Vision

Introduction to Cognitive Science

Our Eyes as a Window to the World

The Myth of ‘Perfect’ Perception

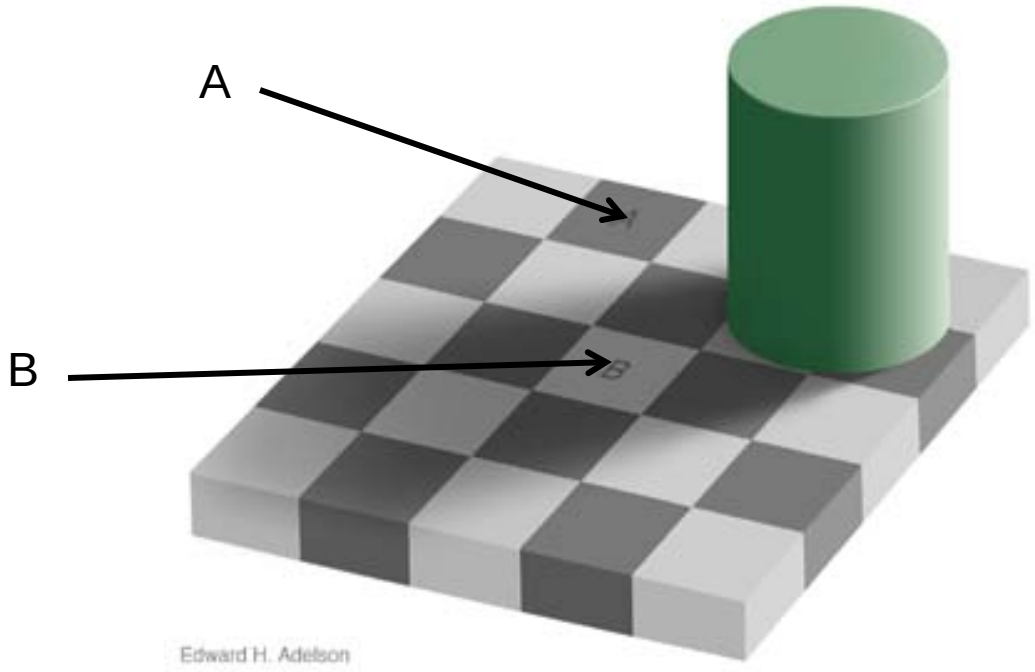
- The myth of perception is that as long as:
  - Our eyes are functioning properly
  - We’re not wearing rose-colored glasses
  - We’re not drugged
  - We’re not subjected to some visual illusion
  - We’re not in the Matrix
  - … (other exceptional/rare situations)
- … we perceive the world *exactly as it is*.
- Wrong!

Our Senses are Limited

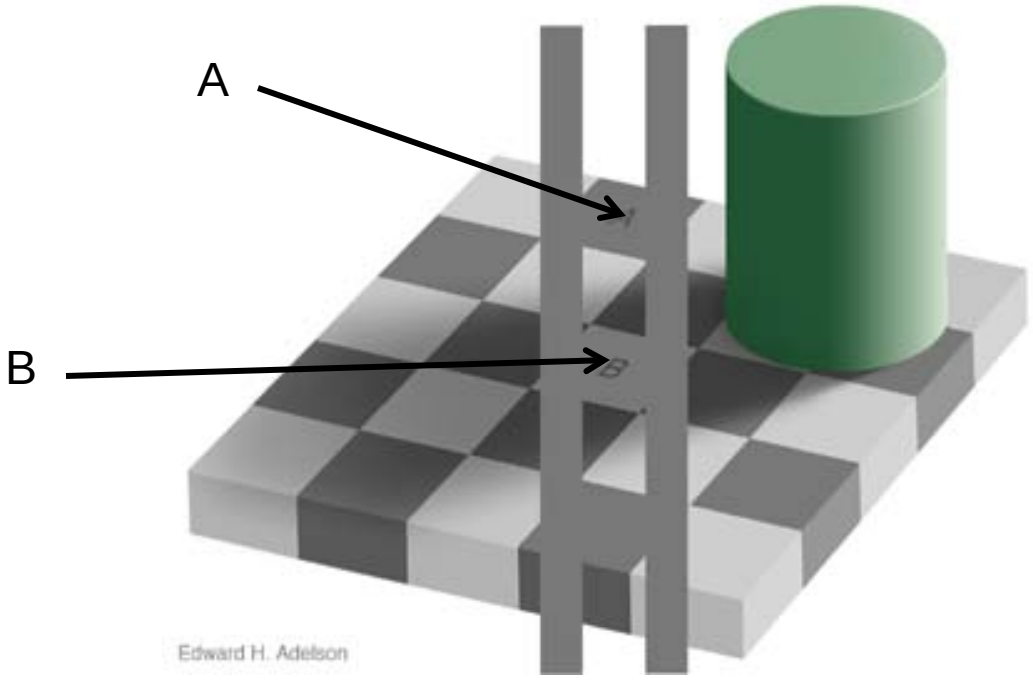
- Our eyes only perceive a very small part of the electromagnetic spectrum; we don't see infrared, ultraviolet, X-Rays, Gamma-Rays, Micro-waves, Radio-Waves, etc.
- Similar for our other senses.

- OK, but what we *do* perceive, is still *exactly as it is*, right?
- Wrong!

The Checker Board Shadow Illusion



The Checker Board Shadow Illusion



Which Perception is ‘Correct’?

- But is it wrong to perceive A to be darker than B?
- If it is a chess board, then A really is darker than B in some real physical sense, even if the raw stimulus happens to be the same.
- And that’s just it: perception is an *interpretation* of the raw sensory stimuli.
- Perception = f(sensory stimuli)

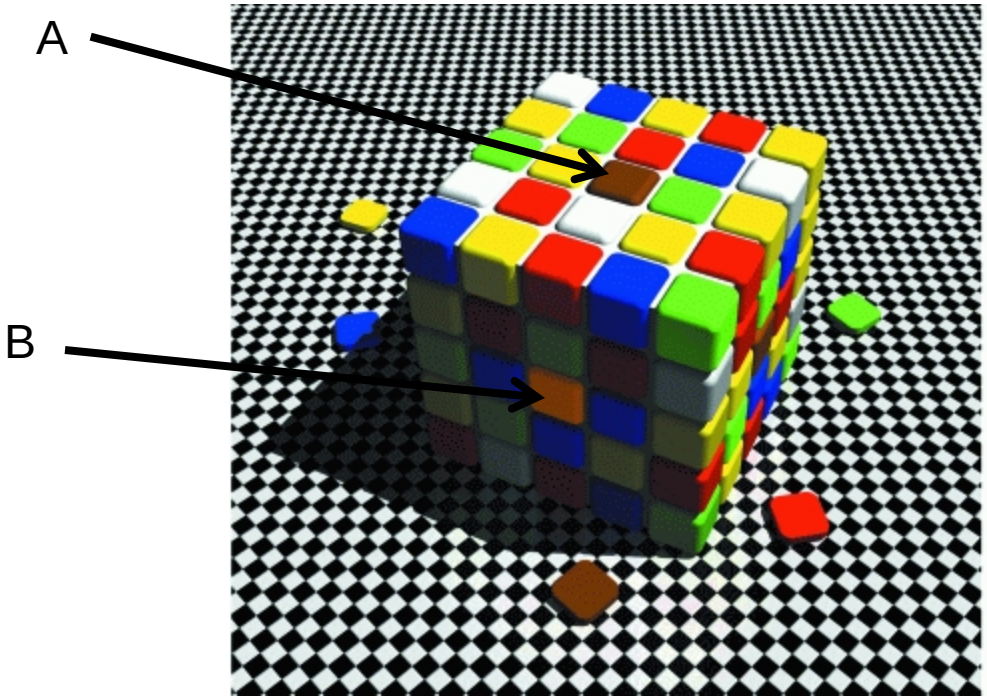
Perception as an Inversion Problem

- We can look at perception as an *inversion problem*: our mind/brain has to figure out what is going on in the world ‘out there’, given the raw incoming sensory stimuli.
- But this inversion problem is inherently *underspecified*: that is, at all times, an infinite number of scenarios can produce the incoming sensory stimuli we get.
- Hence, we have to make a guess, i.e. perception is inherently inferential (and non-deductively so!)

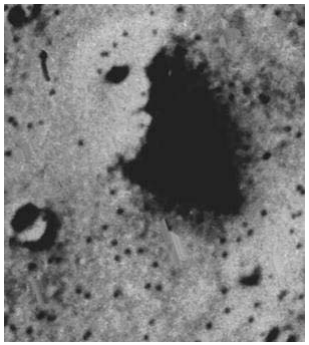
The Blind Spot

X X

Color Contrast

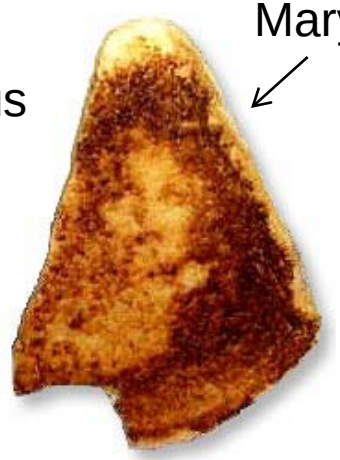


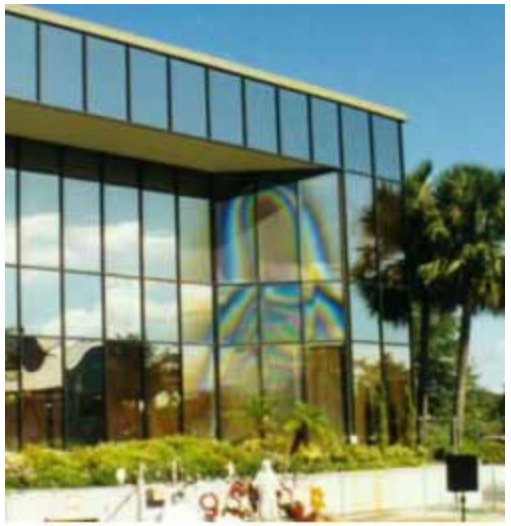
Devils, Angels, Aliens, Jesus, Mary, and … Illinois!



Mary

Jesus





Clearwater, FL USA

[Jesus Pan]

[Local Jesus]

Expectations, Fears, and Wishes

- What we perceive is effected by:
  - Our expectations
    - Corridor Experiment
  - Our fears
    - Monsters under the bed
    - Person in the shadow
  - Our wishes
    - N-Rays, St. Nick
  - And probably many other states of mind

Background Knowledge

- Background knowledge/beliefs drive perception (‘believing is seeing’)
  - Color constancy: Cut-outs of trees and donkeys (made of same color material, but perceived as different color, in accordance to our beliefs)
  - Size constancy: even though approaching objects would seem to get bigger and bigger going by the mere raw 2D projection on retina, we never perceive them as growing in size, because we know (or at least based on background experience have strong reasons to believe) that they remain the same size.

Perception and Concepts

- We see faces in lots of things because faces are important in our lives.
- Indeed, our interpretation of incoming stimuli is in terms of concepts: faces, people, trees, tables, chairs, etc.
- Concepts are useful ways to think about the world ... but are they ‘real’?
- (if a tree falls in a forest …)

Pattern ‘Bias’

- It is important for our survival to recognize patterns in nature. Patterns allow us to make predictions and control our lives.
- Moreover, it is better to be safe than sorry: better to recognize a non-face as a face than not to recognize a face!
- Hence, our brains have ‘pattern bias’: when there may or may not be a pattern, it is likely to say that there is a pattern.

Summary: Perception is Constructive!

- *At all times*, how we perceive things is a construction of our mind (brain).
- Perception = f(raw sensory input, attention, beliefs, expectations, …)
- A very complex function!
- It is estimated that about a third of our brain is dedicated to visual processing alone!

Why Perception Doesn't Feel Constructive

- Two big reasons:
  - We are, under normal circumstances, getting a lot 'right'
    - I rarely get contradicted in my constructions of reality and how I subsequently act on that
  - We consistently perceive the world in the same way (we're stuck in Plato's cave)

Object Perception

Object Perception and Object Recognition

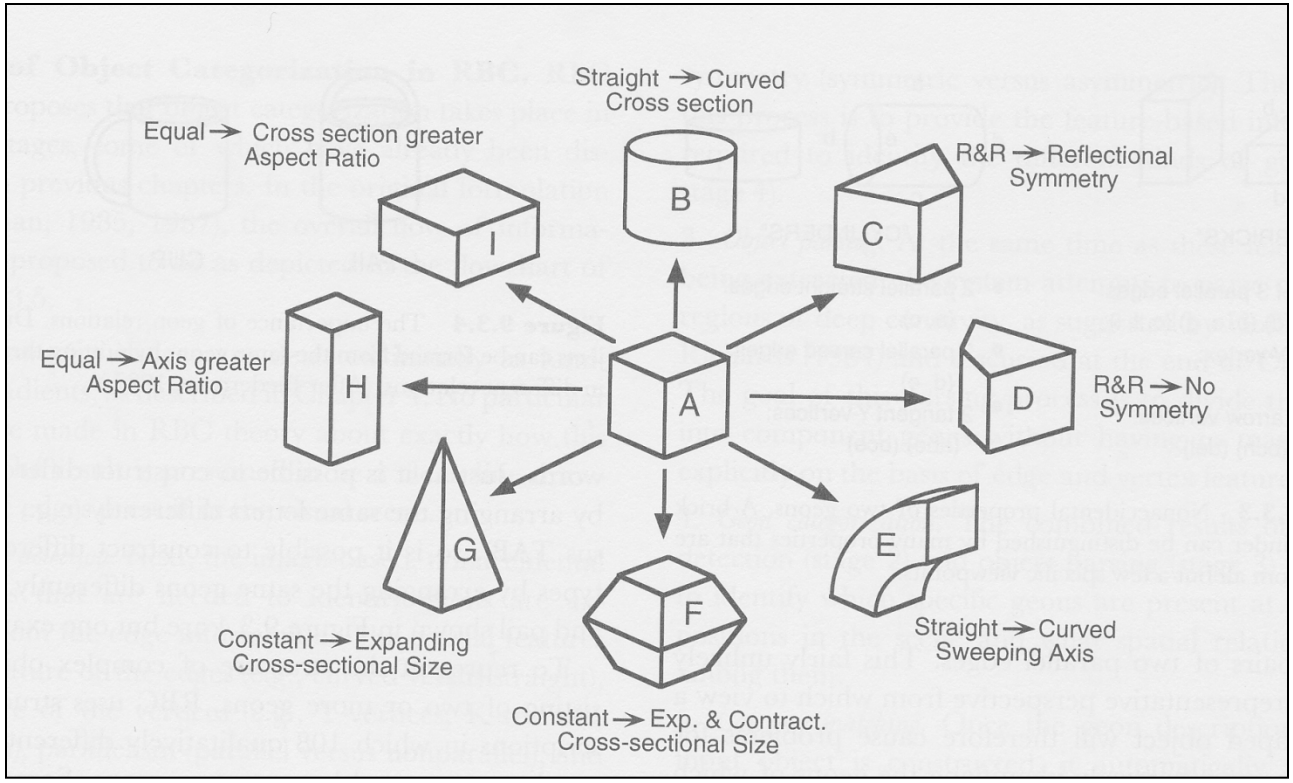
- Normally, object perception is talked about as ‘object recognition’.
- However, as we just saw, perception is constructive: it parses up the world in a way that is, in some way, useful.
- Indeed, rather than ‘tables’ and ‘chairs’ being ‘out there’, and that we subsequently ‘recognize’ as such, it may be better to simply talk about ‘stuff’ being out there, that we parse (perceive) as ‘tables’ and ‘chairs’.
- That said, from now on I’ll talk about recognition.

The Problem of Object Recognition

- We accurately, and effortlessly, recognize objects in our environment even though the ‘raw’ image on the retina changes *dramatically*, due to variations in:
  - Position
  - Size
  - Color
  - Surface texture
  - Viewpoint
  - Exemplars of a given class
- How do we do that?!?

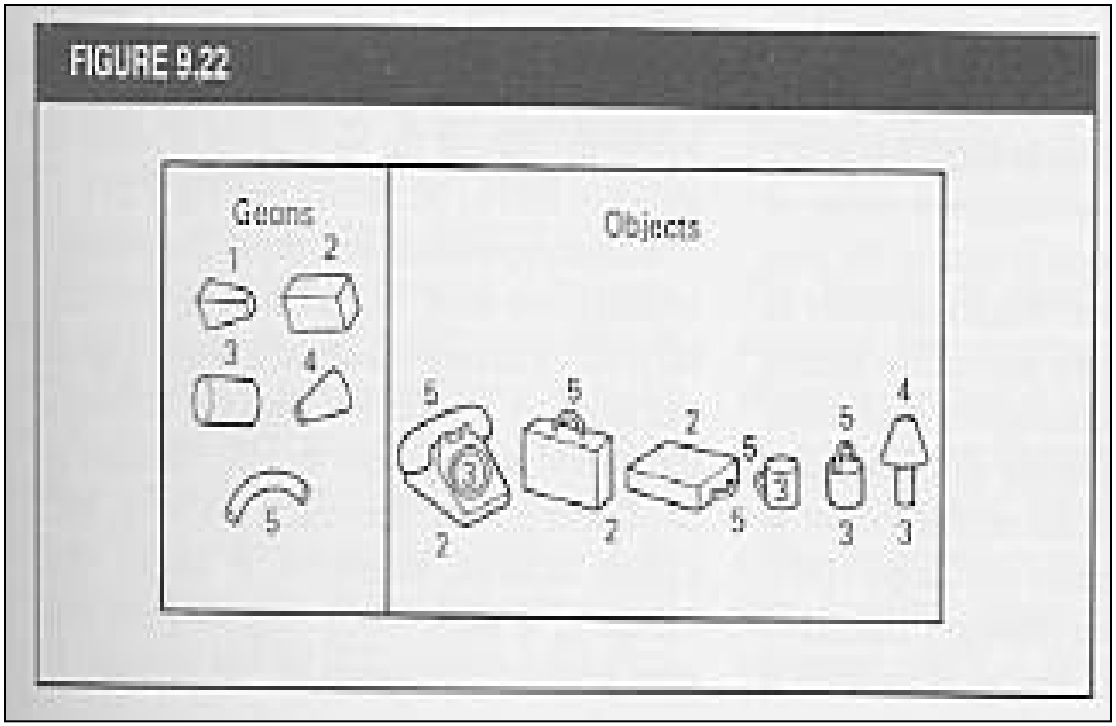
Recognition-by-Components

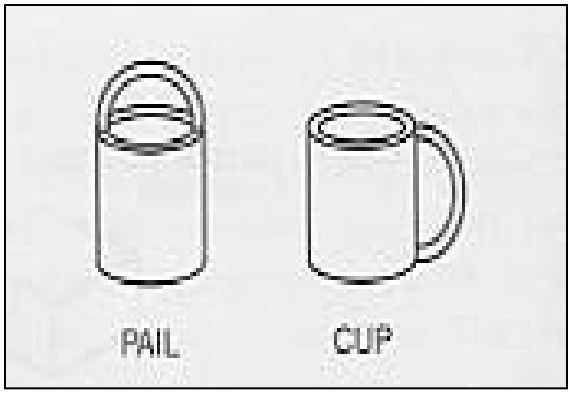
- Biederman (inspired by Marr) defined 36 geons that qualitatively differ along 5 dimensions (cross-sectional curvature, symmetry, axis curvature, size variation, aspect ratio)



Recognition-by-Components

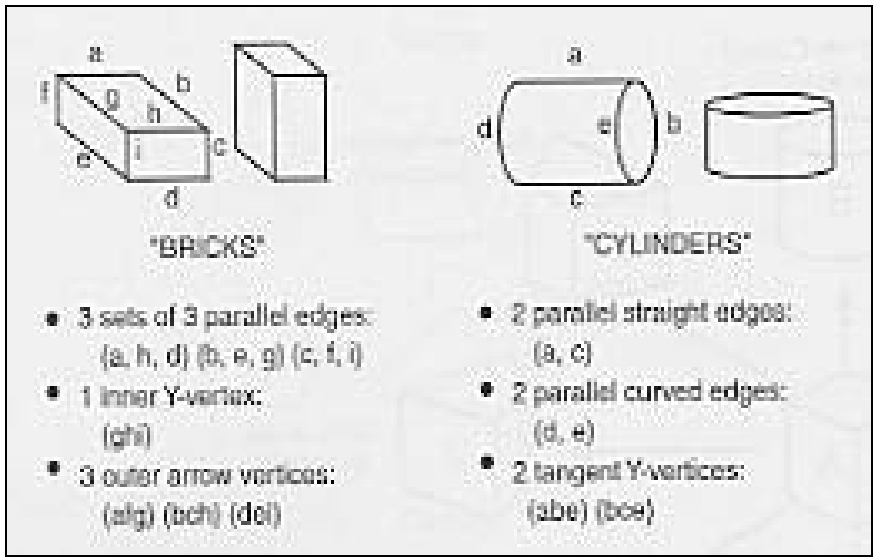
- Many (most? All?) objects can be specified as spatial arrangements of primitive volumetric components, called geons (geometric ions)
- The 36 geons can generate over 150 million 3-geon objects (like an alphabet)



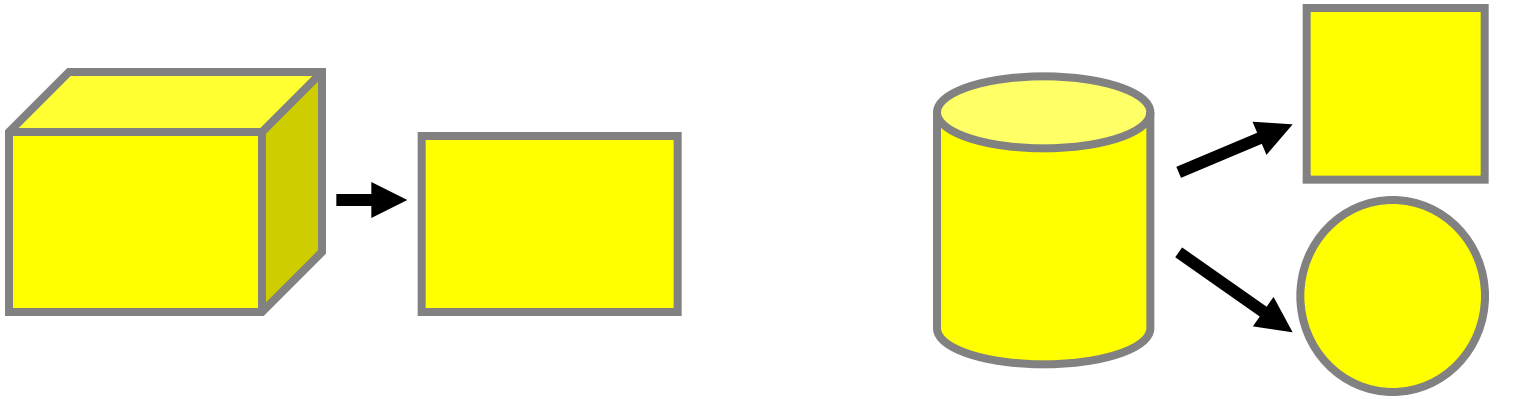


View invariance of geons

- each different geon has its own key properties at the level of the 2D full primal sketch – geons can be identified directly from image
- these “non-accidental” properties are invariant over different views…



- …with a few exceptions (called “accidental views” – highly unstable)



Space Perception

Space Perception

- Object recognition is about perceiving ‘what’
- Space perception is about perceiving ‘where’

Cues

- Another myth about perception: Many people think that space perception is (solely) due to the fact that we have two eyes.
- However, if you close one eye, you can still perceive 3D quite well, e.g. you have little problems navigating your environment. How is that possible?
- It turns out that there are *many* cues that we use to perceive space:
  - Convergence
  - Accommodation
  - Occlusion
  - Etc.

Types of Cues

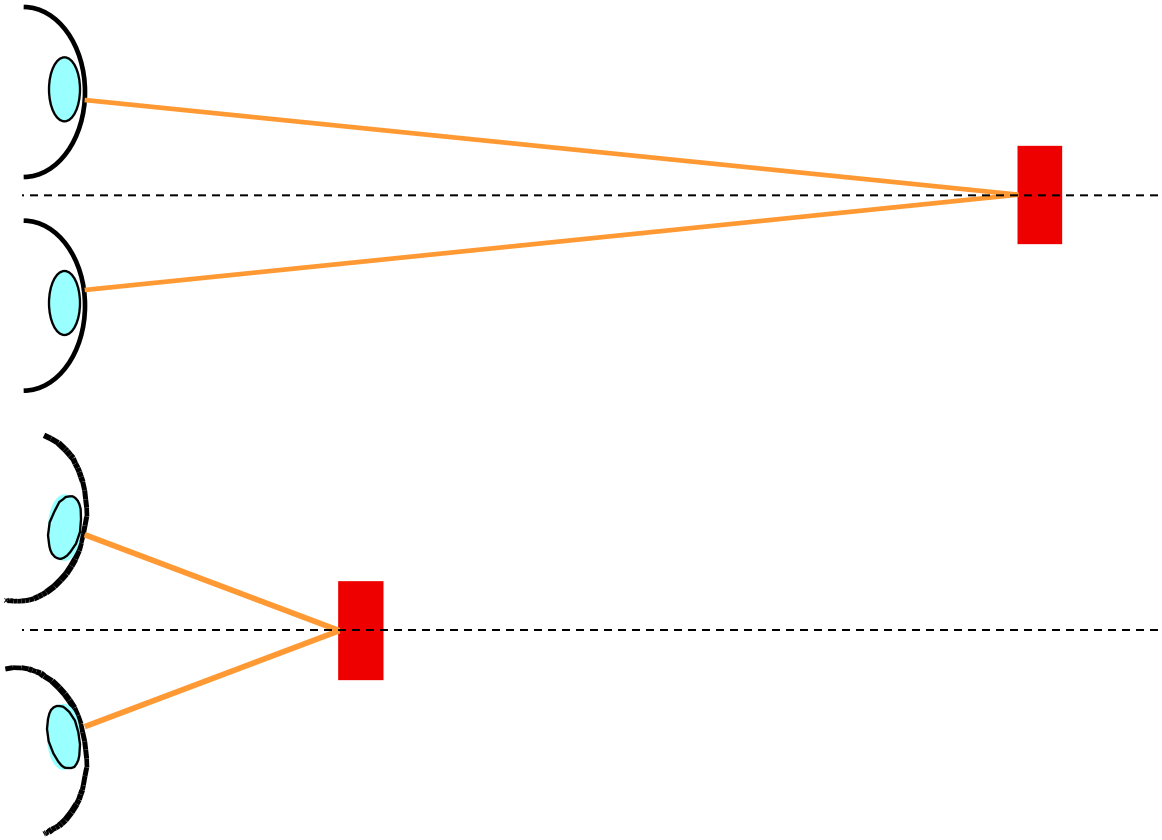
- What does the cue require?
  - Monocular or Binocular?
    - Does the cue require one eye or two eyes?
  - Optical or Ocular? (External or Internal?)
    - Is the cue related to the visual image (optical) or is it related to the physiology of our eyes (ocular)?
  - Static or Dynamic?
    - Does the cue work for a single ‘snapshot’, or does it need the image to change?

Information Present in Cue

- What does the cue tell us?
  - Absolute or Relative?
    - Does the cue tell us exactly where (e.g. how far away) some object is located, or does it tell us where that object is in relation to other objects?
  - Quantitative or qualitative?
    - If a cue is relative, does it say that something is, for example, exactly twice as far way as something else, or does it merely tell us that something is further away than something else? Note: all absolute cues are quantitative
  - What is the ‘effective range’ of the cue?
    - Some cues only work well for close by objects, whereas others work better with a far distance.

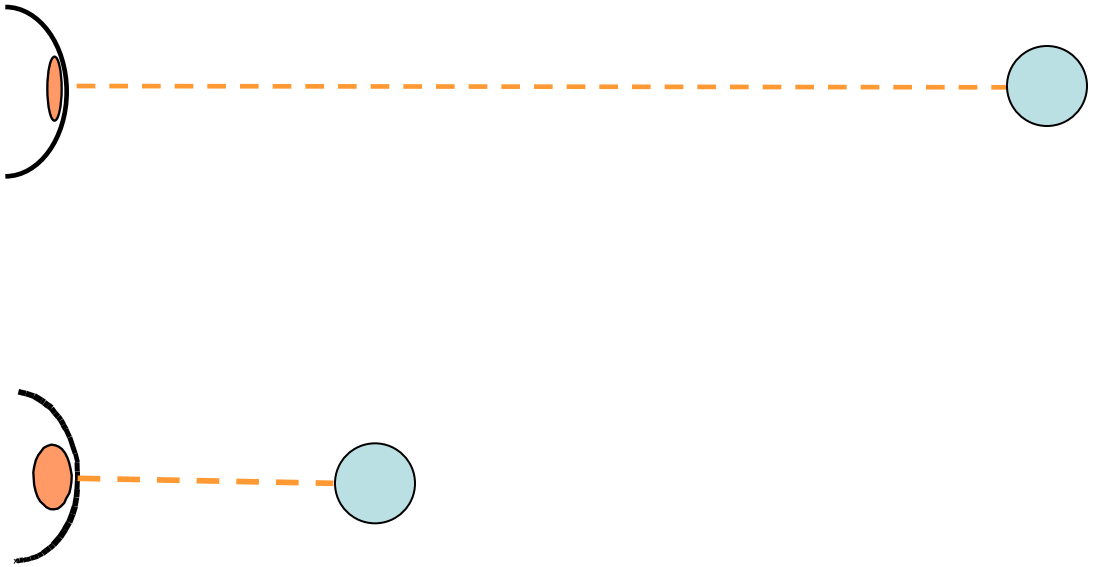
Convergence

- Eyes rotate to fixate objects of interest
- Angle of inward looking relative to head midline varies systematically with distance
- So, how much the muscles have to rotate the eyes provides a clue as to how far the object is



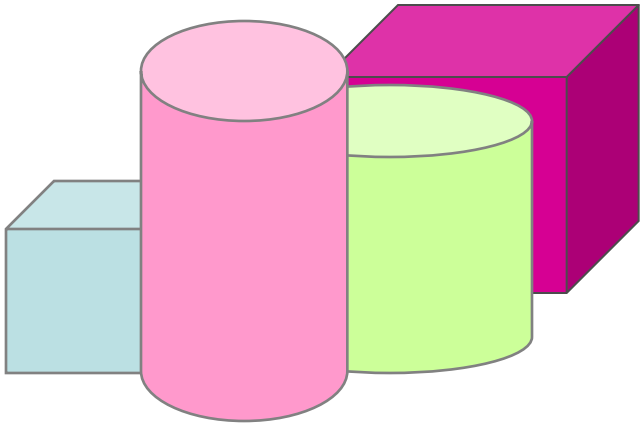
Accommodation

- Lens bulges as we focus on nearer objects
- Tension in ciliary muscles varies systematically with distance



Occlusion

- Closer objects hide or partially hide distant objects $\Rightarrow$ hidden objects are seen as more distant

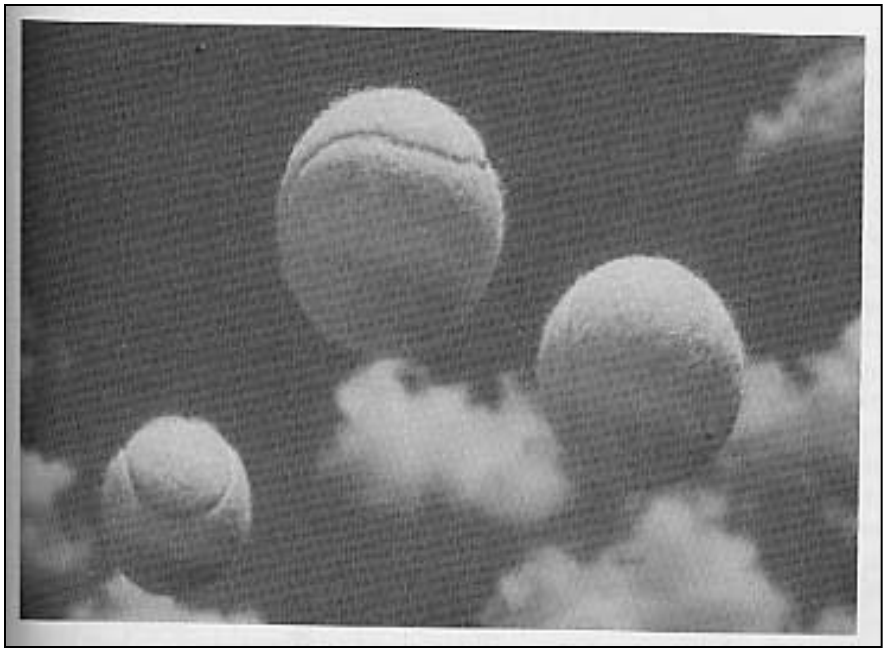


Familiar Size

- Familiar size: If an object's size is known, then the object's distance can be determined

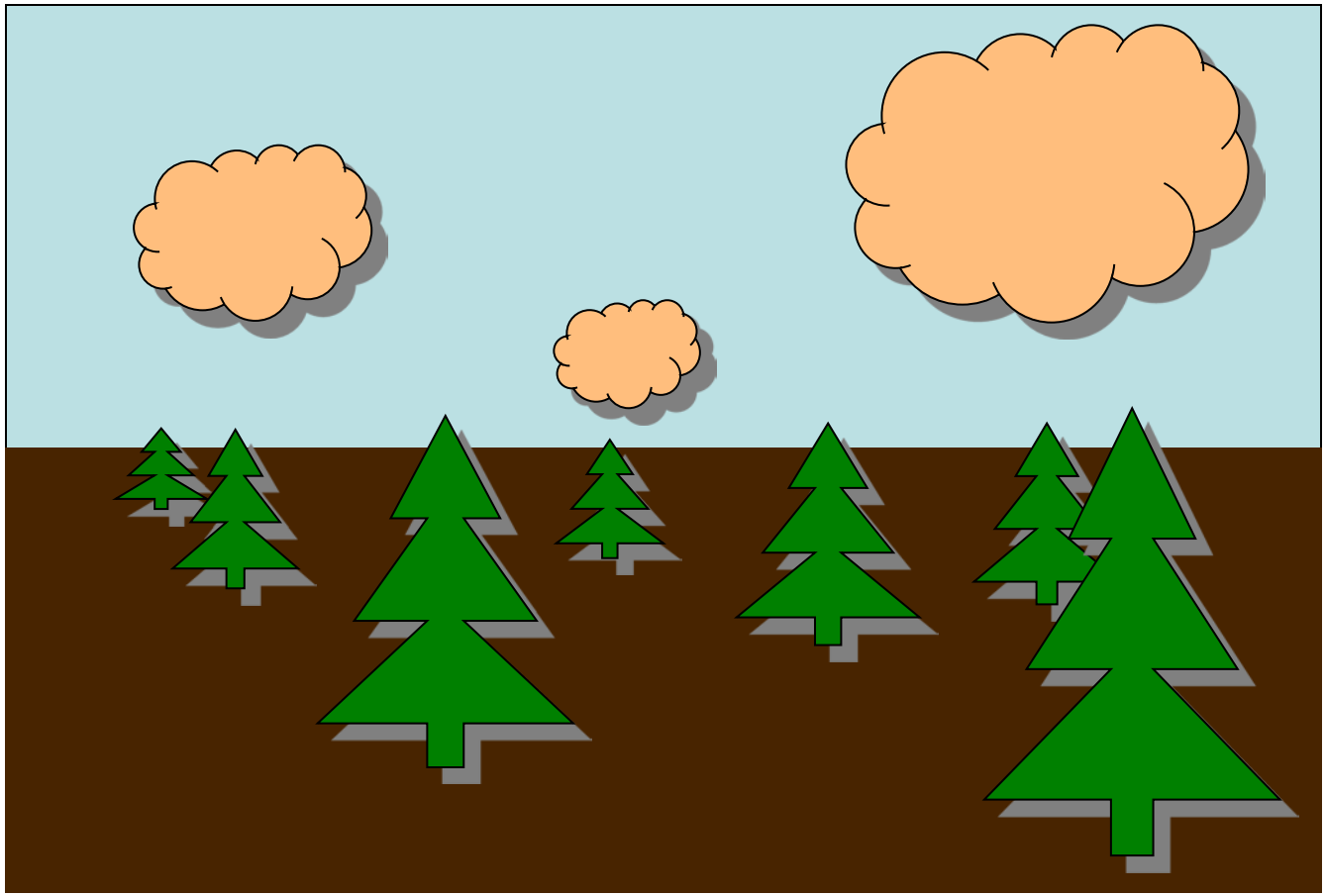
Relative Size

- Relative size: If two objects of known similar sizes occupy different visual angles, then the object that occupies the larger visual angle must be closer



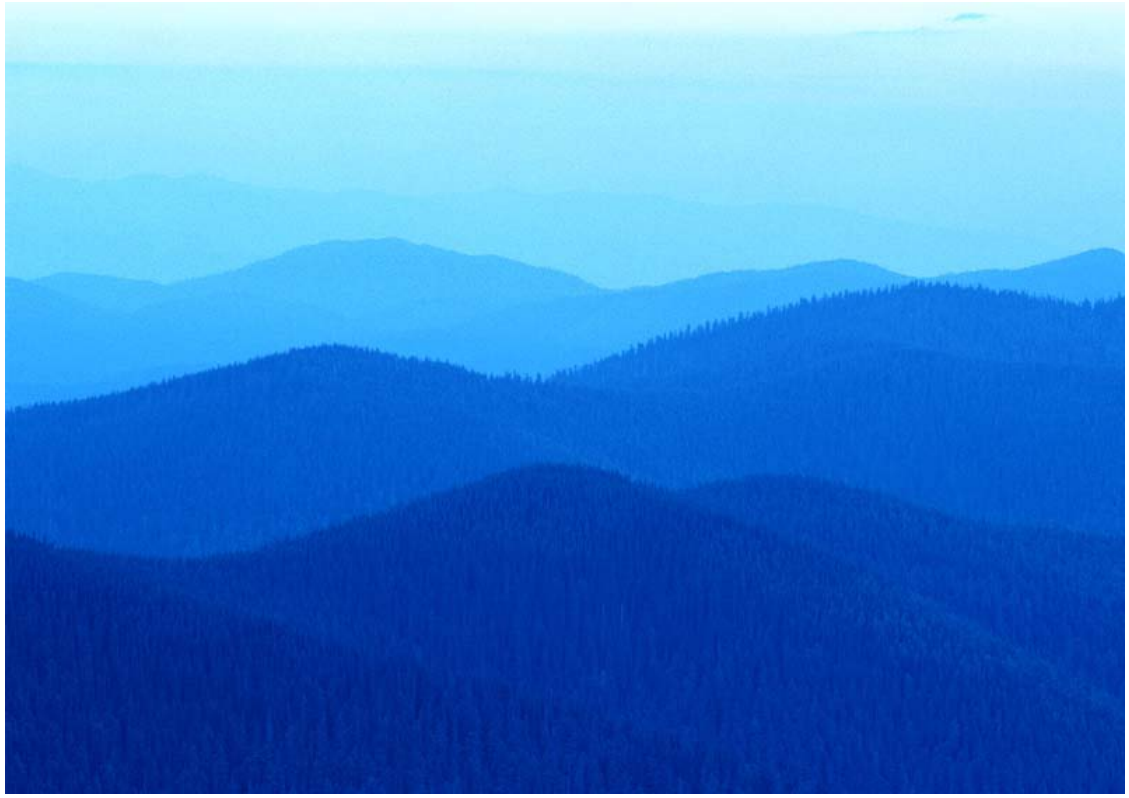
Height in the visual field

- Distant objects are higher in the visual field $\Rightarrow$ objects that are higher in the visual field (or more generally, closer to the horizon) are perceived as more distant



Atmospheric Perspective

- Particles of dust, water, and pollution cause us to see distant objects as less sharp $\Rightarrow$ Fuzzy objects are perceived as being more distant (note how in this example light and dark gives us a big clue)

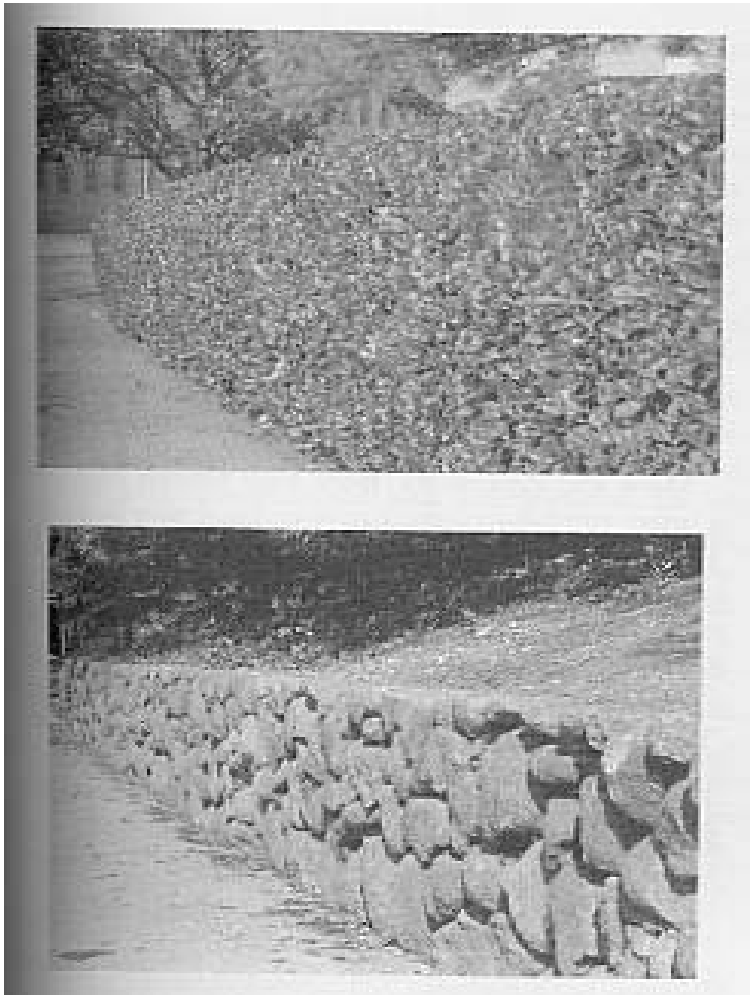


Linear perspective

- **Lines that are parallel in the scene converge with distance on the image plane**

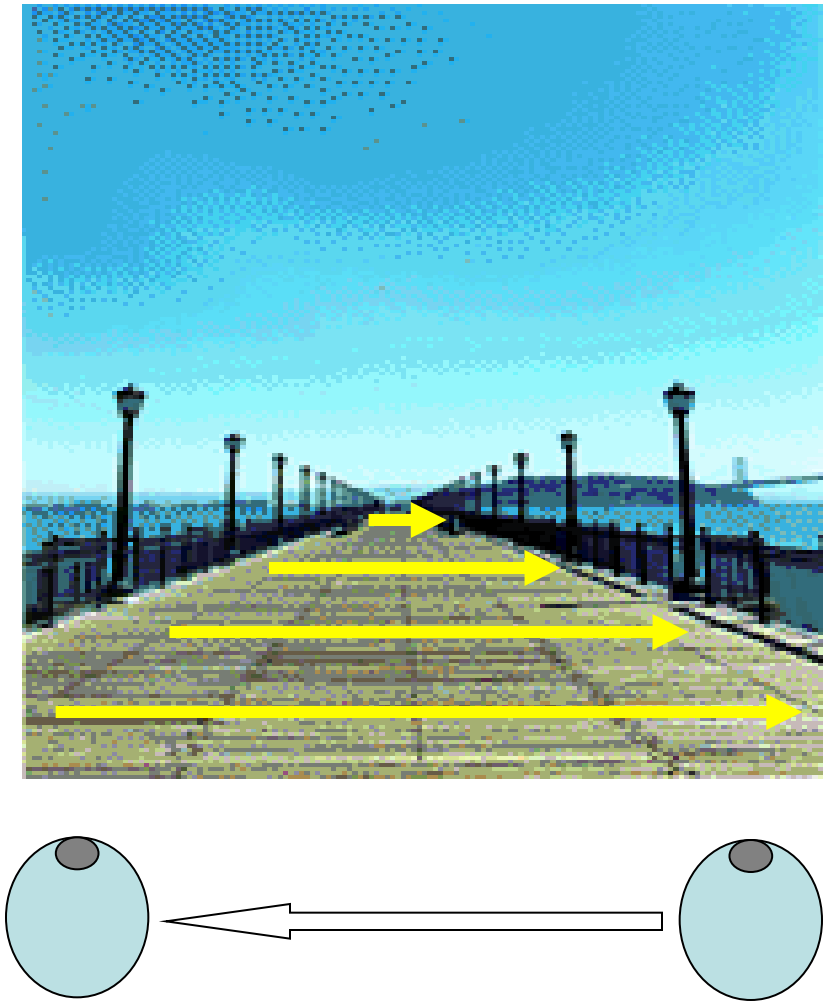
Texture gradient

- **Elements that are equally spaced in a scene occupy smaller visual angles as distance increases ⇒ surfaces with elements that occupy smaller and smaller visual angles are perceived as receding in depth**

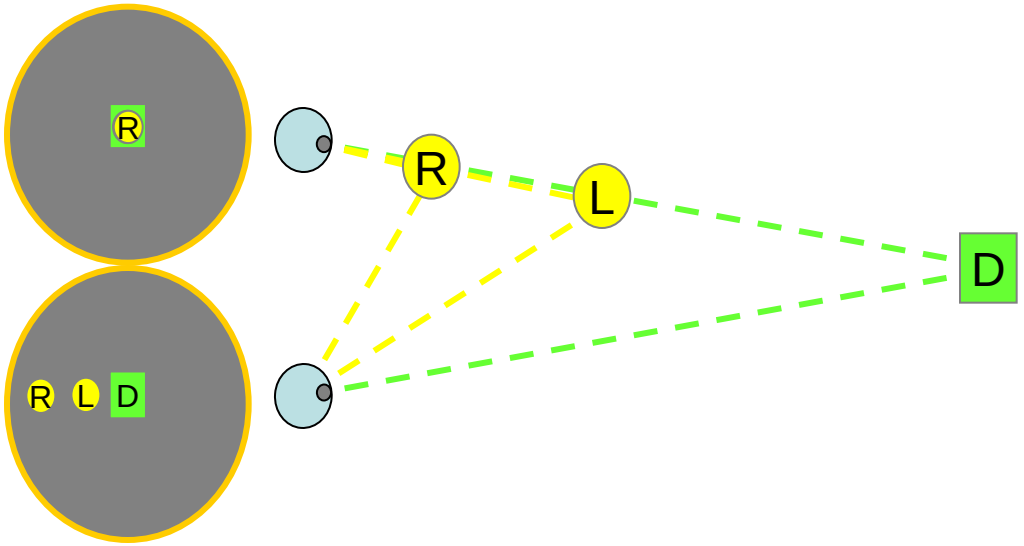


Motion Parallax

- Distant objects move more slowly (i.e., have lower optical velocity) than nearby objects

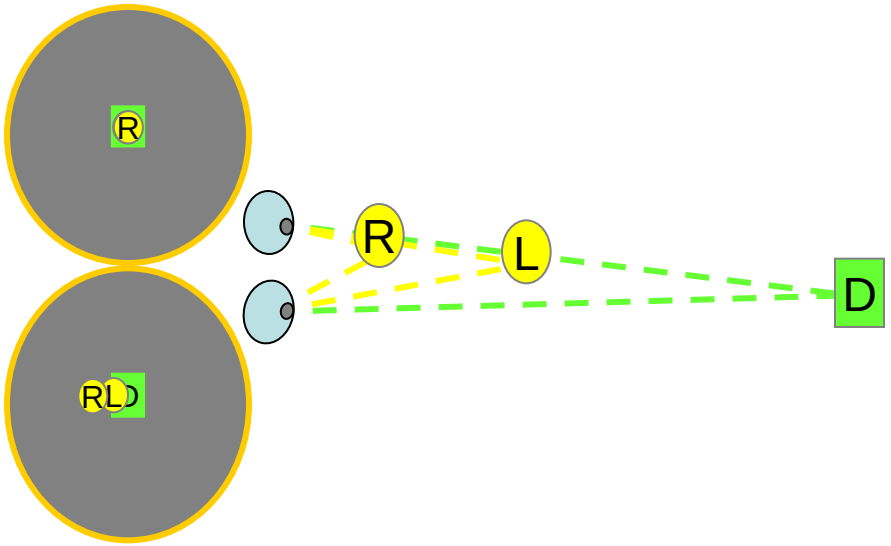


Stereopsis

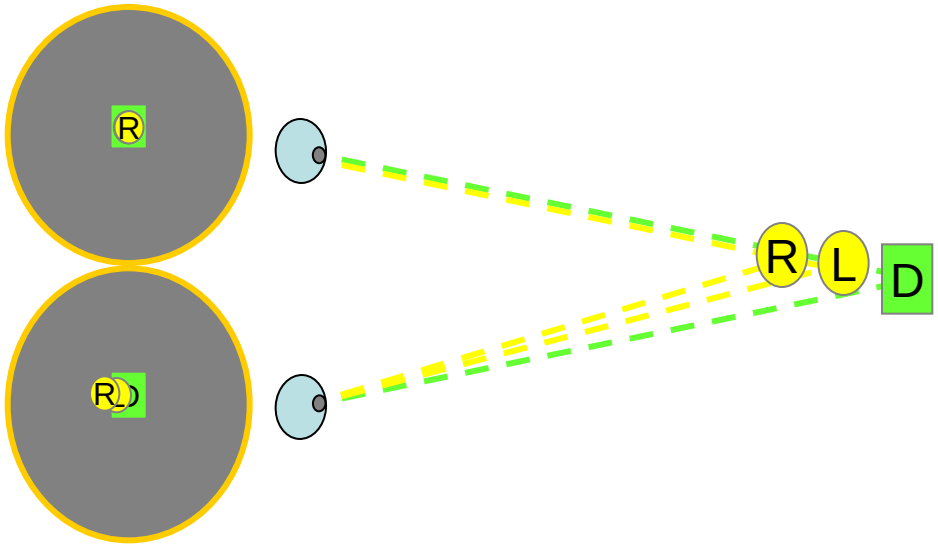


- Hold L index finger at arm's length
- Hold R index finger at ½ arm's length
- Close R eye and align both fingers with distant point (D)
- Fixate distant pont
- Close L eye and open R eye. Fixate distant point again.

Binocular disparity: the difference in the images on the two eyes due to:



separation of two eyes in space



difference in depth of objects in scene