01 - Visual perception

Enrico Puppo
Department of Computer Science, Bioengineering, Robotics and Systems Engineering
University of Genova

90529 Data Visualization
24 September 2020

https://2020.aulaweb.unige.it/course/view.php?id=4293

Credits:
material in these slides is partially taken from
- T. Munzner, University of British Columbia
- A. Lex, University of Utah

other credits in the slides

Perception and Cognition

- Perception
  - Identification and interpretation of sensory information
  - From the physical stimulus to recognising information
  - Shaped by learning, memory, expectation

- Cognition
  - The processing of information, applying knowledge

Hear someone speak: Perception

Understand the language and the words: Cognition

Perception vs Cognition

Perception (optical)

- Not conscious, pre-attentive
- Eye, optical nerve, visual cortex
- First processing, faster
- Reflexes
- Basic perception
- (edges, planes)

Cognition

- Conscious, attentive
- Somatosensory association, higher mental functions (frontal cortex)
- Complex processing, slower
- Thoughts
- Recognising objects
- Relations between objects
- Conclusion drawing
- Problem solving
- Learning
- …

Experiment

- Image containing several coloured labels:
  - How many yellow labels are there? (0.5s)
  - How many labels of word "yellow" are there? (2s)

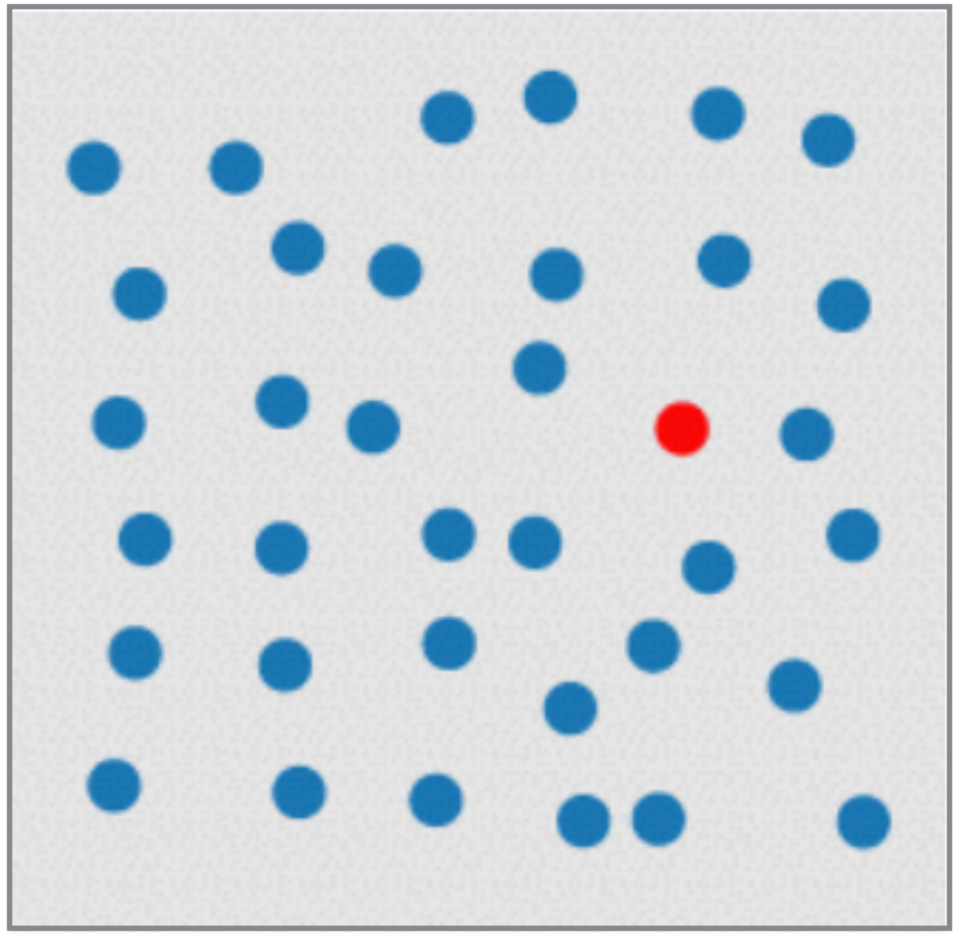
Pre-attentive processing

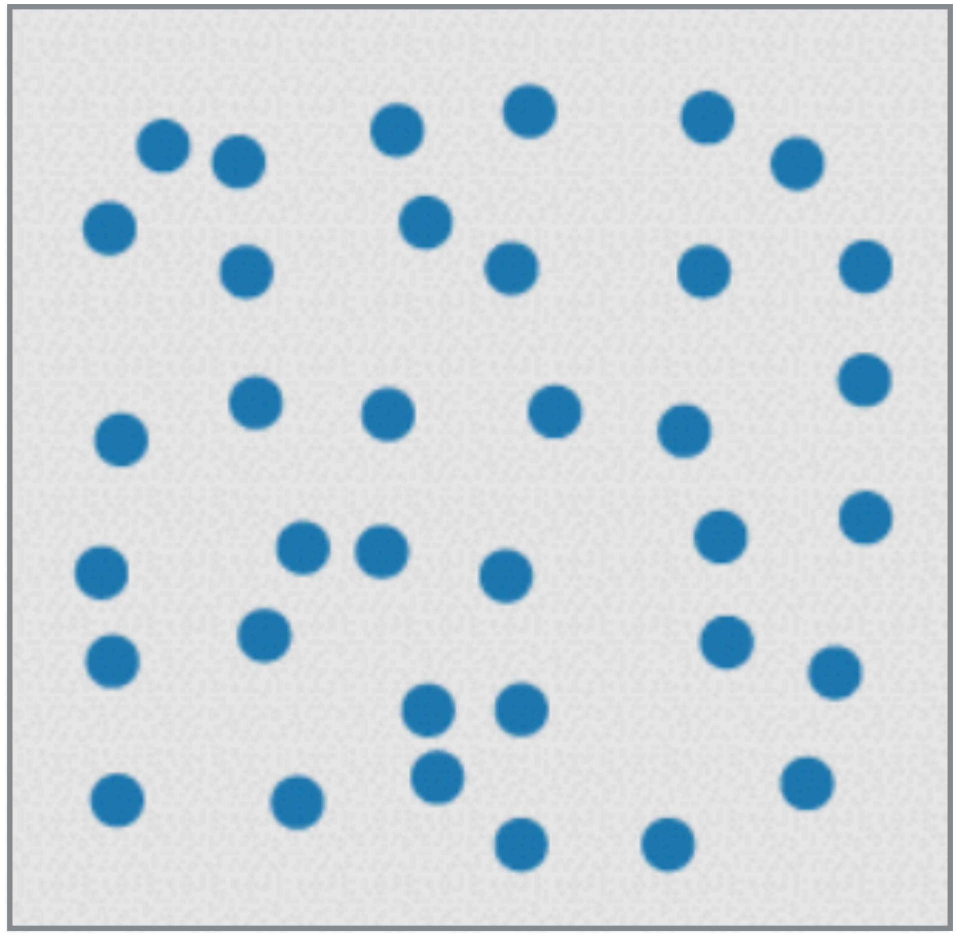
- Properties detected by the low-level visual system
  - Very rapid (200-250 milliseconds)
  - Very accurate
  - Processed in parallel
- Happens before focused attention
- Independent on the number of distractors

- As opposed to:
  - sequential search
  - cognitive process
  - needs attention

Pre-attentive processing

- Spot intruder in one of the two sets (0.2s)

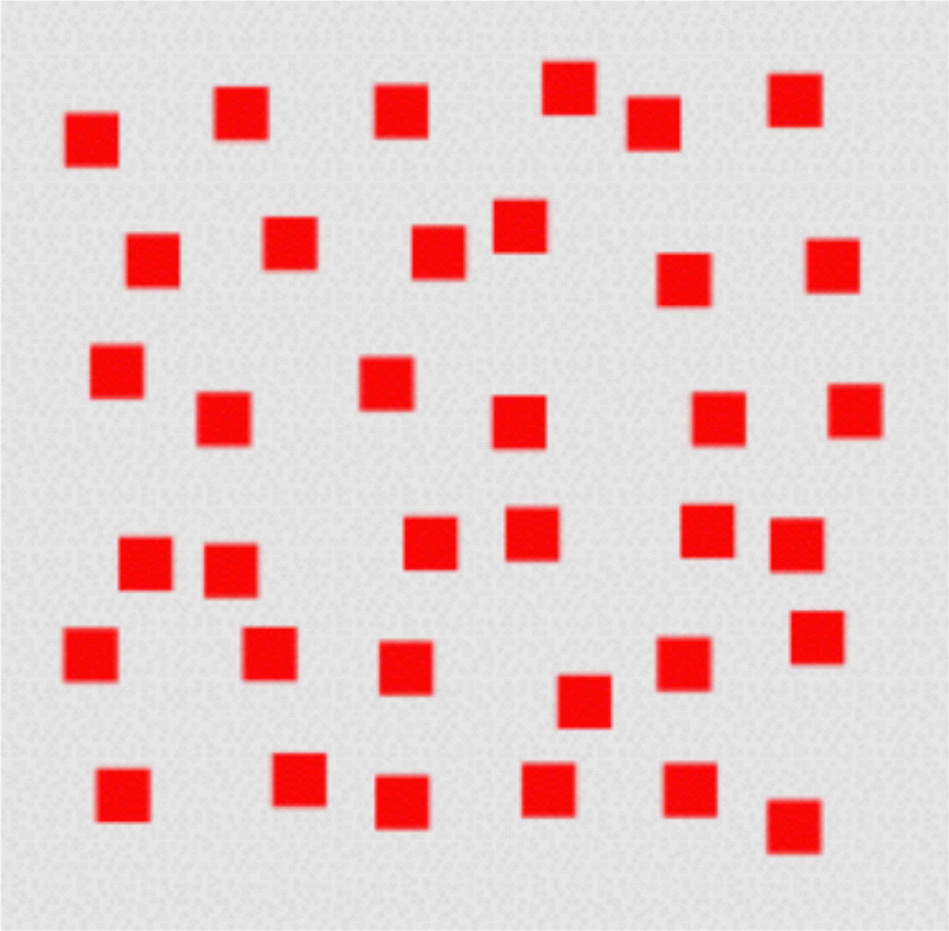


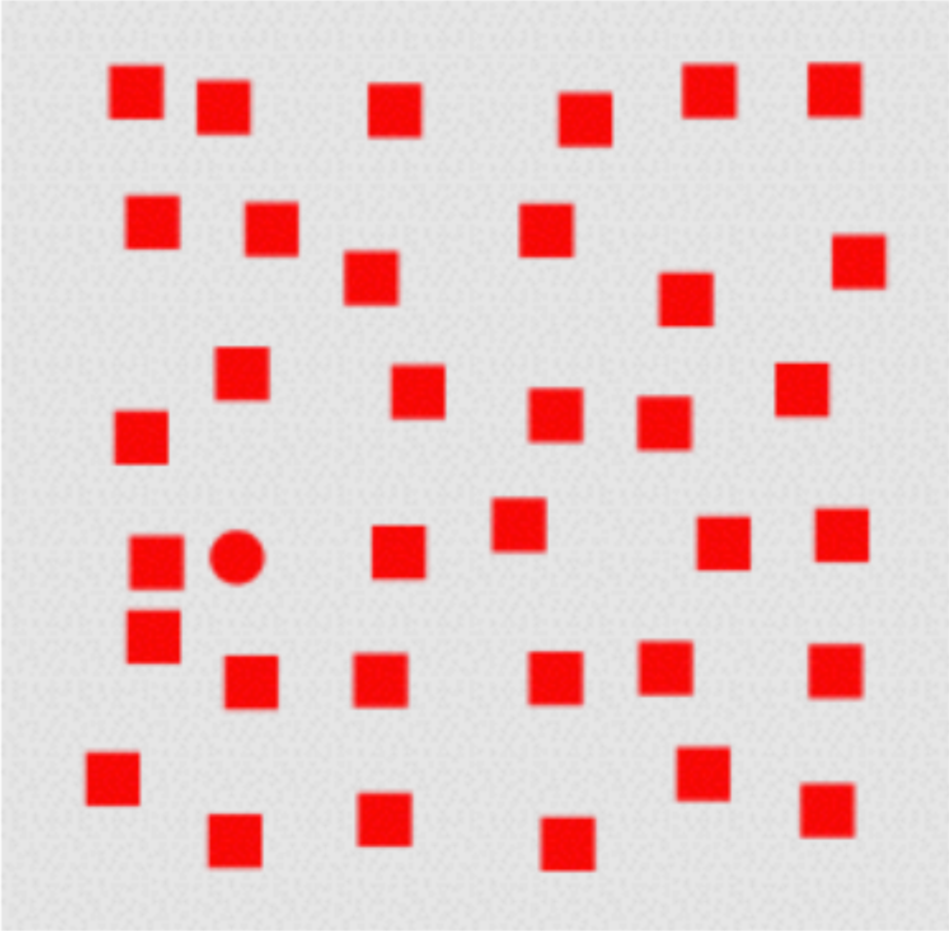


- Difference in hue: extremely effective

Pre-attentive processing

- Spot intruder in one of the two sets (0.5s)

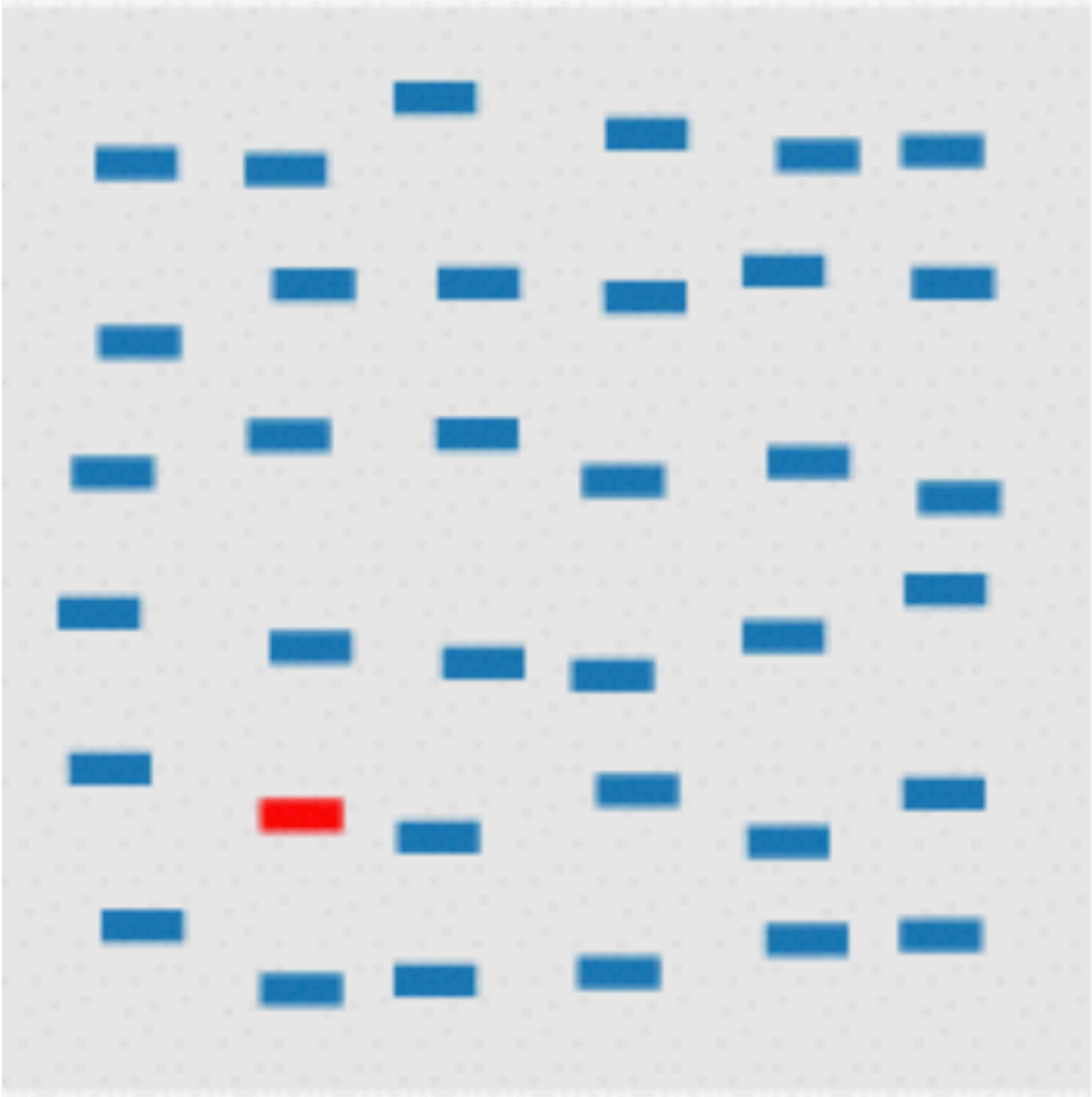




- Difference in shape: less effective

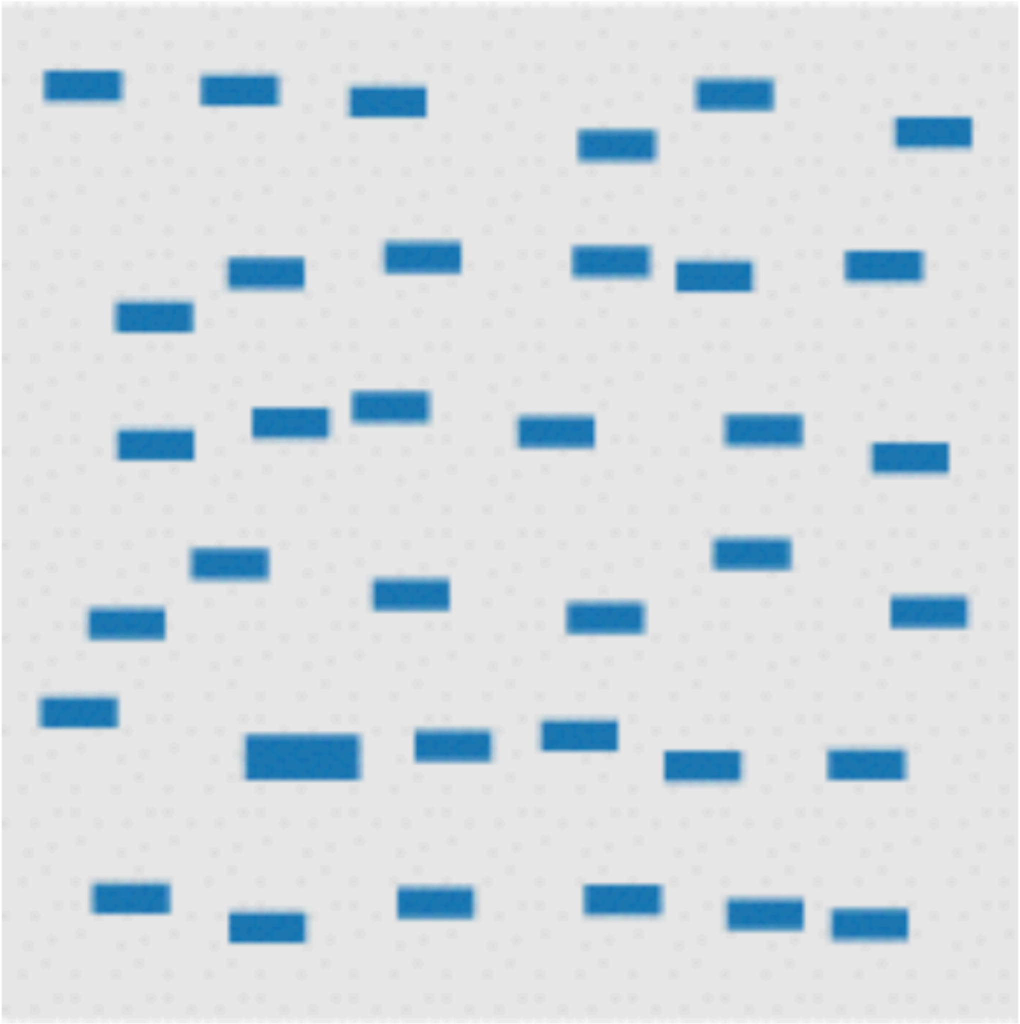
Pre-attentive processing

- Difference in orientation: less effective



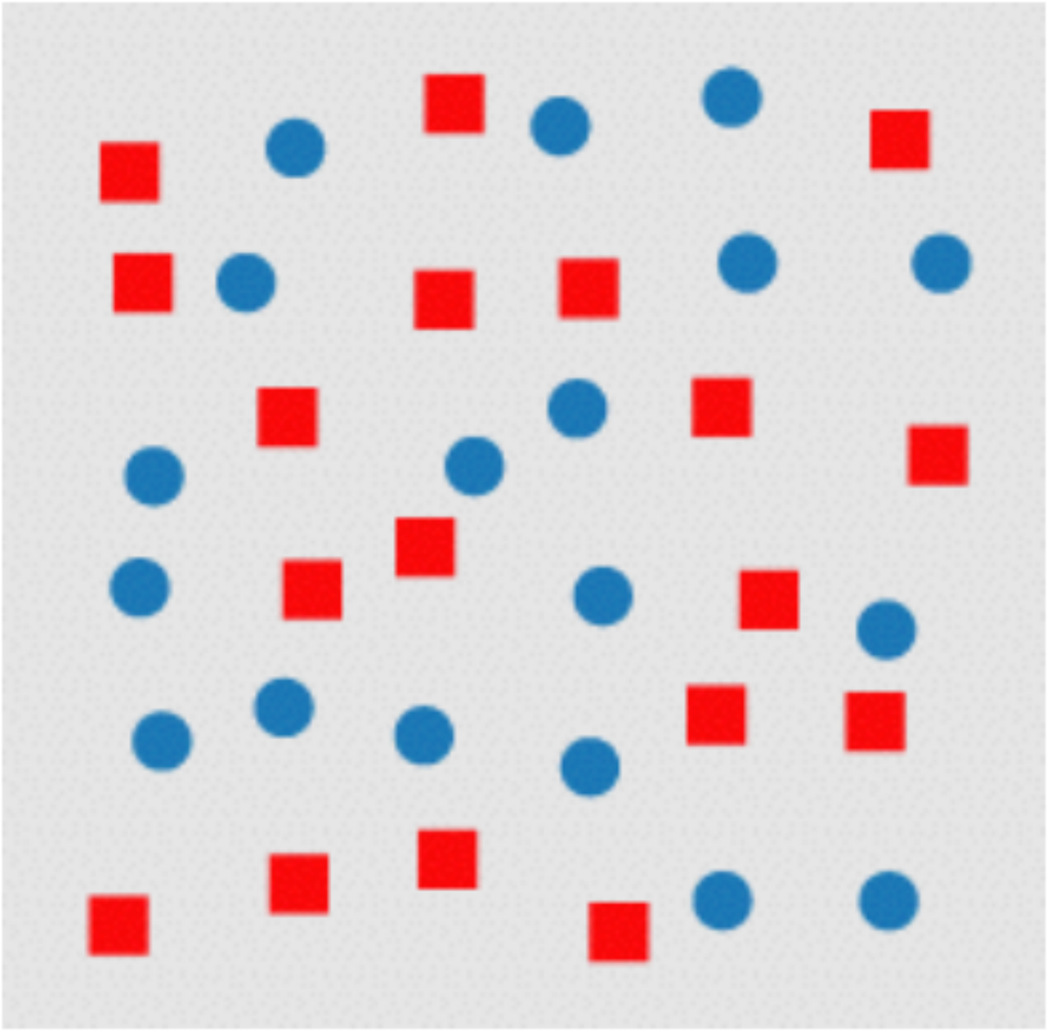
Pre-attentive processing

- Difference in closure / size : even less effective

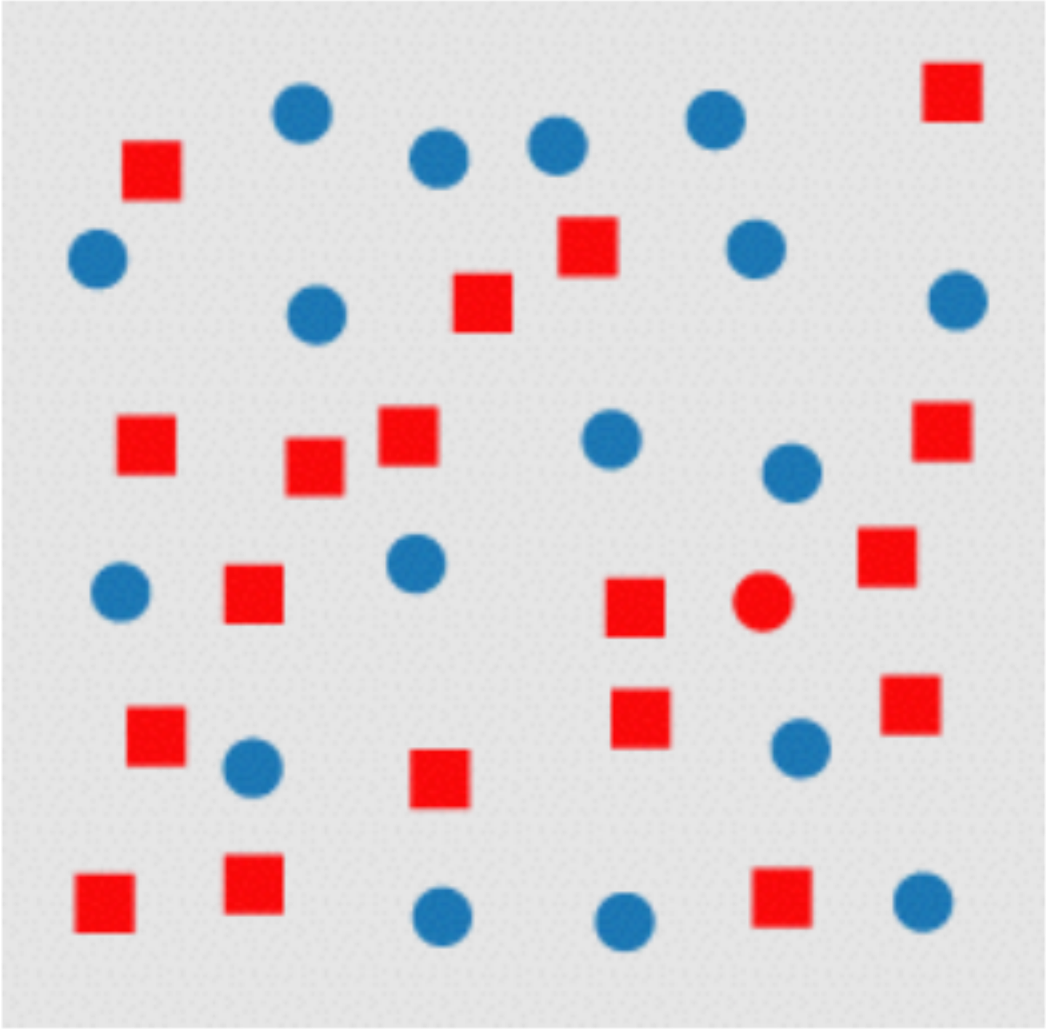


Pre-attentive processing

- No combination of different propreties!



Red squares
and blue circles
are easily taken apart



Red circle hard to spot
as distractor objects
have both properties

Pre-attentive processing

Tasks:

- target detection
  - detect the presence or absence of a target
- boundary detection
  - detect a texture boundary between two groups of elements
- region tracking
  - track one or more elements as they move in time and space
- counting and estimation
  - count / estimate the number of elements in a given class

Pre-attentive processing

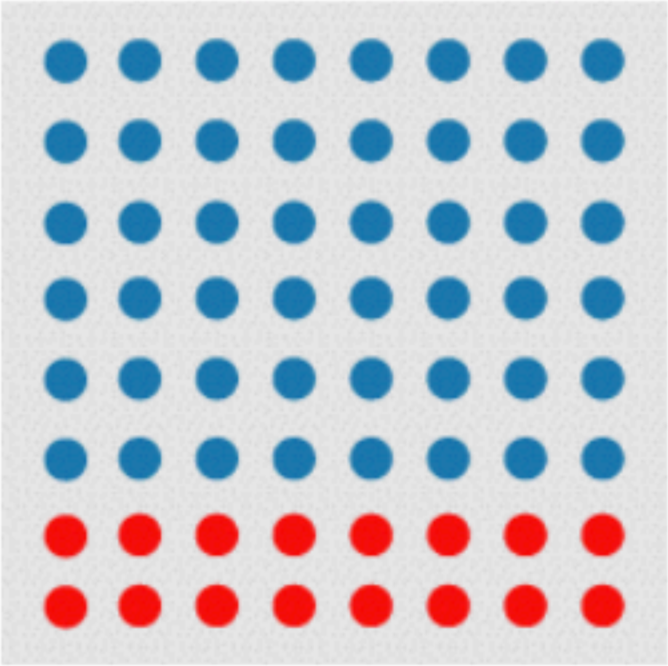
Tasks:

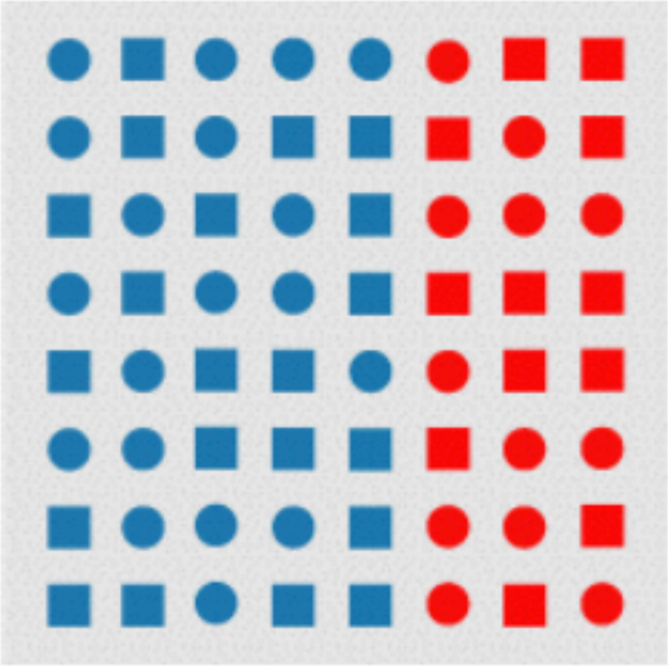
Number Estimation

Pre-attentive processing

Feature hierarchy:

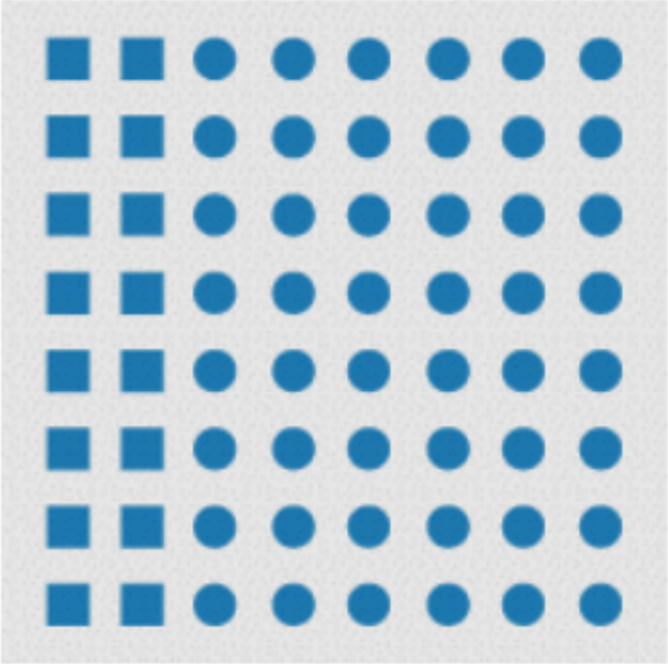
Hue
no distractors

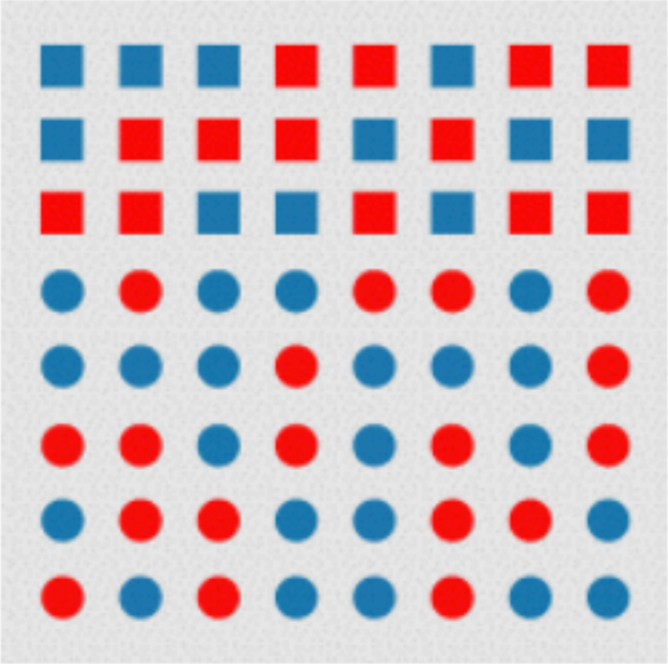




Hue
shape distractors

Shape
no distractors

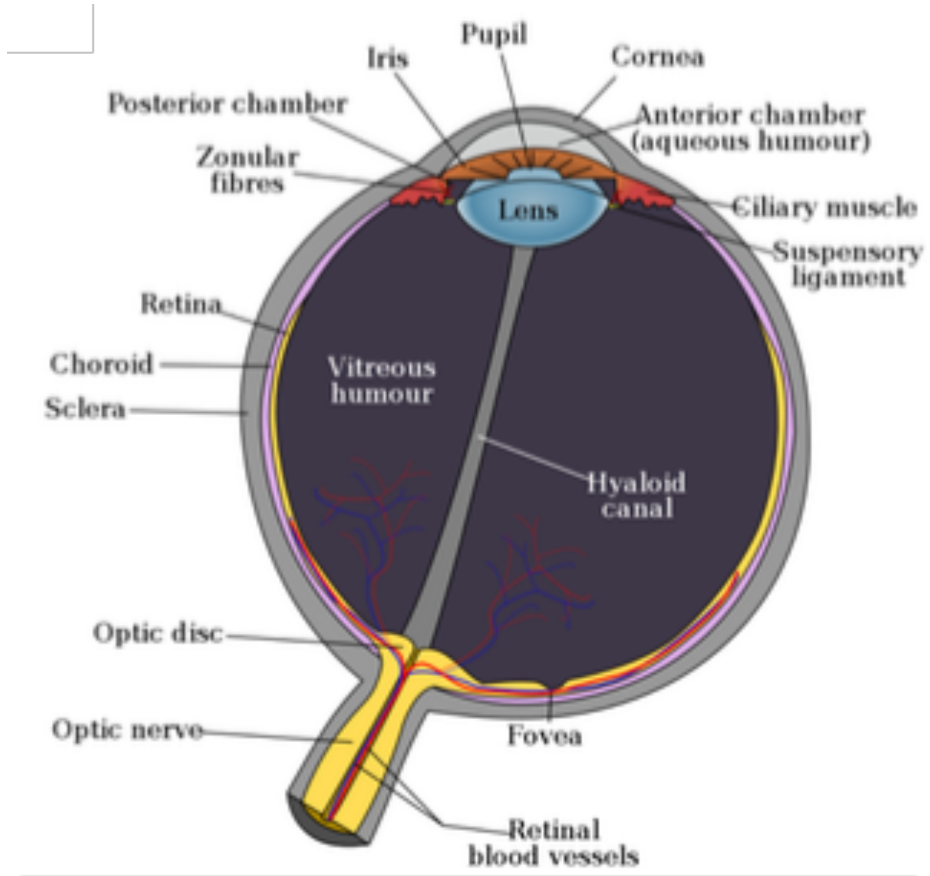


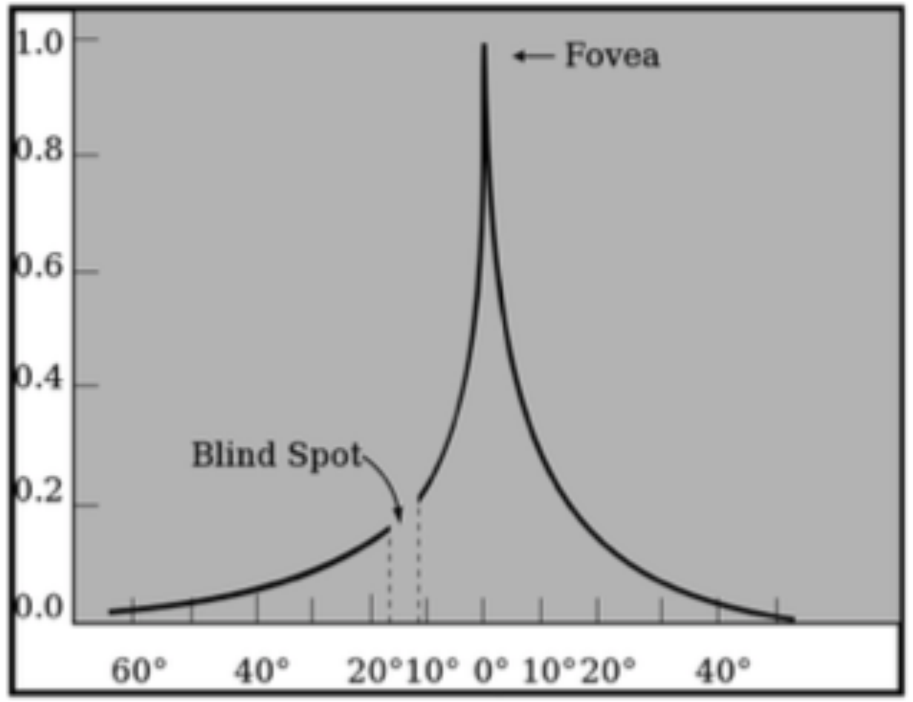


Shape
hue distractors

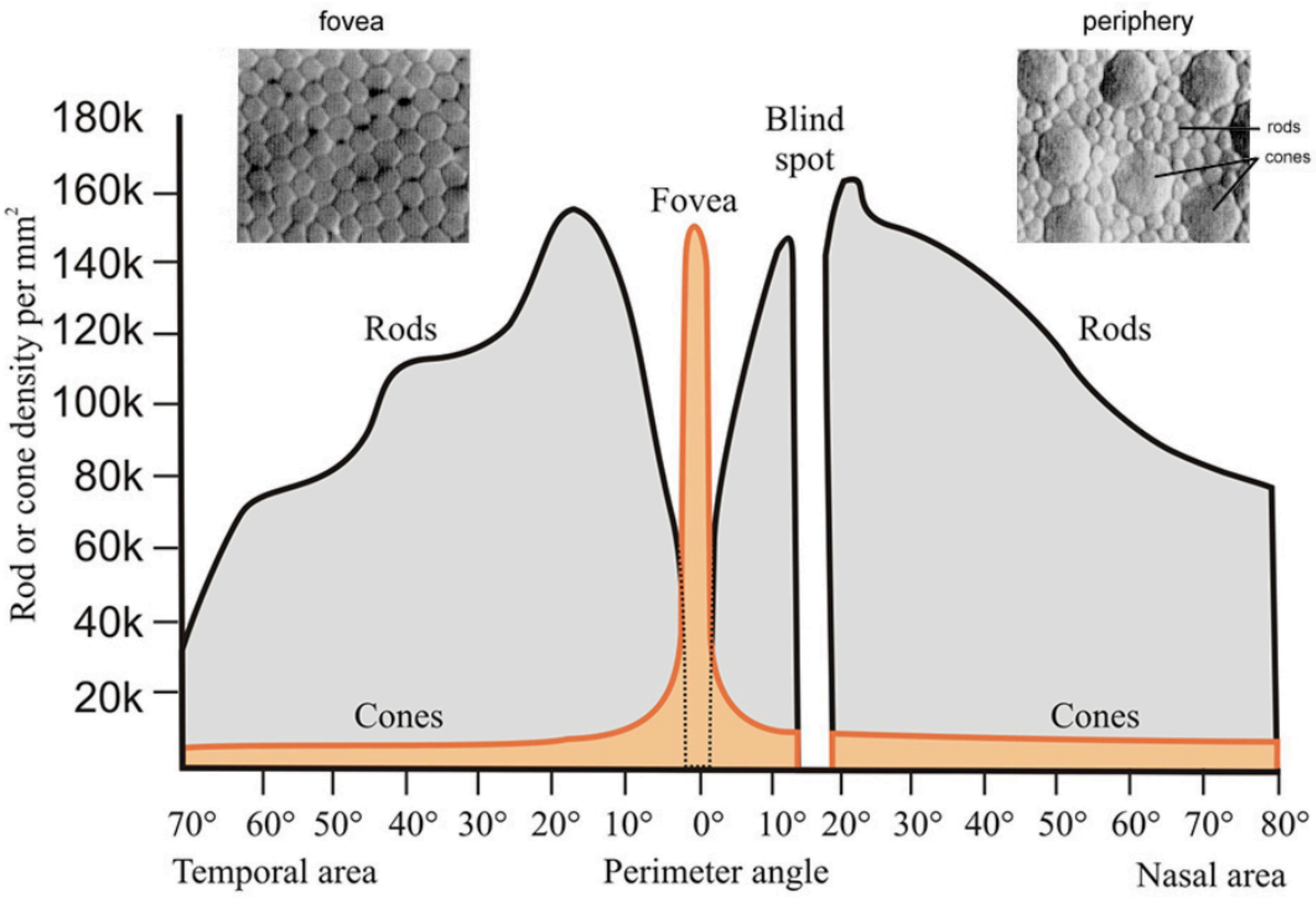
The human visual system

- Cones:
  - color vision
  - dense in the center
  - 5-6 millions
- Fovea:
  - central zone of retina
  - only cones, 27 times the density
  - responsible for central vision
- Rods:
  - b/w vision
  - peripheral vision
  - 120 millions



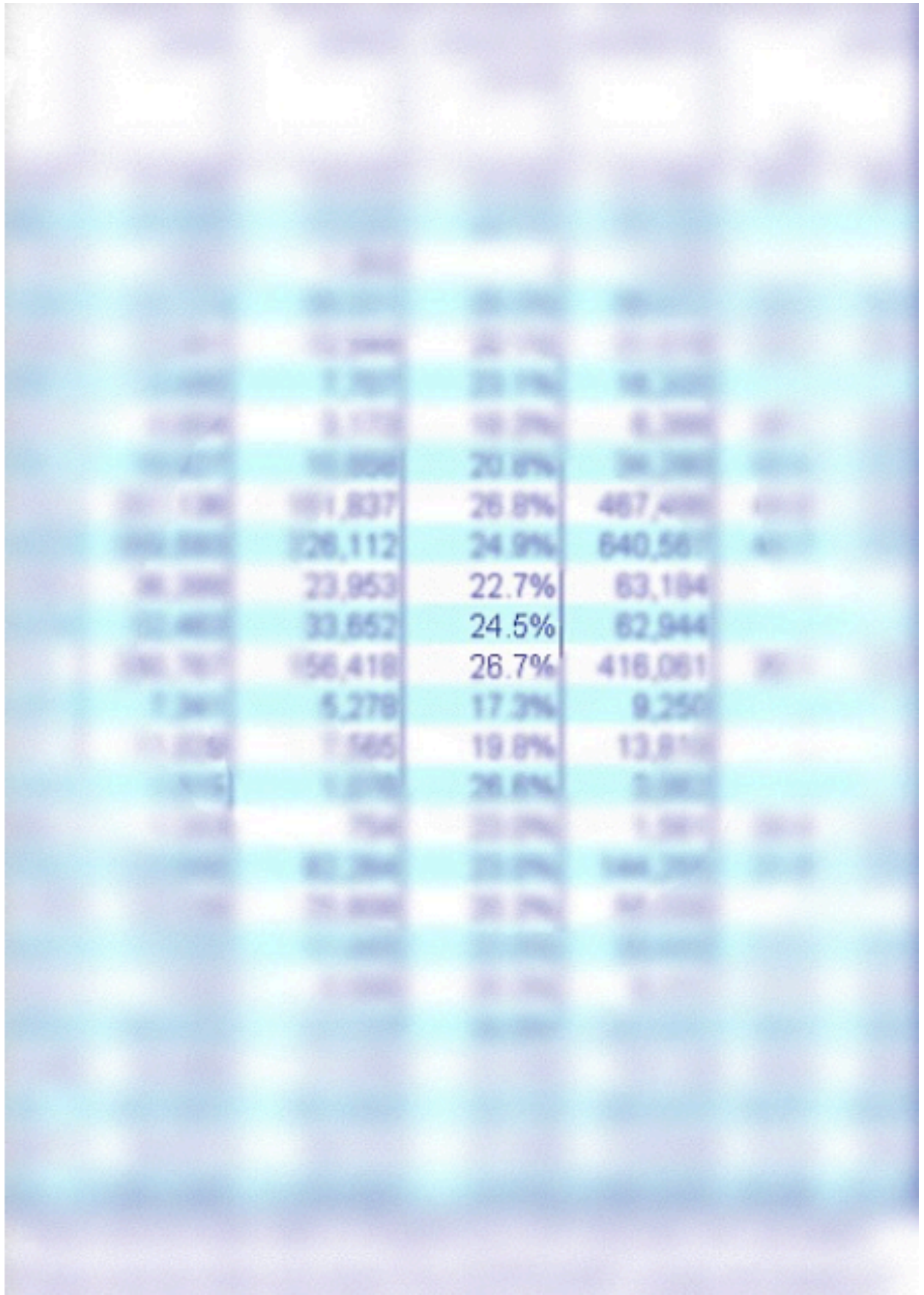


The human visual system



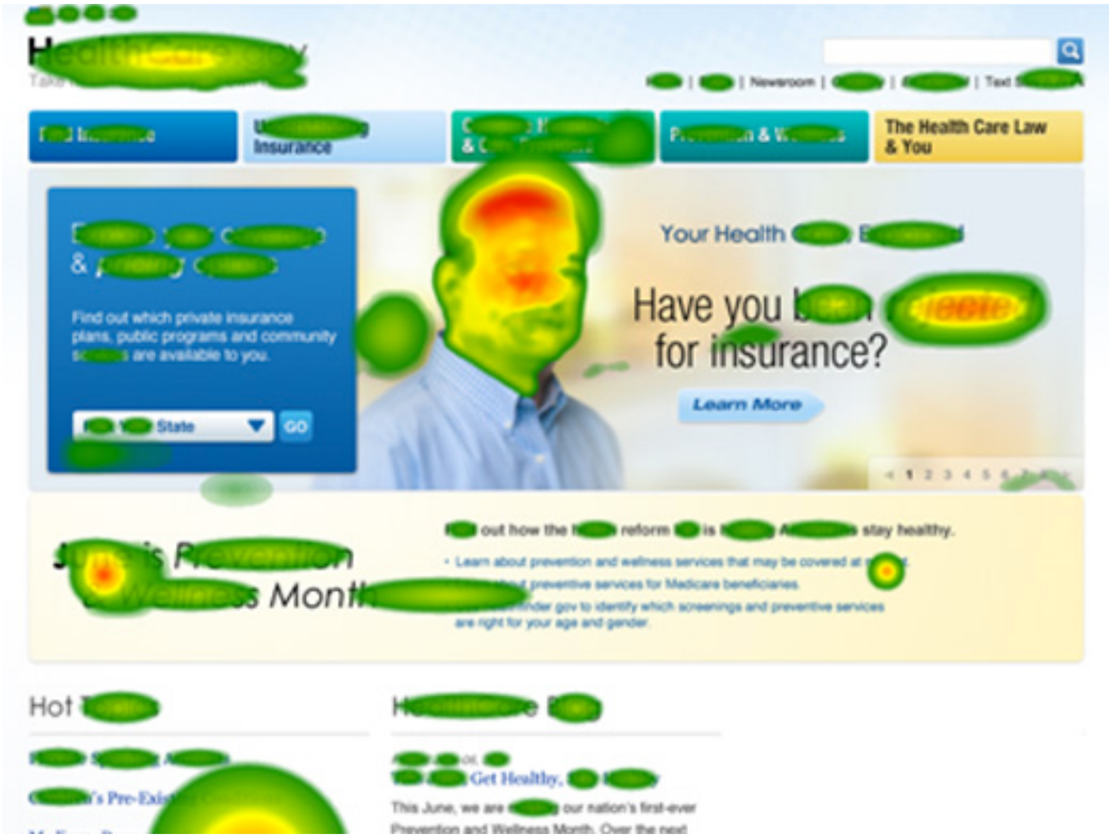
The human visual system

- The eye is not like a camera
- At each instant in time, our vision is sharp only in a very small area around the fovea
- All the rest is badly blurred and vaguely coloured
- Areas outside the fovea provide just contextual information



The human visual system

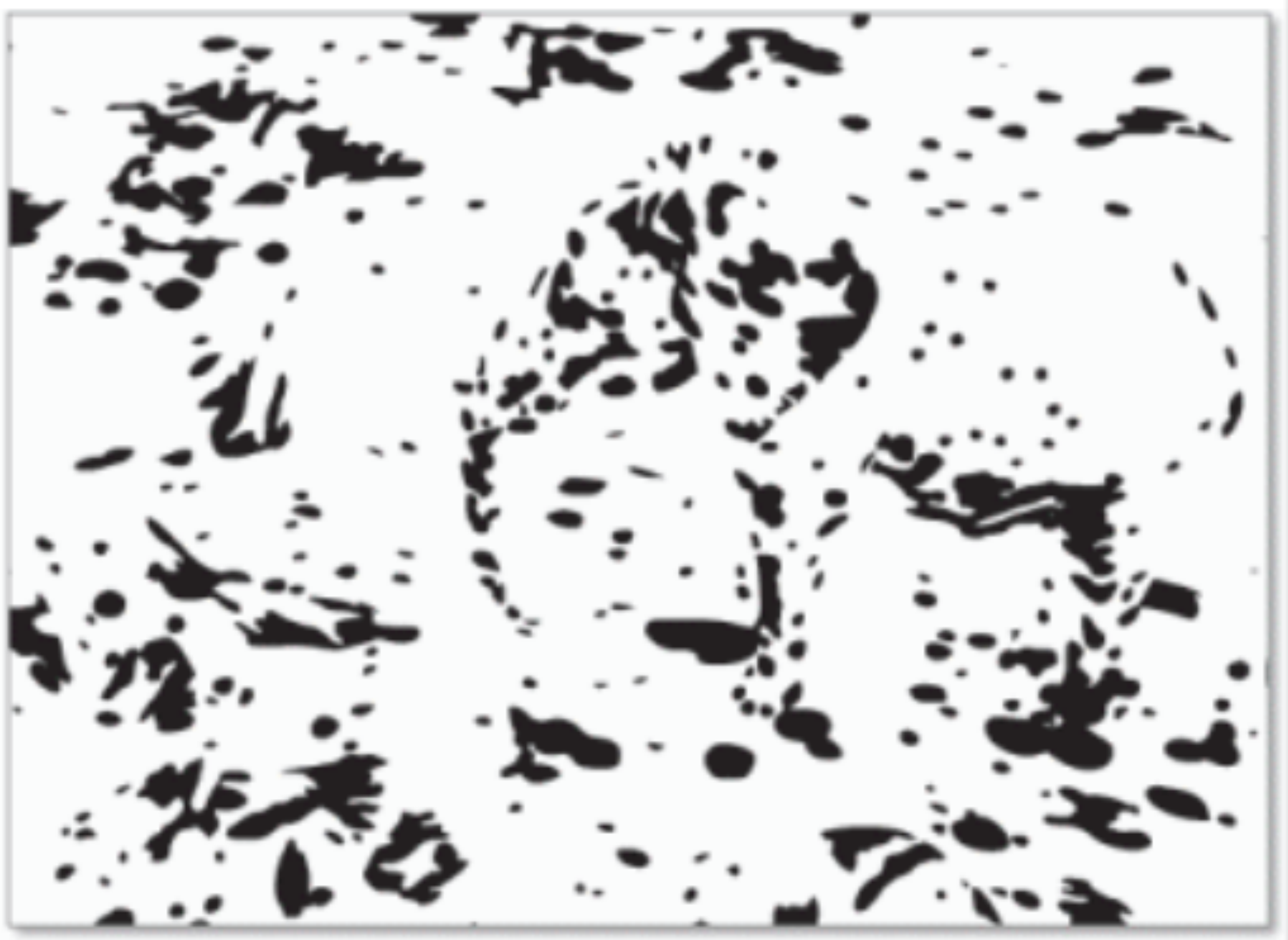
- Vision works as a sequence of *fixations* and *saccades*:
  - Fixation: maintaining gaze on a single location (200-600 ms)
  - Saccade: moving between different locations (20-100 ms)
- Our mental images are a form of integration of several snapshots focused at different spots, as obtained with fixation
- Where do we fixate our gaze?
  - where there's something attractive
  - where we know there's something we want to look at



Emergence images: perceptual hysteresis

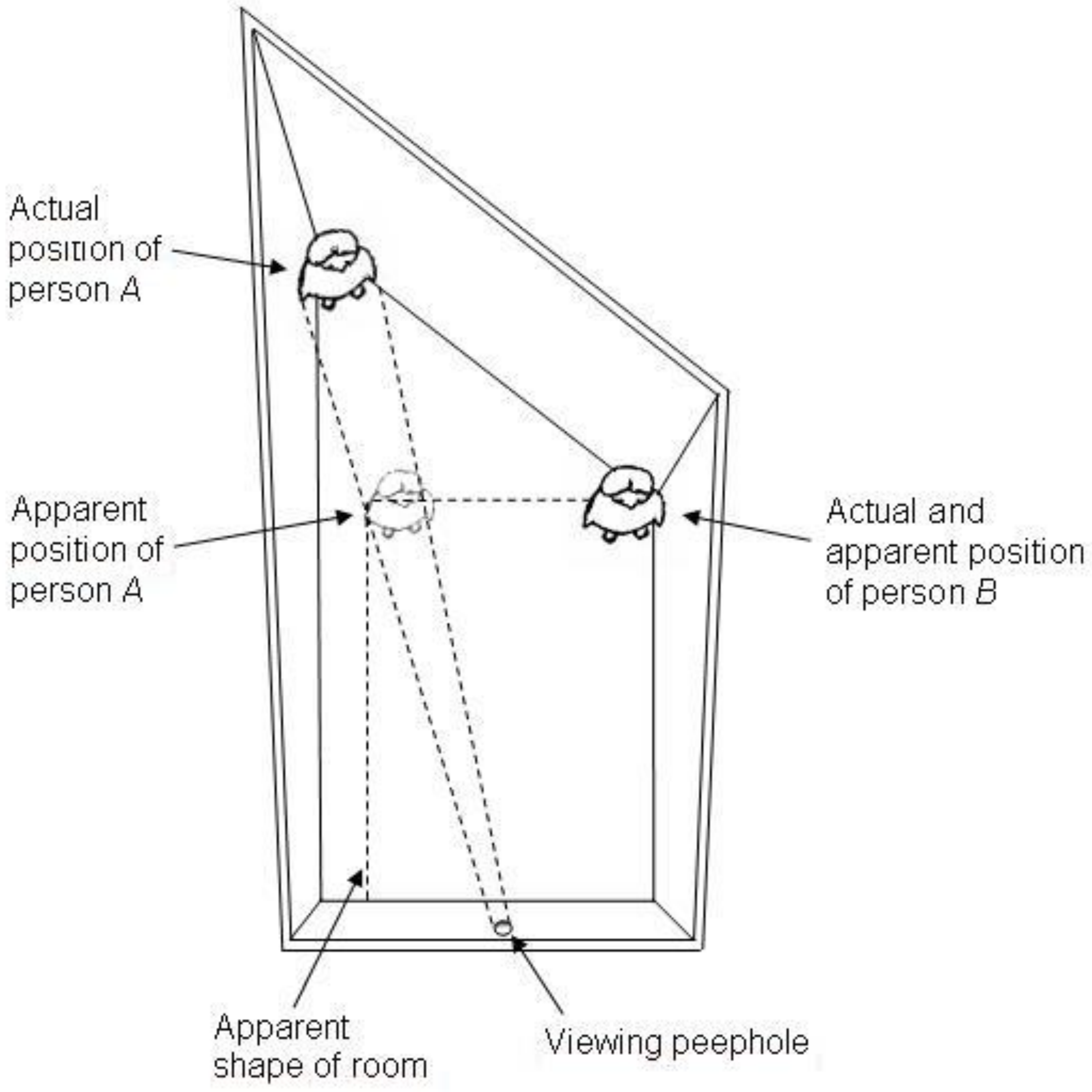
- Human visual system is very good at aggregating information from seemingly meaningless pieces
- What we see depends on our priors: goals and expectations

Vision is built top-down from the input



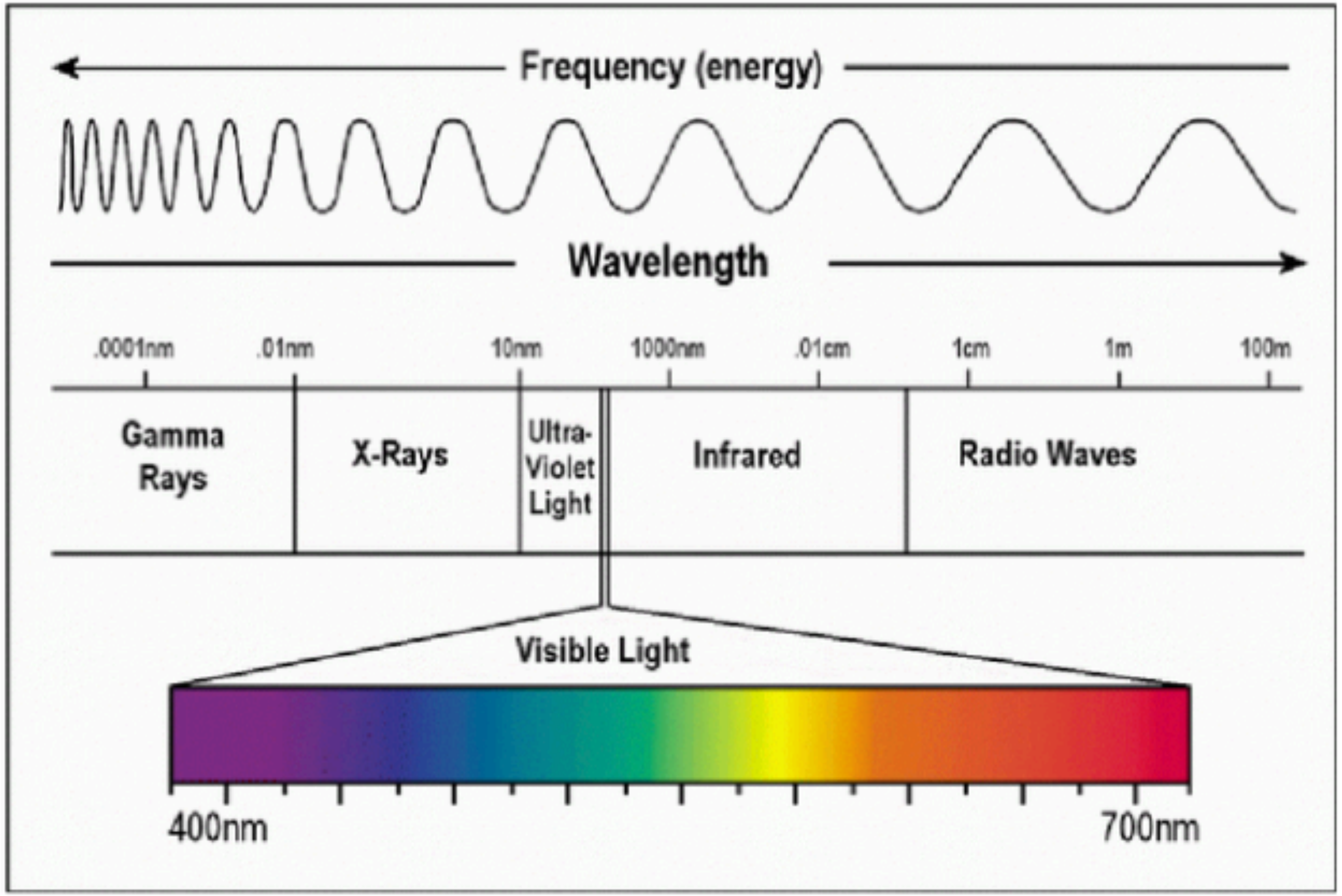
Ames room illusion

- https://www.youtube.com/watch?v=Ttd0YjXF0no



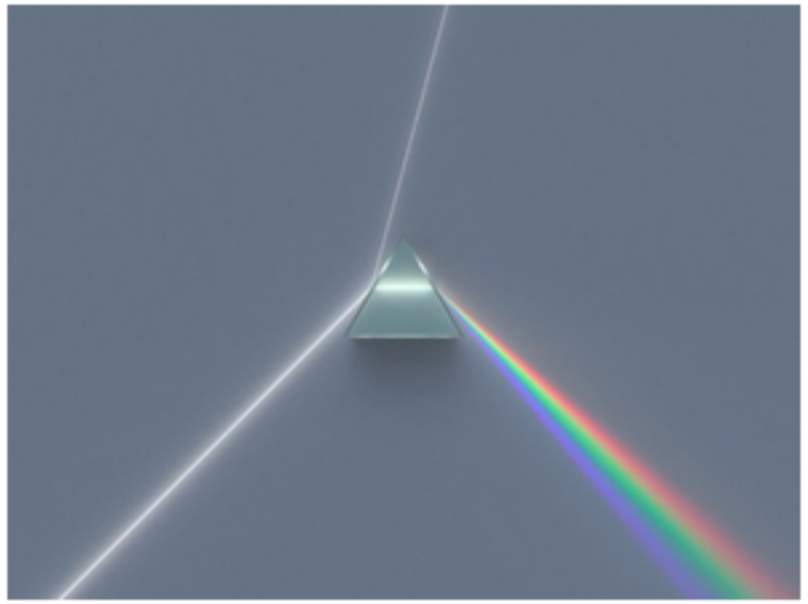
Color vision

- Electromagnetic spectrum: .0001 nm - 100 m wavelength



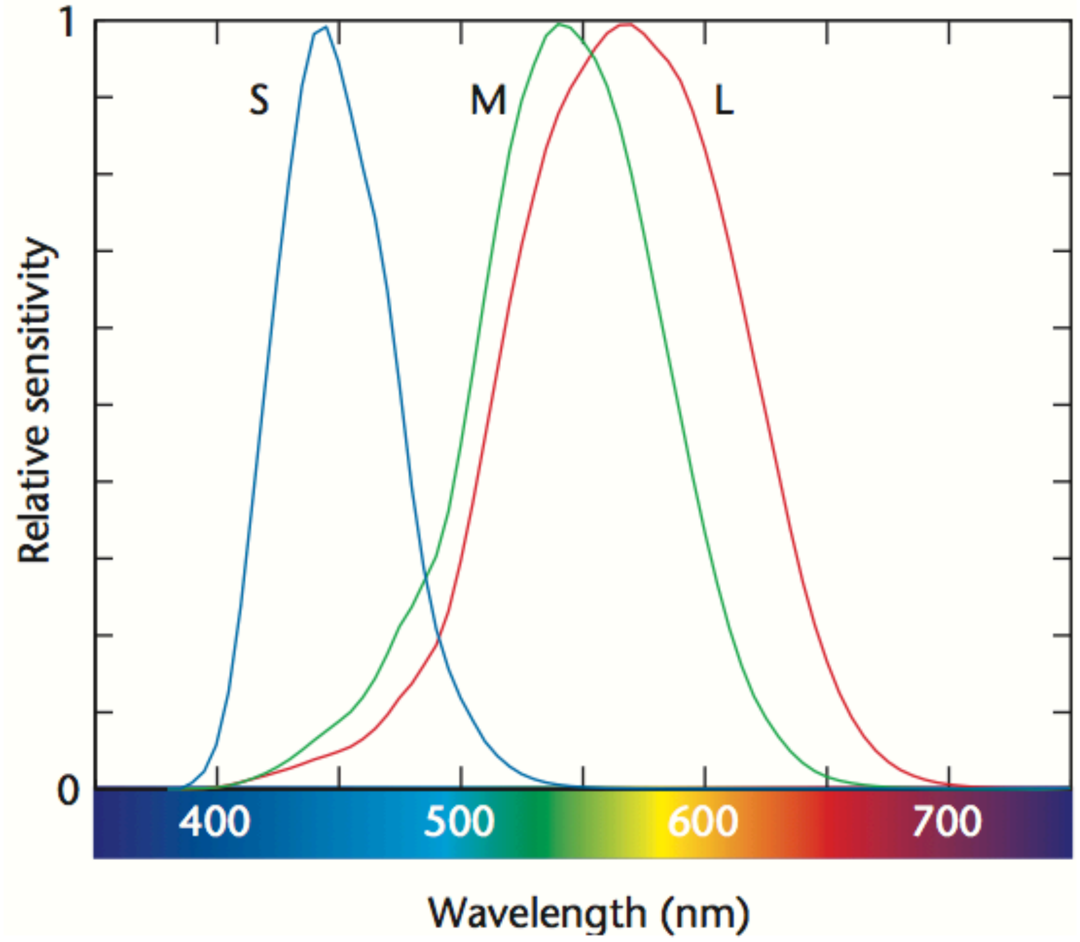
Color vision

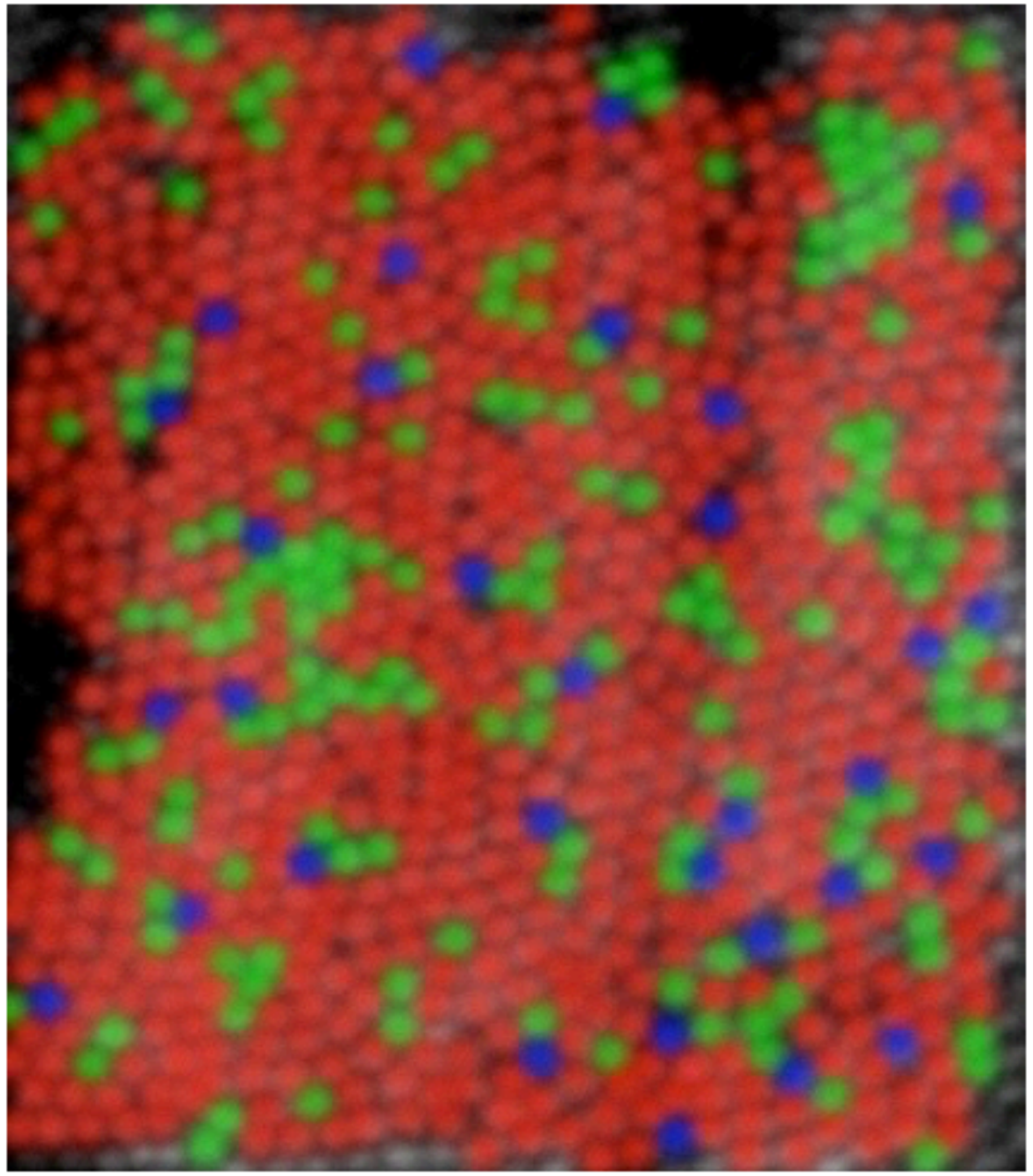
- Visible part of electromagnetic spectrum
  - 390-750 nm
- Spectral colors
  - evoked by a single wavelength
- Other colors: unsaturated colors
  - Mix of multiple wavelengths
  - Gray scale (achromatic colors): mix of all visible wavelengths



Color vision

- Three different types of cones:
  - present with different densities
  - each type sensitive to a different range of frequencies in the visible spectrum
  - called L, M, S from the zones of spectrum at which they have highest response

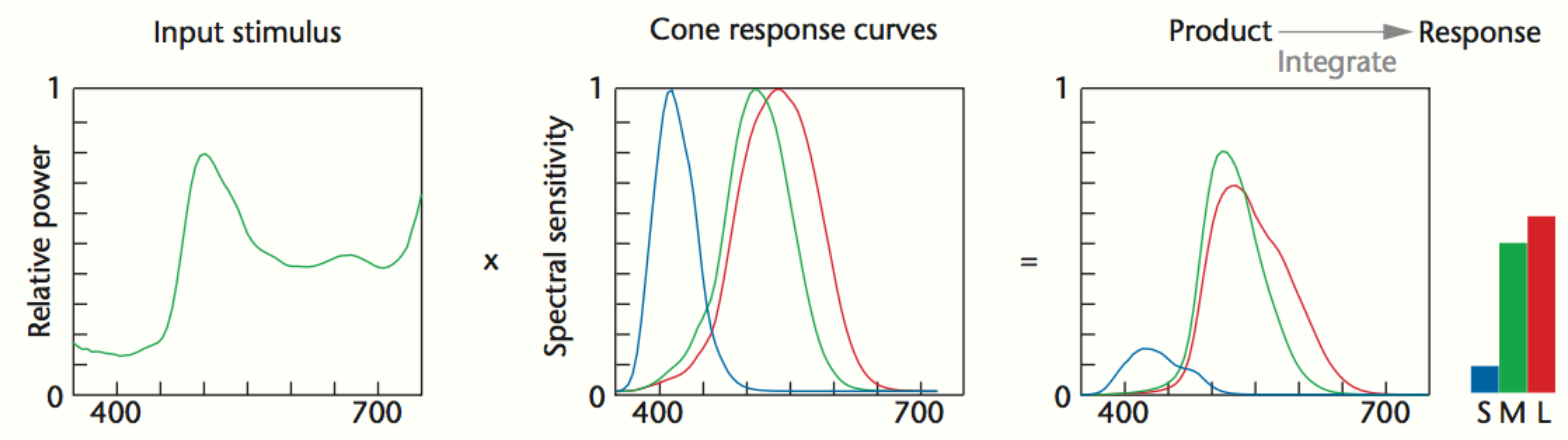




R 63% - G 31% - B 6%

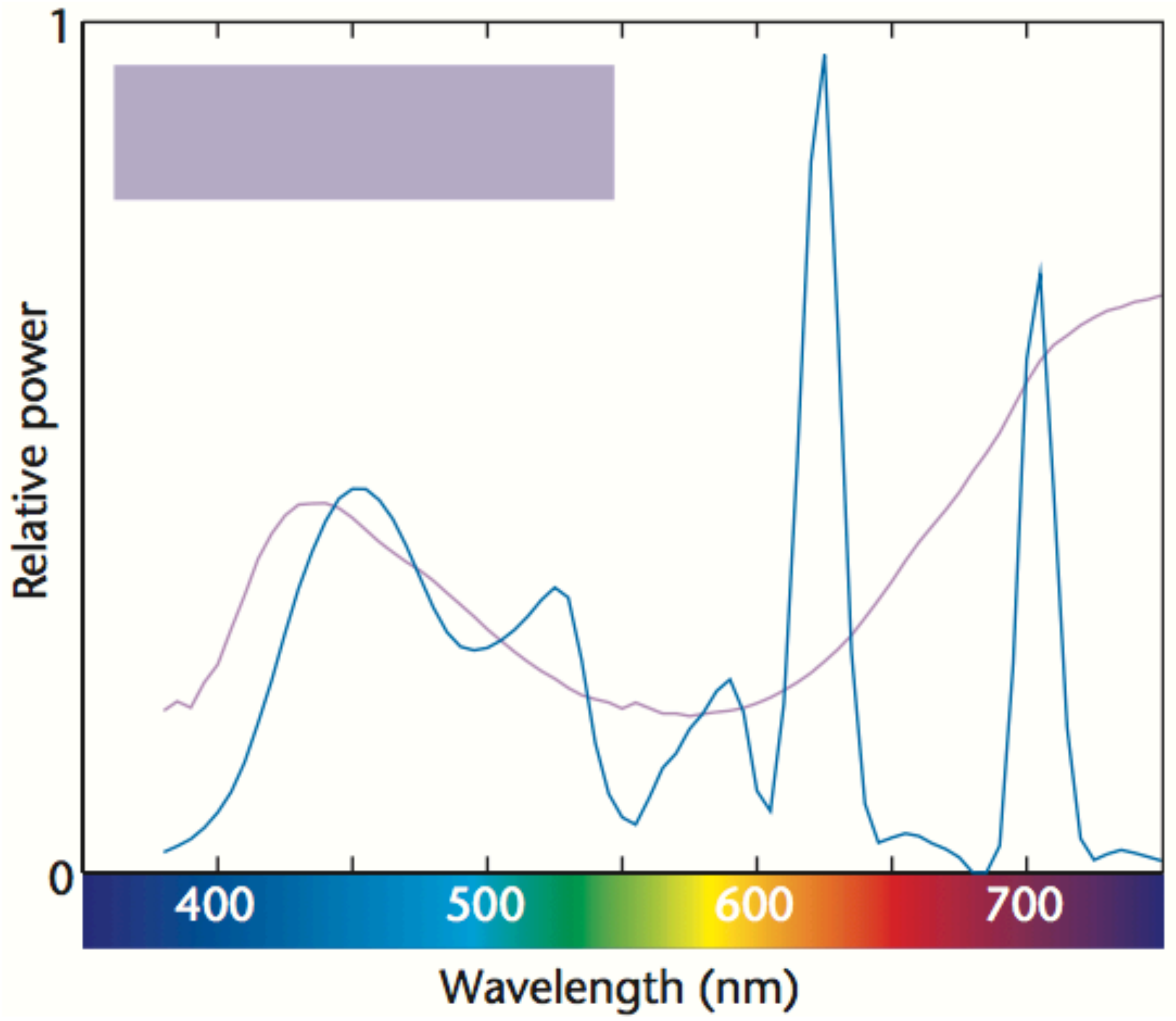
Color vision

- Trichromacy: all spectra can be reduced to precisely three values without loss of information with respect to the visual system.
- Metamerism: any spectra that create the same trichromatic response are indistinguishable.



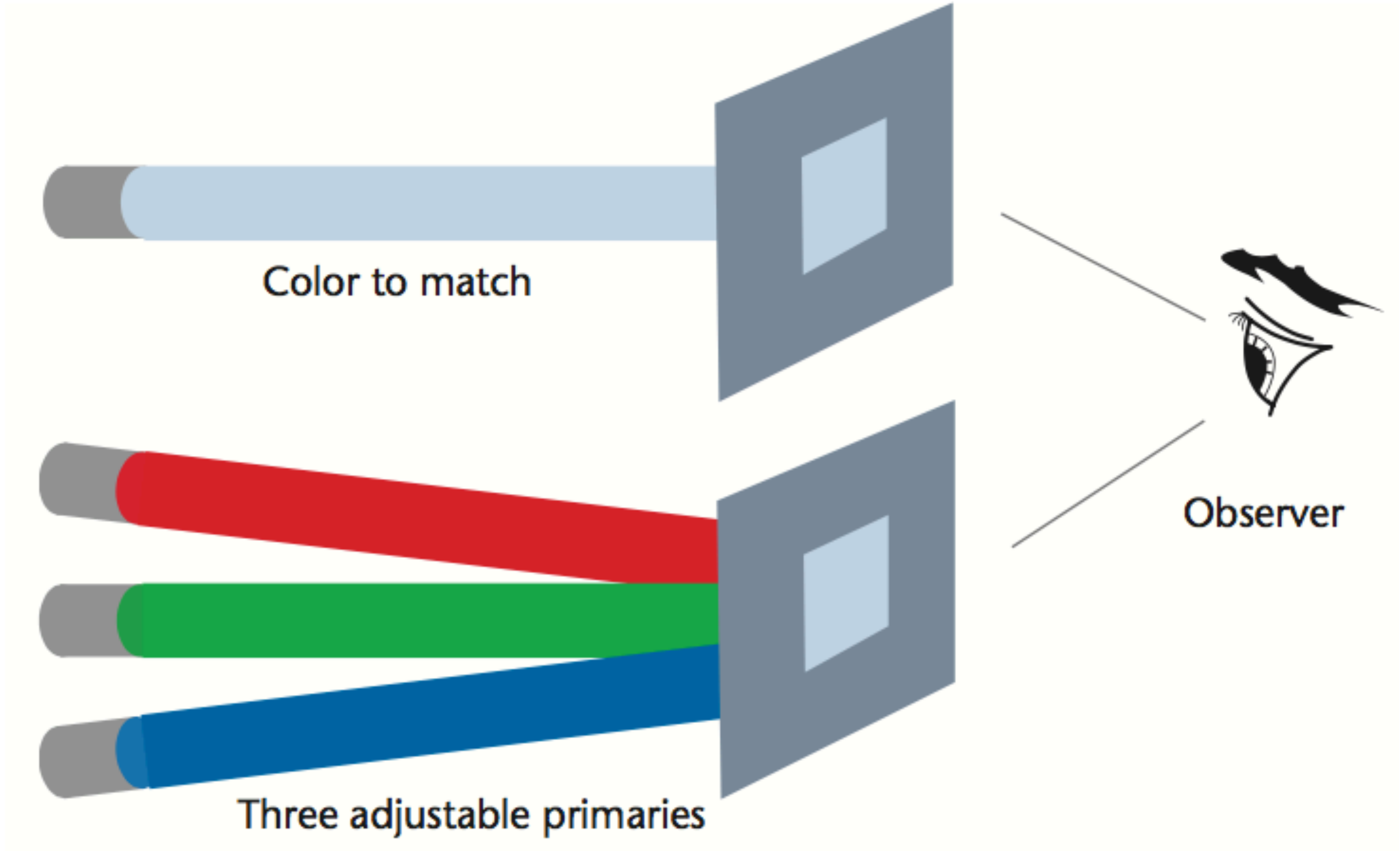
Color vision

- Metamerism:
  - brain sees only cone response
  - different spectra appear the same



Color vision

- Color matching experiment:
  - Adjust the levels of three different light sources (e.g., R, G, B) to match a color in the spectrum
  - The levels of the three sources encode the given color with respect to such sources
- Repeat with many subjects & many colors
  - Obtain a discrete encoding map

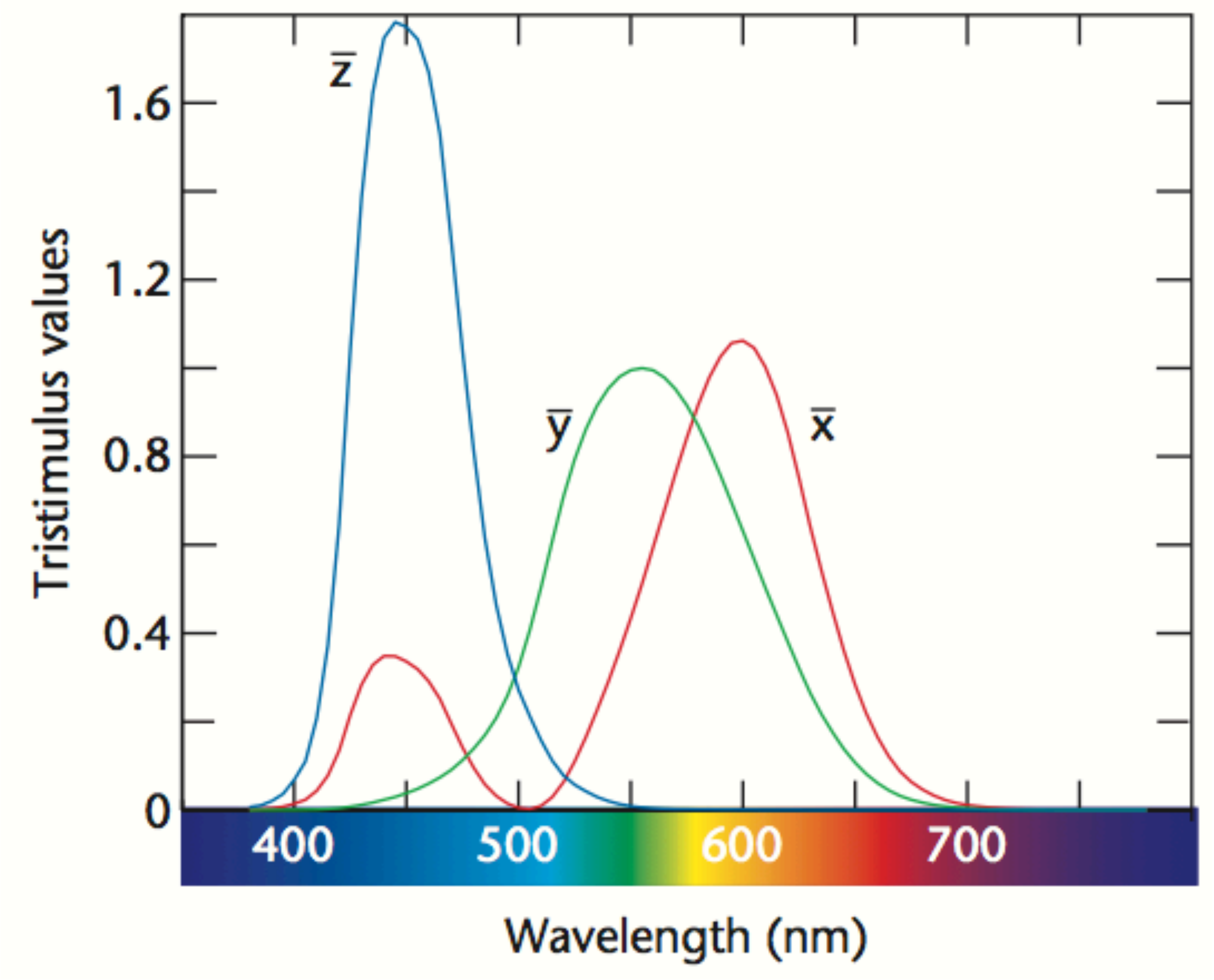


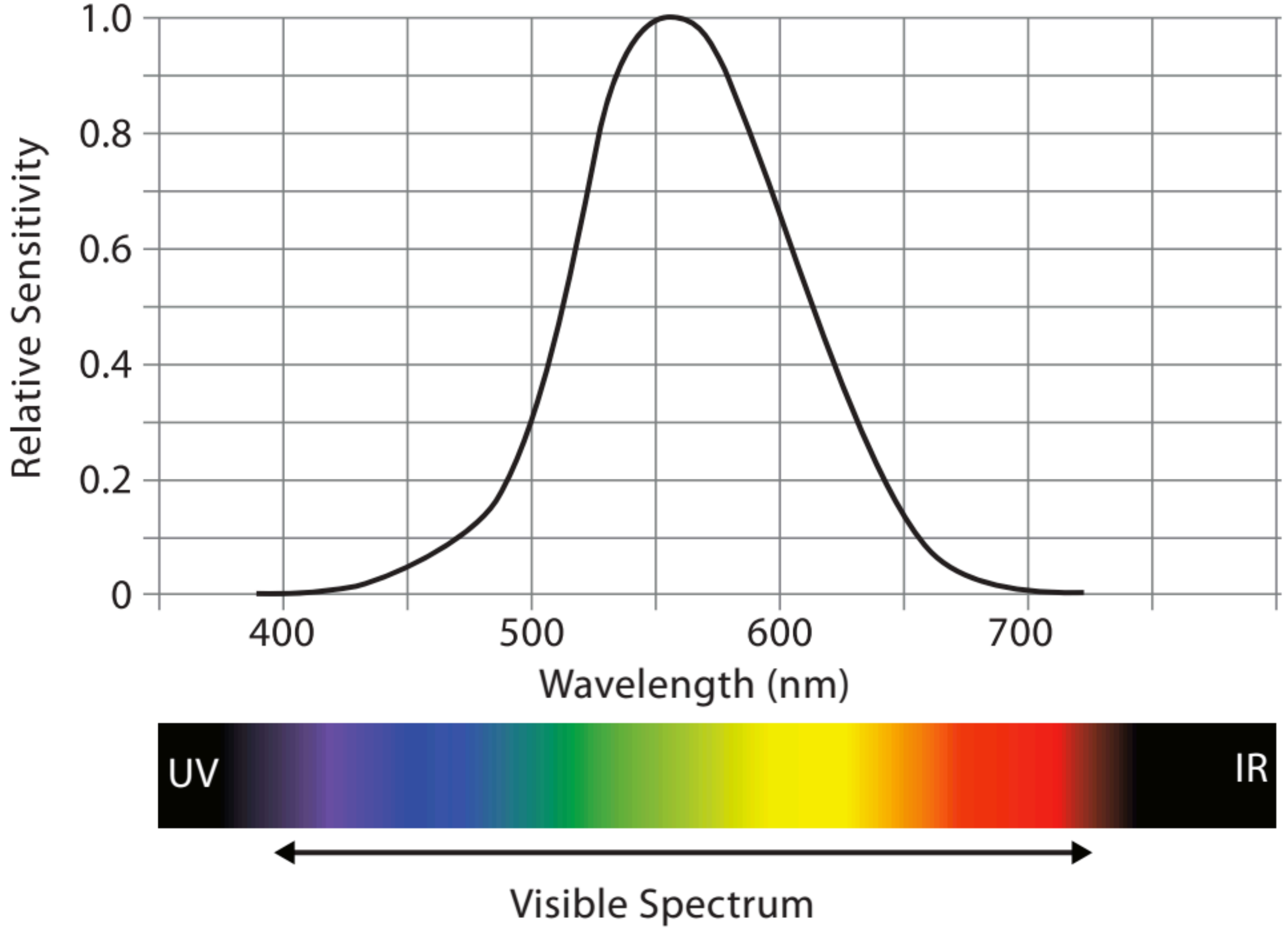
Color vision

- Grassman's law: additive property of tri-stimuli
  - if tri-stimulus RGB_1 matches spectrum S_1 and tri-stimulus RGB_2 matches spectrum S_2, then tri-stimulus RGB_1+RGB_2 matches spectrum S_1+S_2
- Consequence: on the basis of a finite set of matches we can encode an infinite set of colors
- In practice:
  - select three monochrome light sources, e.g.: R 645 nm, G 526 nm, B 444 nm
  - sample the visible spectrum and match tri-stimuli for each sample
  - interpolate color matching functions on the three channels

Color vision

- Color matching function: the CIE tri-stimulus values XYZ
  – results from algebraic manipulation of measured matches
  – value Y set to match perceived brightness (luminance)

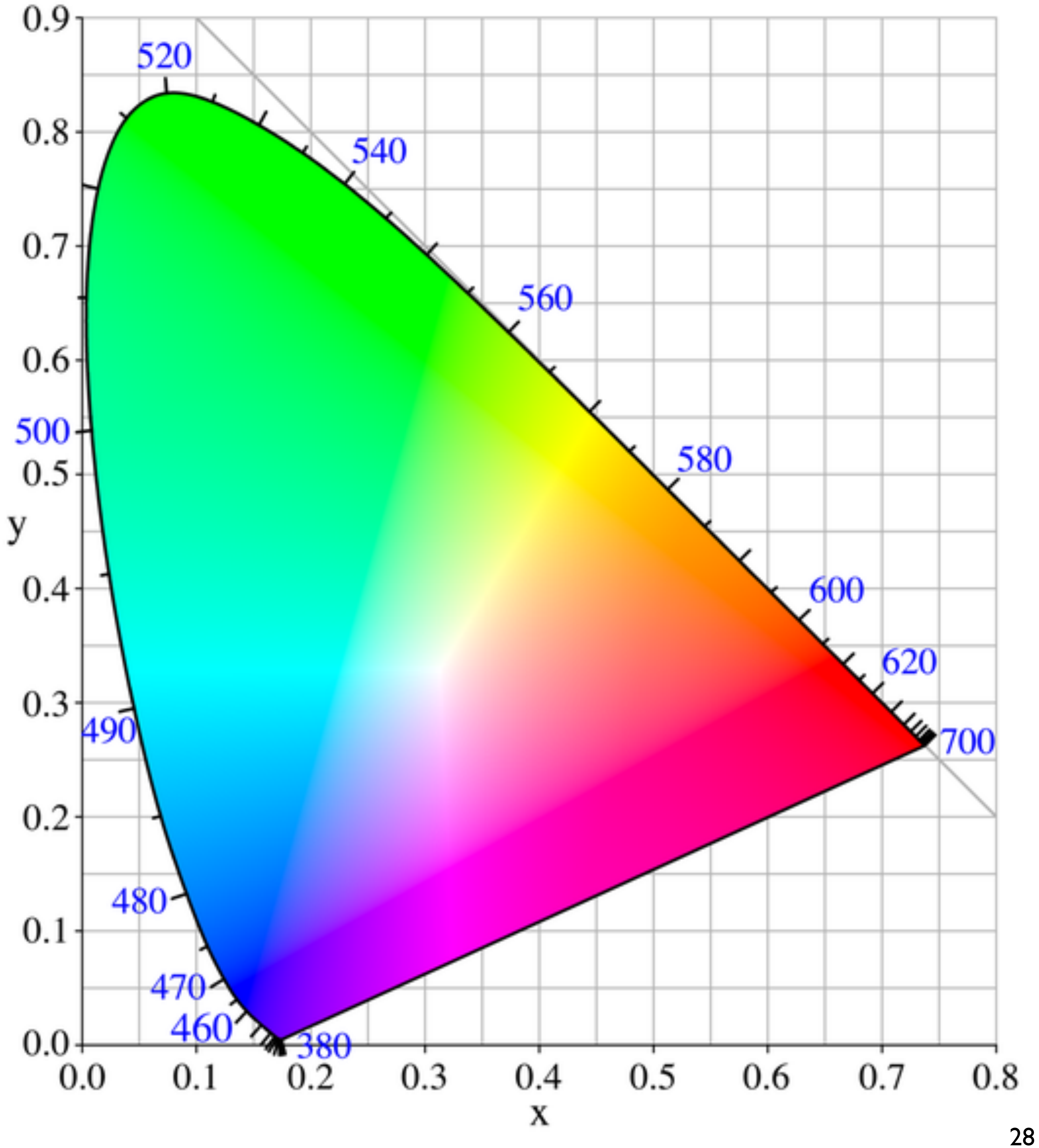




Color vision

- CIE chromaticity diagram
  - factor our *luminance*
  - x = X/(X+Y+Z)   y = Y/(X+Y+Z)
  - any color can be encoded as xyY
  - *chromaticity* encoded as (x,y)

  - pure saturated colors on the horseshoe-shaped path
  - the straight line below is called the *purple line*
  - the more internal, the less saturated
  - greyscale lies at the central point



Color vision

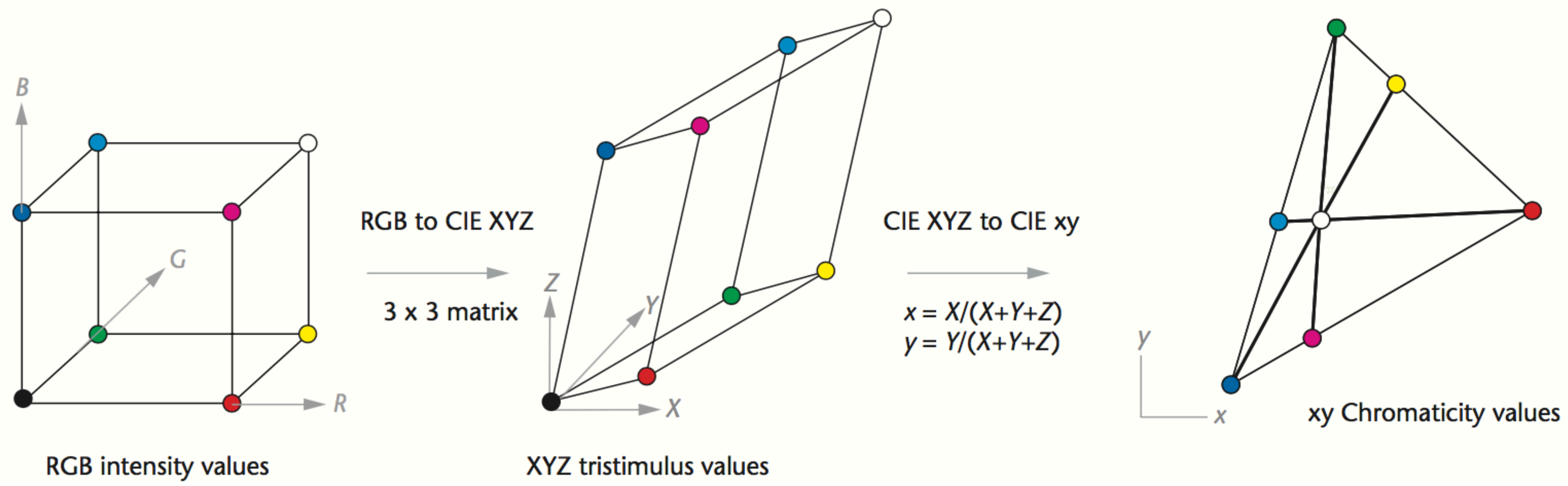
- The RGB space
  - Three sources of light R, G, B, adjustable in the range [0,1]
  - Possible value span a cube in a 3D space RGB
  - Map RGB space to XYZ space by determining the tri-stimulus for each base color

$$\begin{bmatrix} R & G & B \end{bmatrix} M = \begin{bmatrix} X & Y & Z \end{bmatrix}$$

$$M = \begin{bmatrix} X_R & Y_R & Z_R \\ X_G & Y_G & Z_G \\ X_B & Y_B & Z_B \end{bmatrix}$$

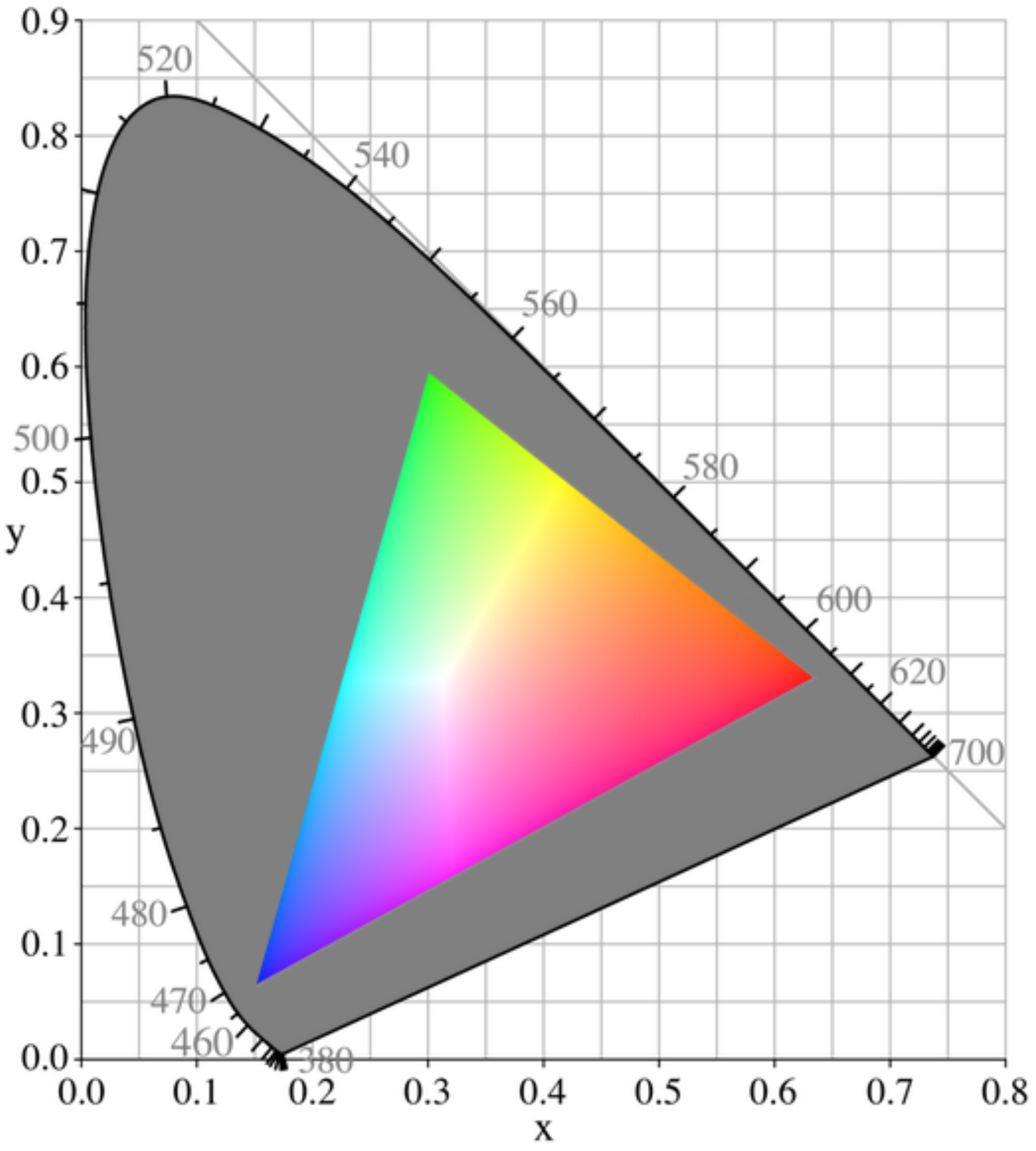
Color vision

- The RGB space maps to a triangle in the chromaticity diagram



Color vision

- The portion of chromaticity diagram spanned by the image of a color space is its *Gamut*
- The gamut of RGB space is rather small
- RGB is *not* safe to interpolate
  - don't know what to expect when interpolating between (R_1,G_1,B_1) and (R_2,G_2,B_2)

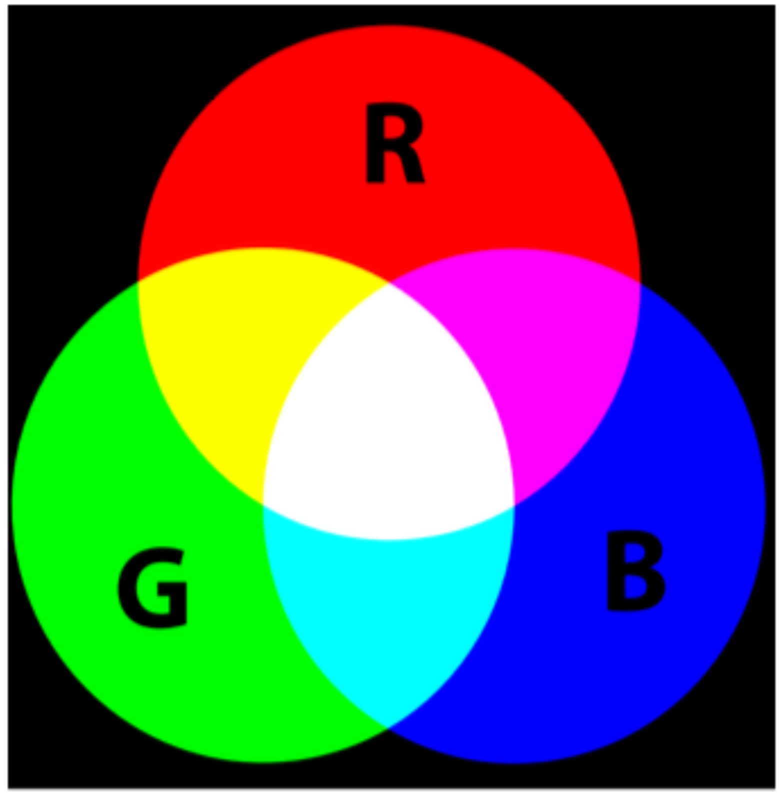


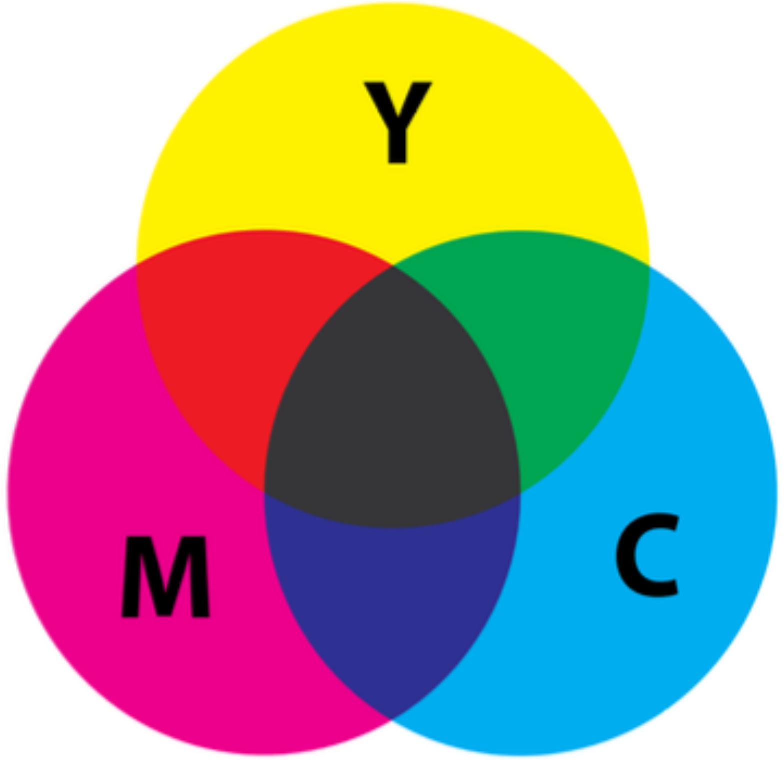
Color vision

Color composition

- Three primary colors can be combined to obtain all others
- Additive composition: add colors by combining lights of a certain wavelength
- Subtractive composition: subtract colors from a reflective surface using pigments that absorb light of a certain wavelength

- RGB: additive
- CMY / CMYK: subtractive



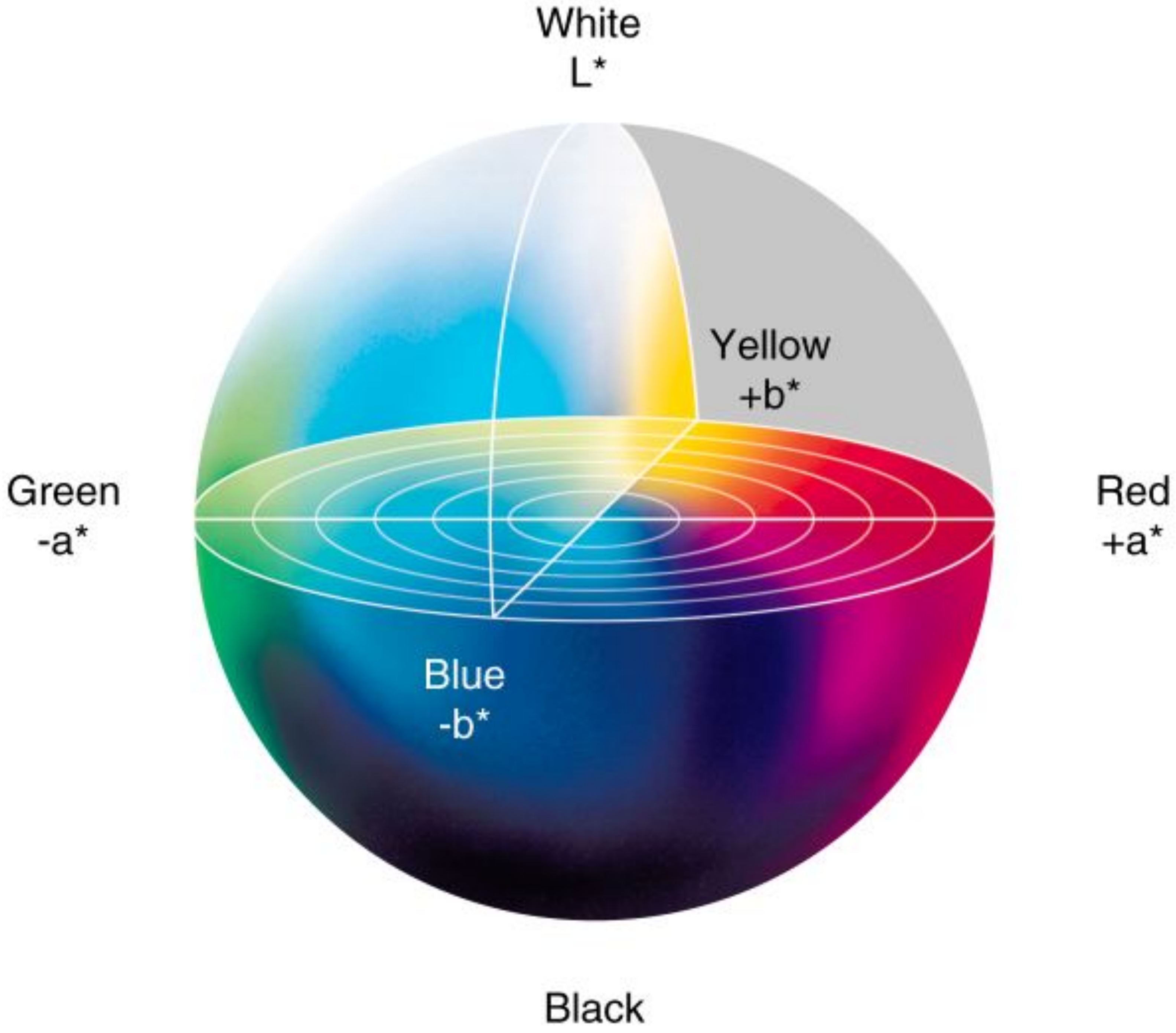


Color vision

CIELab L*a*b*

- Non-linear transformation from XYZ and reference white
- L* encodes luminance in range [0,100]
- a* encodes hue in green-red range
- b* encodes hue in blue-yellow range

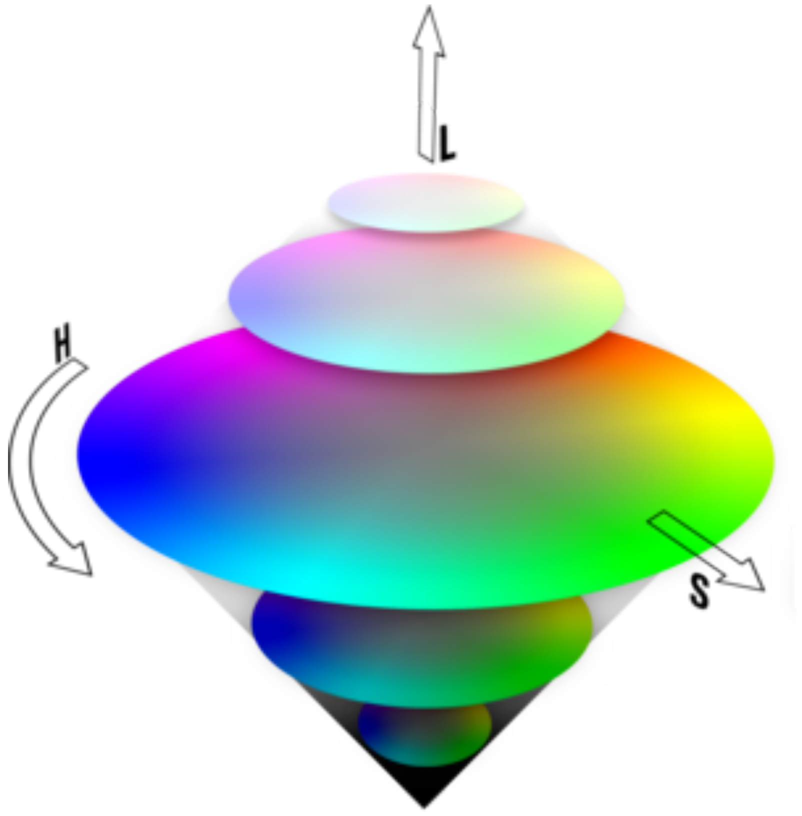
- Perceptually linear: safe to interpolate

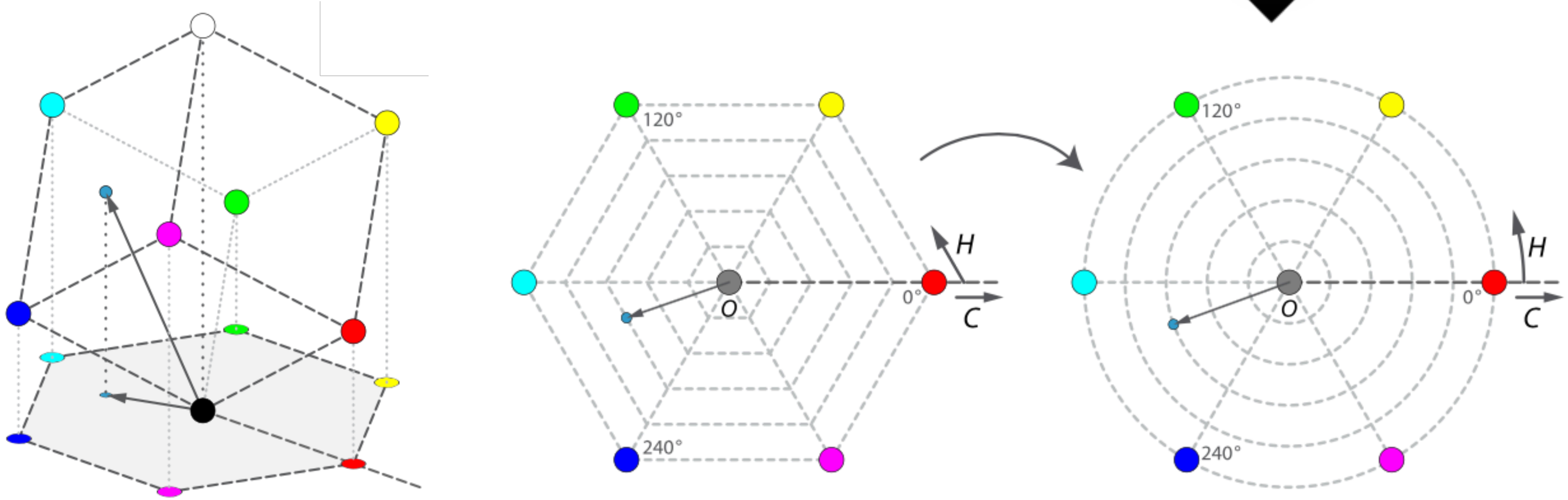


Color vision

HSL

- Simple transformation of RGB
- Separates lightness from hue and saturation
- Only pseudo-perceptual: lightness is not true luminance
- Can be interpolated, with care





Color vision

Transparency:

- colors can be defined to have a certain degree of transparency / opacity
- transparency is encoded in a separate channel, called the *alpha* channel
- alpha in range [0,1] ↔ [transparent, opaque]

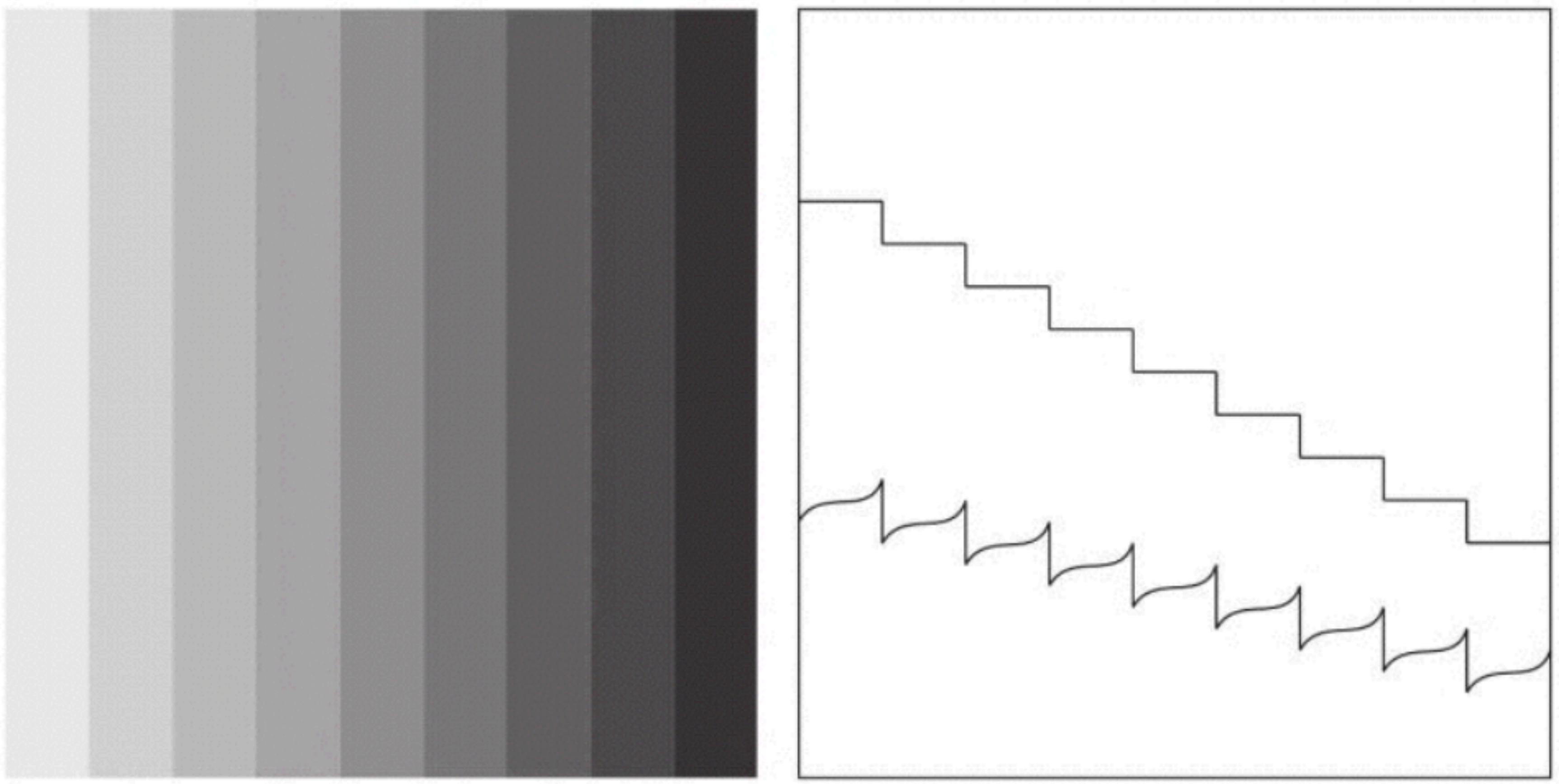
- useful when colors overlap
- result depends on order of overlap:

$$C = \alpha_{front}\, C_{front} + (1 - \alpha_{front})\, C_{back}$$

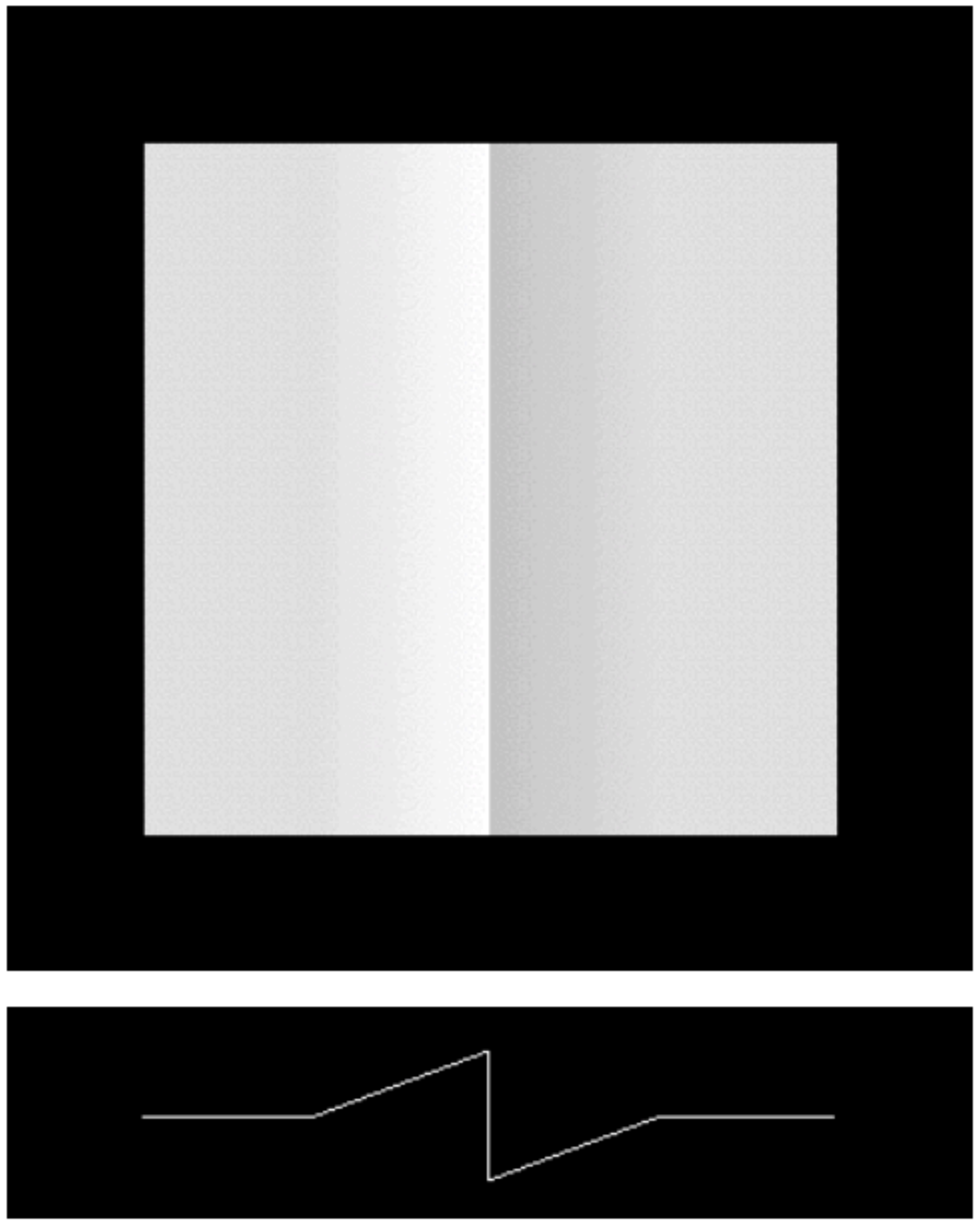
Red: $\alpha = 0.7$
Blue: $\alpha = 0.3$

Perceptual illusions - contrast enhancement

- Luminance: measured amount of light
- Brightness: perceived amount of light
  - – dependent on background
  - – visual system distorts luminance to enhance contrast

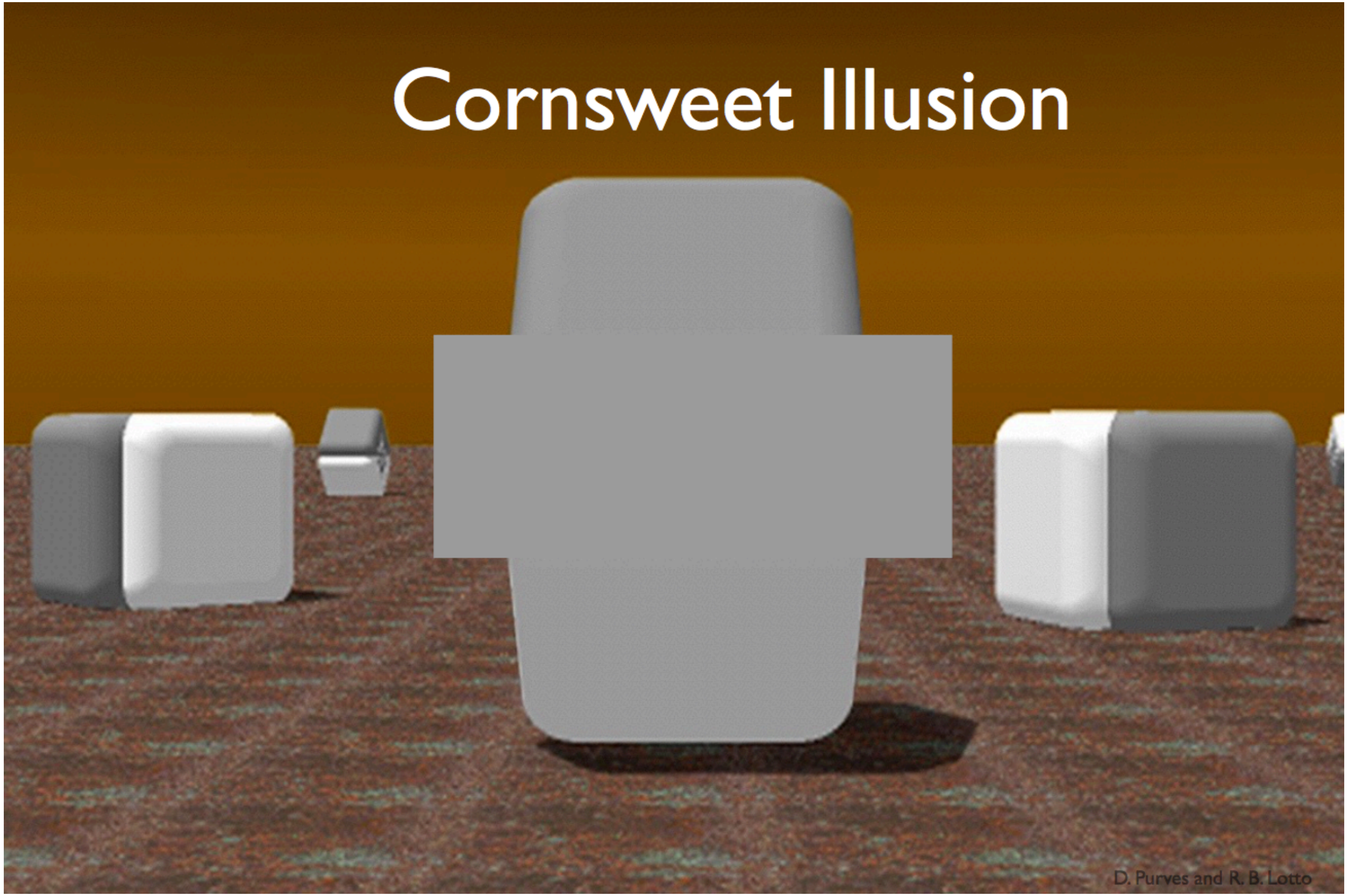


Chevreul illusion



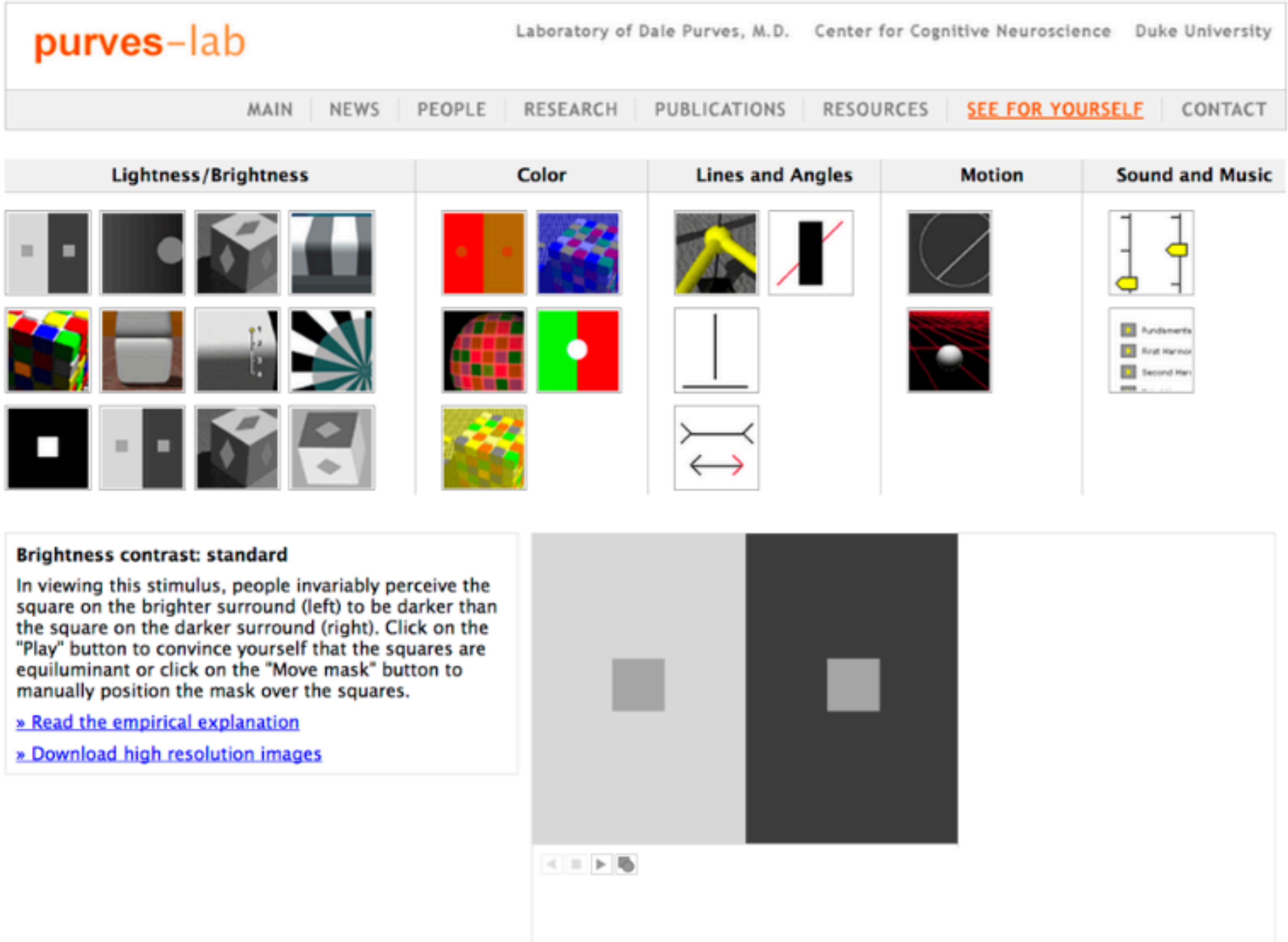
Cornsweet effect

Perceptual illusions - contrast enhancement



Perceptual illusions

- Several demos on-line:
  - http://purveslab.net/see-for-yourself/

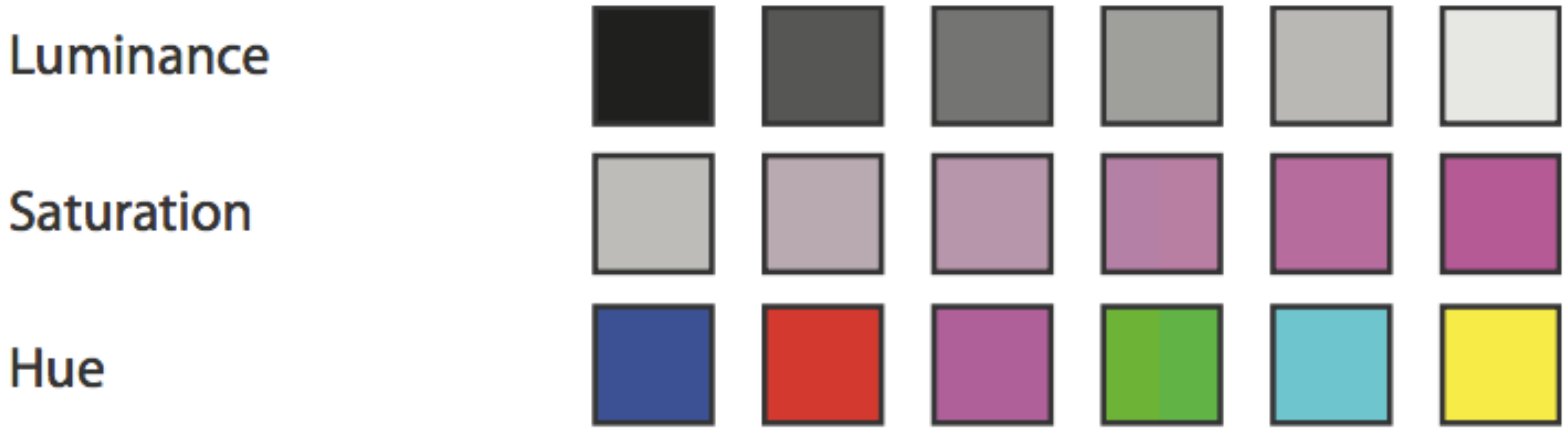


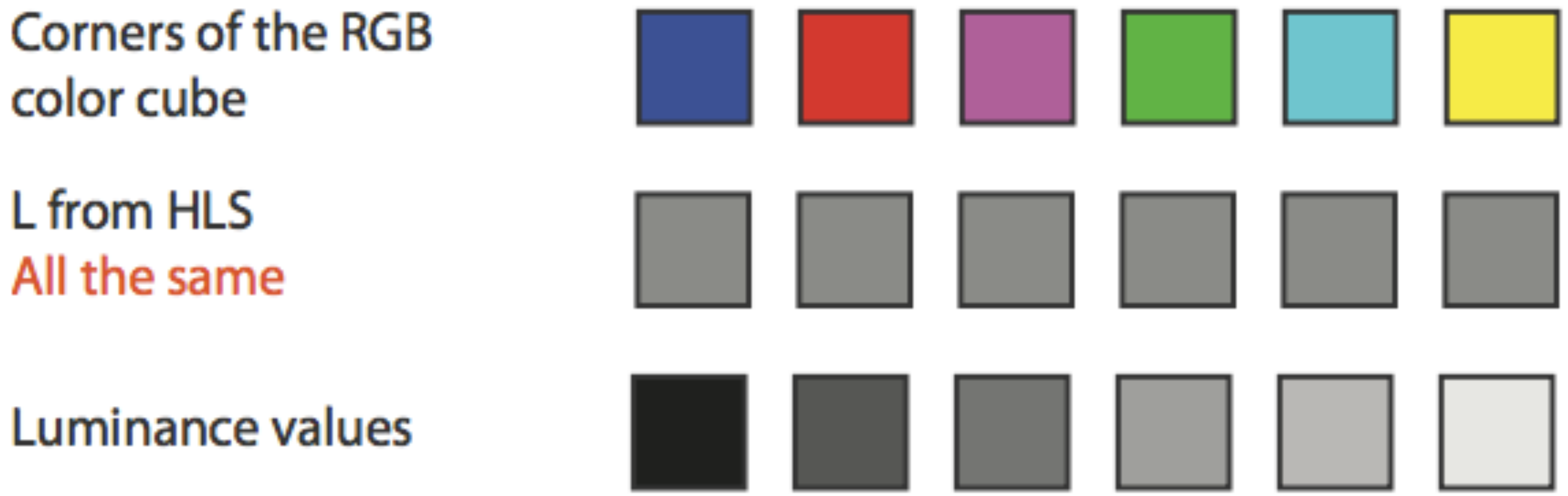
Color in Vis

Luminance, saturation hue

- 3 channels for encoding attributes
  - identity for categorical
    - hue
  - magnitude for ordered
    - luminance
    - saturation

- Be careful with RGB and HSL!





Color in Vis

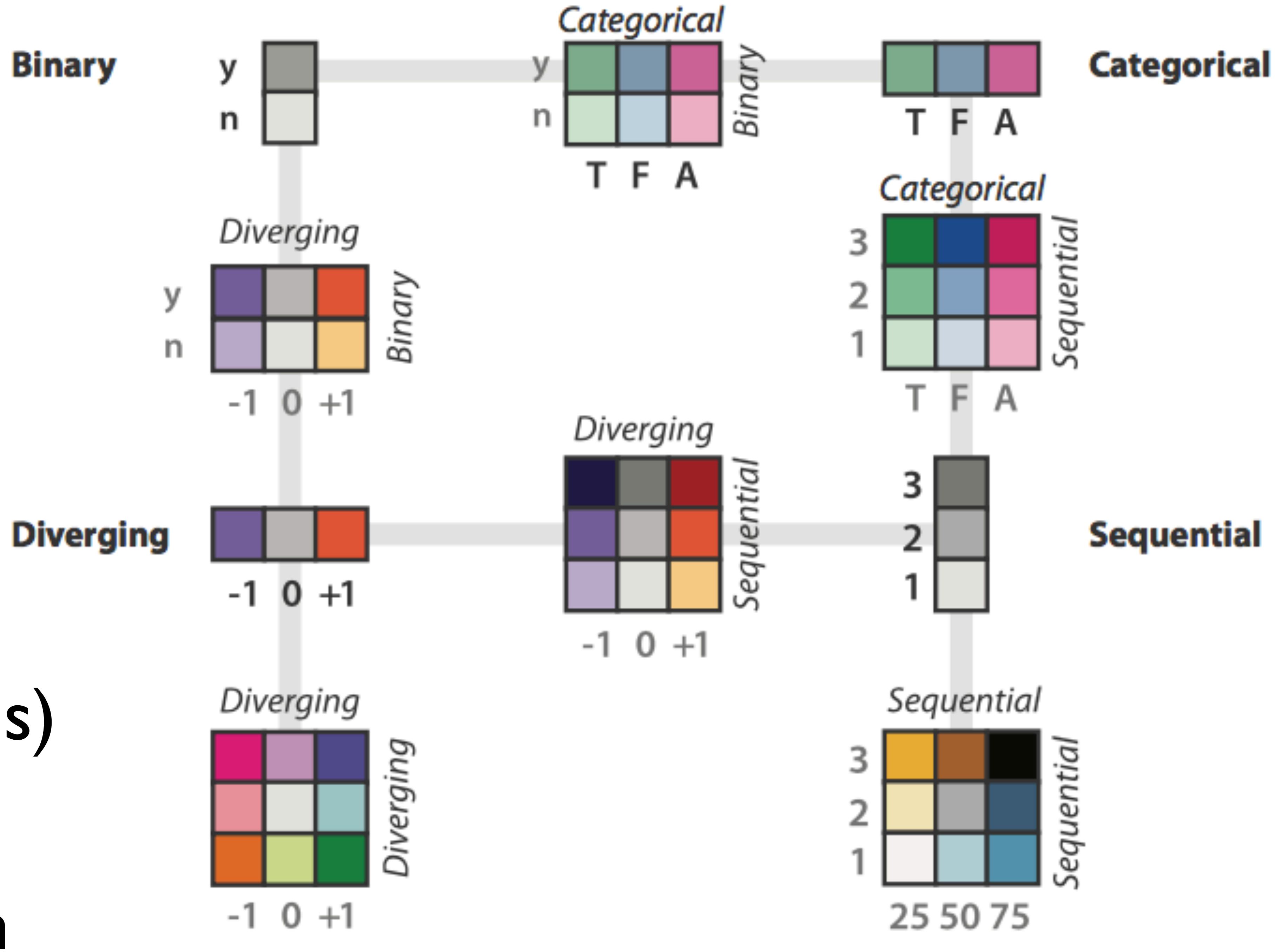
Colormaps

→ Categorical

→ Ordered

→ *Sequential* → *Diverging*

→ Bivariate



Categorical limits (non contiguous)

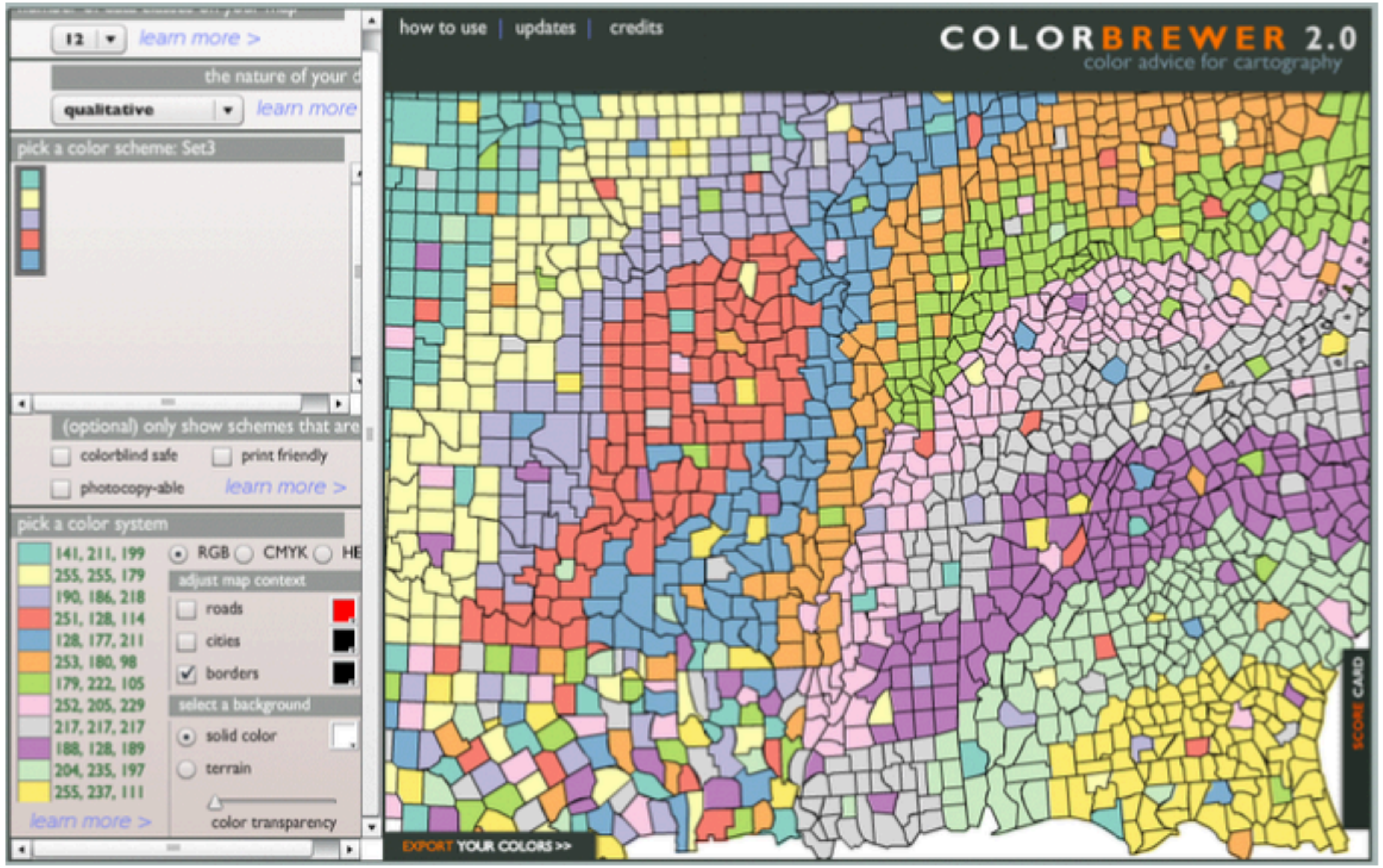
- 6-12 bins hue
- 3-4 bins luminance, saturation
- size affects salience: use saturation inversely proportional to size

Color in Vis

Categorical channel: example of labeling

- *Colorbrewer*: very useful tool to obtain color schemes

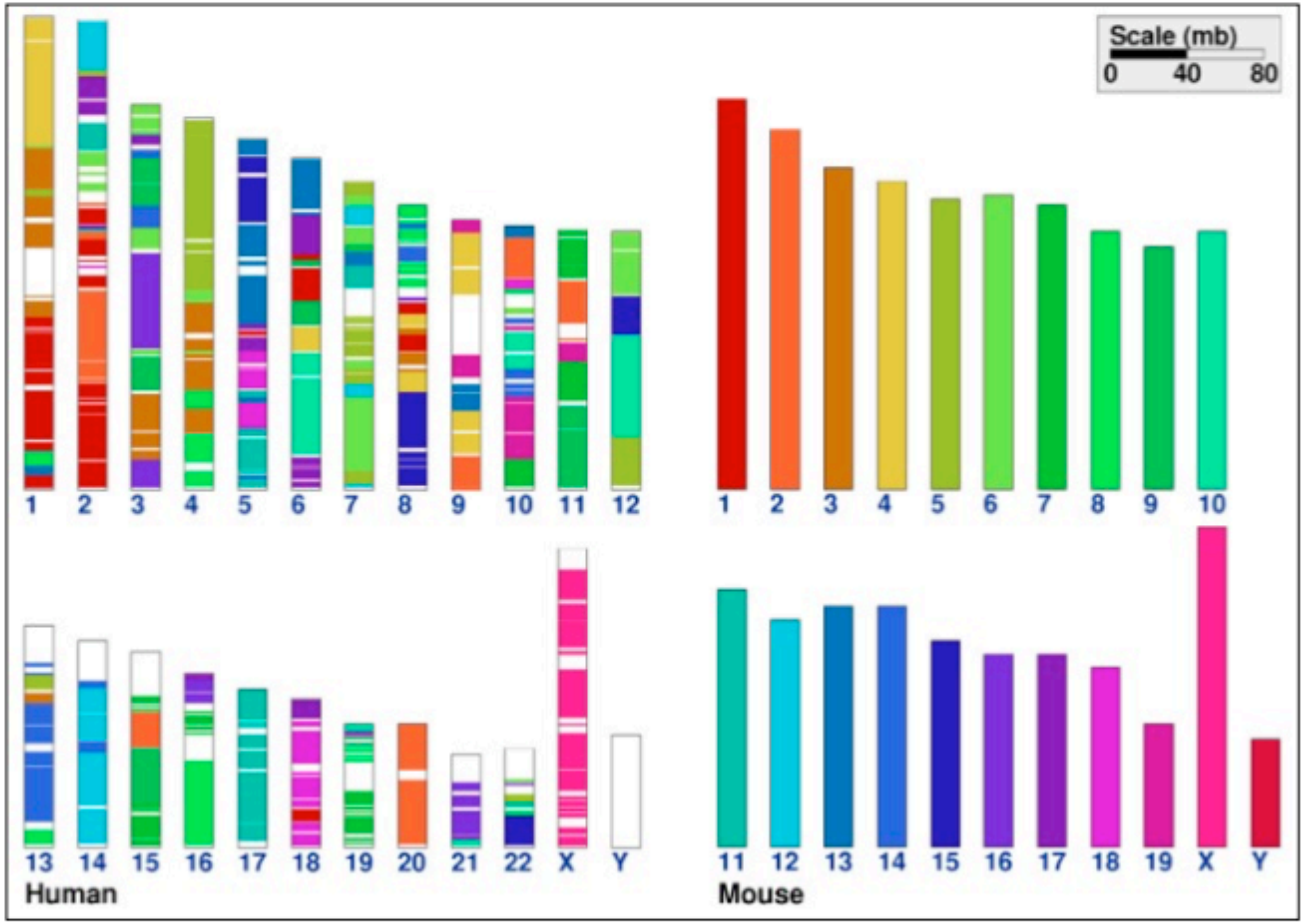
<u>*http://colorbrewer2.org*</u>



Carefully designed color scheme for 12 colors [colorbrewer]

Color in Vis

Discriminability constraints for categorical channel: example



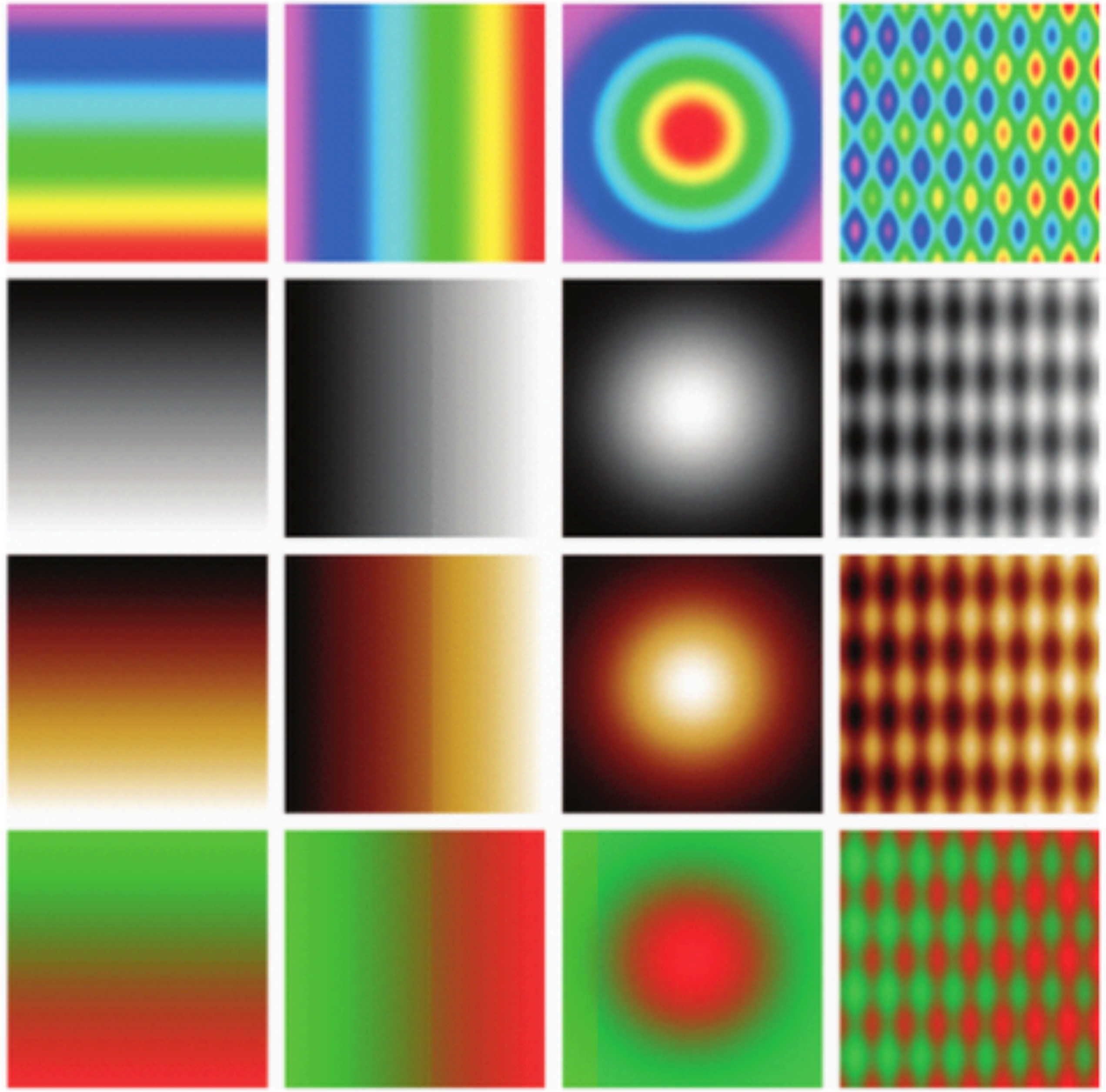
[Cinteny: flexible analysis and visualisation of synteny in genome rearrangements in multiple organisms. Sinha & Meller. BMC Bioinformatics, 8:82, 2007]

Color in Vis

Ordered / quantitative channel: *transfer function*

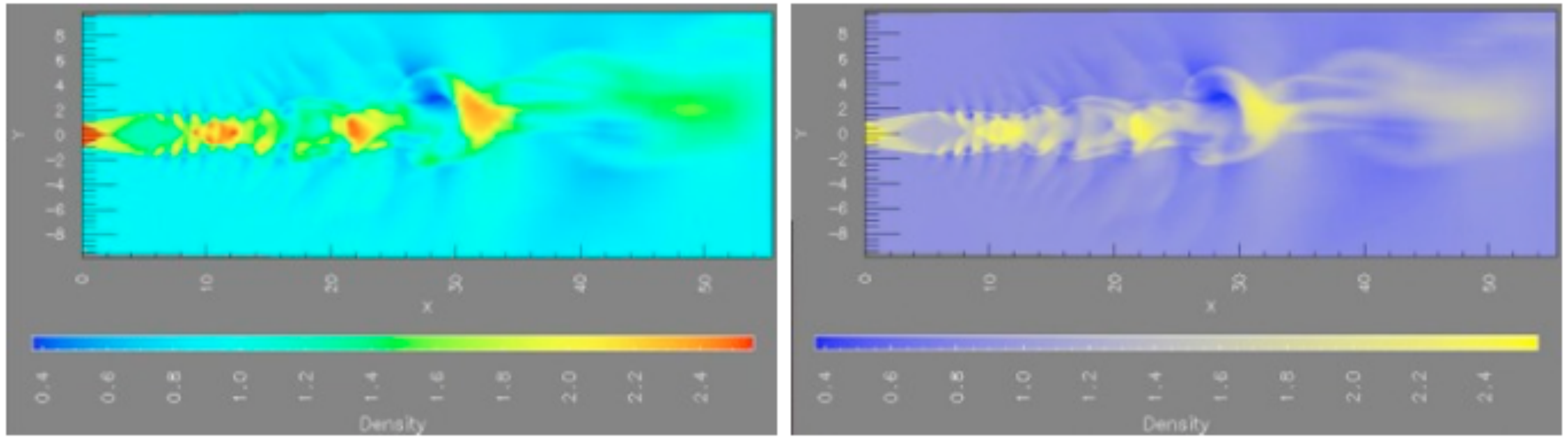
- Map a range of values [min, max] to a scale of colors
- May be either discrete (binned) or continuous

- Luminance works well
- Saturation works but not as good
- Don't use hue! ordering hues is not intuitive

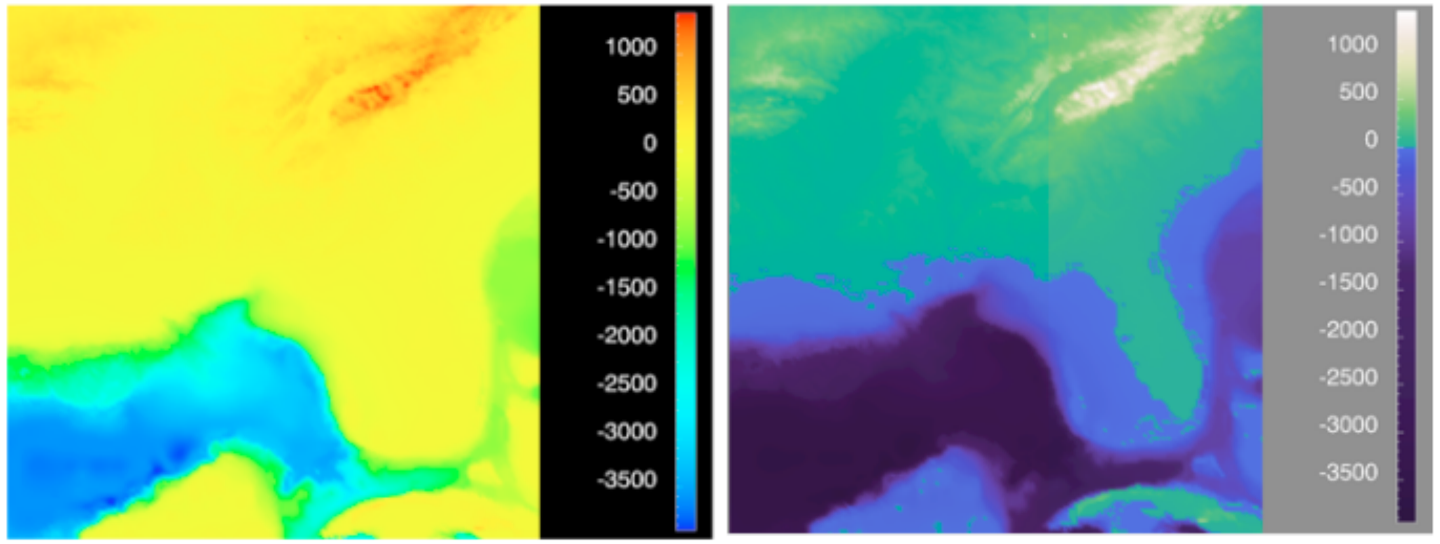


Color in Vis

Ordered color: Rainbow is a poor default



- problems:
  - perceptually unordered
  - perceptually non-linear
- benefit:
  - fine-grained structure visible and nameable
- Alternatives
  - few hues for large-scale structure
  - combine hues with monotonically increasing luminance
  - segmented rainbows are good for categorical, may be used for binned



Change blindness

- Details on an image cannot be remembered across separate scenes
  - except in areas with focused attention!
- Interruption (e.g. a blink, eye saccade or blank screen) amplifies this effect
- Not failure of vision system
- Failure due to inappropriate attentional guidance

Change blindness

Change blindness

Change blindness

Change blindness

Attention blindness

Pre-attentive processing and Change blindness in Vis

“studying a display may offer no assistance in searching for specific data values. In this scenario, methods that draw attention to areas of potential interest within a display (i.e., preattentive methods) would be critical in allowing viewers to rapidly and accurately explore their data.”

“The presence of change blindness in our visual system has important implications for visualization. The images we produce are normally novel for our viewers, so <u>prior expectations cannot be used</u> to guide their analyses. Instead, we strive to <u>direct the eye</u>, and therefore the mind, <u>to areas of interest or importance</u> within a visualization. This ability forms the first step towards enabling a viewer to abstract details that will persist over subsequent images.”

C.G. Healey

Takeaways

- To find meaning in what we see we must selectively pay attention to what is important
- Low-level vision is driven by object features rather than a conscious effort where to look (e.g., pre-attentive processing)
- Attention is driven by preexisting knowledge, expectations, and goals stored in long-term memory
- Knowing about perception is important for Vis and UI design
  - how to choose your colors
  - how to draw attention
  - how to minimize risk of overlooking

To Do

- Homework 0 - Mandatory! Evaluated! Strict deadline!:
  - Fill in the enrollment survey (two mins)
  - Deadline one week from now

- To read to study this class
  - Book VAD: Color theory and Colormaps (Ch 10.2, 10.3)
  - Book VAD: Channel Effectiveness; Relative vs. Absolute Judgements (Ch 5.5, 5.6)
  - Paper: M.C. Stone, Representing colors as three numbers

    http://www.stonesc.com/pubs/Stone%20CGA%2007-2005.pdf
- To read for the next class
  - Book VAD Ch. 2: Data abstraction