COGS 101A: Sensation and Perception

Virginia R. de Sa

Department of Cognitive Science

UCSD

Lectures 10 and 11 (combined notes):

Models of Grouping and Object Perception

Course Information

- Class web page: http://cogsci.ucsd.edu/ desa/101a/index.html
- Professor: Virginia de Sa
  - ⋆ I'm usually in Chemistry Research Building (CRB) 214 (also office in CSB 164)
  - ⋆ Office Hours: Monday 5-6pm
  - ⋆ email: desa at ucsd
  - ⋆ Research: Perception and Learning in Humans and Machines

For your Assistance

TAS:

- Jelena Jovanovic OH: Wed 2-3pm CSB 225
- Katherine DeLong OH: Thurs noon-1pm CSB 131

IAS:

- Jennifer Becker OH: Fri 10-11am CSB 114
- Lydia Wood OH: Mon 12-1pm CSB 114

Course Goals

- To appreciate the difficulty of sensory perception
- To learn about sensory perception at several levels of analysis
- To see similarities across the sensory modalities
- To become more attuned to multi-sensory interactions

Grading Information

- 25% each for 2 midterms
- 32% comprehensive final
- 3% each for 6 lab reports - due at the end of the lab
- Bonus for participating in a psych or cogsci experiment AND writing a paragraph description of the study

You are responsible for knowing the lecture material and the assigned readings. Read the readings before class and ask questions in class.

Academic Dishonesty

The University policy is linked off the course web page.

You will all have to sign a form in section

For this class:

- Labs are done in small groups but writeups must be in your own words
- There is no collaboration on midterms and final exam

Midterm results

- Low 20 out of 70
- High 68.5 out of 70
- Mean 70%
- Median 73% exit

Common Mistakes I saw

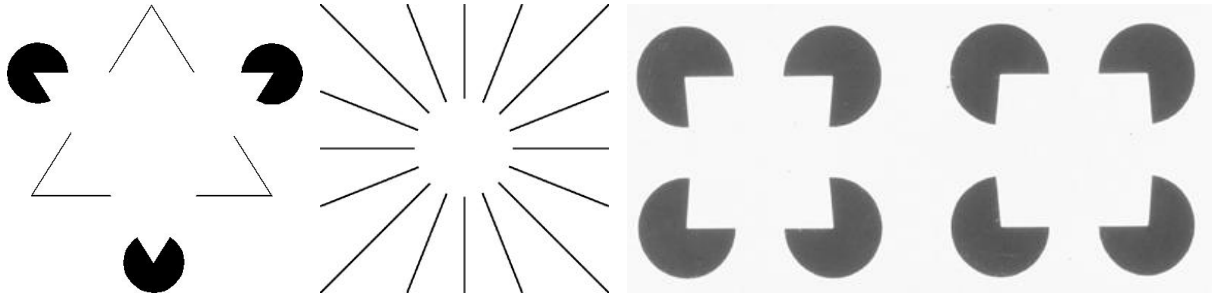
- receptive field
- contrast sensitivity function
- Mach bands question (off-center cells)
- What do you expect for V2 receptive fields?

Gestalt Psychology

The **Gestalt** movement in Psychology believes *the whole is different than the sum of its parts*

Motivated by findings such as

- Apparent motion from two spots of light we can get motion demo vary ISI for apparent motion demo
- Illusory contours

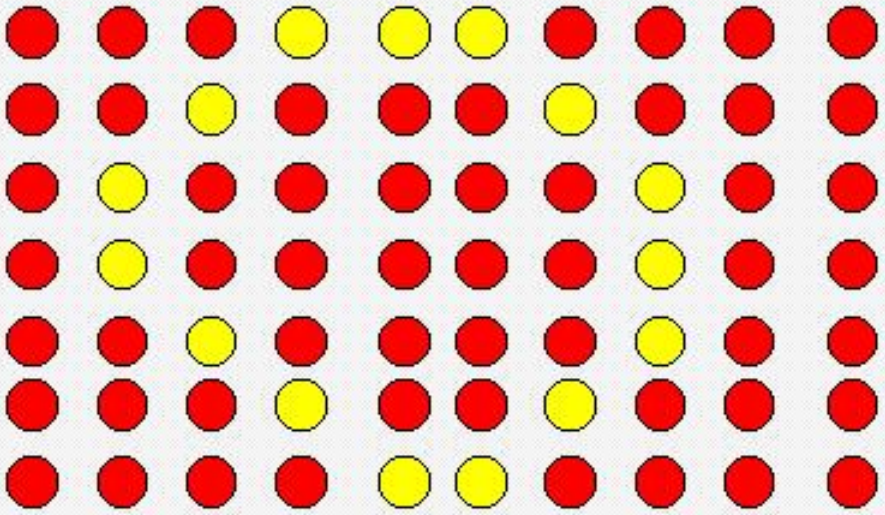


Gestalt laws of Perceptual Organization

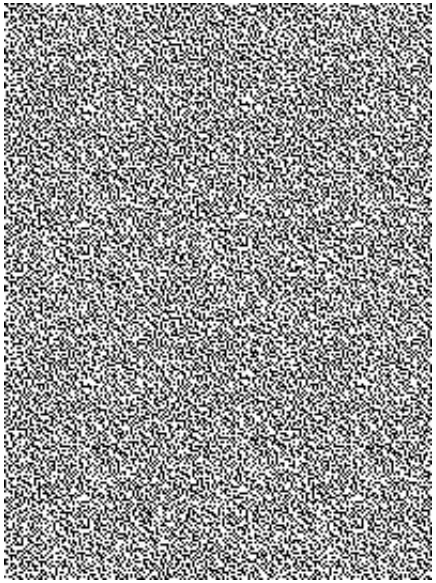
- Law of Pragnanz (Law of good figure) aka Law of Simplicity *Every stimulus pattern is seen in such a way that the resulting structure is as simple as possible*
- Law of Similarity *Similar things are grouped together*
- Law of Good Continuation *Points that, when connected, result in straight or smoothly curving lines are seen as belonging together, and the lines tend to be seen in such a way as to follow the smoothest path*
- Law of Proximity or Nearness *Things that are near to each other appear to be grouped together*
- Law of Common Fate *Things that are moving in the same direction appear to be grouped together*
- Law of Meaningfulness of Familiarity *things are more likely to form groups if the groups appear familiar or meaningful*

Law of Similarity

Items that look similar will be seen as parts of the same form



Law of Common Fate

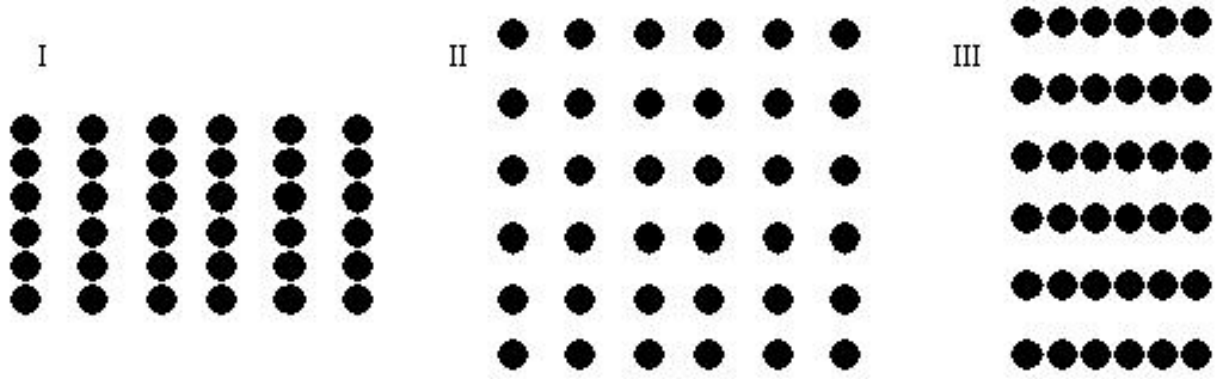


http://www.tutkie.tut.ac.jp/ mich/kitazaki.hm.html

Law of Pragnanz

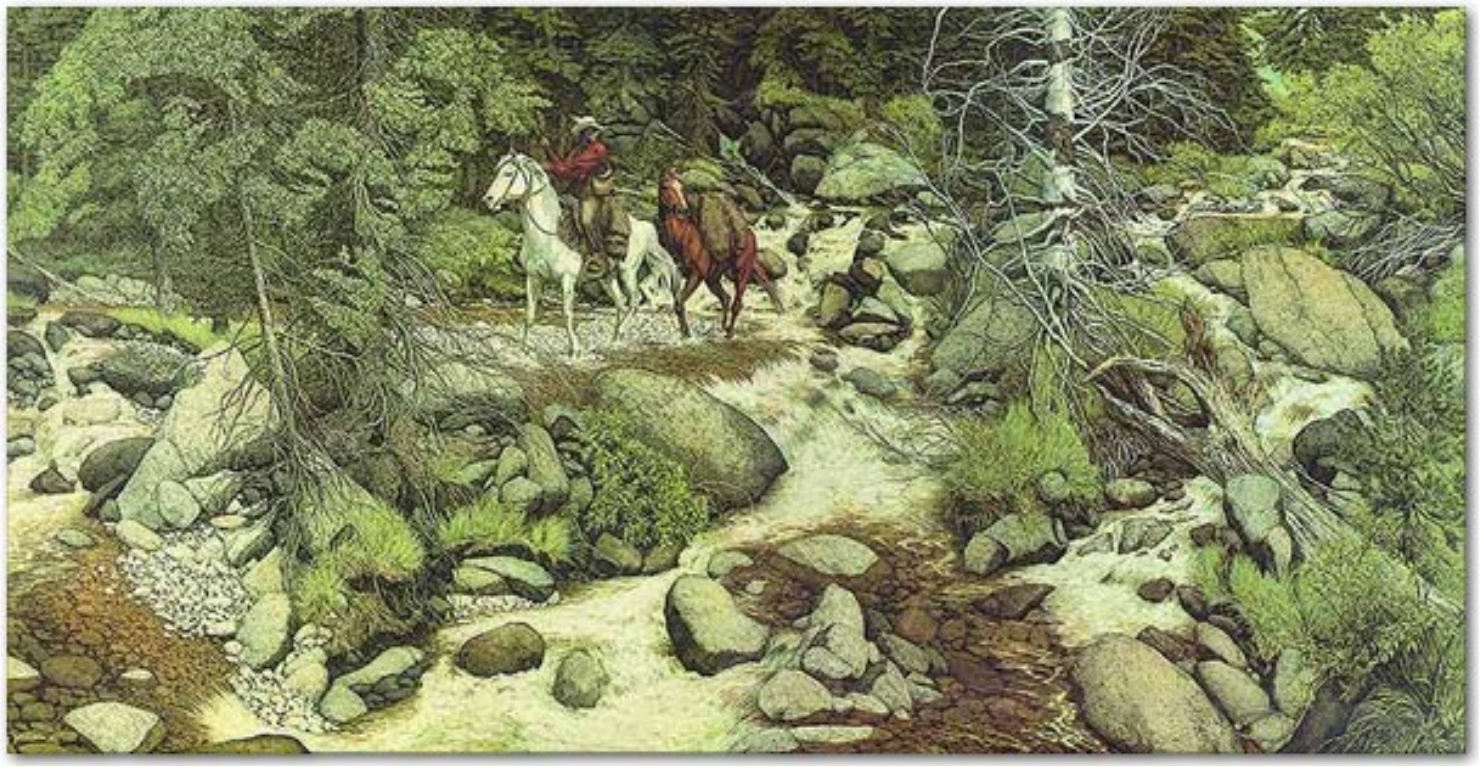
We see this as a square and triangle, not as a combination of strange shapes

Law of Proximity



http://www.cc.gatech.edu/classes/cs6751_97_winter/Topics/human-cap/process.html

Law of Meaningfulness



by Bev Dolittle (1985) http://bighornprints.com/prints/foresthaseyes.jpg

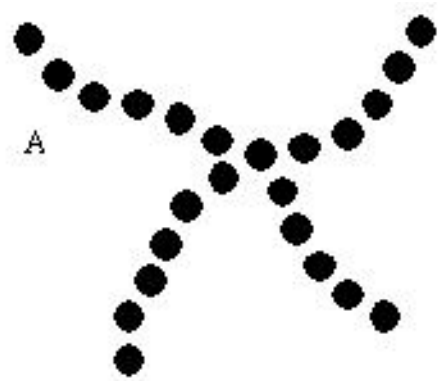
Law of Meaningfulness

This image is also from Beverly Doolittle

Law of Meaningfulness

The original image was created by R.C. James. This image was taken from Andy Wilson's home page which was scanned from David Marr's book Vision.

Law of Good Continutation



http://www.cc.gatech.edu/classes/cs6751_97_winter/Topics/human-cap/process.html

Amusing Example from a Photography Magazine

OW! That must hurt!

Gestalt laws of Perceptual Organization

- Law of Pragnanz (Law of good figure) aka Law of Simplicity *Every stimulus pattern is seen in such a way that the resulting structure is as simple as possible*
- Law of Similarity *Similar things are grouped together*
- Law of Good Continuation *Points that, when connected, result in straight or smoothly curving lines are seen as belonging together, and the lines tend to be seen in such a way as to follow the smoothest path*
- Law of Proximity or Nearness *Things that are near to each other appear to be grouped together*
- Law of Common Fate *Things that are moving in the same direction appear to be grouped together*
- Law of Meaningfulness of Familiarity *things are more likely to form groups if the groups appear familiar or meaningful*

Gestalt Laws are Heuristics

Gestalt Laws are not really laws (that must be obeyed) but **heuristics** (rules of thumb that are often obeyed)

Using heuristics allows us to usually find the right solution much faster than would be required without, but we don't always get the right solution

Palmer and Rocks additional grouping Principles

- Principle of common region *Elements within the same region of space are grouped together*
- Principle of element connectedness *Things that are physically connected are perceived as a unit*
- Principle of synchrony *visual events that occur at the same time will be perceived as going together* (like Principle of common fate, but they don't have to move or change in the same way)

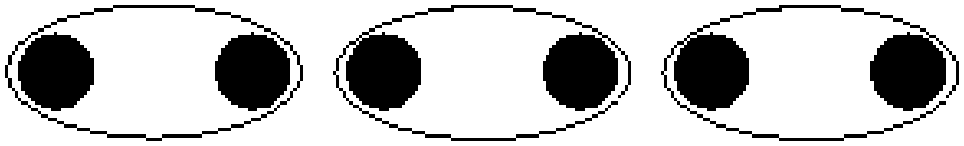
Palmer and Rocks additional grouping Principles

(after Figure 5.18 in the text)

Palmer and Rocks additional grouping Principles

(after Figure 5.18 in the text)

Principle of Common Region



(after Figure 5.18 in the text)

Principle of Connectedness

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

Principle of Synchrony

(after Figure 5.18 in the text)

How can we quantitatively assess these grouping rules?

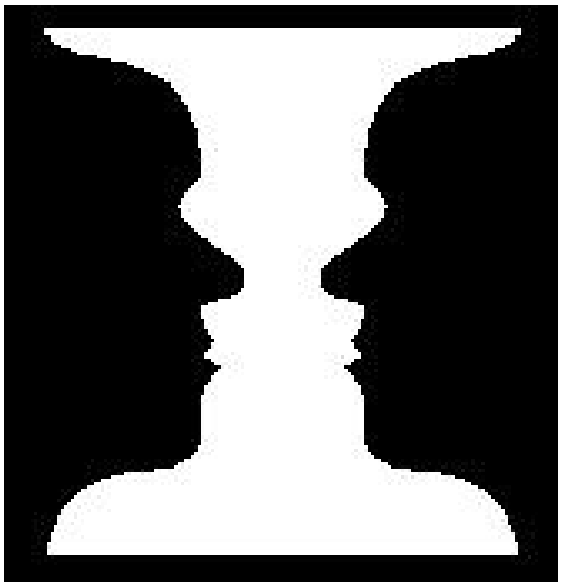
repetition discrimination task -find two adjacent identical shapes and press appropriate key (to match repeating shape)

(after Figure 5.19 in the text)

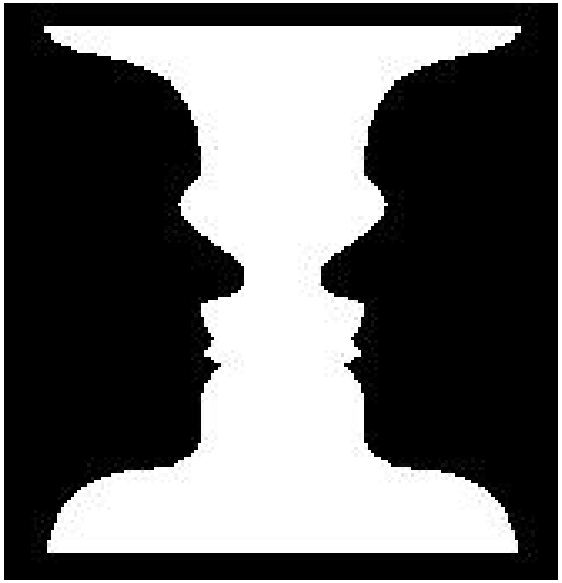
Figure-Ground segregation

How do we segregate an object from its background? This is the problem of **figure-ground segregation**

reversible figure-ground (which is the figure)



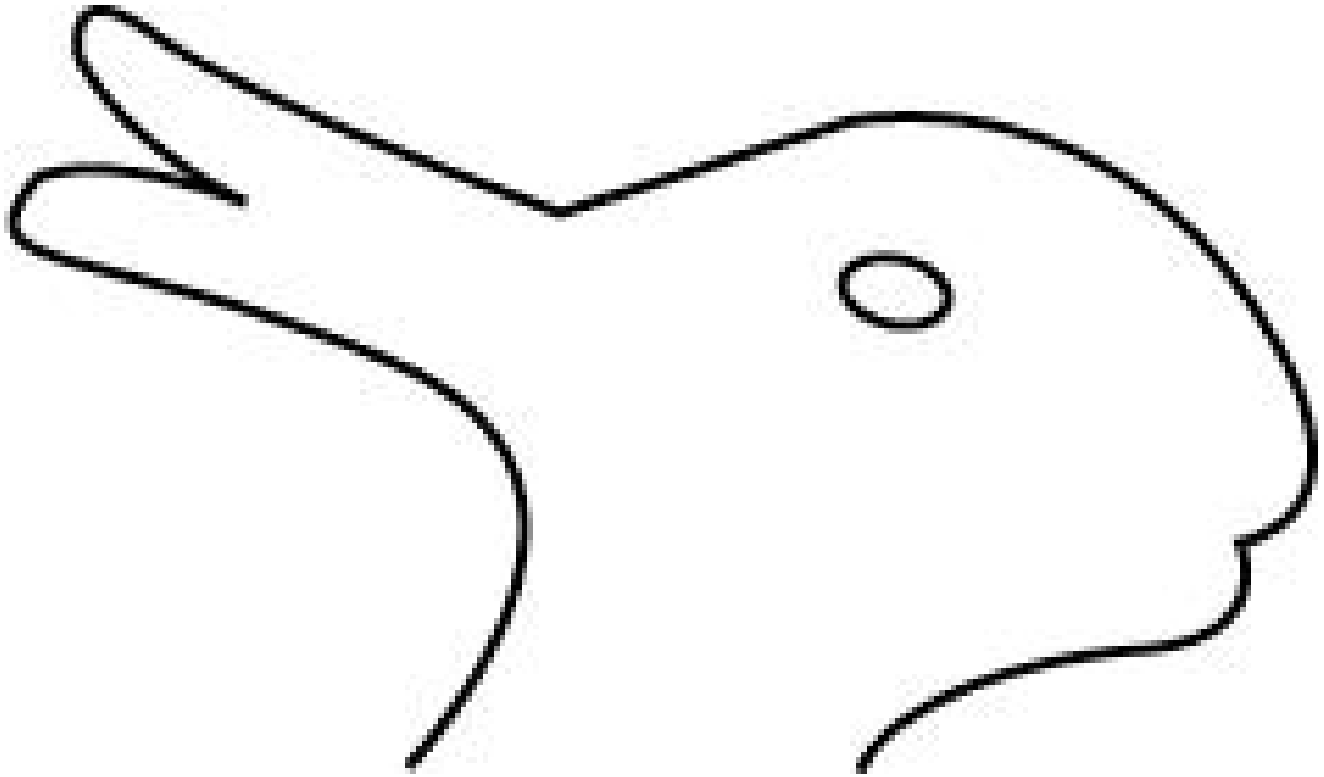
Properties of figure-ground segregation



- The figure is more memorable (priming studies)
- The figure is seen as in front of the ground
- The ground is seen as unformed material
- The separating contour seems to belong to the figure

Note you can't see both figures as when you see the vase as a figure the face like shapes are unformed background (and vice-versa)

More ambiguous figures

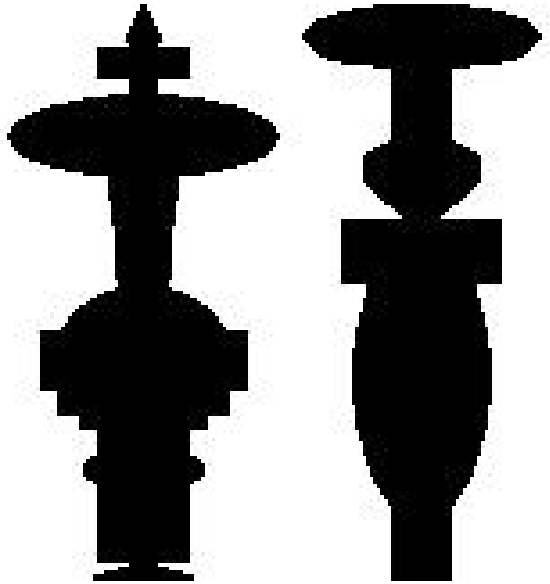


More ambiguous figures

More ambiguous figures

Factors that influence which part will be figure

Symmetrical areas tend to be seen as figure



Factors that influence which part will be figure

Symmetrical areas tend to be seen as figure

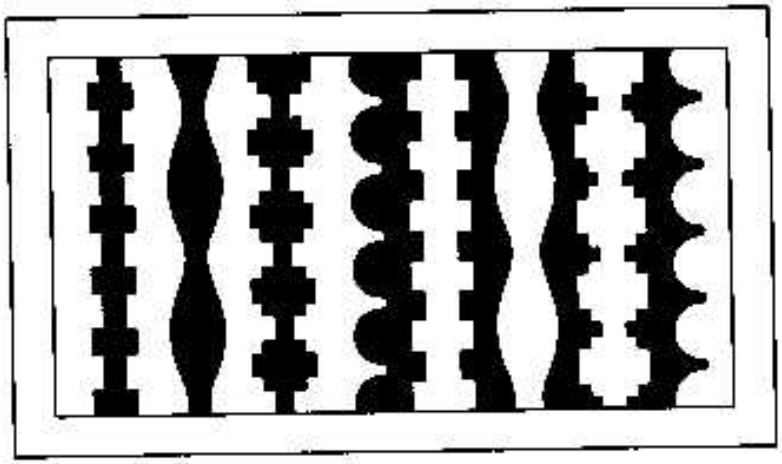
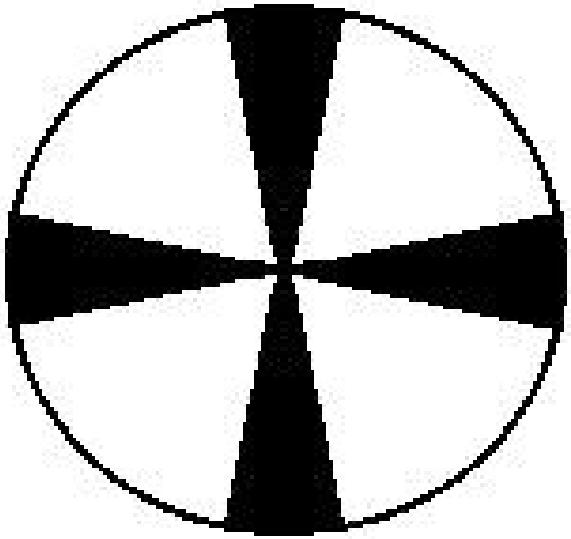


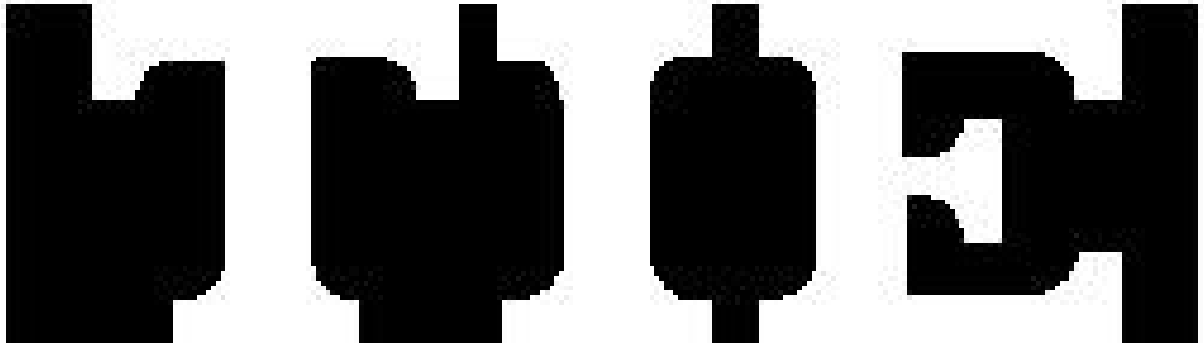
Figure 7.25
Symmetry and figure ground. Look to the left and to the right, and observe which colors become figure and which become ground. (Adapted from Hochberg, 1971.)

Relatively smaller areas are more likely seen as figure



Vertical or Horizontal orientations are more likely seen as figure (draw on board)

Meaningful objects (especially letters/words) are more likely seen as figure



Models of Object Recognition

Marr's computational theory

In 1982, David Marr published a very influential book $Vision$

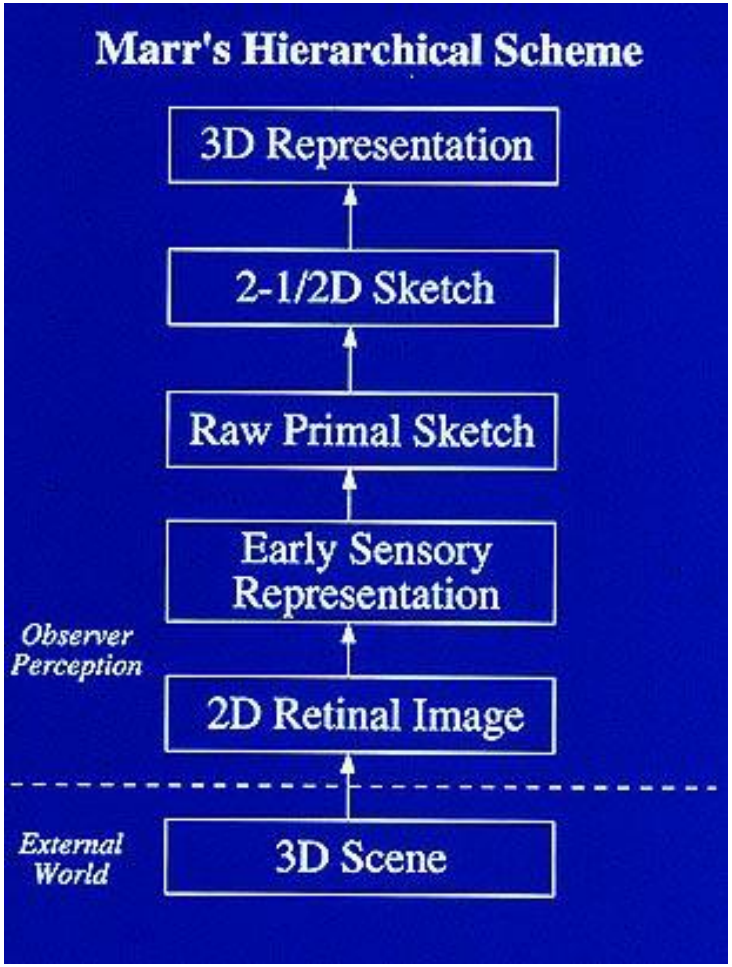
Marr's computational theory

- First look for edges and simple shapes (primitives) **Primal sketch**
  - ⋆ use **natural constraints** to tell object edges from illumination edges (object edges sharp, illumination edges fuzzy)
- similar features are grouped to give information on surfaces and their layouts $2\frac{1}{2}$-**D sketch**
- **3-D sketch** full 3-D representation

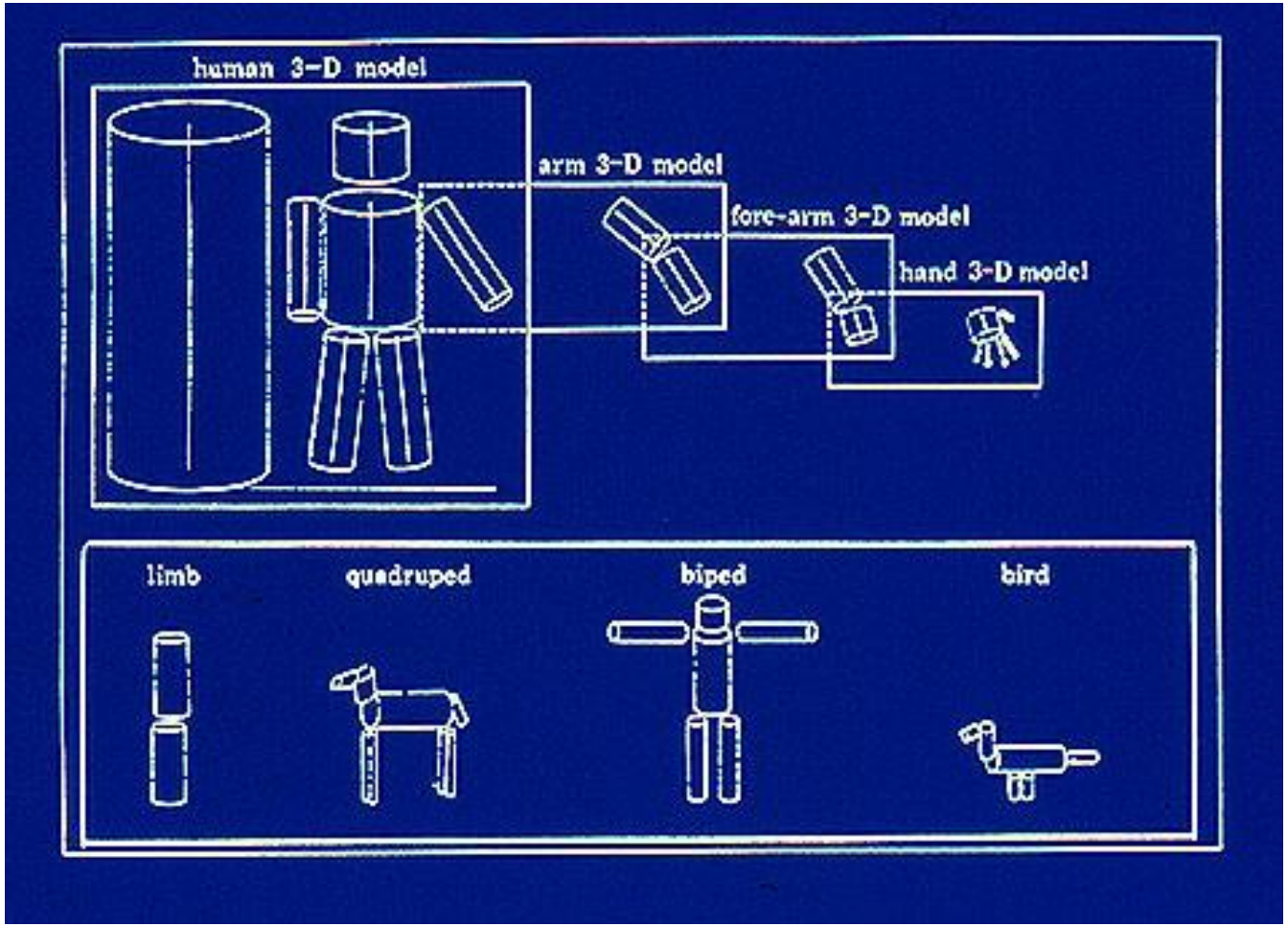
Note:

The primal sketch does not reach conscious perception

Motivated by knowledge of physiology (edge detecting neurons, termination detectors)


Marr's Hierarchical Scheme
3D Representation
2-1/2D Sketch
Raw Primal Sketch
Early Sensory Representation
Observer Perception
2D Retinal Image
External World
3D Scene

Marr's computational theory



Treisman's Feature Integration Theory (FIT) theory

Features of objects are processed separately (pre-attentively) and then recombined (focused attention)

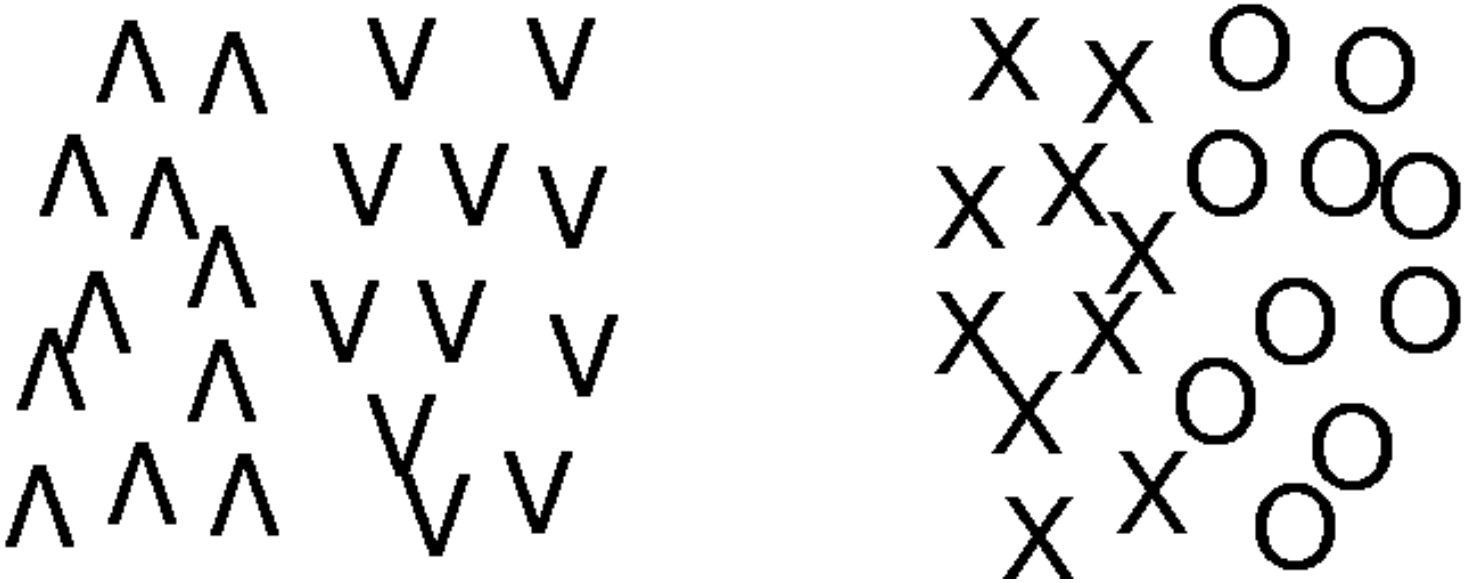
Example features (color,orientation, curvature)

Once combined we can perceive an object

Preattentive features determined based on popout experiments (texture segmentation, visual search)

Also based on experiments with **illusory conjunctions**

Popout Texture Segmentation



If the texture boundary pops out, it has different features that can be identified pre-attentively

Popout Texture Segmentation

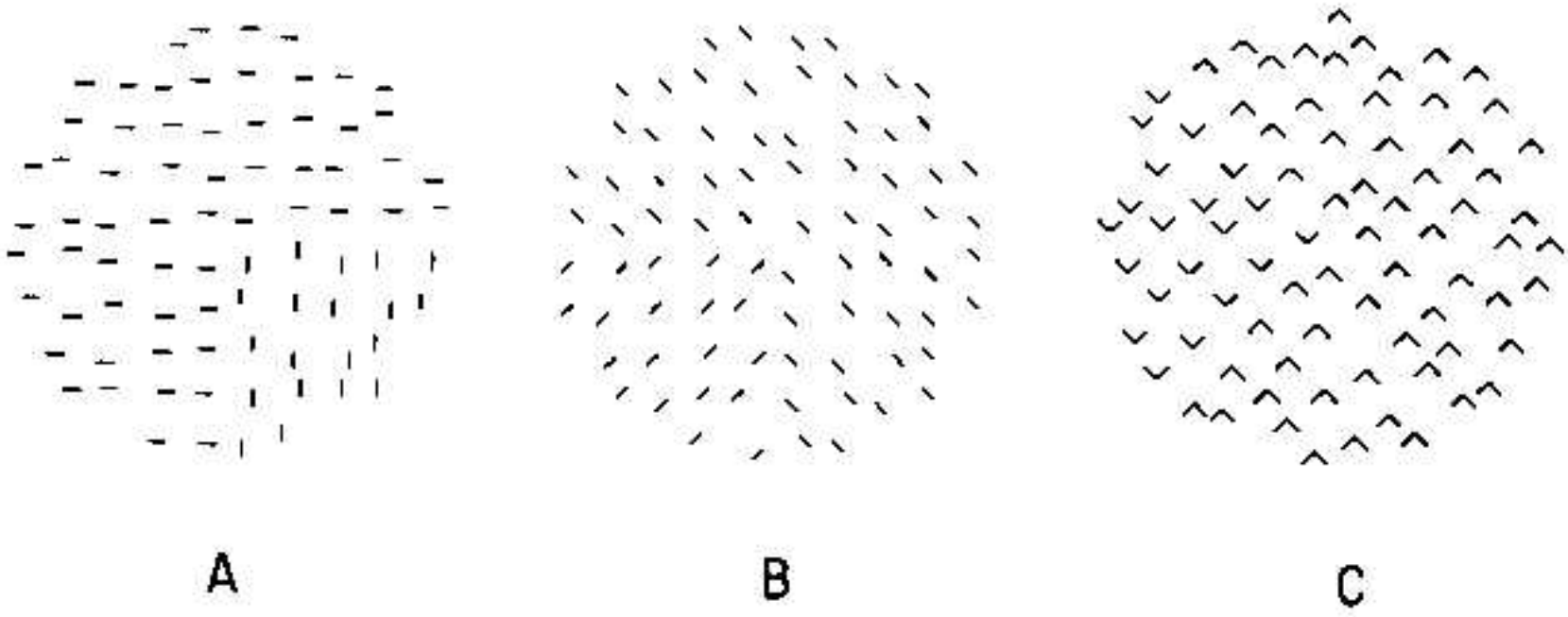
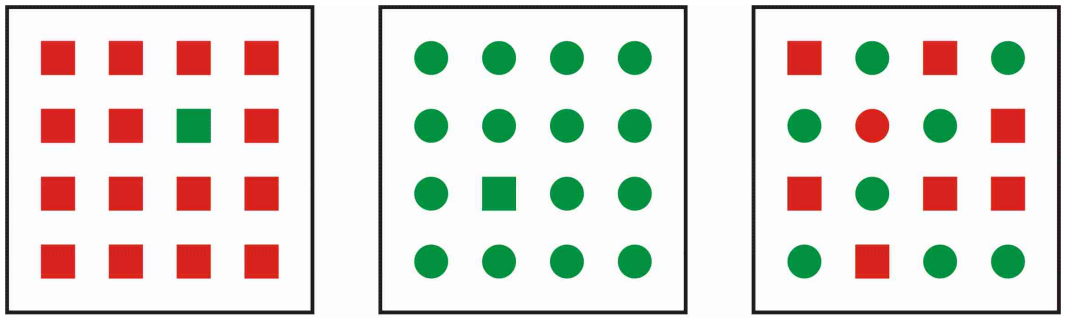
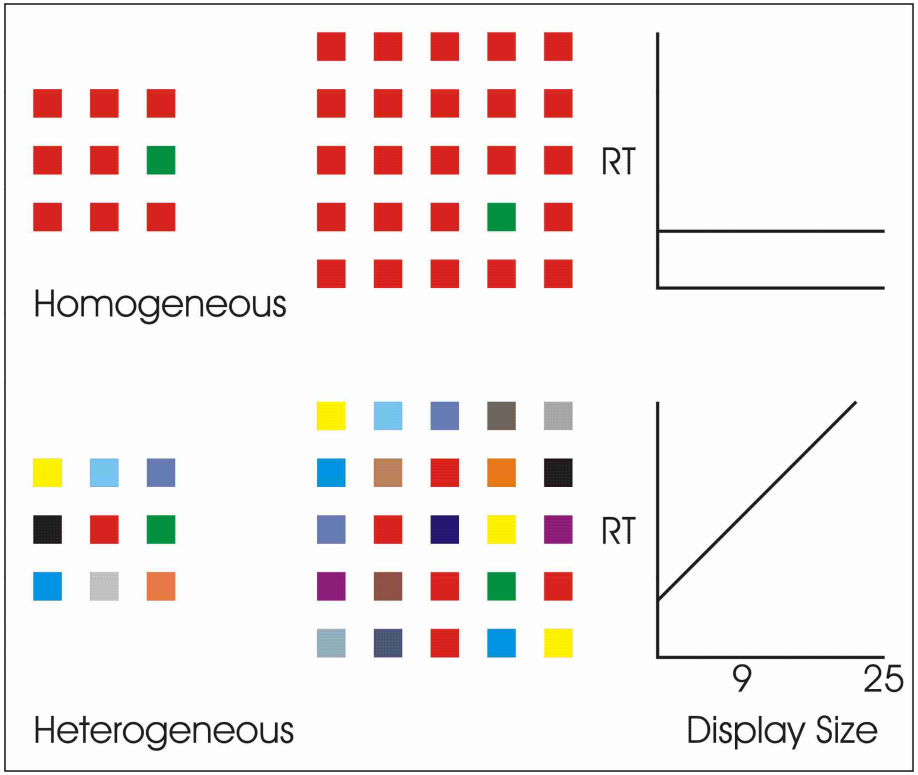


FIG. 6.6. The task is to locate the region of the field containing the disparate elements. These panels show how elements group on the basis of similar line slope to make the task easy in Panels A and B but difficult in Panel C. (Adapted from Olson & Attneave, 1970.)

Feature vs Conjunction Search

If target differs from all distractors along one dimension you get popout. If target differs only in the conjunction of features, you get serial search.





(above 3 images from www.owlnet.rice.edu/ psych351/Images/)

Treisman basic features

Visual Search paradigm

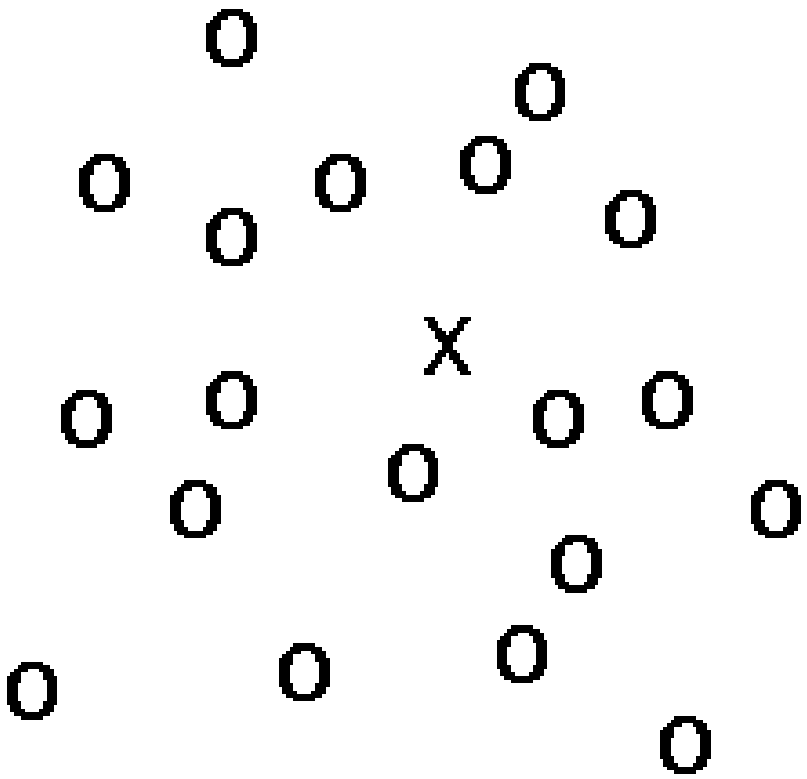
Find a particular element

e.g. Find the Black O

serial search - when search time increases with elements in the display

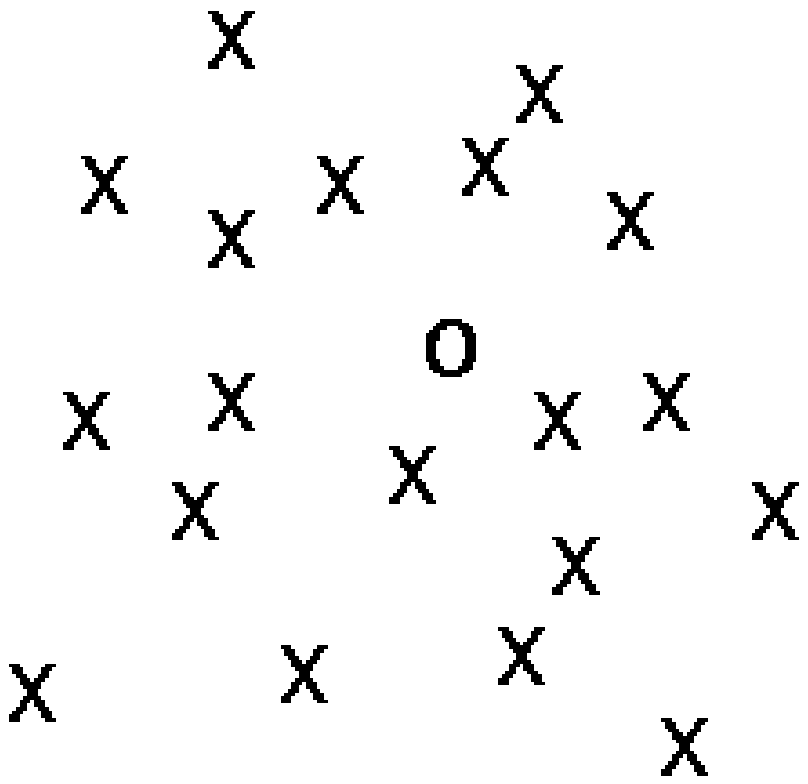
popout - when search time is almost constant for any number of elements in display

Treisman search paradigms



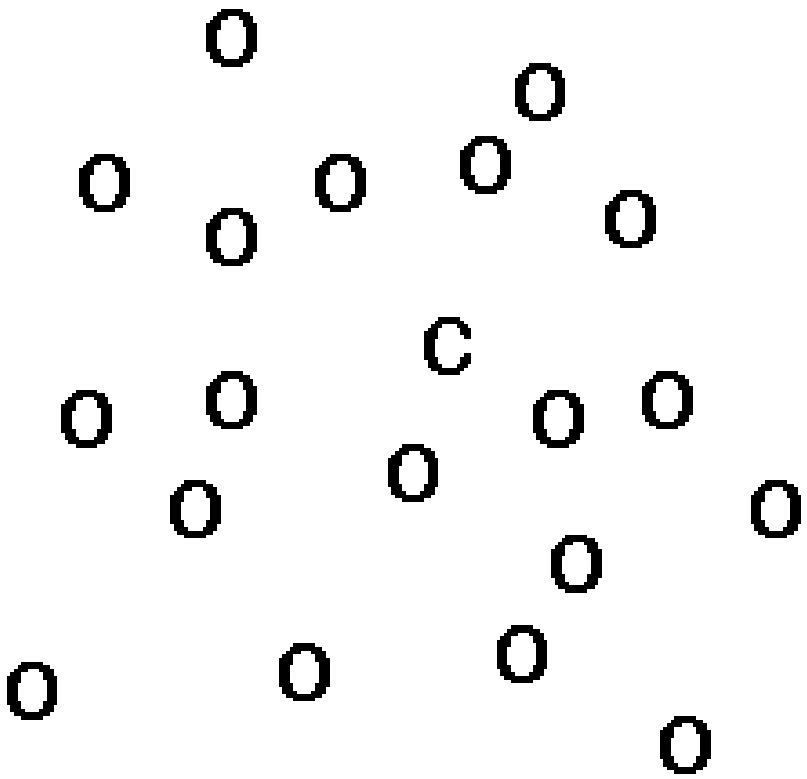
A popout search

Treisman search paradigms



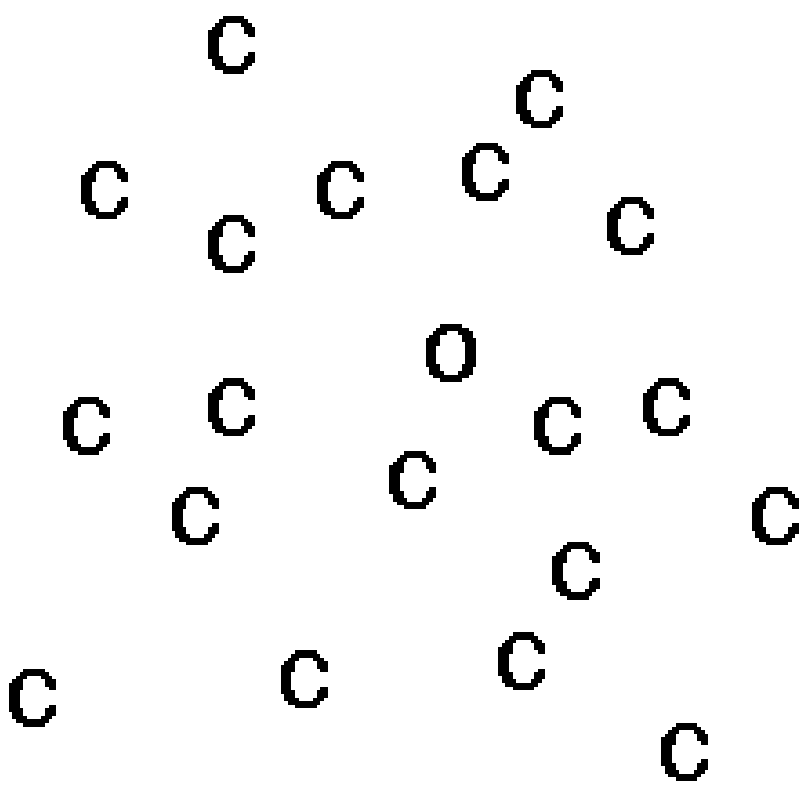
A popout search

Treisman search paradigms



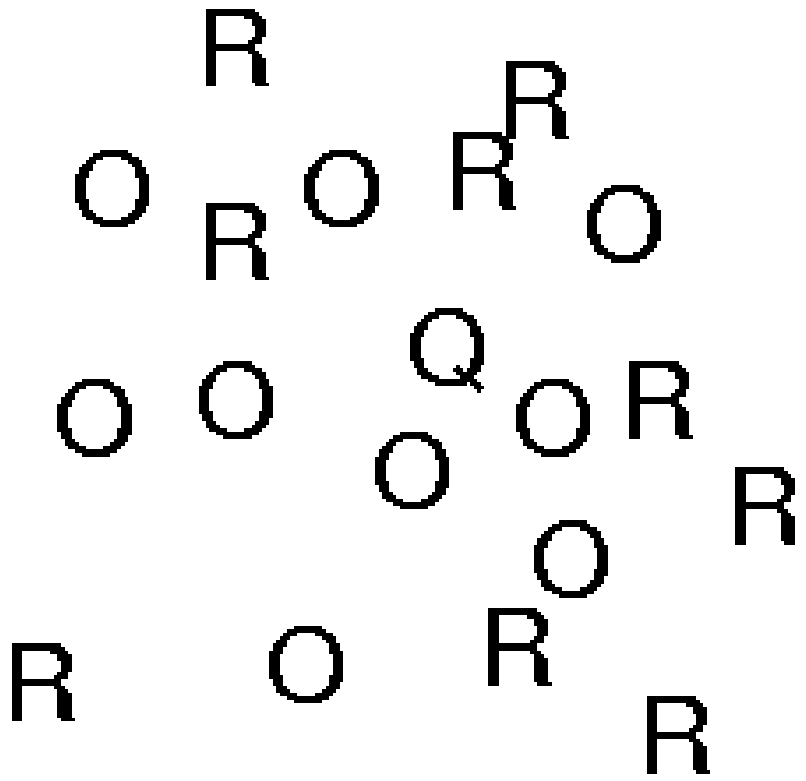
A popout search

Treisman search paradigms



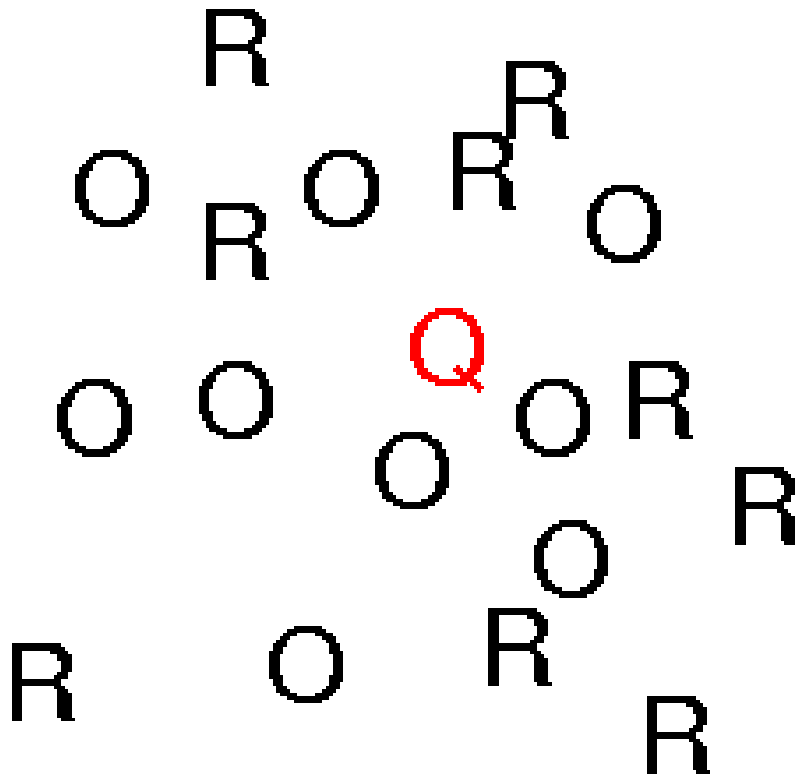
(Subtle point) Slightly less popout indicates that lack of a gap may be less salient than a gap

Treisman search paradigms



No popout (conjunction search)

Treisman search paradigms



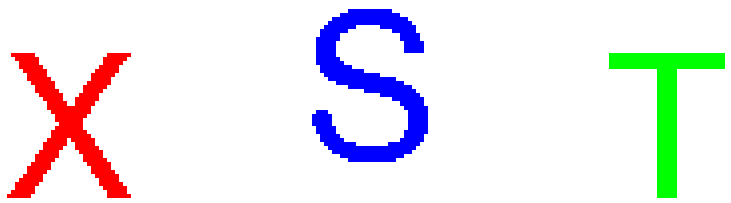
Popout again by feature search for red

Other evidence for separate feature analysis

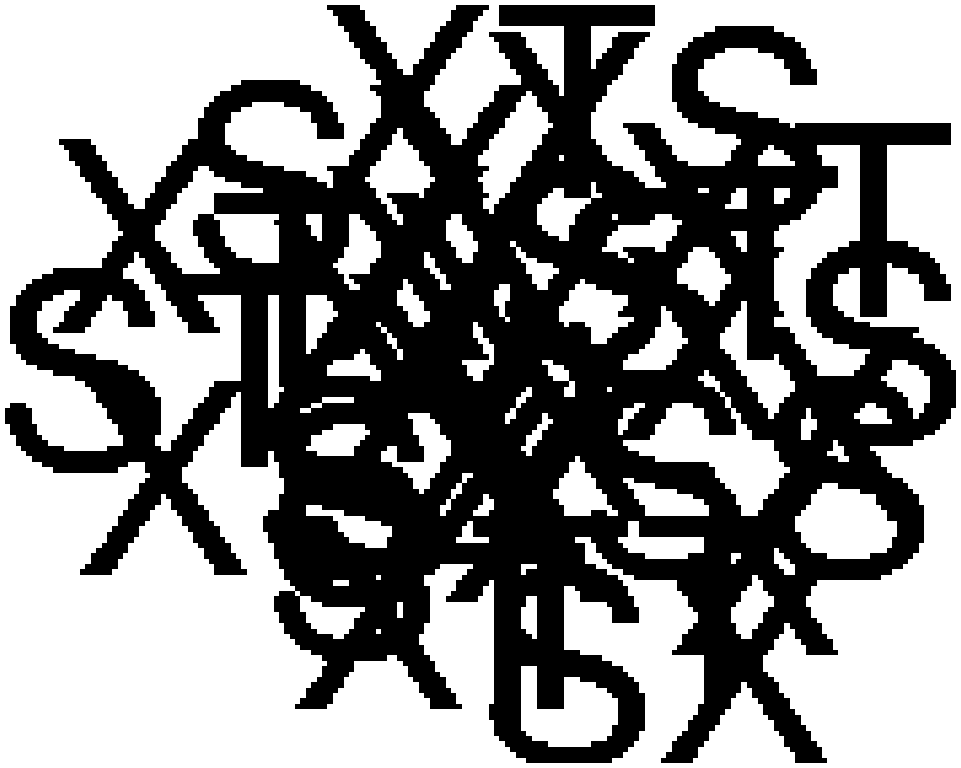
Illusory conjunctions – When fast presented stimuli (200msec presentation time - followed by mask to prevent afterimages)

Remember: press fast twice

Illusory Conjunctions



Illusory Conjunctions



Illusory Conjunctions

Many subjects will pair up the colors and letters wrong

Illusory Conjunctions

Many subjects will pair up the colors and letters wrong

e.g. report Red S, Green T, Blue X when there was actually Red X, Blue S, Green T

FIT theory

- features are processed independently (and without attention)
- with focused attention features at a particular location are combined

Illusory conjunctions arise when there is insufficient time to combine features

Popout occurs when targets differ in a single feature

Serial search occurs when targets differ only in their combination of features

Biederman's Recognition by Components

Biederman's Recognition by Components model proposes that we recognize objects by recognizing spatial combinations of 36 volumetric primitives called **geons**

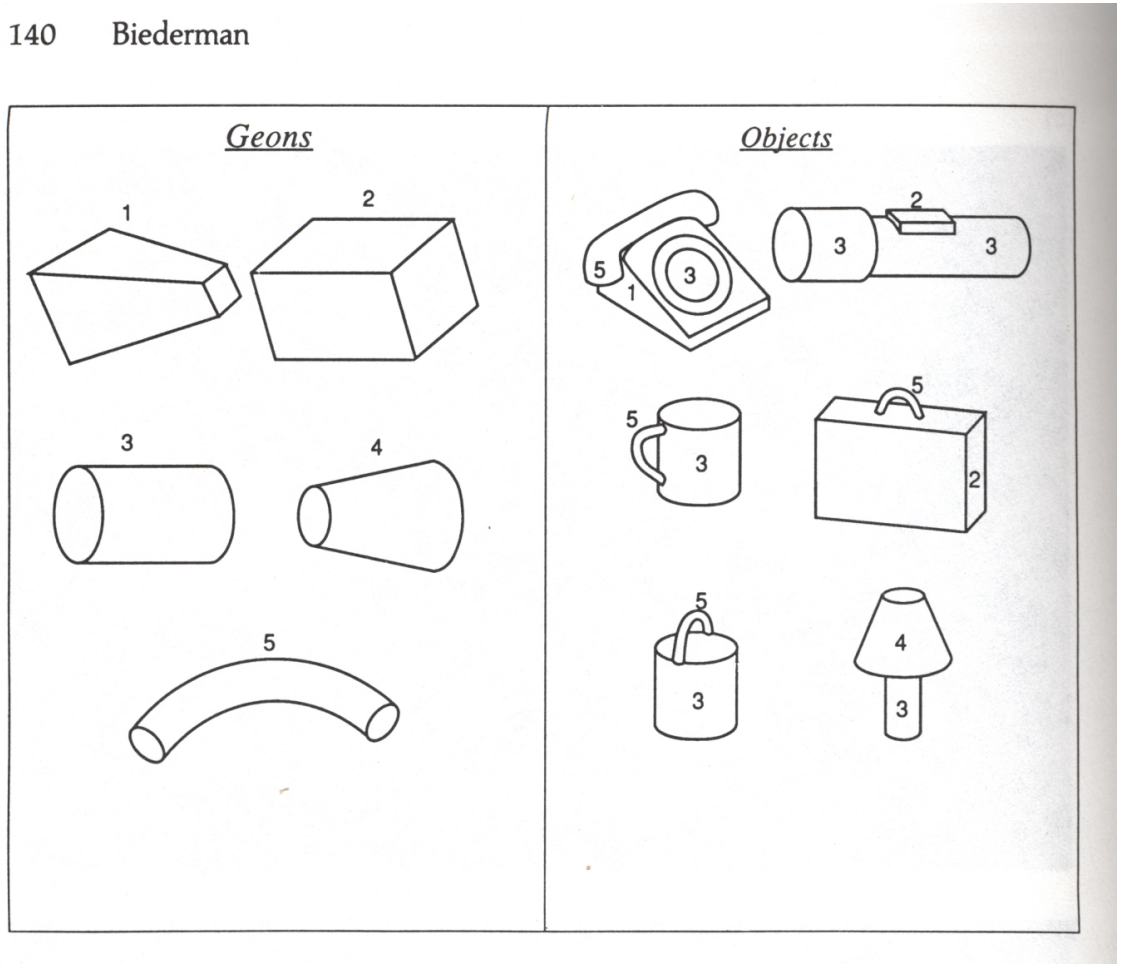
Geons have **view invariance**, are **discriminable** from other geons, and **resistant to visual noise**

(e.g. geon 4 on top of geon 3 is a lamp)

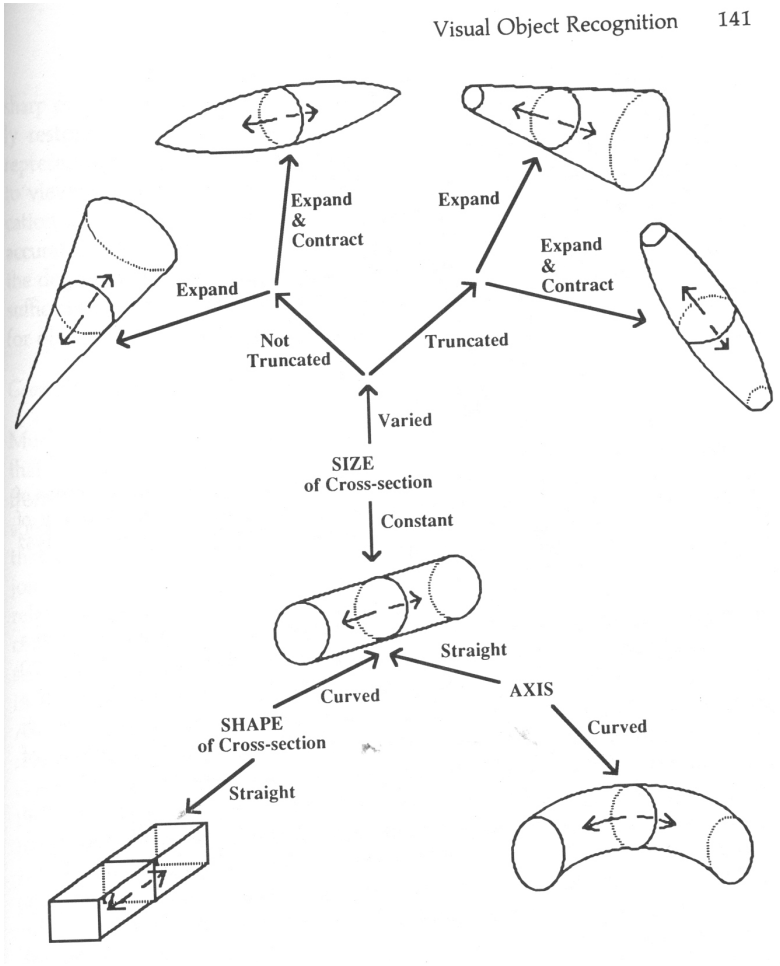
Not able to model how we distinguish different items within a class (e.g. different faces, different birds)

Example Geons

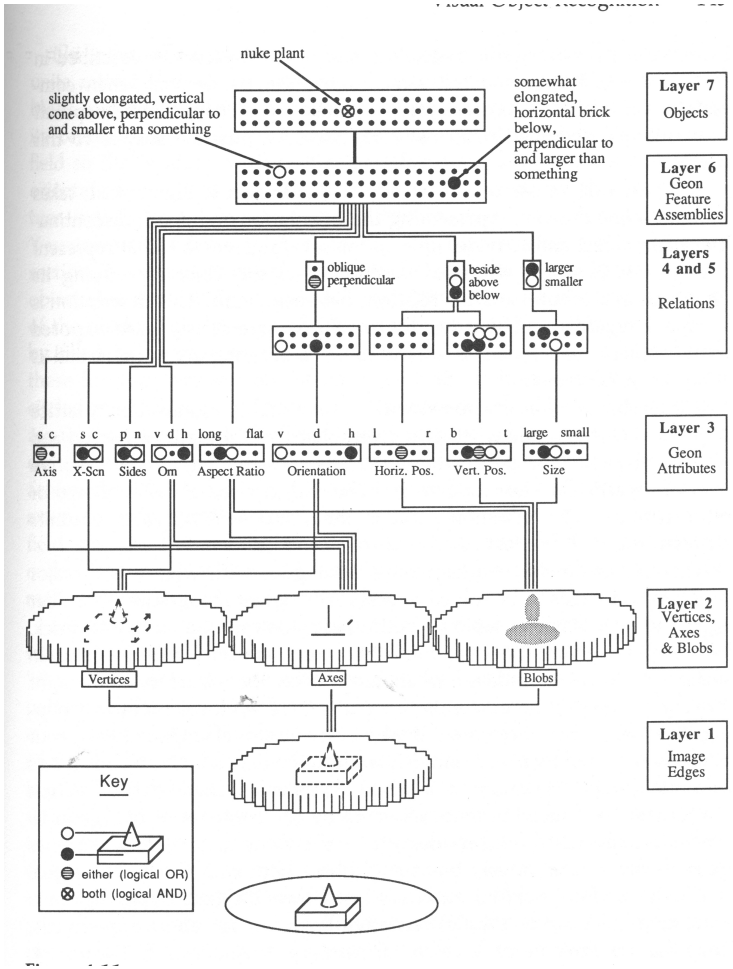
Objects are made up of Geons



How Geons are related



Hummel and Biederman's computational model of object recognition



How were geons determined?

Psychophysical studies show that when you obscure geon shapes, subjects are less able to recognize objects.

Can do experiments where you change geons or just arrangement between them and see how recognition performance changes.

Bottom-up vs Top-down models

Marr's model is a bottom-up model - It considers how the input could be processed to obtain the objects identity

Biederman's model is also a bottom-up model - spatial relationships between simple features are computed to obtain primitives (Y vertex, arrow vertex,parallel lines...) which are combined to get Geons whose spatial relationship defines the object

We have considered Treisman's model as a bottom-up model - We compute features and then combine them but Treisman "carrot, lake, tire" study shows that there are **top-down** effects also

Why?

Bottom-up vs Top-down models

Marr's model is a bottom-up model - It considers how the input could be processed to obtain the objects identity

Biederman's model is also a bottom-up model - spatial relationships between simple features are computed to obtain primitives (Y vertex, arrow vertex,parallel lines...) which are combined to get Geons whose spatial relationship defines the object

We have considered Treisman's model as a bottom-up model - We compute features and then combine them but Treisman "carrot, lake, tire" study shows that there are **top-down** effects also

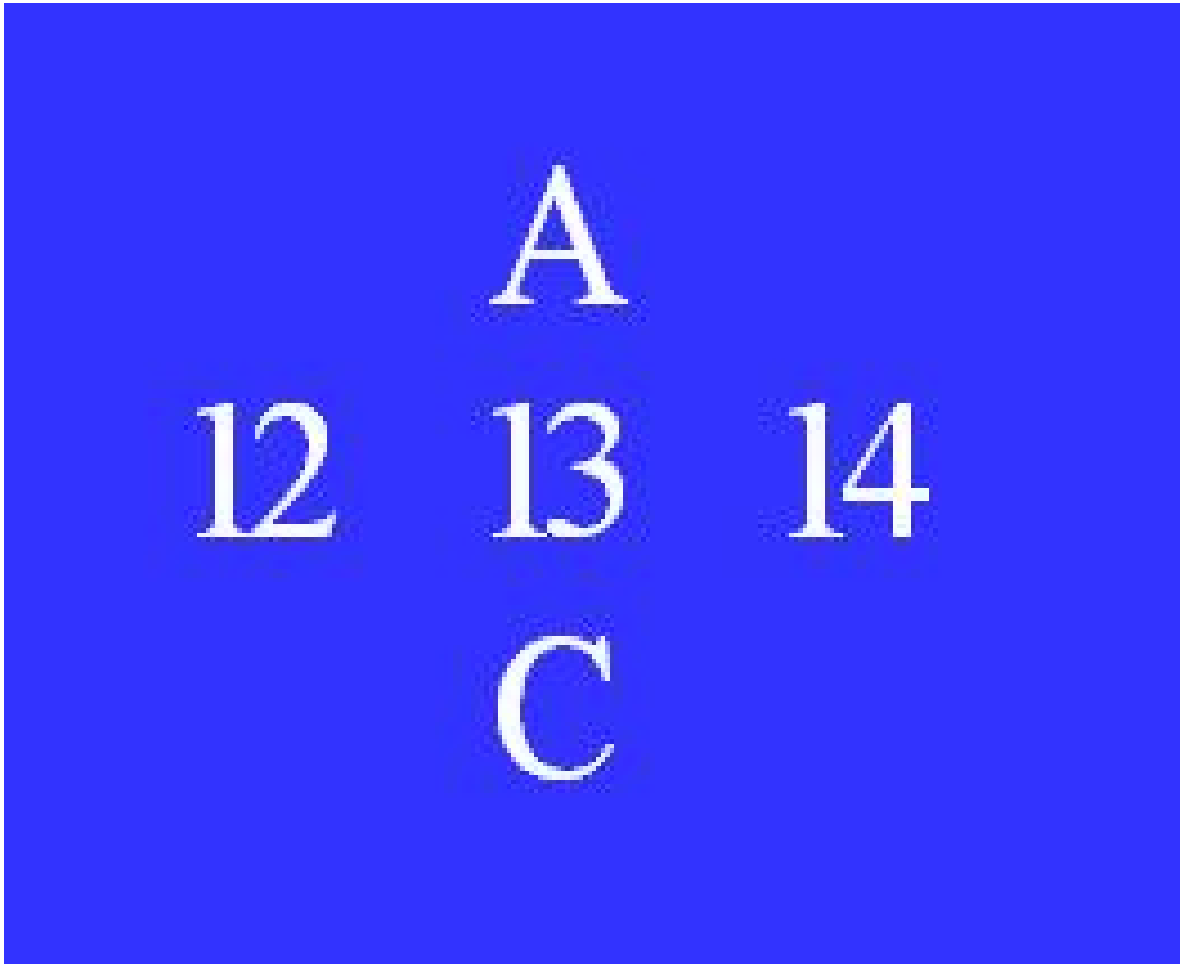
Why?

Because the identical visual input gives rise to different perceptions based on prior exposure

Top-down effects in object recognition

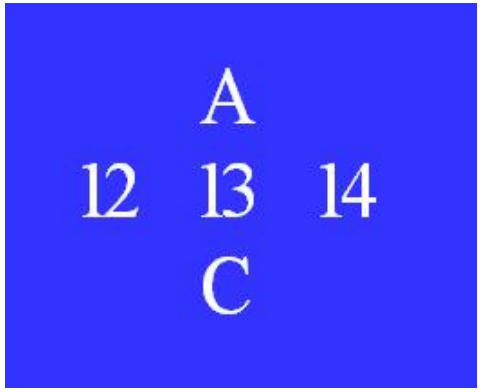
Ambiguous figures are another occasion where the same bottom-up visual input can be seen in two different ways. Your perception can also be biased by your top-down expectations.

Top-down Effects: Object identification influenced by context



http://www1.cs.columbia.edu/~paley/spring03/assignments/HWtmp/sz184/pattern.ht

Top-down Effects: Object identification influenced by context



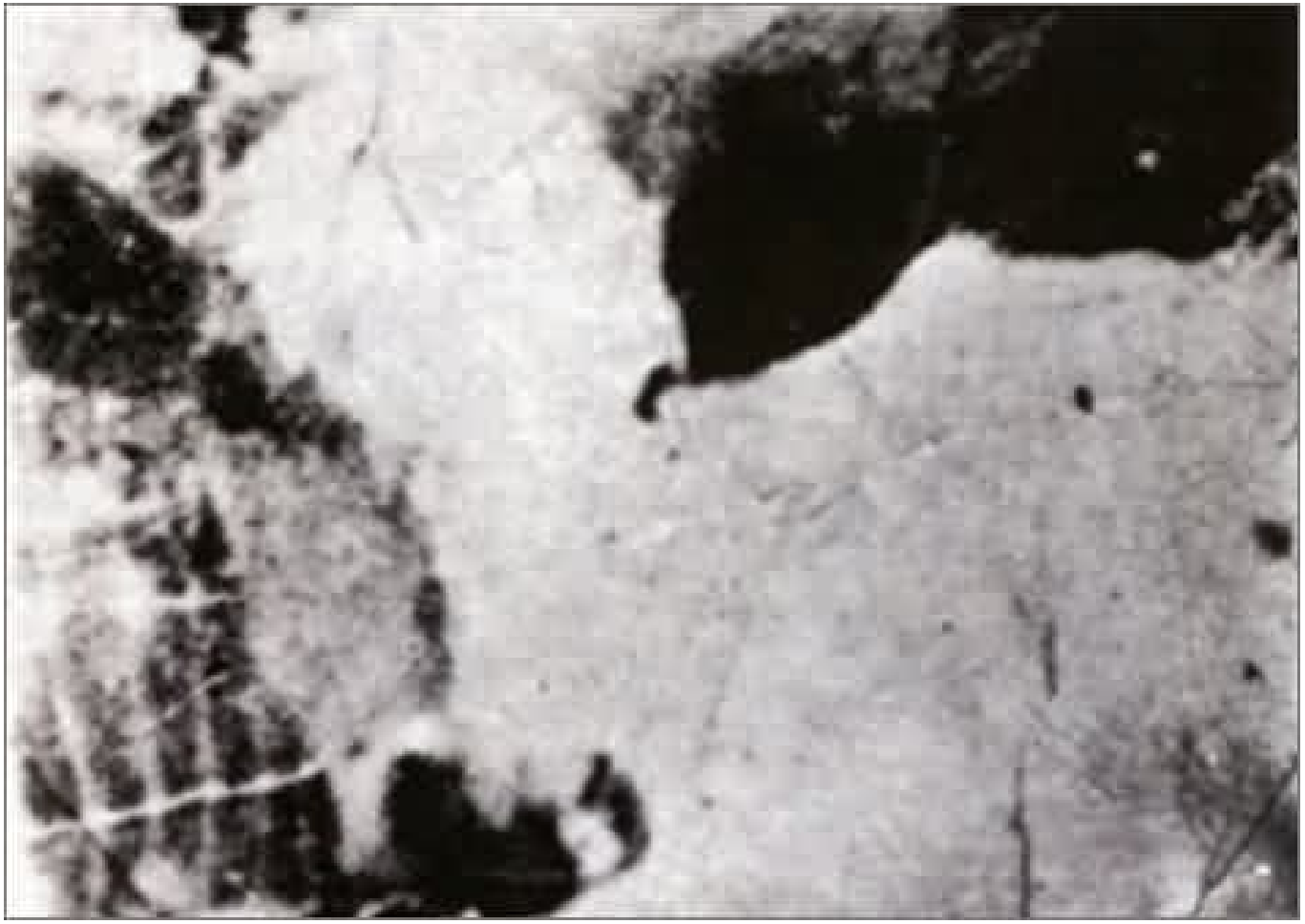
http://www1.cs.columbia.edu/~paley/spring03/assignments/HWtmp/sz184/pattern.ht

Can also be influenced by priming. If you have been primed for letters, you will tend to see a B.

Top-down expectations influence your perception

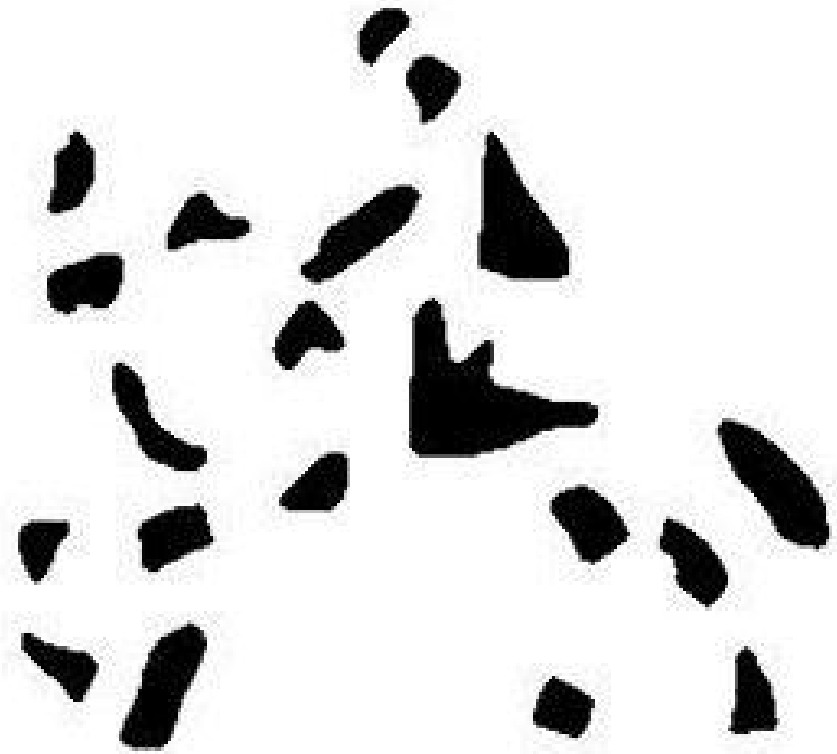
Acknowledgement: Image supplied by Amazing Art

Top-down expectations influence your perception



Acknowledgement: Image supplied by Amazing Art

Top-down expectations influence your perception



Acknowledgement: Image supplied by Amazing Art

Top-down expectations influence your perception

Acknowledgement: Image supplied by Amazing Art

The Thatcher Illusion – influence from learned experience with (upright) faces


[Thompson 1980,Perception 9 483-484]

The Thatcher Illusion – influence from learned experience with (upright) faces


[Thompson 1980,Perception 9 483-484]

This illusion was first described by Thompson in 1980. I got this from http://www.essex.ac.uk/psychology/visual/thatcher.html

Influence from learned experience is specific

[Sinha and Poggio Nature 1996, 384 p 404]

Influence from learned experience is specific

[Sinha and Poggio Perception 2002, 31(1)]
http://perceptionweb.com/perc0102/sinha.html

Top-down effects in object recognition

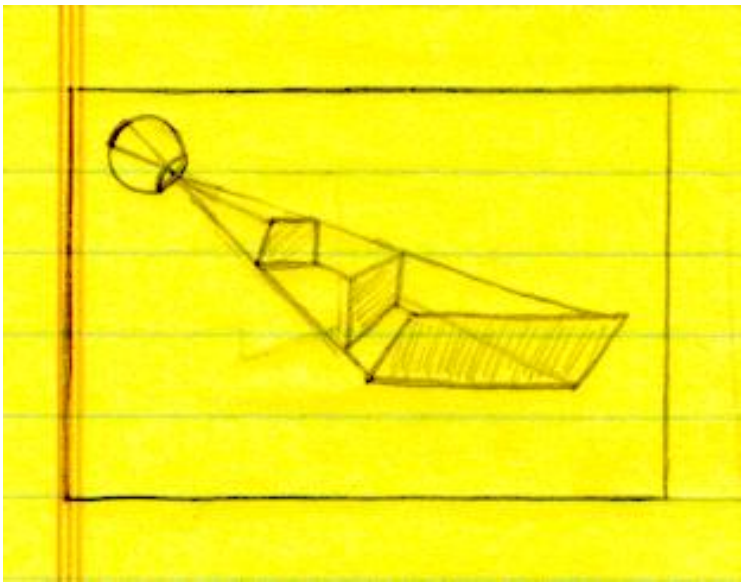
In fact the brain has to make assumptions in order to see 3-D scenes.

Why is object recognition hard

There are an infinite number of 3-D scenes that can lead to the same retinal image

(e.g. see Figure 5.46 in the textbook)

Consider the Penrose Triangle. When viewed from one direction we see what looks like an "impossible triangle"



http://www.graphics.cornell.edu/ jaf/projects/pn/space.html

But when viewed from another angle, we see the true shape. However it is still hard to "see" that shape when viewed from the **accidental** viewpoint. Your brain

makes the assumption that you are not viewing from an accidental viewpoint. (This is related to the Gestalt Law of Continuity)

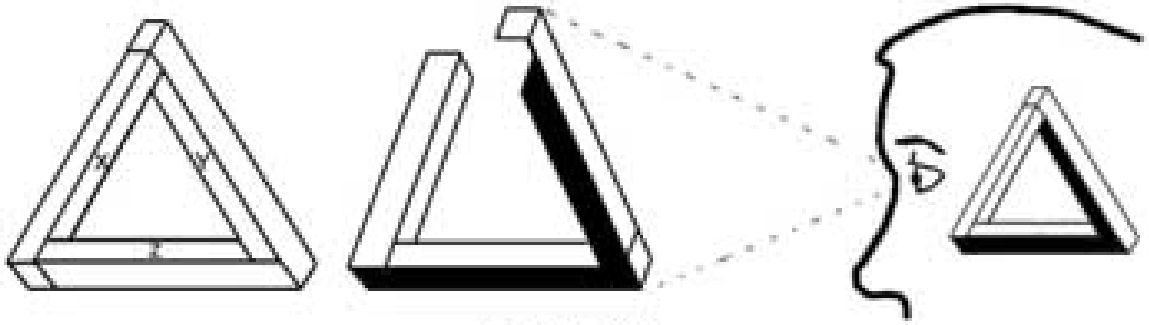


Figure 20

http://collections.ic.gc.ca/science/francais/bio/optique2.html

Vision is Hard - distinguishing edges from shadows from object edges

Vision is Hard - distinguishing edges from shadows from object edges

This is the earlier image with just the edge information. It is difficult to recognize the person because shadow and object edges are confused.

Newly-sighted adults "see but don't see" – Object recognition is hard

"Having often forgot which was the Cat, and which the Dog, he was asham'd to ask; but catching the Cat (which he knew by feeling) he was observ'd to look at her steadfastly and then setting her down, said, So Puss! I shall know you another Time'." [Cheselden, 1728]

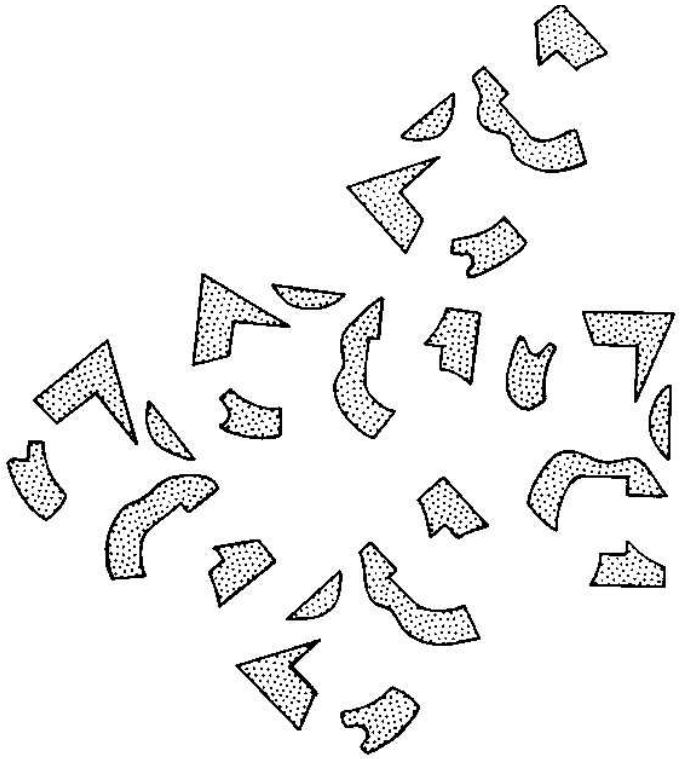
"When ... the experiment was made of giving her a silver pencil case and a large key to examine with her hands; she discriminated and knew each distinctly; but when they were placed on the table, side by side, through she distinguished each with her eye, yet she could not tell which was the pencil case and which was the key." [Wardrop 1827]

"Thus, for patient TG, telling a circle from a square, or either from a triangle was very difficult; he had to stare at the angles, one at a time, engaging in what we have called "scanning", to do it." [Valvo 1971]

The brain uses many top-down heuristics

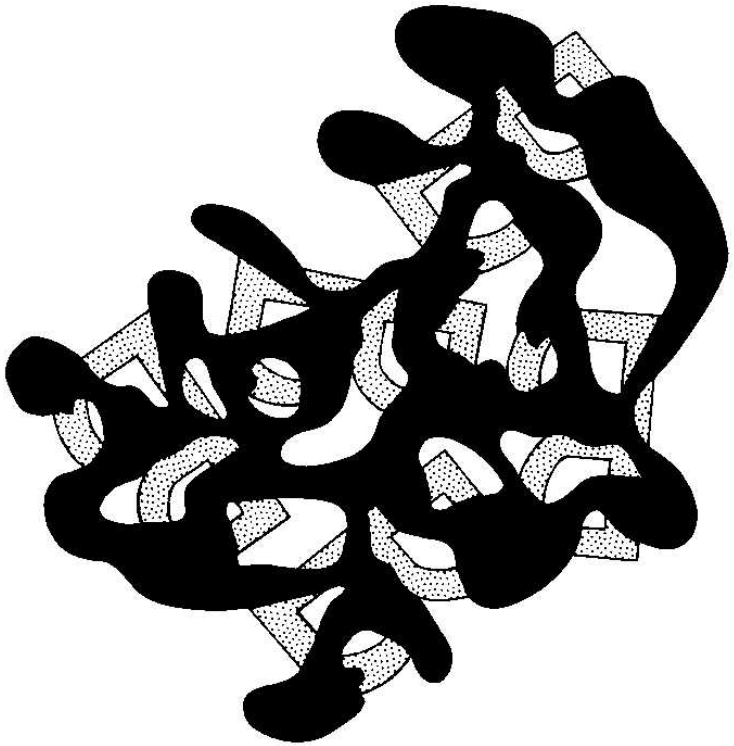
In addition to the Gestalt heursitics, we will consider the **occlusion heuristic**, **light from above heuristic**

Occlusion heuristic



(Bregman 1981)

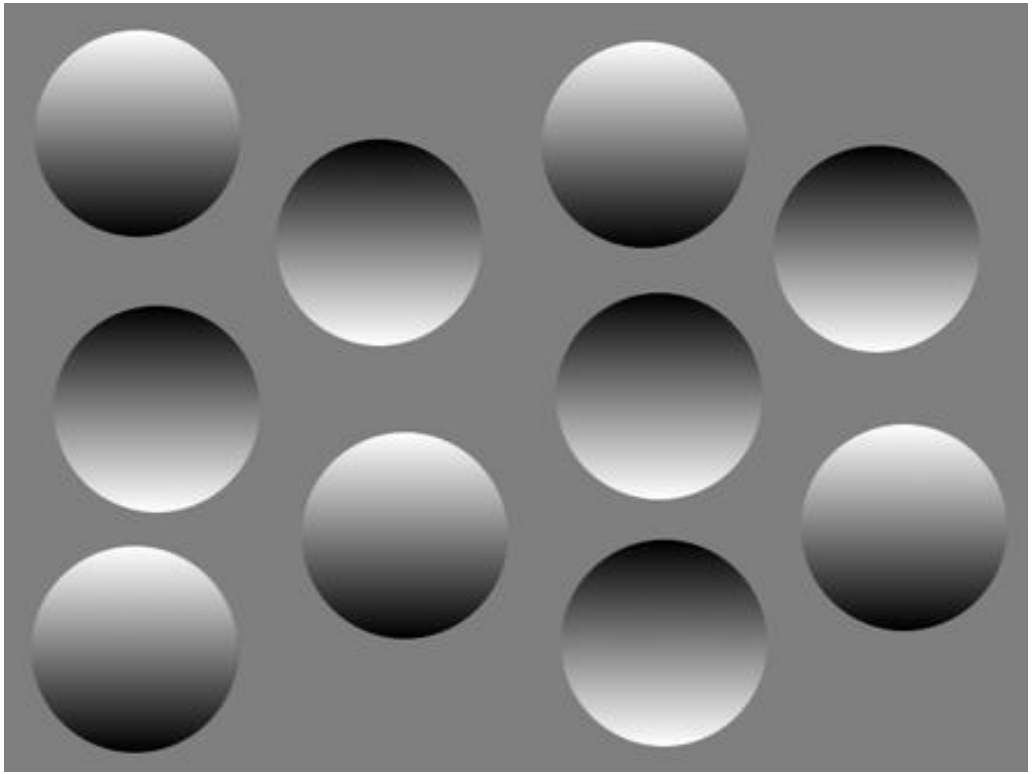
Occlusion heuristic



(Bregman 1981)

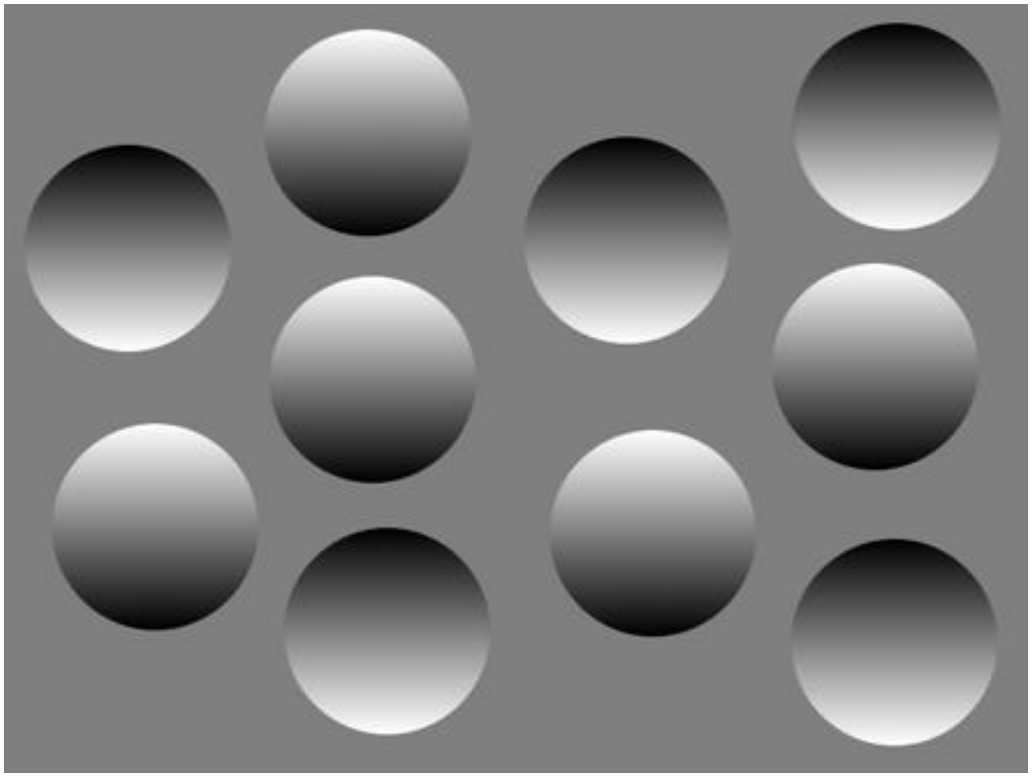
We see occluded objects as continuing between the occluder

Experience with lighting



http://www.graphics.cornell.edu/ jaf/projects/pn/space.html

Experience with lighting

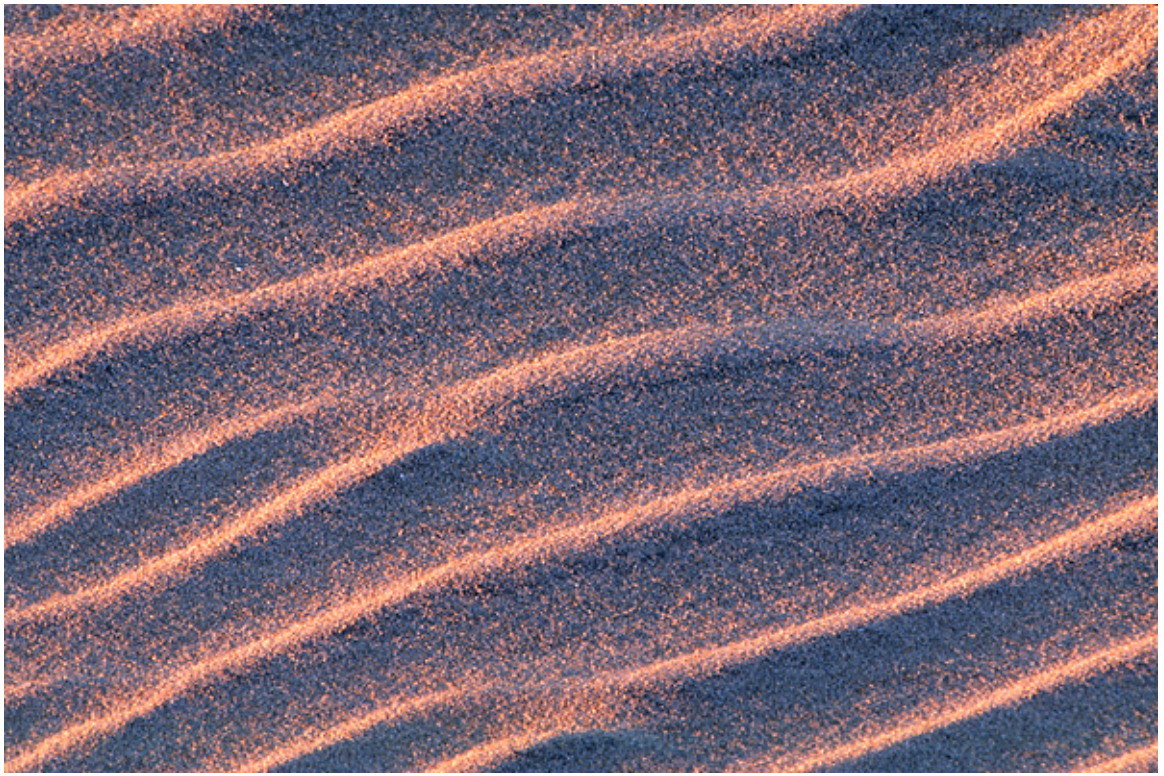


http://www.graphics.cornell.edu/ jaf/projects/pn/space.html Same image upside down (note the difference in perception of 3-D shape)

Experience with lighting

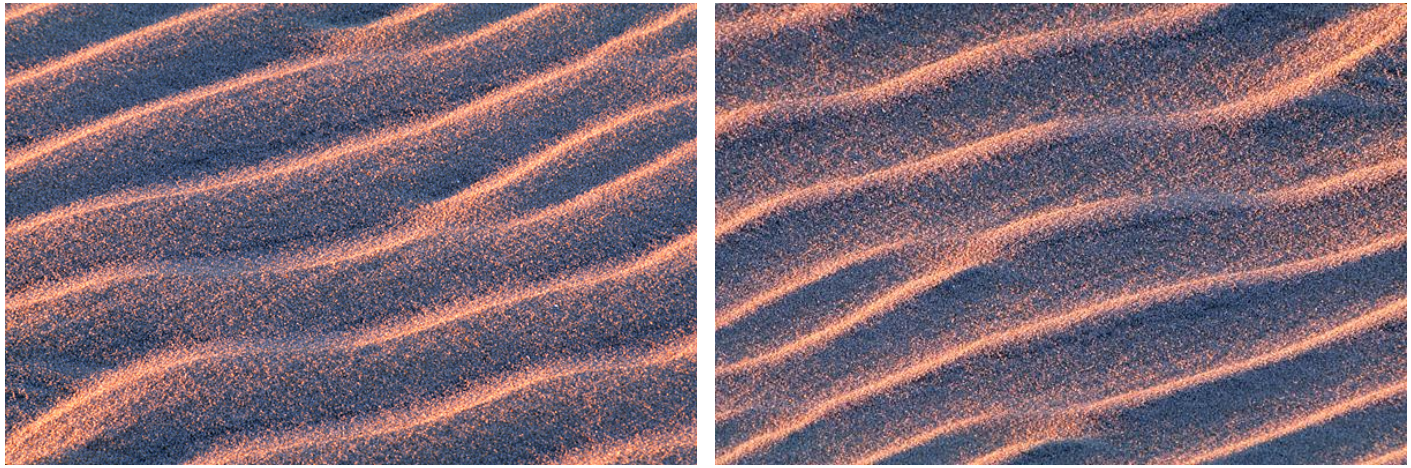
NC1097: Sand ridges at sunrise, Pea Island National Wildlife Refuge, the Outer Banks, NC

Experience with lighting



same picture upside down

Experience with lighting



Different perceptions because we assume lighting comes from above

Influence from learned experience with lighting

http://www.princeton.edu/ ftong/

Influence from learned experience with lighting

http://www.princeton.edu/ ftong/

Experience with lighting

http://www.graphics.cornell.edu/ jaf/projects/pn/space.html Who is this?

Experience with lighting

http://www.graphics.cornell.edu/ jaf/projects/pn/space.html Who is this? (same figure with contrast reversed)

Other evidence for top-down effects

Palmer (1975) experiment

- Show a scene and then briefly flash a picture of a related or unrelated item
- related items were correctly identified 80% of the time, unrelated 40% of the time

Treisman illusory conjunction experiment

- Show an orange triangle a blue ellipse and a black o
- normally you get some illusory conjunctions, but
- illusory conjunctions occur less when subjects were told "carrot,lake and tire"

Seeing with Sound - Using your auditory system to "see" objects

vOICe page

Your text mentions Bach-Y-Rita's Optacon vision to tactile transduction. There is also a vision to auditory system that you can try out.

Rather than train a computer to do visual recognition (which has proved extremely hard), these projects are working on using other modalities of the human brain to train them to recognize visual signals.

This is similar in spirit to the palm pilot script idea. It's too hard to get character recognition working, lets get humans to learn how to make it easy for us. The human brain is the best pattern recognition device.

Next Class

Chapter 9 (Perception and Action) Review for Midterm 2 – Bring any questions you have