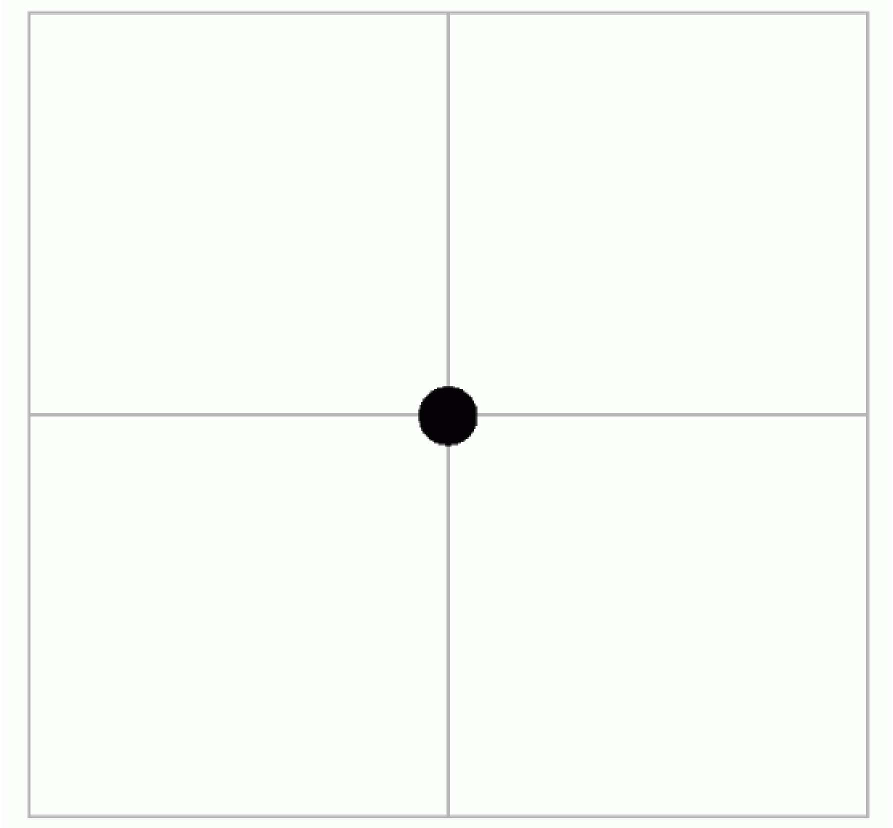
- color (cones) in the center
- grey (rods) outside, too
- more grey
- more green
- blind spot



Successive Contrast

Focus on the black circle for a few seconds, then switch to one of the white fields.

What do you see?

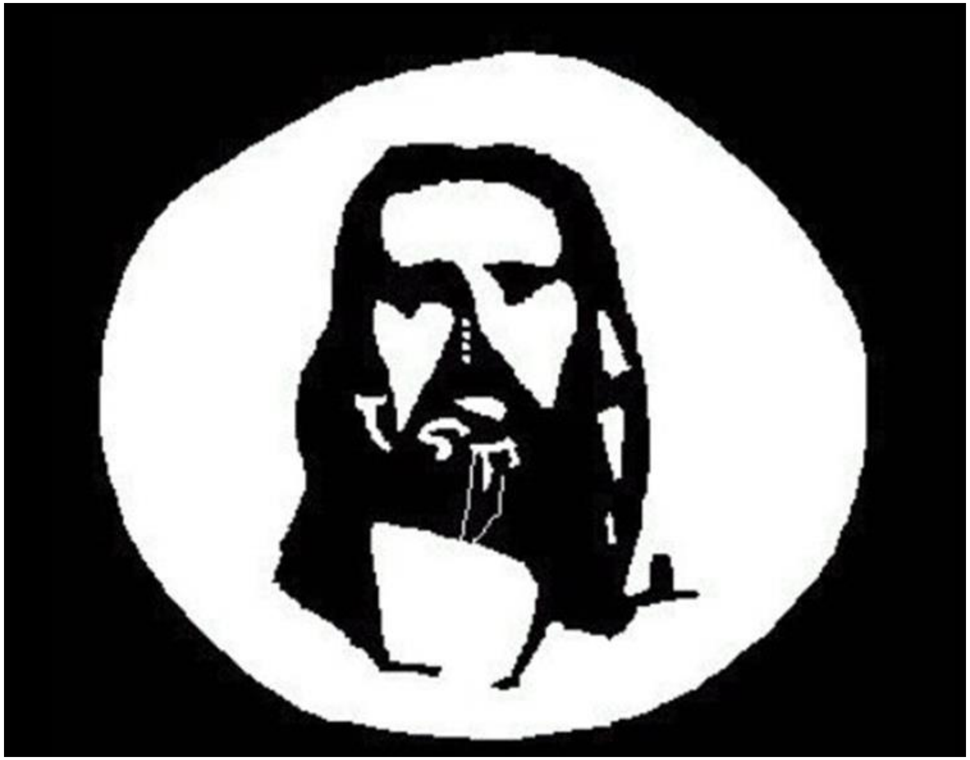


Recall This Optical Illusion

Follow the instructions:
1) Relax and concentrate on the 4 small dots in the middle of the picture for about. 30-40 secs.
2) Then, take a look at a wall near you (any smooth, single coloured surface)
3) You will see a circle of light developing
4) Start blinking your eyes a couple of times and you will see a figure emerging…
5) What do you see? Moreover, who do you see?

Inverted...

Is this what you saw?



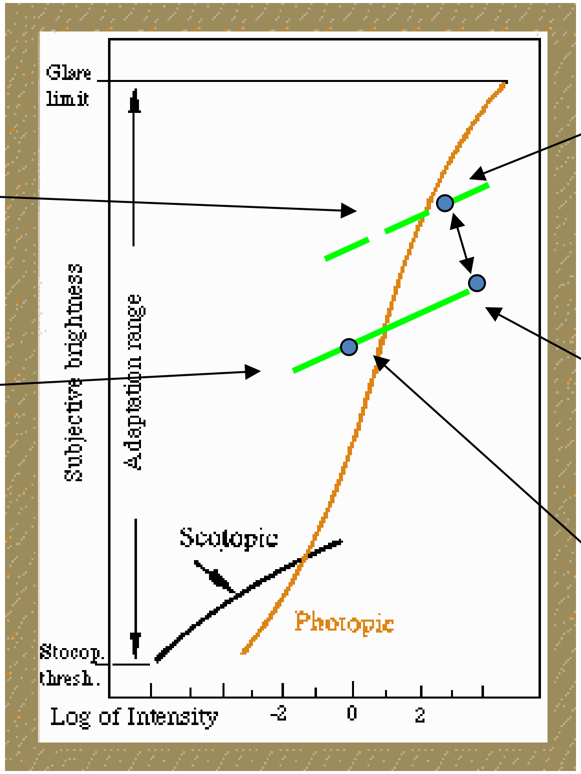
Explanation

While the retina can perceive a high range of intensities, it cannot handle all simultaneously

- at any given time, each region adapts to a small intensity range determined by the local intensity
- that is why you have to wait a while when you step from a bright into a dark room (say, a dark movie theater from a brightly lit lobby)

after moving the eye: eventually adapted range

currently adapted range

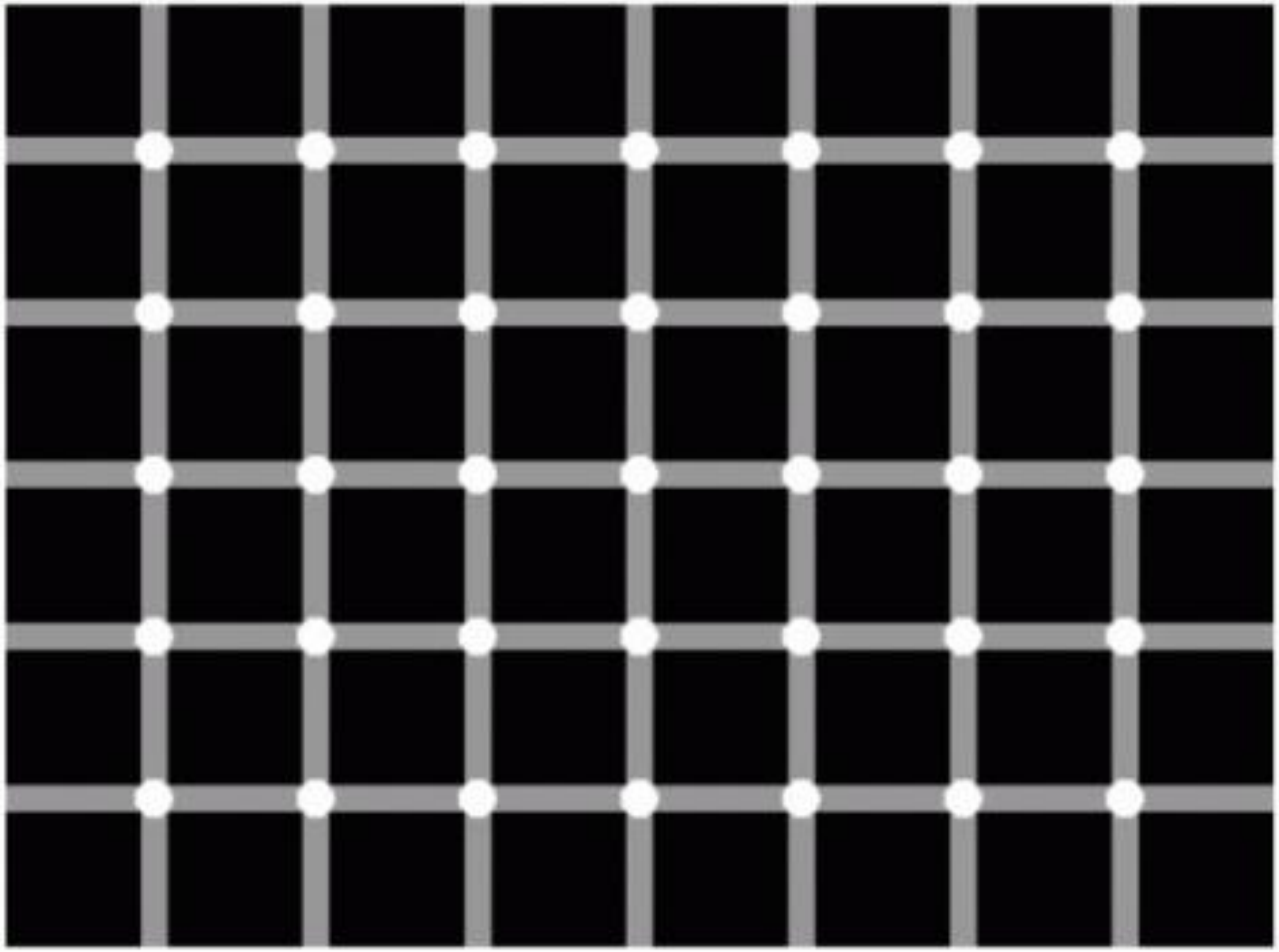


eventually the bright area intensity is unsaturated, matches neighborhood (which was already adapted here before)

after moving the eye: new bright area saturates intensity perception

current dark area in picture falls here

Herman Grid Illusion

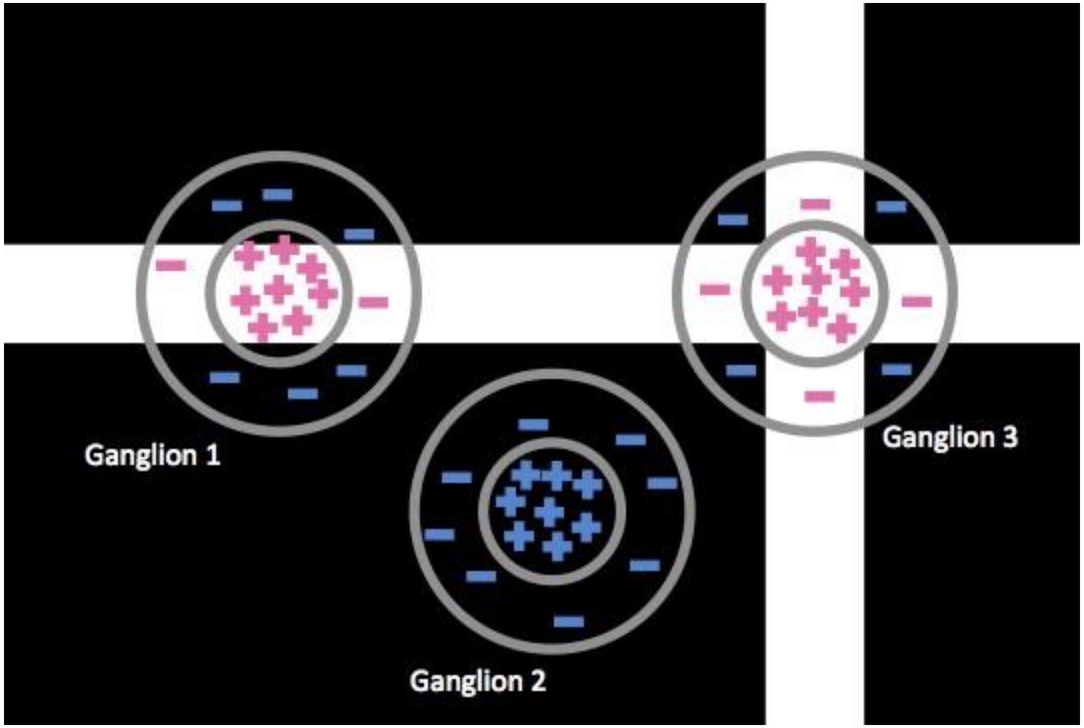


Explanation

The reason lies in the center/surround organization of the Ganglion cells in the receptive field

<u>Ganglion 1</u>
10/16 inputs exposed to light
- 8 are excitatory
- 2 inhibitory
- -> 6 stimulated

<u>Ganglion 2</u>
no exposure
-> no stimulation



<u>Ganglion 3</u>
12/16 inputs exposed to light
- 8 are excitatory
- 4 inhibitory
- -> 4 stimulated

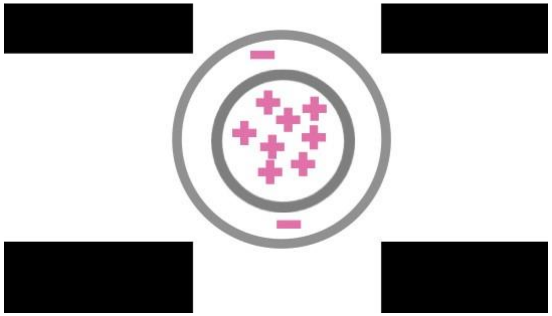
Ganglion 1 senses brighter than ganglion 3
-> that’s why the line intersections appear grey

But….

Why do the dark spots disappear as soon as you look directly at them?

It's because:

- our central vision is sharp and clear, allowing us to resolve details with great accuracy
- ganglion cells close to the fovea have a very small receptive field, with fewer inhibitory inputs
- therefore, at the fovea, there is less inhibition of the center by the surround, and the dark spots disappear

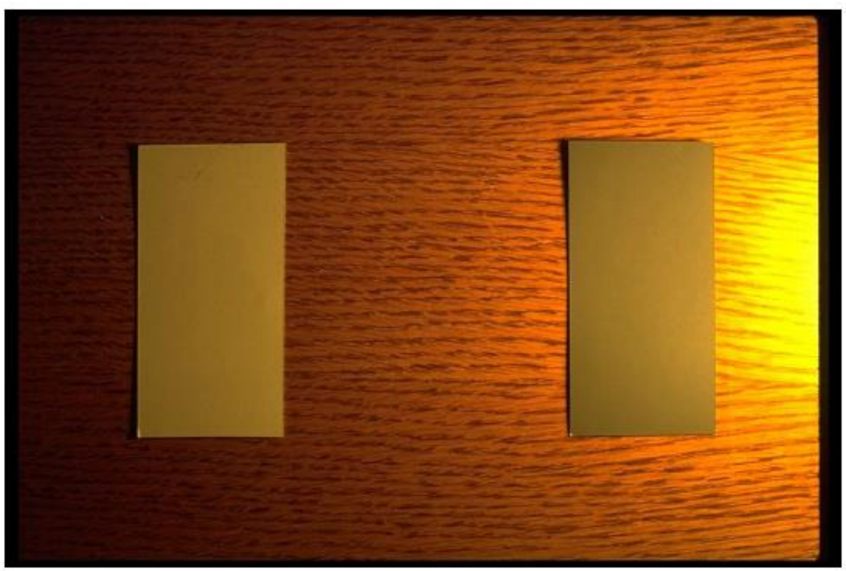


Read more here

Dynamic Range Contrast

Local adaption level varies, which changes the relative contrast of the objects in the local scene

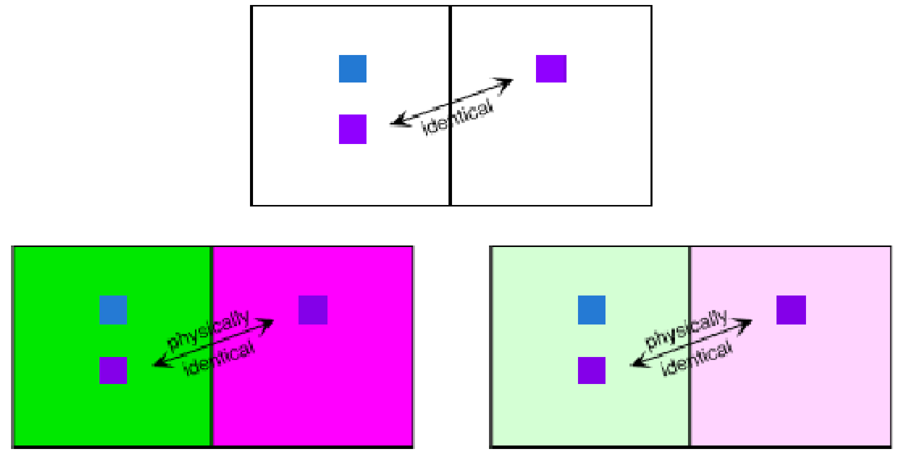
Are these two strips the same or different?



Surround Matters
Especially for Colors

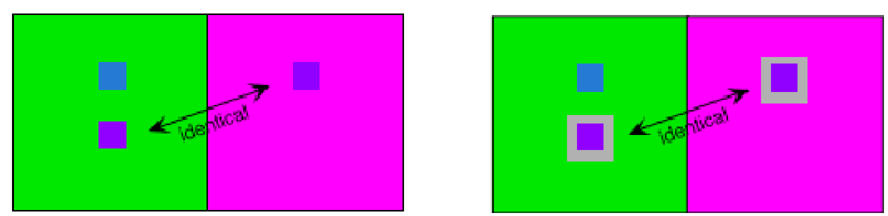
Compare these three panels

- white background
- saturated background
- non-saturated background



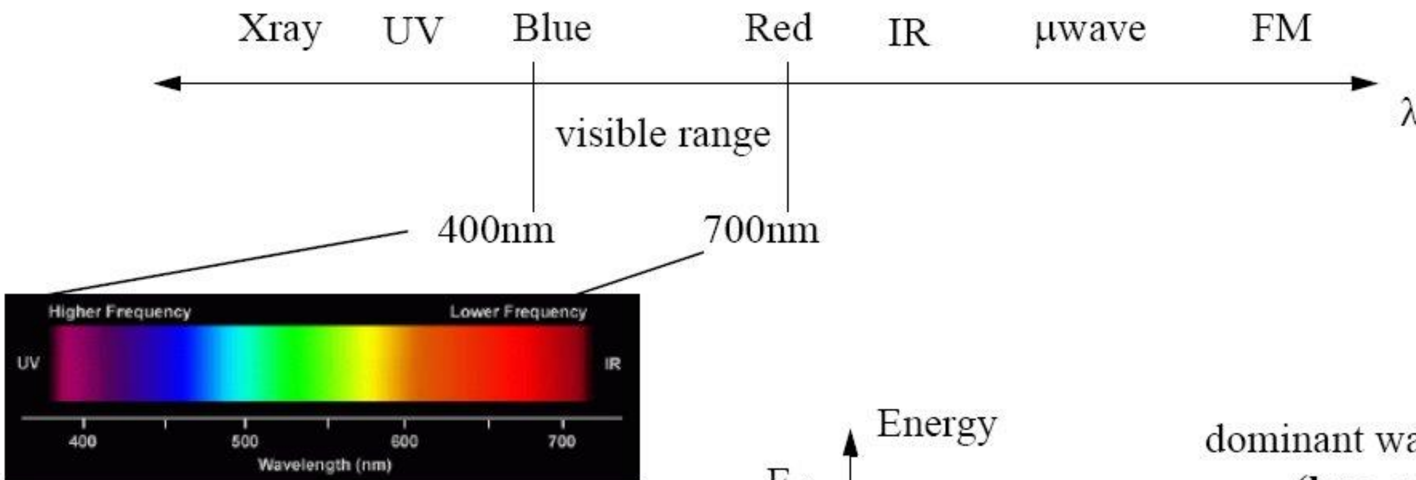
Guidelines:

- use saturated colors sparingly
- they may cause undesired effects
- neutral borders can help

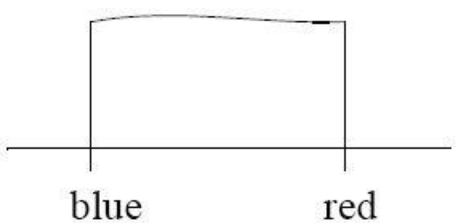


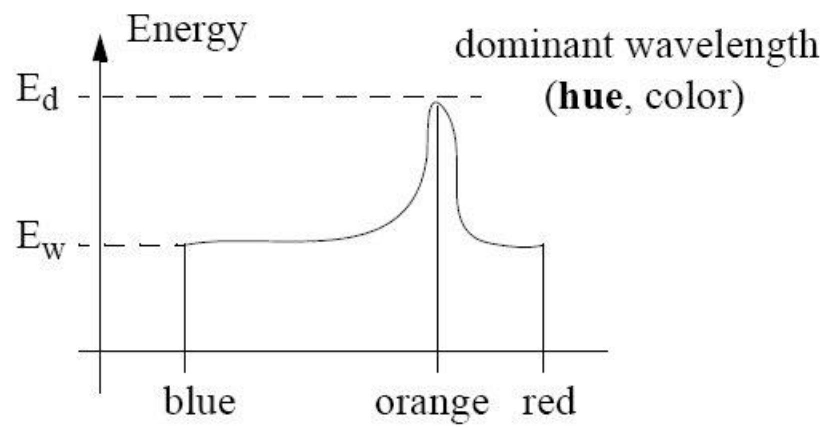
Spectrum of Wavelengths

Spectrum:



white light: all wavelengths



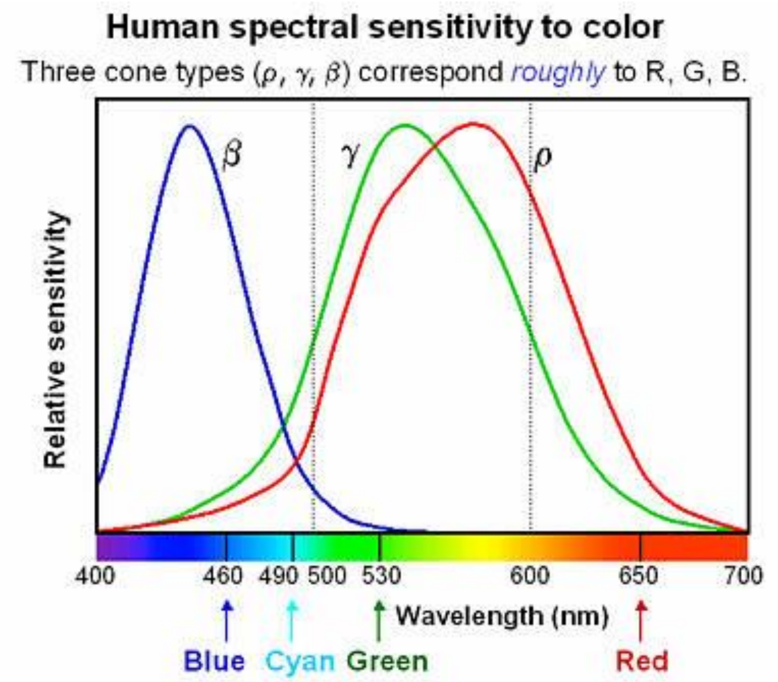


$E(\lambda)$: overall intensity = **brightness** or **luminance**

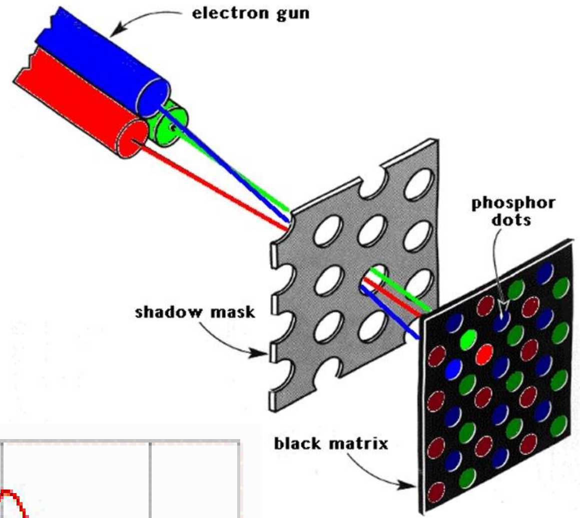
ratio of E_d vs. E_w = **saturation**

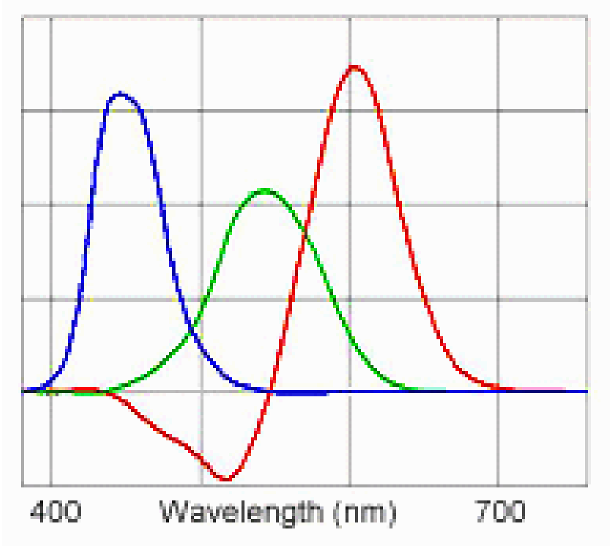
The human eye differentiates about 300 hues and 100-150 luminance variations

Perception Curves



human color sensitivity curves

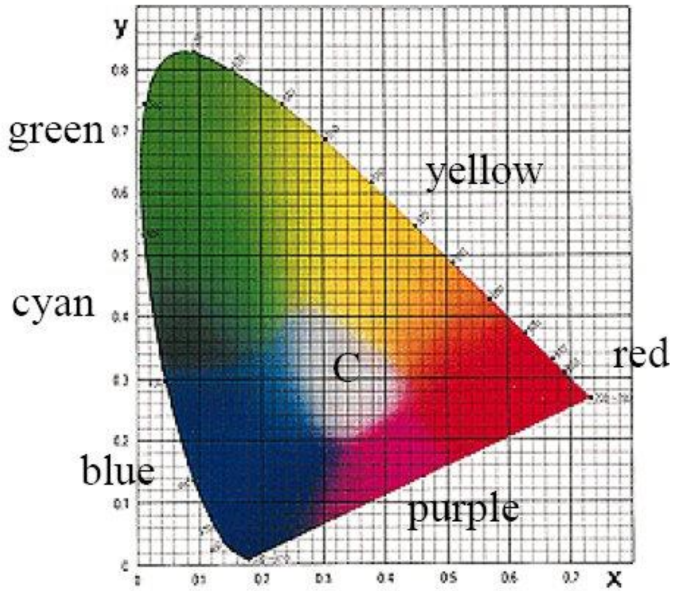




color generation by mixing RGB primaries

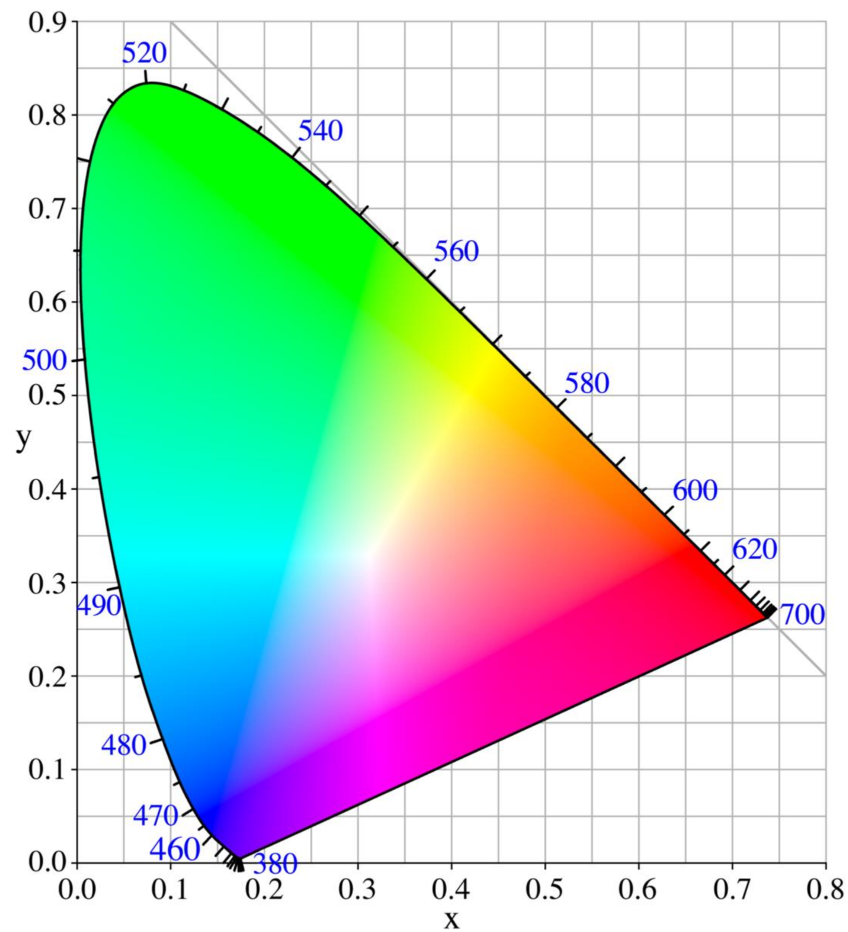
Perceptual Color Spaces

- Instead of R, G, B primaries it uses X, Y, Z primaries
- Normalizing for luminance and perceptive distance yields the CIE chromaticity diagram (1931)

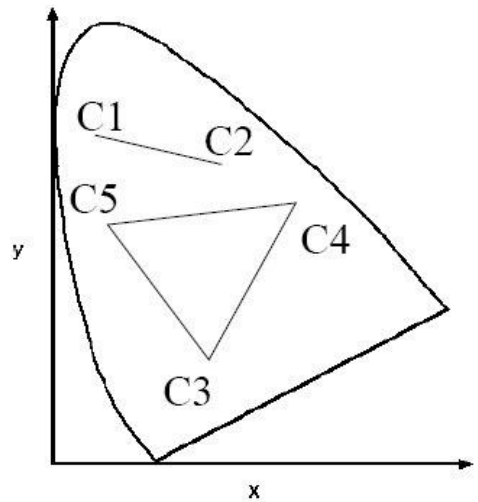


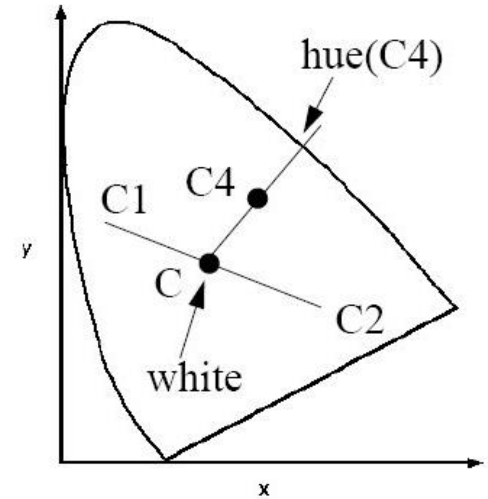
- Points on the boundary are the pure spectrum colors (from red to blue)
- Note: the purple line (joining blue and red) is not part of the visible spectrum of pure colors
- Interior points represent all visible colors (equidistant colors cause equal perceptive difference)
- Point ‘C’ is the white-light position

So, Can You Generate All Visible Colors with Three Primaries?



The CIE Chromaticity Diagram





Color gamuts:

- all colors on the line C1-C2 can be generated by mixing proper amounts of C1 and C2
- all colors within the triangle C3-C4-C5 can be generated by mixing amounts of C3, C4, C5
- the triangle defined by the primaries C3, C4, C5 defines the gamut of the monitor

Notice: no triangle can encompass all visible colors in the CIE → modern monitors are unable to display all visible colors

-

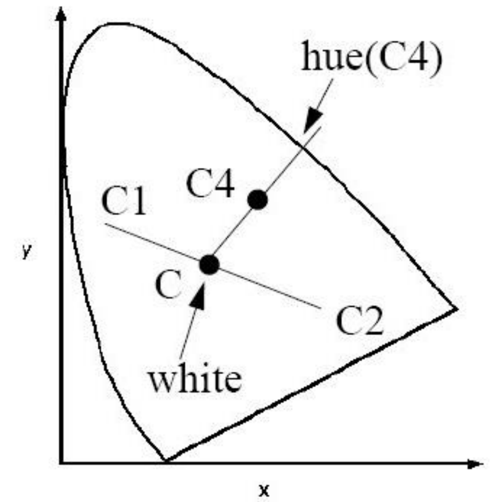
Complementary colors:

- C1, C2 are complementary when the gamut line C1-C2 goes through the white point C
- we can create white light by mixing appropriate amounts of C1 and C2
- also, we can create C1 by subtracting some amount of C2 from white light

Pure color (hue) of a color:

- Extending line C4-C to the border yields the hue of C4

The CIE Chromaticity Diagram



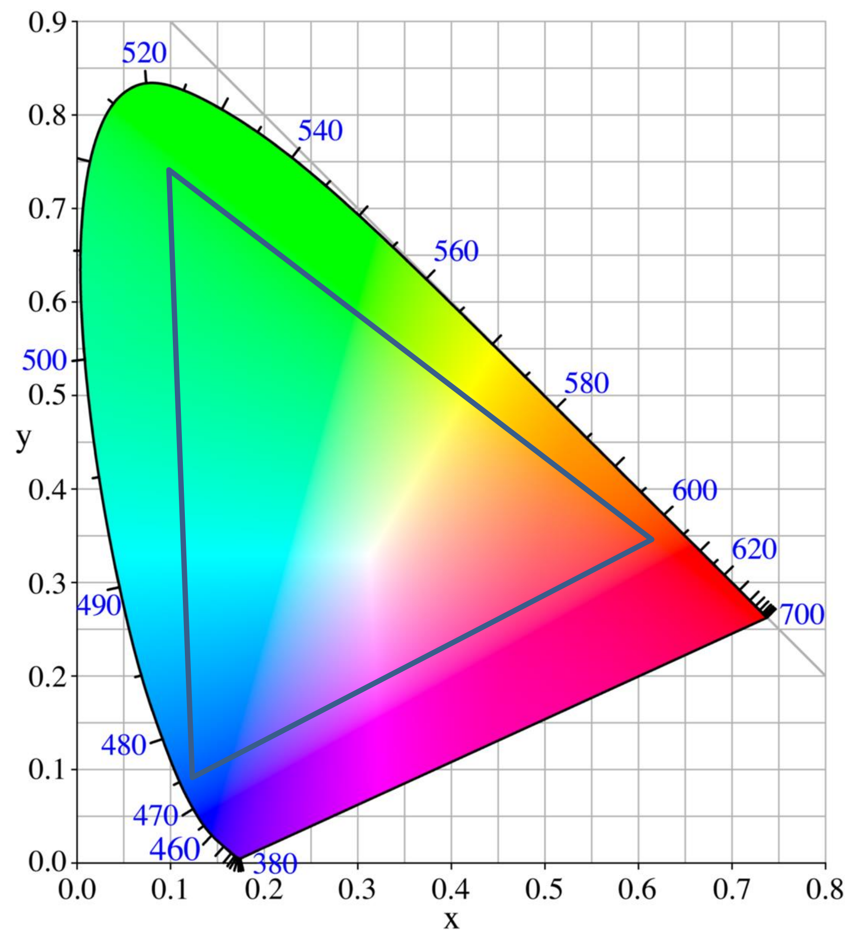
Complementary colors:

- C1, C2 are complementary when the gamut line C1-C2 goes through the white point C
- we can create white light by mixing appropriate amounts of C1 and C2
- also, we can create C1 by subtracting some amount of C2 from white light

Pure color (hue) of a color:

- Extending line C4-C to the border yields the hue of C4

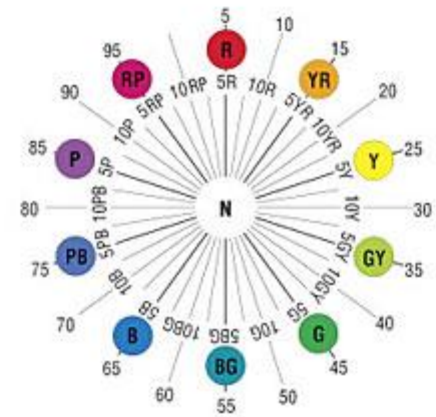
So, Can You Generate All Visible Colors with Three Primaries?

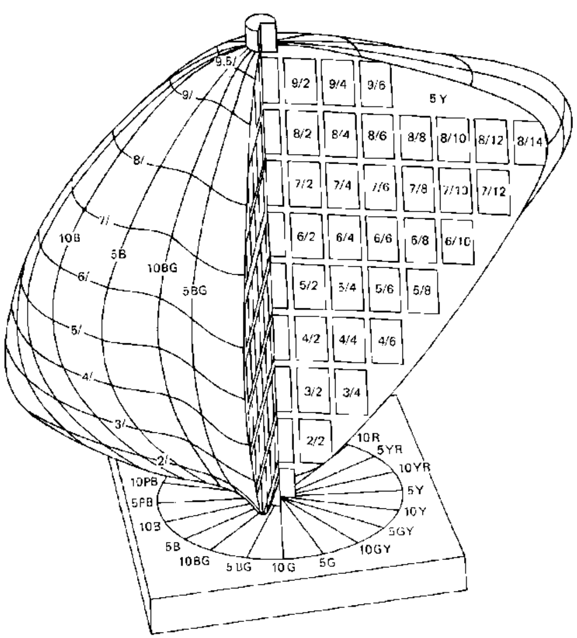


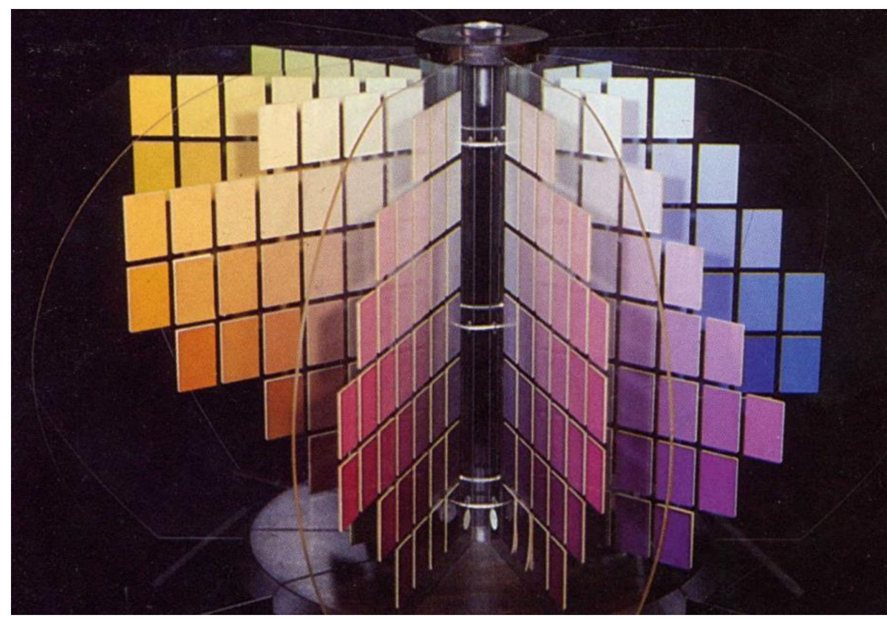
The Munsell Perceptual Color Space

The (irregularly shaped) Munsell tree has 3 axes:

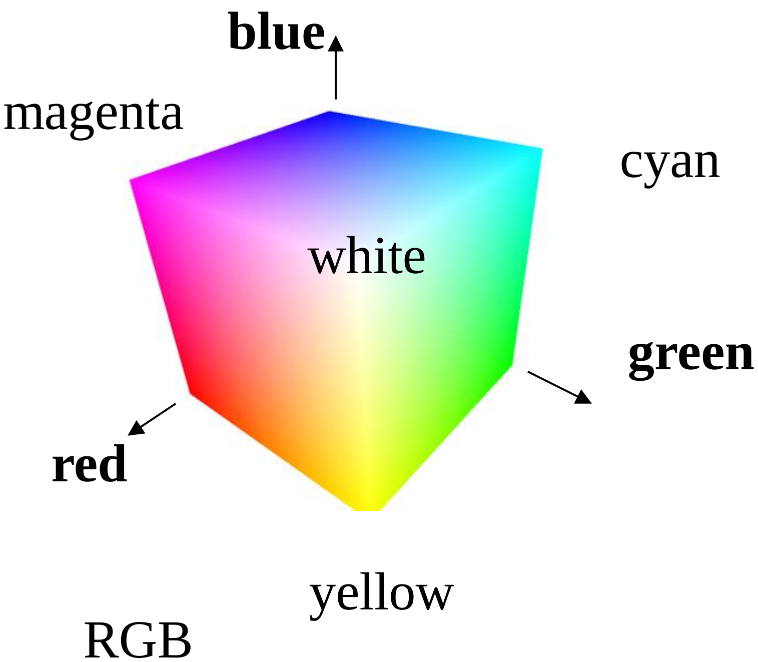
- chroma (saturation): distance from the core (values 0-30, with fluorescent colors having the maximum 30)
- value (brightness): vertical axis (0– 10 (white))
- hue: 10 principal hues (R, YR, Y, GY, G, BG, B, PB, P, RP)

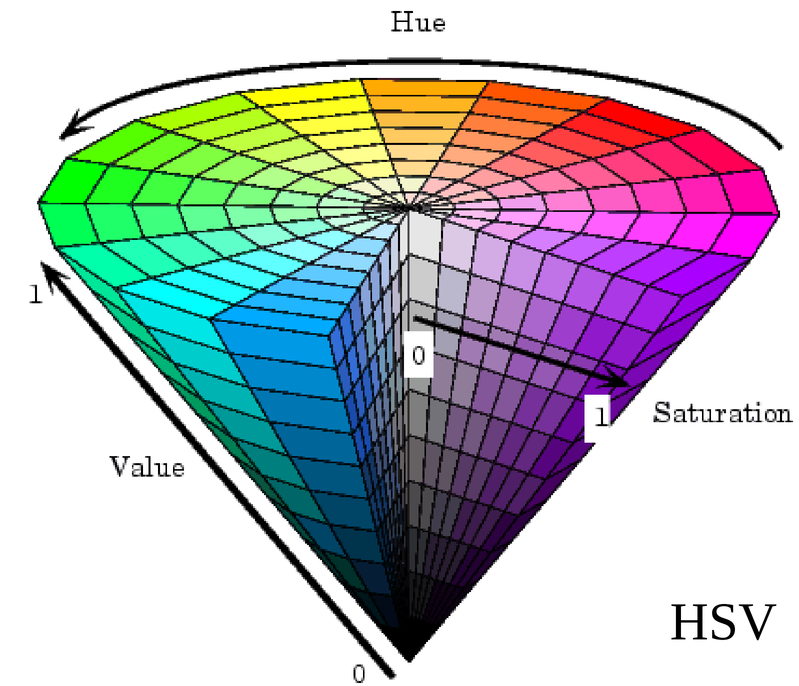






Non-Perceptual Color Spaces



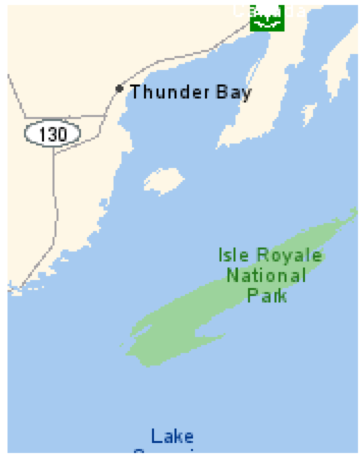


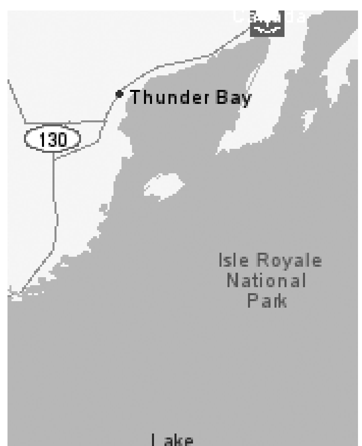
How to convert from RGB to HSV?

Contrast Revisited

Difference of brightness in adjacent regions of the image

- grey-level (luminance) contrast
- color contrast





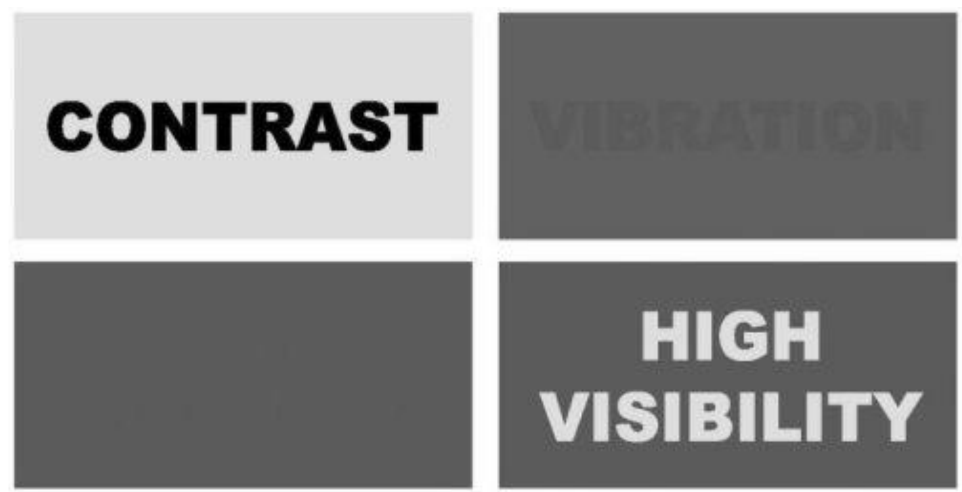
Contrast Needs Brightness Diff.

Munsell tree – unwrapped

- ignore the red circles
- look at what heights the longest rows are

color

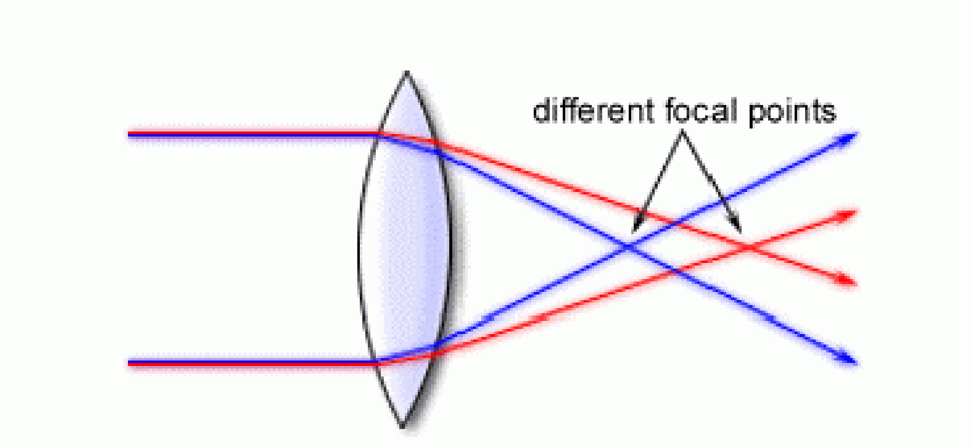
same image in grey-scale (brightness only)



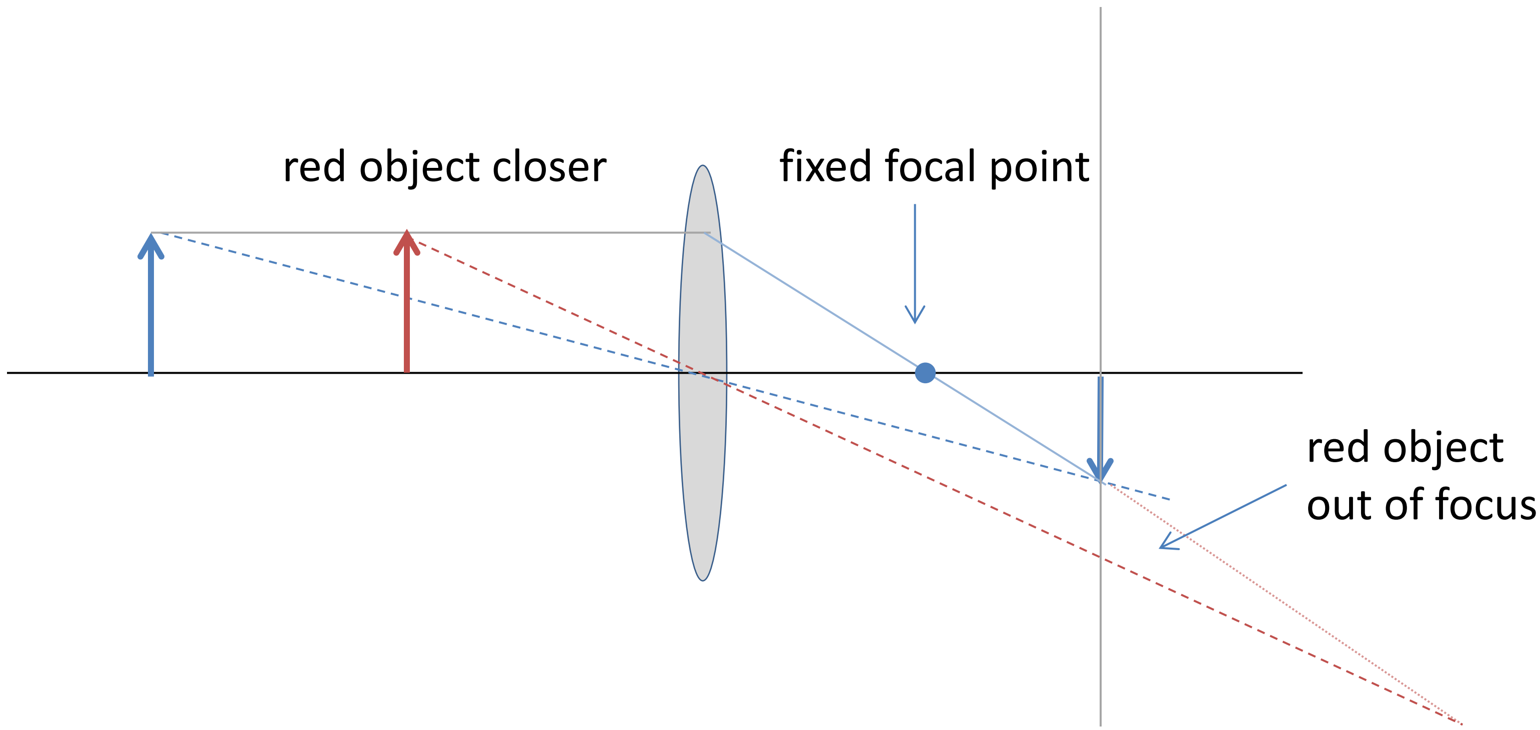
Chromatic Aberration

Different wavelengths of light are focused at different distances within the eye

- short-wavelength blue light is refracted more than long-wavelength red light
- focusing on a red patch, an adjacent blue patch will be significantly out of focus
- the human eye has no correction for chromatic aberration



Compare: Far vs. Near Objects



would have to shift focal point to the left to bring red object into focus
in the case before, all objects were in the same plane but the focal point changed
the blurring effect is equivalent → need to change focal point to gain focus for either color
this is tiring to the eye and causes the problems

This is really painful

This is better

Wiring: The Visual Pathways

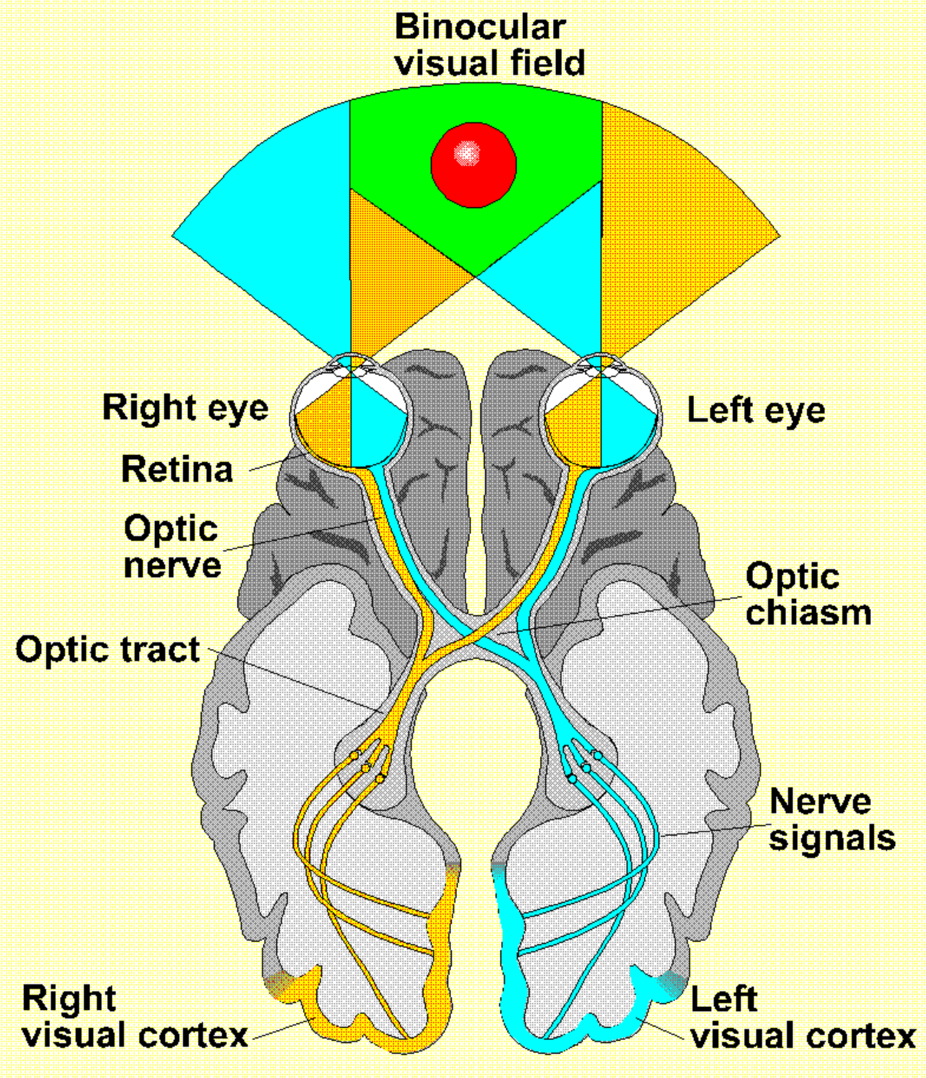
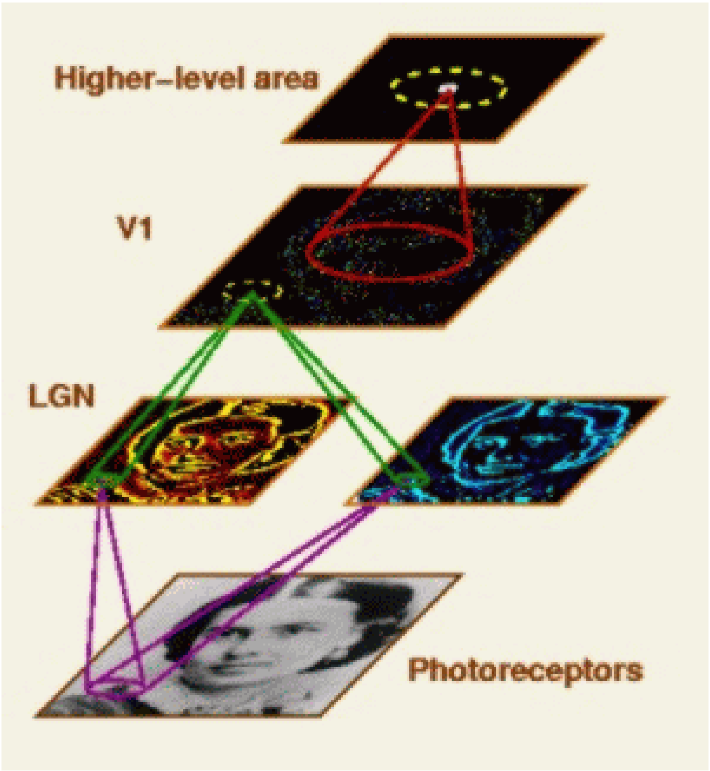


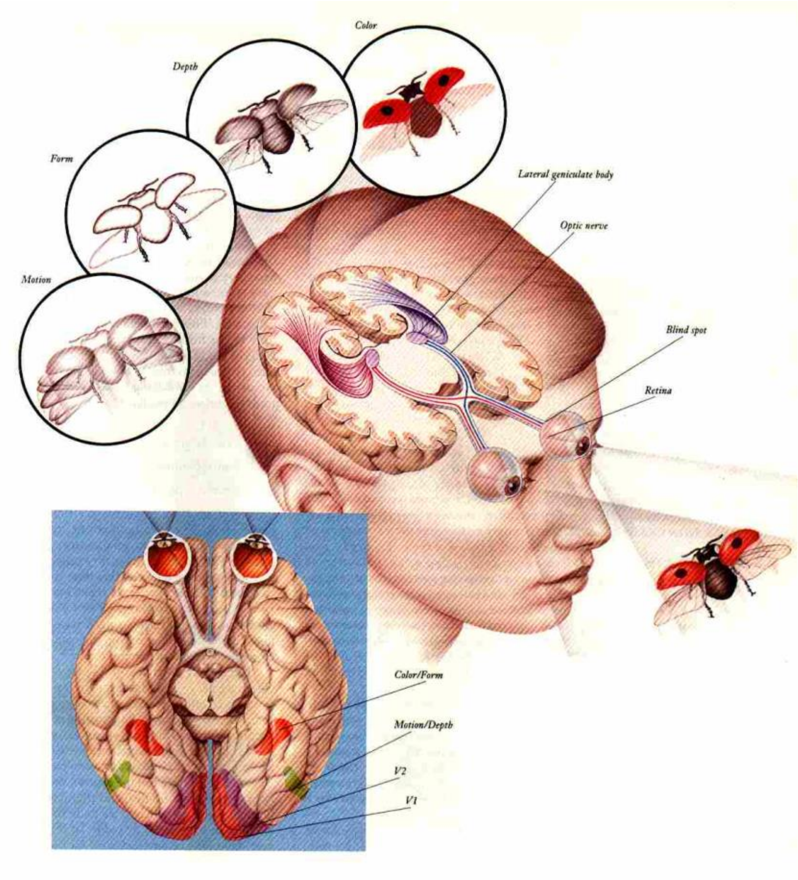
Figure 10: Binocular vision, showing visual pathways in the brain

Processing Unit: The Visual Cortex

Visual cortex breaks input up into different aspects:

- color, shape, motion, depth

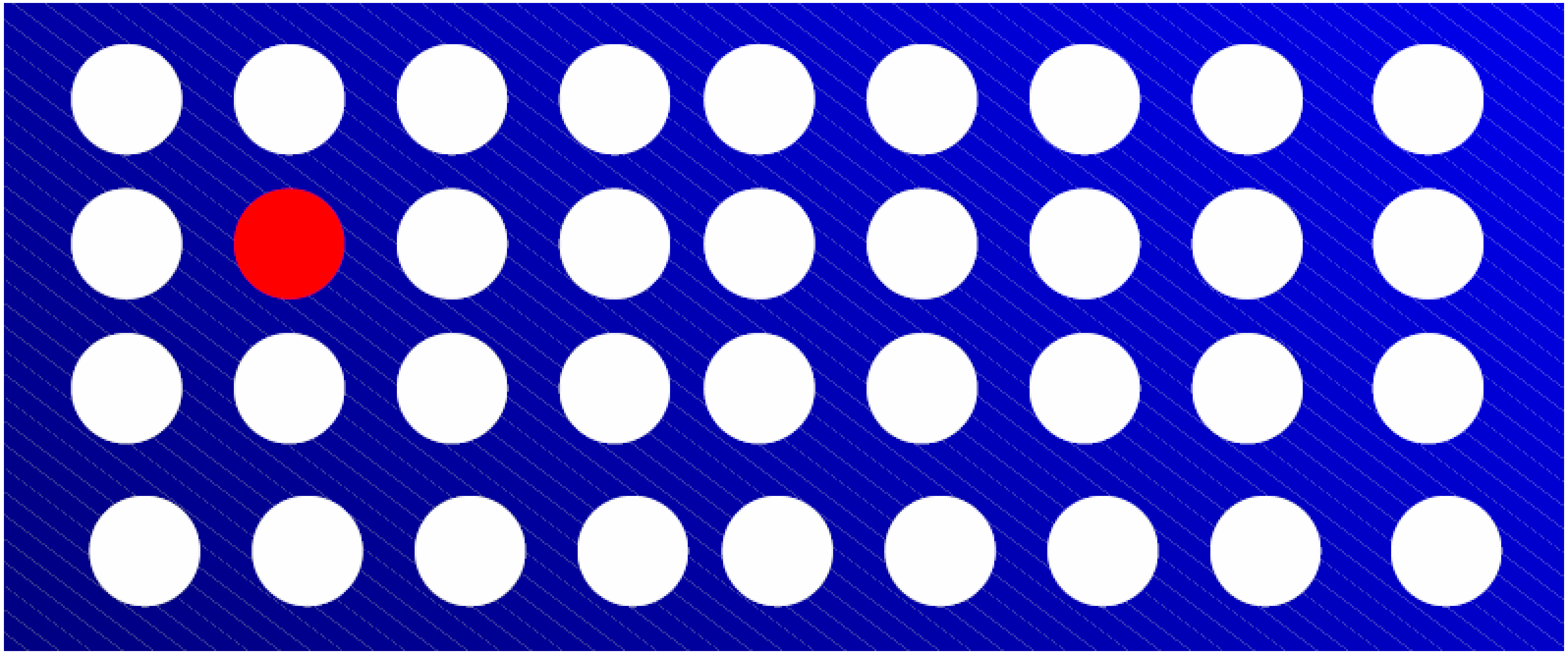




Pre-Attentive Processing

If you want it or not: some features are always detected

And fast – within 200 ms or less



Pre-Attentive Processing

Why is it so fast?

Well, because 50% of the brain is dedicated to vision

Vision is a MASSIVELY parallel processor dedicated to

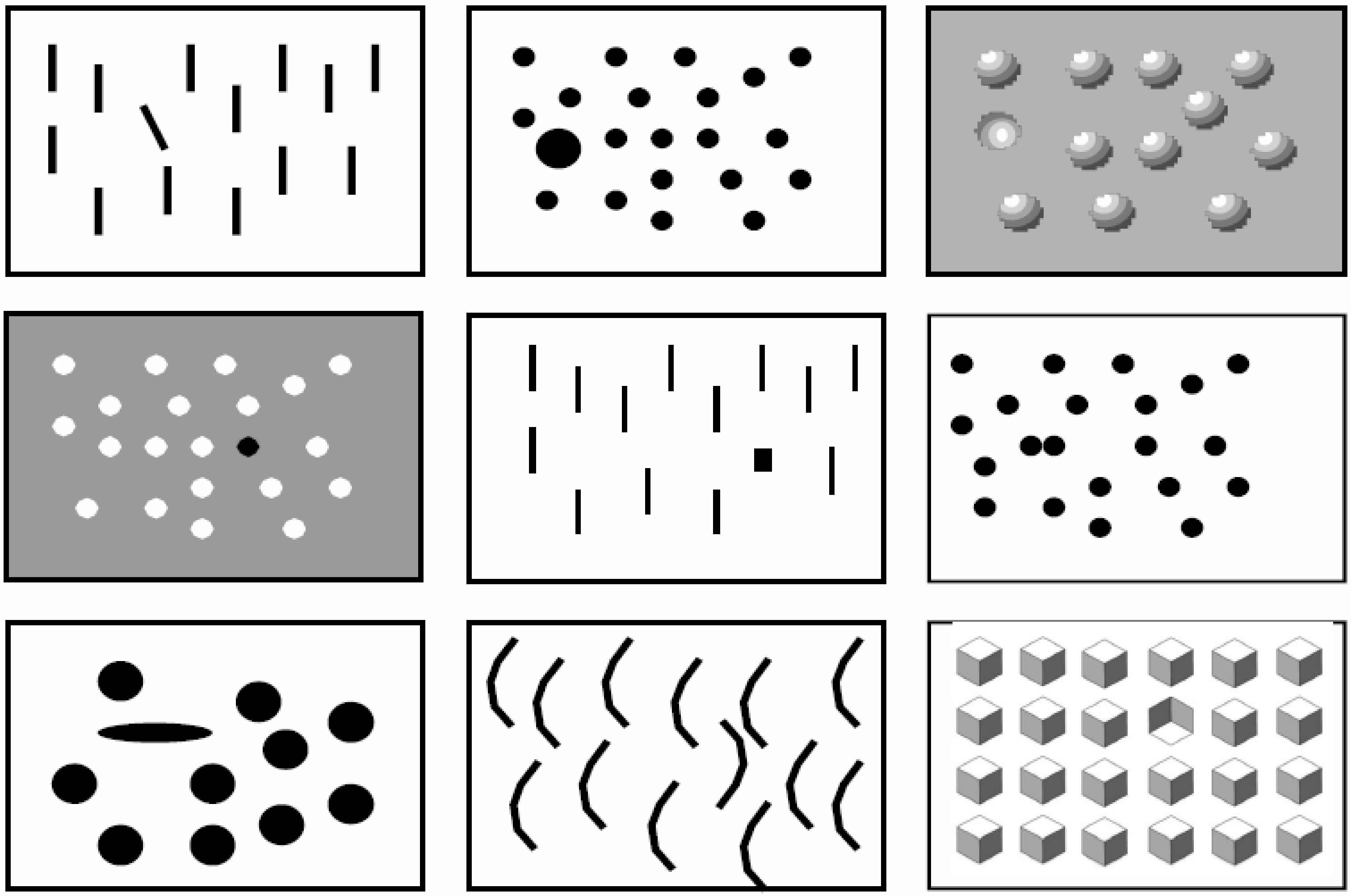
- detect
- analyze
- recognize
- reason with

visual input

Pre-Attentive Processing

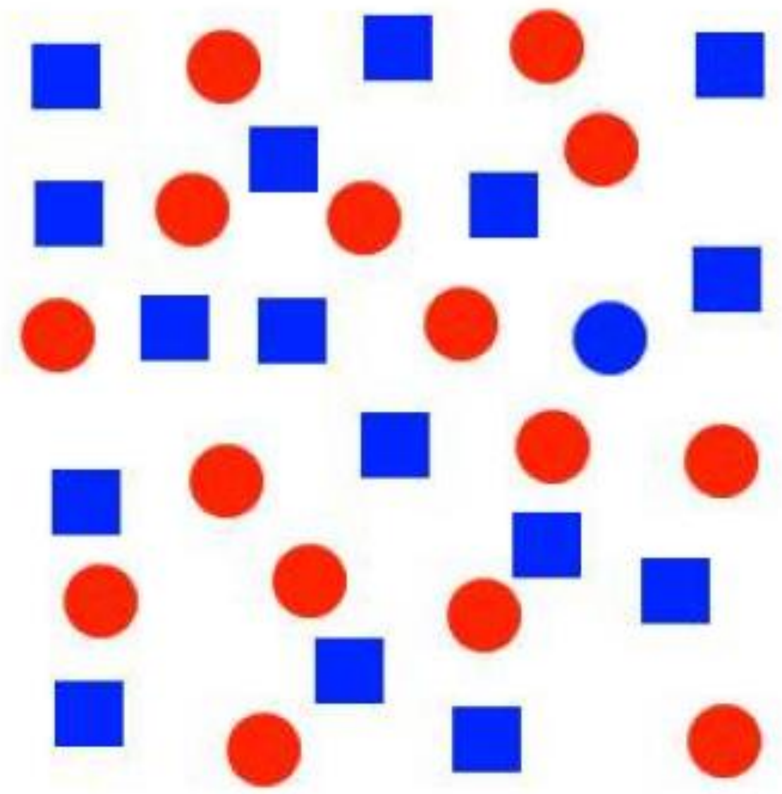
Sensitivity to differences in:

- color, orientation, size, shape, motion, shading, 3D depth, …



Pre-Attentive Processing

But there are limits: conjunctions don't work well



quick: find the blue circle

Pre-Attentive Processing

Some features/cues are stronger than others:

Look at the chart and say the COLOUR not the word

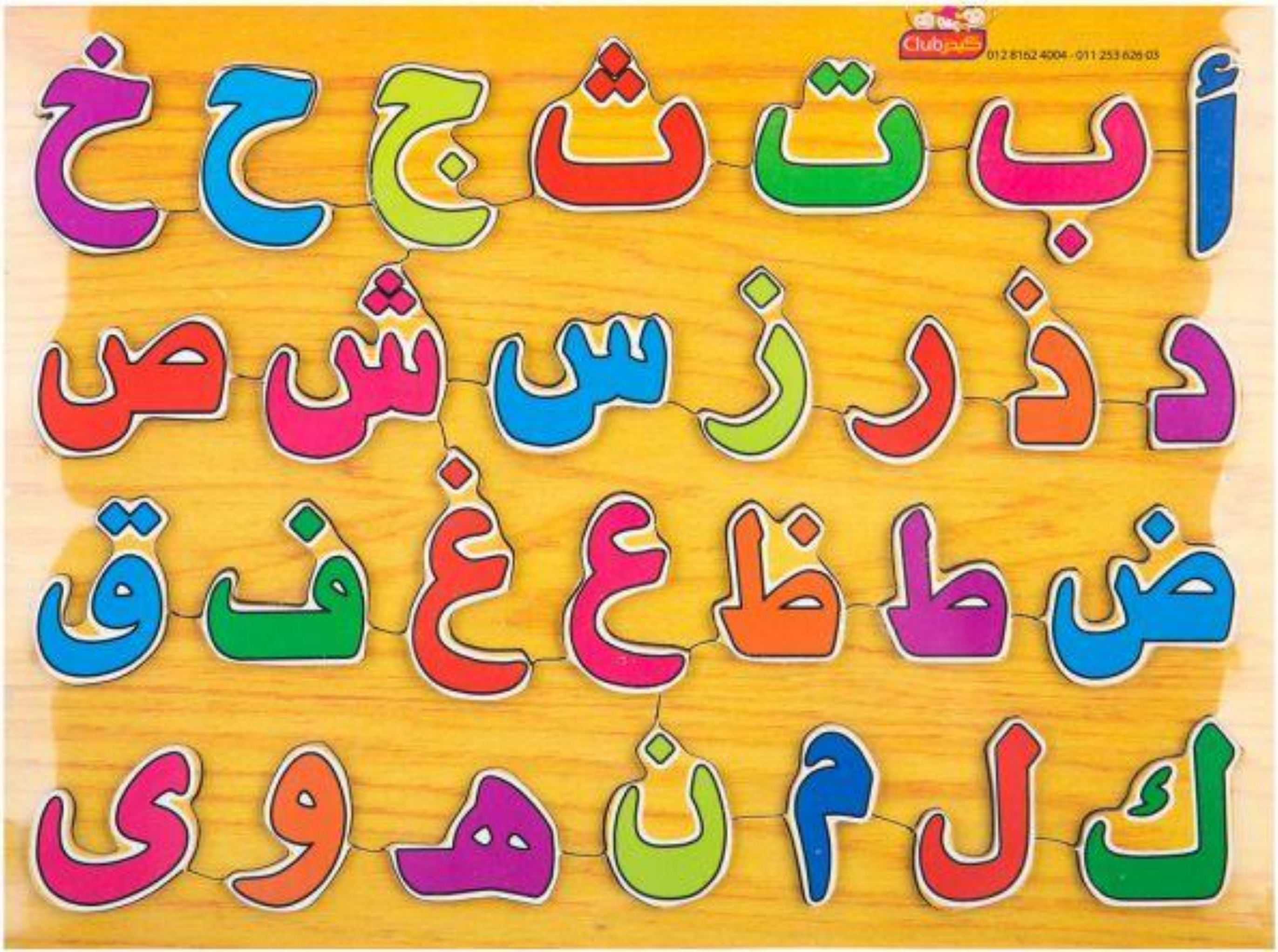
YELLOW BLUE ORANGE
BLACK RED GREEN
PURPLE YELLOW RED
ORANGE GREEN BLACK
BLUE RED PURPLE
GREEN BLUE ORANGE

Left – Right Conflict

Your right brain tries to say the colour but your left brain insists on reading the word.

Pre-Attentive Processing

Now try this (the left brain takes a break)



Pre-Attentive Processing

Words are patterns, which form strong pre-attentive feature

- this would have been different if this had been done in Arabic

There are limits, however

- let’s see the next experiment

Pre-Attentive Processing

Reading 1

Aoccdrnig to a rscheearch at an Elingsh uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht frist and lsat ltteer is at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae we do not raed ervey lteter by it slef but the wrod as a wlohe

Pre-Attentive Processing

Now, is tihs ture? Raed on….

Pre-Attentive Processing

Reading 2

Anidroccg to crad cniyrrag lcitsiugnis planoissefors at an uemannd, utisreviny in Bsitirh Cibmuloa, and crartnoy to the duoibus cmials of the ueticnd rcraeseh, a slpmie, macinahcel ioisrevnn of ianretnl cretcarahs araepps sneiciffut to csufnoe the eadyrevy oekoolnr

Pre-Attentive Processing

Reading 2

According to card carrying linguistics

professionals at an unnamed, university

in British Columbia, and contrary to the

dubious claims of the uncited research,

a slpmie, macinahcel ioisrevnn of

ianretnl cretcarahs araepps sneiciffut to

csufnoe the eadyrevy oekoolnr

Pre-Attentive Processing

Reading 2

According to card carrying linguistics professionals at an unnamed, university in British Columbia, and contrary to the dubious claims of the uncited research, a simple, mechanical inversion of internal characters appears sufficient to confuse the everyday onlooker

What To Learn From This

The human visual system (HSV) tolerates (visual) noise very well

- it can read the randomly garbled text very well
- machines (equipped with computer vision) are poor at this

Humans have only limited computational capacity

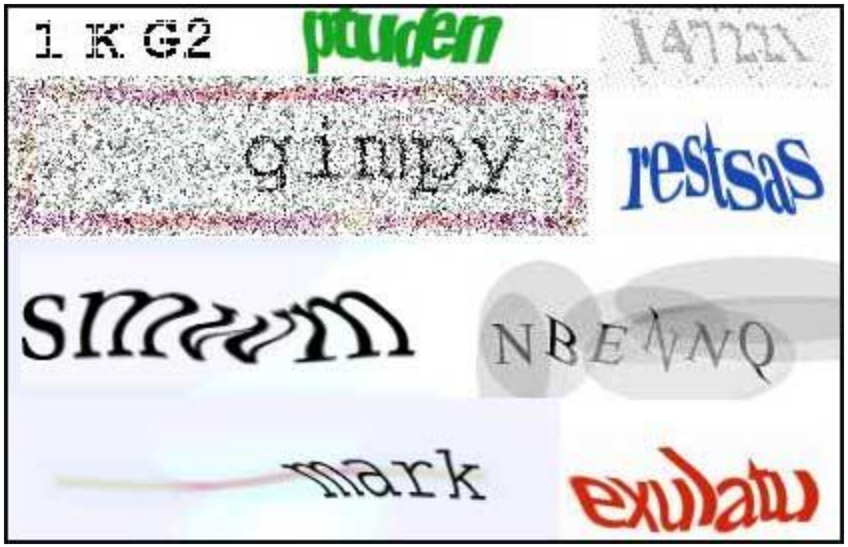
- hard to execute a fixed rule to decipher text
- especially once the text gets longer (7±2 rule of working memory)
- this is where computers excel

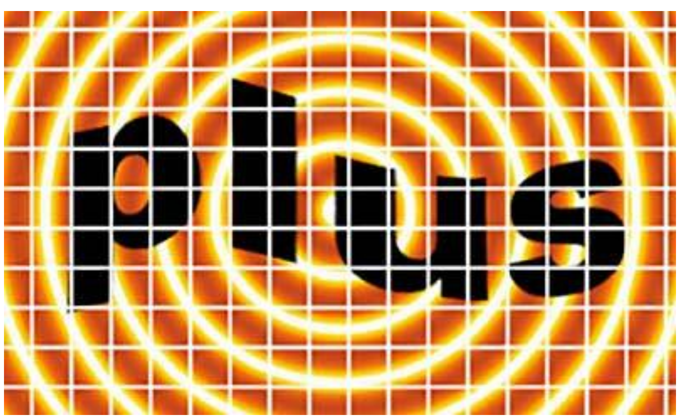
The fact that computers deal poorly with noisy patterns is exploited in CAPTCHA

- **CAPTCHA: C**ompletely **A**utomated **P**ublic **T**uring Test to tell **C**omputers and **H**umans **A**part
- used to ensure that an actual human is interacting with a system
- some examples:
    - creating a new gmail or yahoo account (prevent spammer accounts)
    - submitting files, data, email

CAPTCHA

CAPTCHA: noisy and vastly distorted patterns that are difficult to recognize by machines



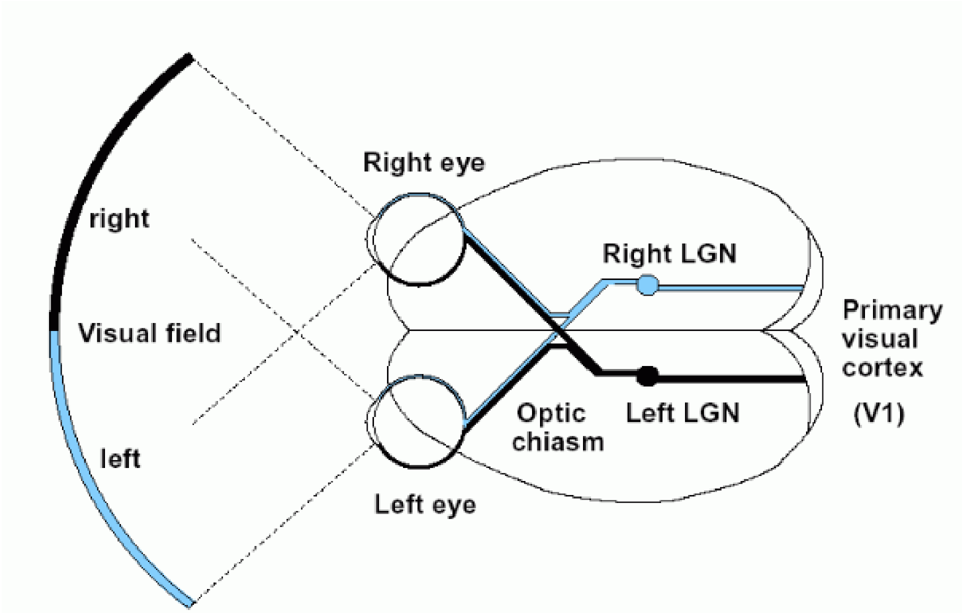


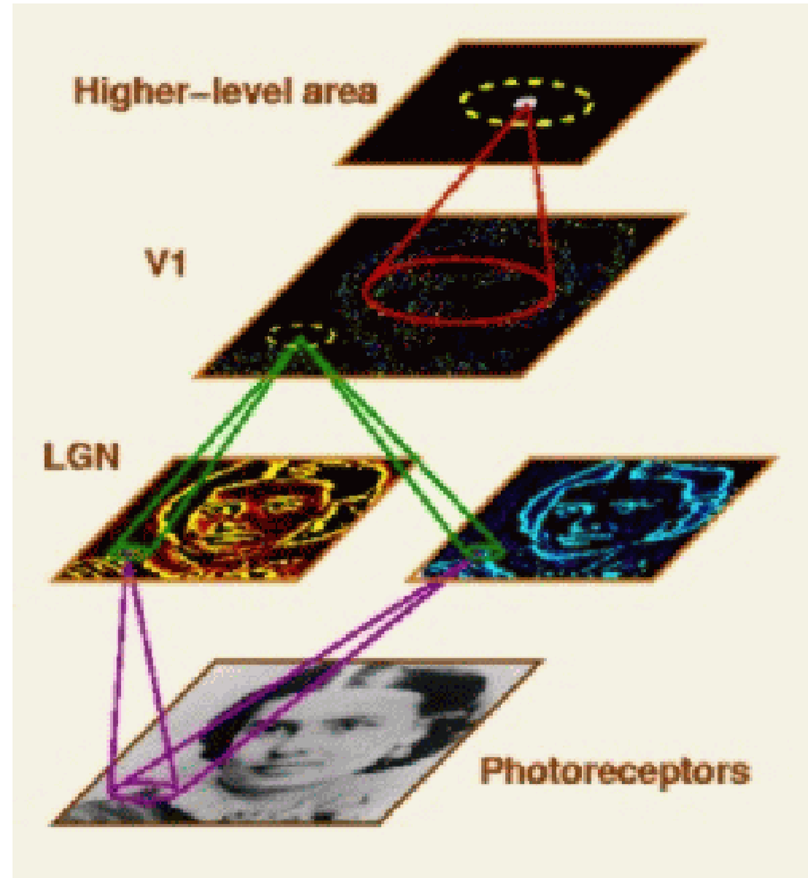
CAPTCHA

But computer vision algorithms have become more sophisticated at CAPTCHA *character* recognition

- the latest approach is *object* recognition

Organization of the Human Brain



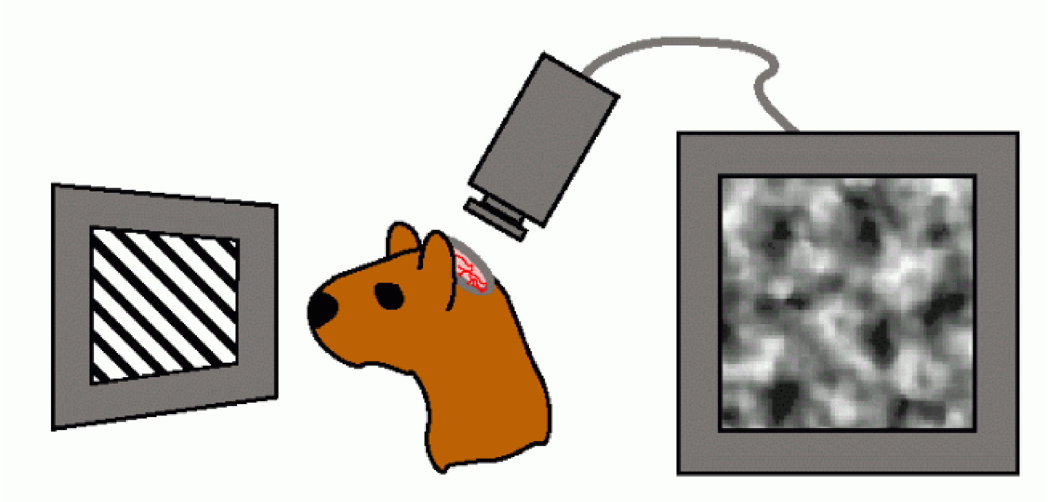


- LGN: left lateral geniculate nucleus of the thalamus
- V1: primary visual cortex

  - a quarter-sized area in the back of the head (the first cortical stage for most visual processing)
- Higher-level areas dedicated to spatial reasoning, associative object recognition, etc.

Measuring Orientation Maps

- Use optical imaging techniques to measure orientation preferences for a large number of neurons
  - remove part of the skull of a laboratory animal, exposing the surface of the visual cortex
  - present visual patterns to the eyes
  - a video camera records either light absorbed by the cortex or light given off by fluorescent chemicals applied to it
  - compare measurements between different stimulus conditions (orientations, temporal, etc)
- See *Topographica* software by Miikkulainen, Bednar, et. al. at University of Texas, Austin
  - java demos available at: http://www.cs.utexas.edu/users/jbednar/demos.html

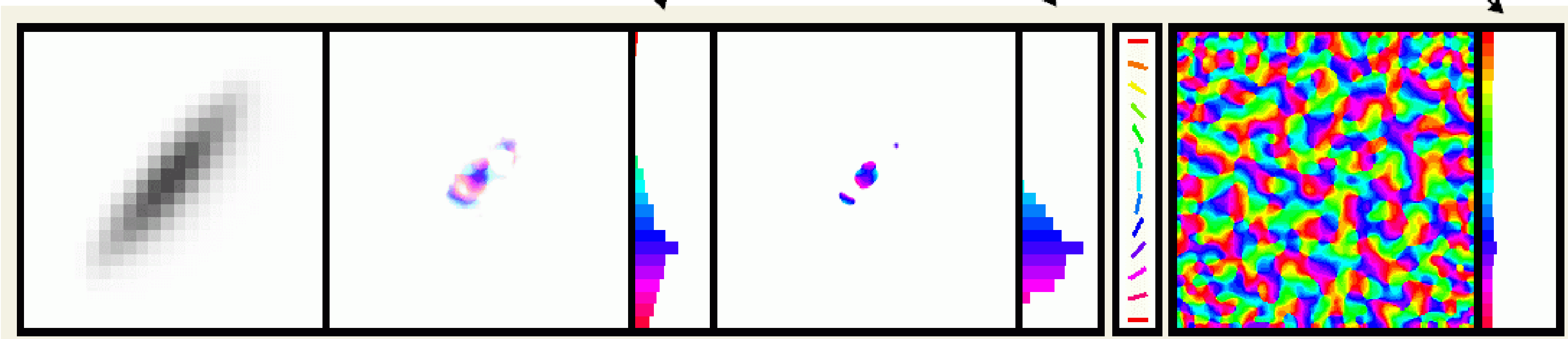


Organization and Sensitivity of the Visual Cortex

orientation histogram of initial response

orientation histogram of settled response

orientation histogram of the organization map



shape presented

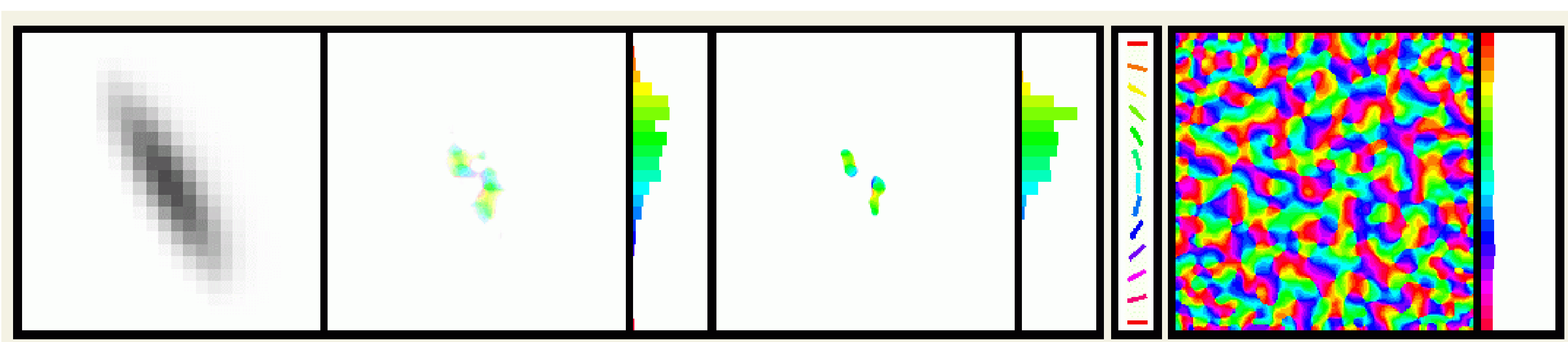
initial activated cortical neurons

settled activated cortical neurons

self-organized orientation map on a flattened 5 x 5 mm area on V1

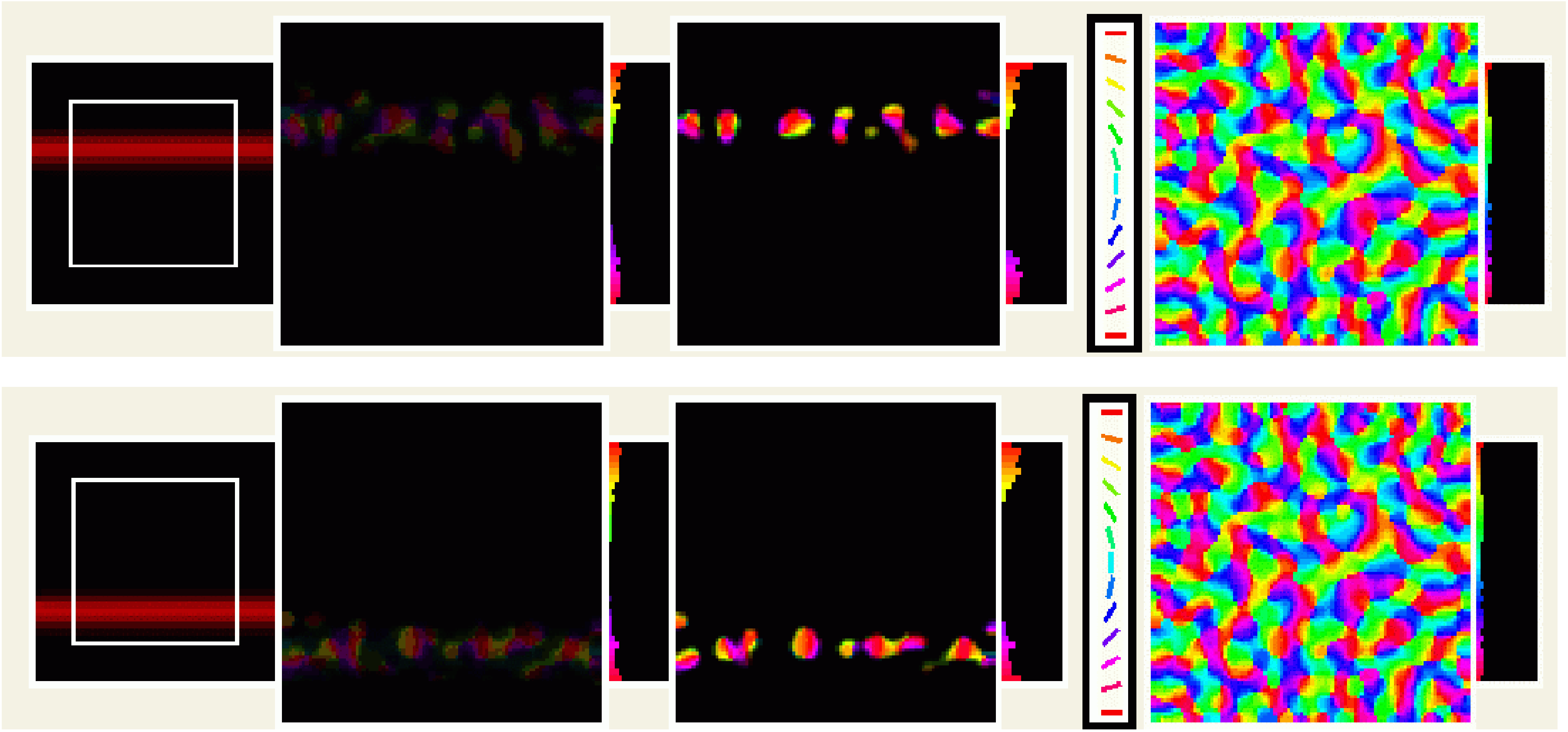
key showing colors assigned to each angle

different orientation of presented shape:

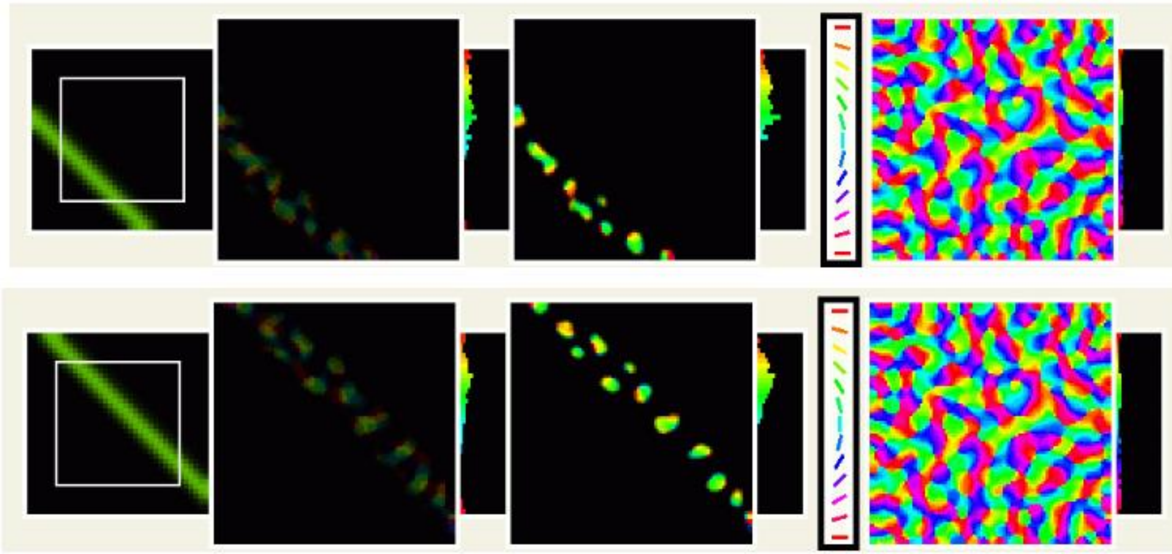


Organization and Sensitivity of the Visual Cortex

- Brain is sensitive to edges (contrast in intensity and color), pre-attentive
- Some more example obtained using *Topographica:*

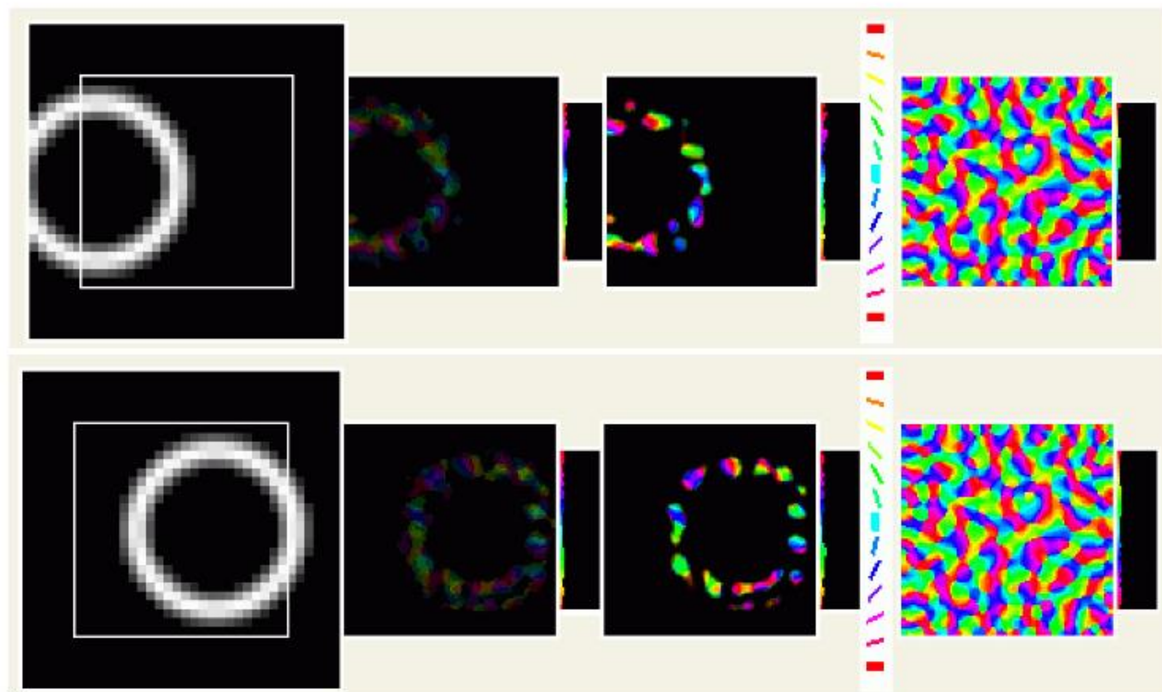


Organization and Sensitivity of the Visual Cortex

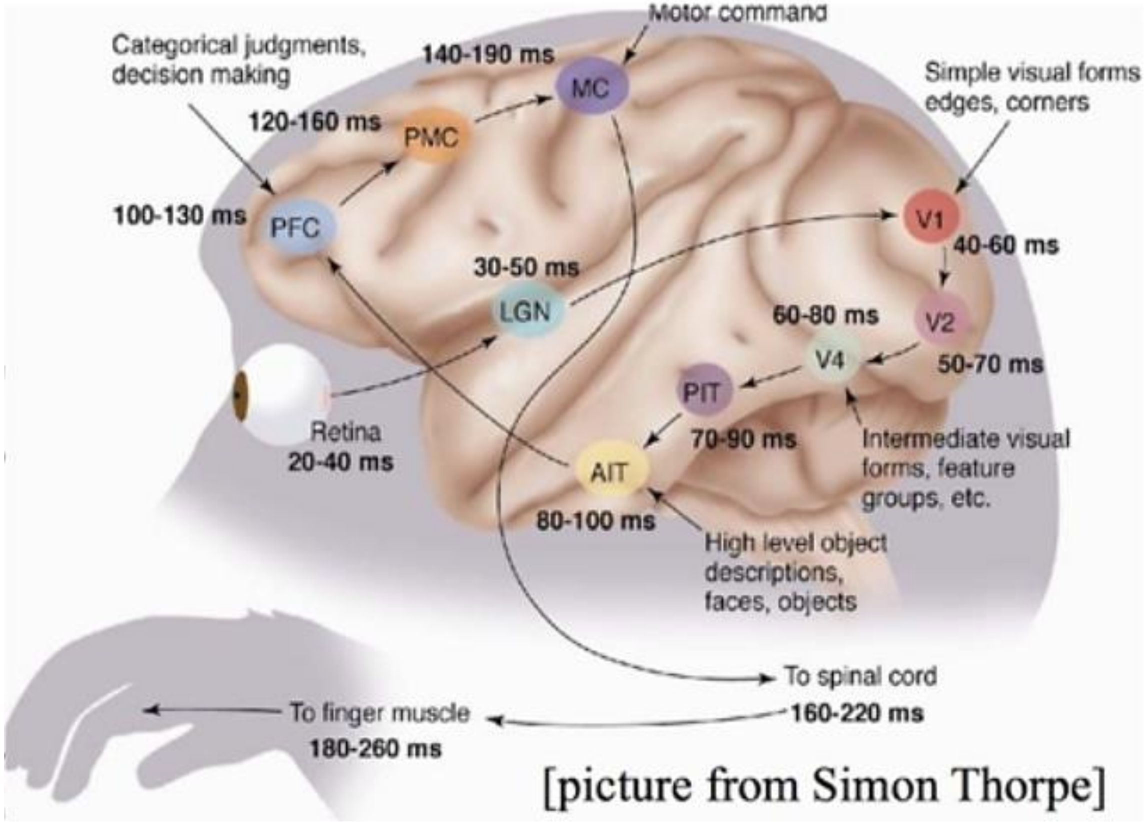


diagonal line

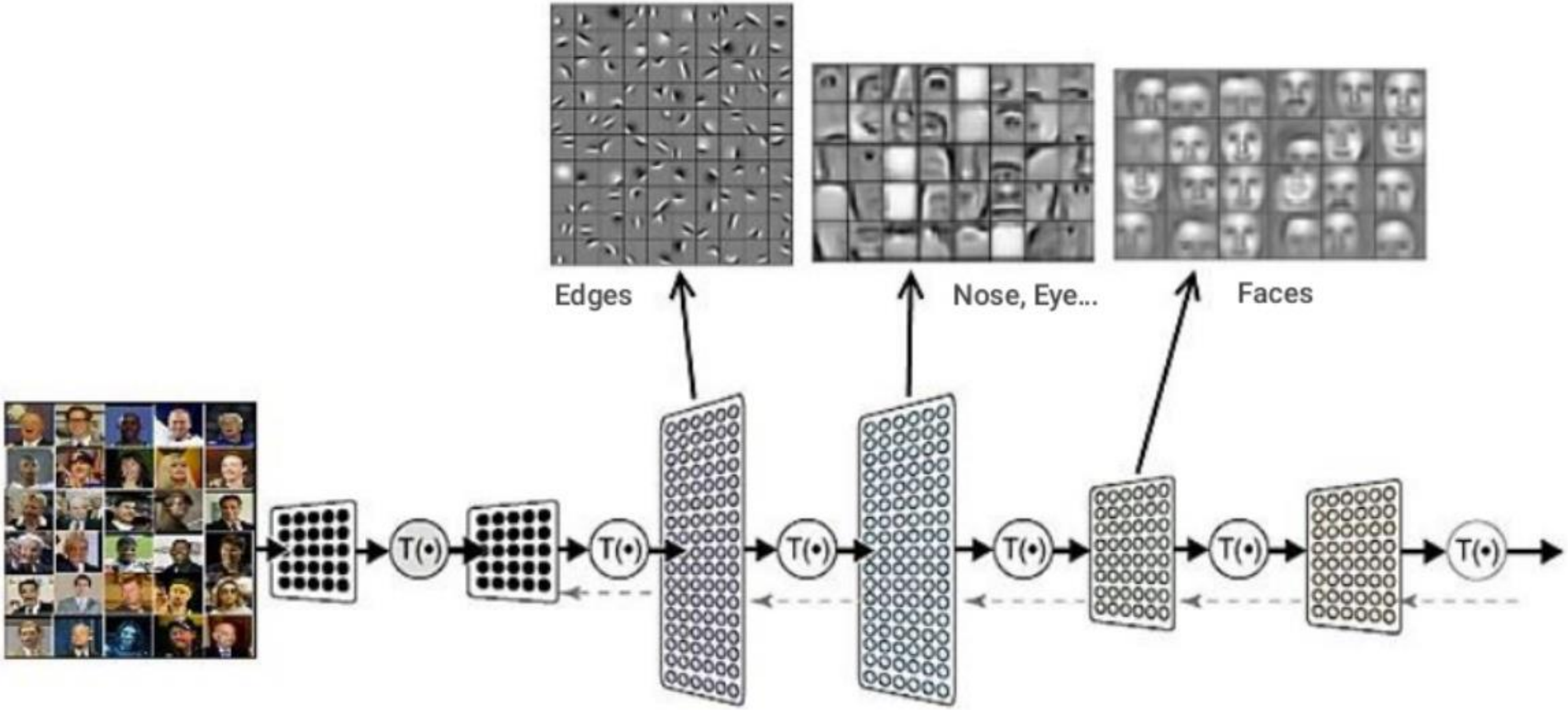
circle



Deep Learning is Inspired by the Brain

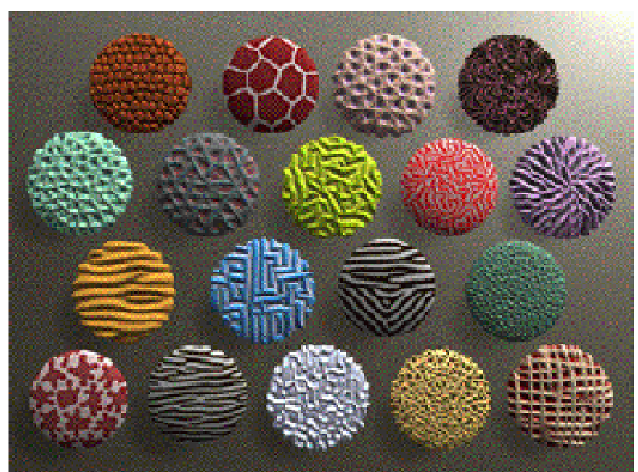


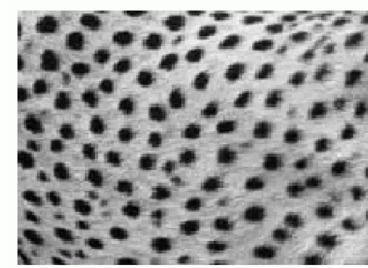
Deep Neural Network: From Local to Global Features

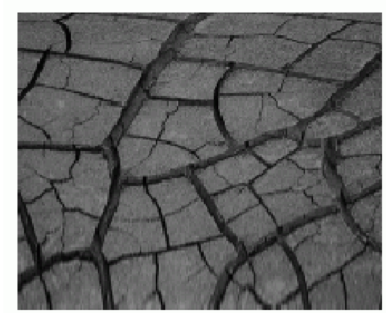


Pre-Attentive Cues With Textures

- A visual texture represents that visual sensation that allows us to pre-attentively differentiate two adjacent, possibly structured parts in our visual field without eye movement
  - visual textures include micro-structures, patterns, profiles, etc.
  - to identify textures, an observations of about 160-200 ms is sufficient (cognitively controlled processes require about 300-400 ms)
- Classification of textures is based on
  - coarseness, contrast, directionality, line-likeness, regularity, roughness
- Textures improve perception of position and orientation
- Texture communicate information about the 3D structure regardless of their coloring



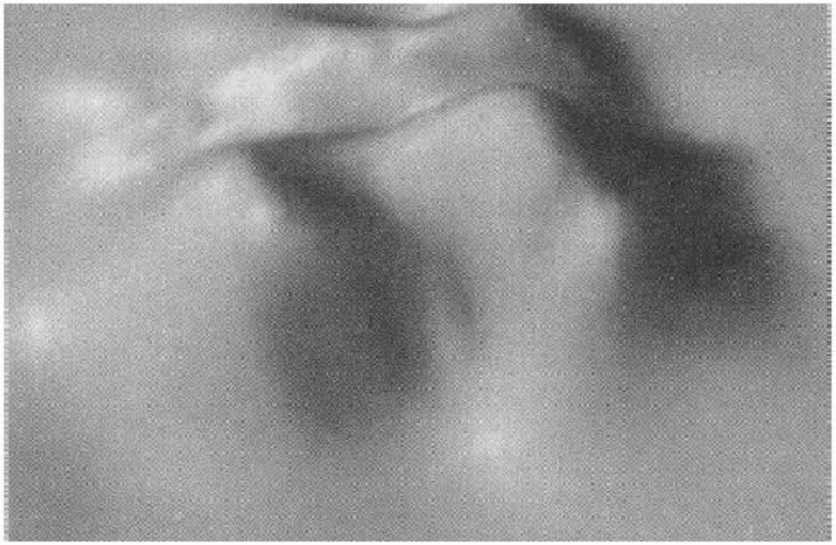
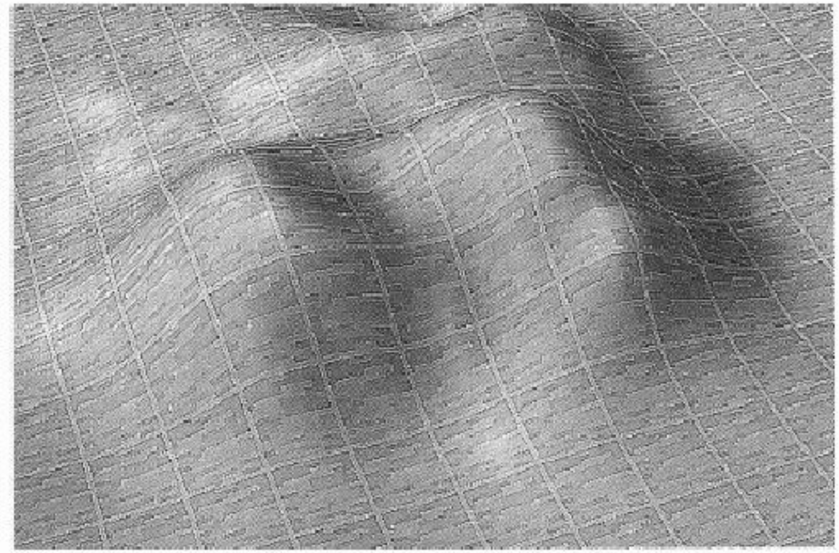




from: Jürgen Döllner

Pre-Attentive Cues With Textures

- Same surface with and without texture

- Textures that do not include information are to be avoided in visualization

    - recall Tufte's aesthetic principle that irrelevant decoration (= chart junk) should be avoided

- Subtle textures for 3D visualizations, however, can be important elements of visual design

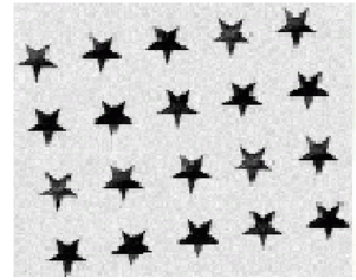
    - see above

Texture Perception

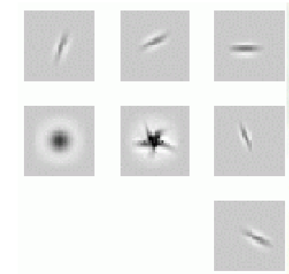
- *Textons*

  - fundamental micro-structures in generic natural images

  - basic elements in pre-attentive visual perception



objects

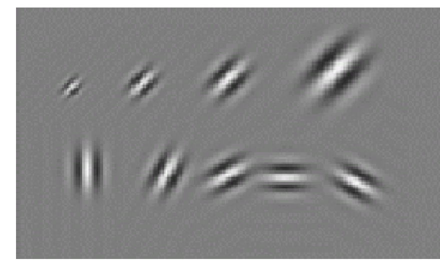


their textons

- Textons can be classified into three general categories:

  1. elongated blobs (line segments, rectangles, ellipses) with specific properties such as hue, orientation, and width, at different level of scales
  2. terminators (end of line segments)
  3. crossings of line segments

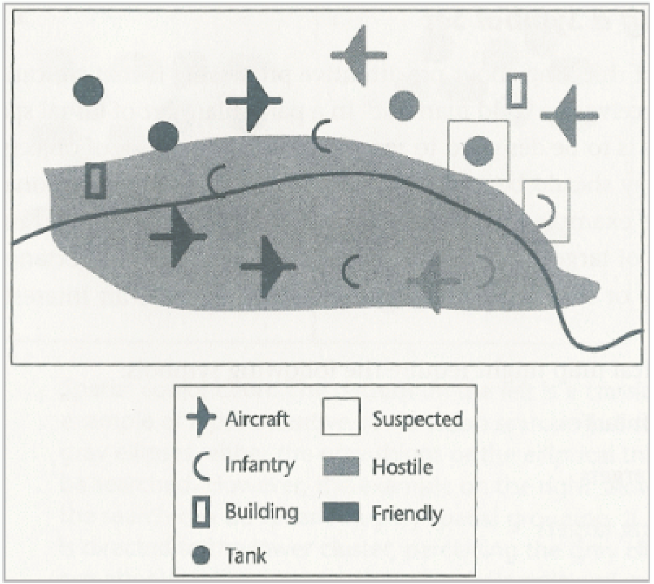
Gabor primitives



- Recall the sensitivity of the neurons in V1

Relation to Symbol and Texture Design

- When designing textures to indicate different regions of a visualization, make sure that the textons are as different as possible
- The same rules apply when designing symbol sets
- Example: A tactical map may require the following symbols:
  - aircraft targets
  - tank targets
  - building targets
  - infantry position targets

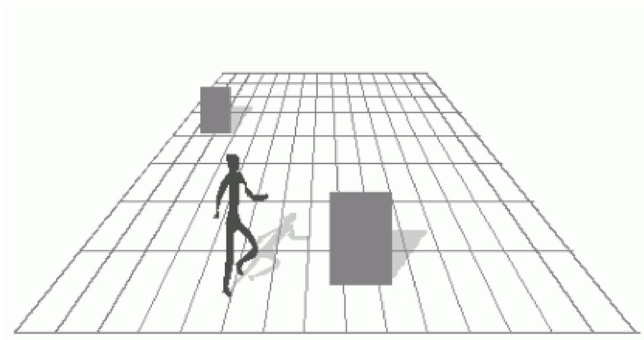


- Each of these target types can be classified as *friendly* or *hostile*
- Targets exist whose presence is suspected but not comfirmed (this uncertainty must be encoded)
- Set of symbols designed to represent different classes of objects
  - symbols should be as distinct as possible with respect to their pre-attentive processing
  - recall: military reconnaissance must occur FAST!

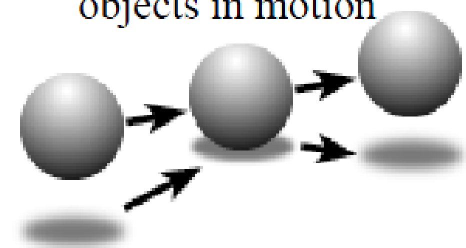
Information Display in 3D: Depth Cues

- 3D display should provide *depth cues*
- Linear perspective:
  - more distant objects become smaller in the image

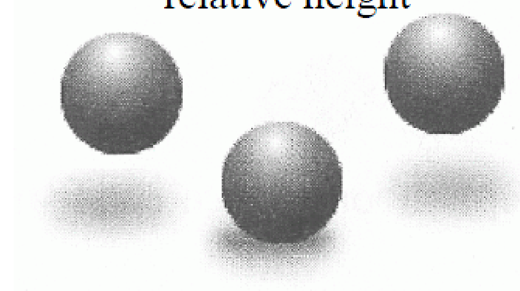
    --> can indicate focus, importance, or ordering
  - elements of a uniform texture become smaller with distance --> can give shape cues
- Shadows:
  - show the relative height of objects above a surface
  - provide strong depth cues for objects in motion
  - can be semi-realistic and still work as a depth cue
- Occlusion:
  - very powerful depth cue

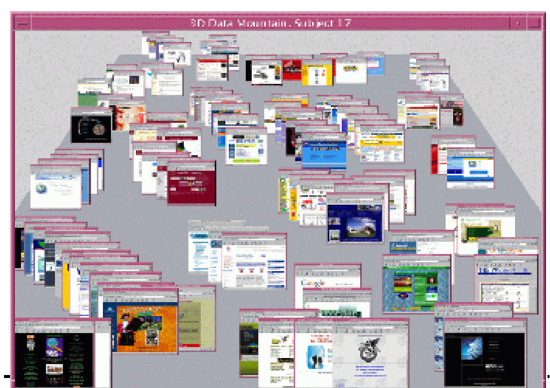


objects in motion



relative height

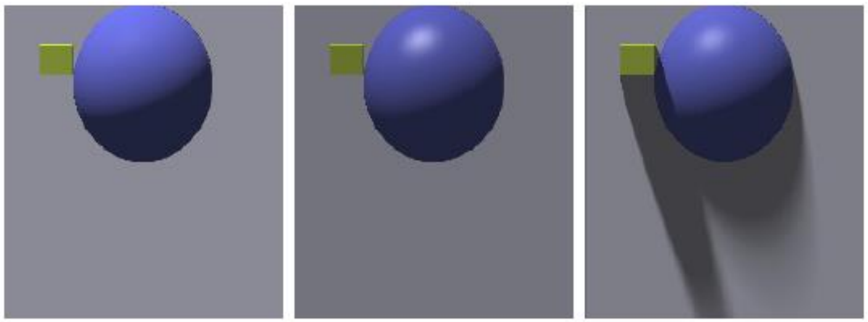




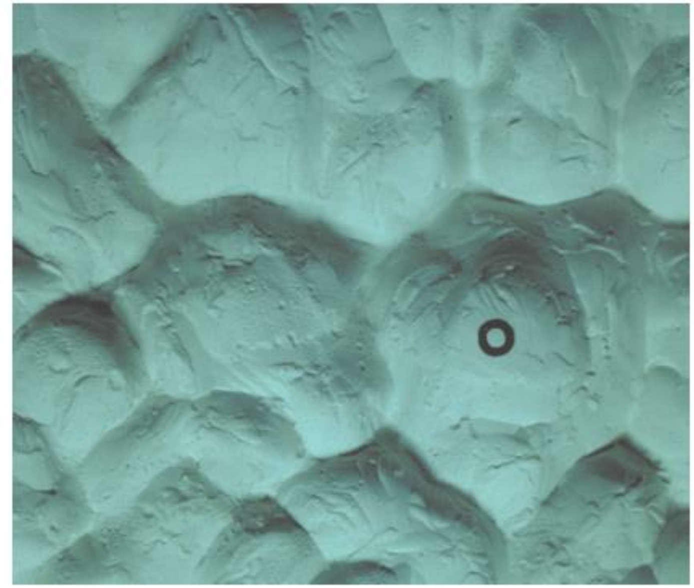
from: Colin Ware

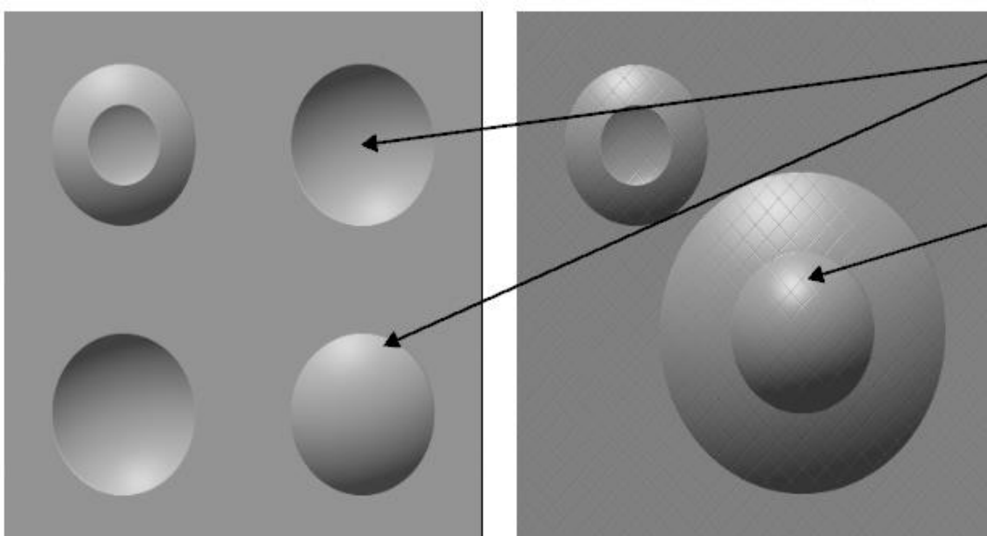
Information Display in 3D: Depth Cues

- Shading:
  - shape cues from shading (*shape-from-shading*)



diffuse + specular + shadows





shape from shading
(hole vs. hill)

specular can reveal fine detail

assume single light source

having more than one light source can lead to confusion

from: Colin Ware

Information Display in 3D: Depth Cues

- Other depth cues:
  - depth of focus
  - motion parallax (structure from motion) --> how objects relate under motion
  - steroscopic depth (binocular displays)
- For fine-scale judgement, for example, threading a needle:
  - stereo is important, and shadows and occlusion
- For large-scale judgement
  - linear perspective, motion parallax, and perspective are important
  - stereo is not so important
- However, for information visualization displays, one may exploit focus to emphasize importance, despite depth relationships