Overview of Vision

Why is vision cool?

Select figures taken from Principles of Neural Science, 4th edition, eds: Kandel, Schwartz & Jessell. McGraw-Hill, (2000). Chapter 25.

The Visual System

We are so familiar with seeing, that it takes a leap of imagination to realize that there are problems to be solved. But consider it. We are given tiny distorted upside-down images in the eyes, and we see separate solid objects in surrounding space. From the patterns of stimulation on the retina, we perceive the world of objects and this is nothing short of a miracle.

- Richard L. Gregory, *Eye and Brain*, 1966

Questions in vision

- How do we see form? Movement? Color? Faces?
- Computers
  - Pattern recognition
  - Neural networks
  - Multiple cameras
- Vs Humans
  - 2 eyes
  - Brain

Visual Perception

- Perception is more than simple image processing
  - 3d perception from 2d images
  - Same object under different conditions/different 2d images
    - Friend walks towards you
      - Size constancy
    - Dim lights
      - Color constancy
- Perception is taking retinal images and creating a coherent stable 3-dimensional world

Gestalt Psychology

- Perception is holistic
  - The whole is greater than the sum of the parts
  - Melody: hear more than the sequence of notes
  - Doesn't matter what key it is in
  - Gestalt is a stable pattern that is constant despite variations in input
- Wertheimer, Koffka, Kohler
- Visual system processes information according to computational rules
- Makes assumptions based on experience and hard-wiring

Pattern Recognition

AMBIGUOUS

Pattern is automatically imposed but flips

Pattern Recognition

SIMILARITY

By color

(also works by size, orientation, motion, etc)

Pattern Recognition

PROXIMITY

Figure-Ground Recognition

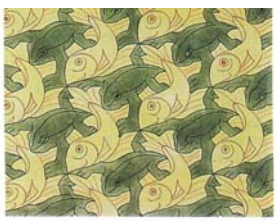
Figure-Ground Recognition

- Some parts of a scene form an object: the figure
- The remainder forms the background
- Dynamic process
  - Can flip: i.e. reversals

Figure-Ground Recognition

Face/Vase

Figure-Ground Recognition



Fish/Frog

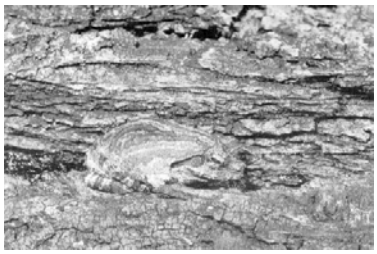
Figure-Ground Recognition

Old Lady/Young Woman

Figure-Ground Recognition

Texture Boundaries

Figure-Ground Recognition



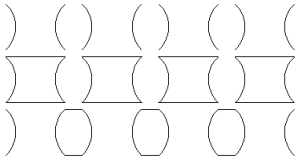
Texture Boundaries

Figure-Ground Recognition

For each aperture, what is the figure?

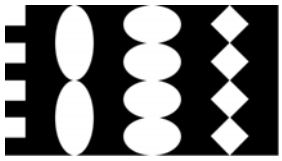
Area - Rubin Disks
Smaller area is usually figure
Larger area is usually ground

Figure-Ground Recognition



Closure
Closed shapes are seen as figure

Figure-Ground Recognition



Symmetry
Symmetric contours are grouped as figures

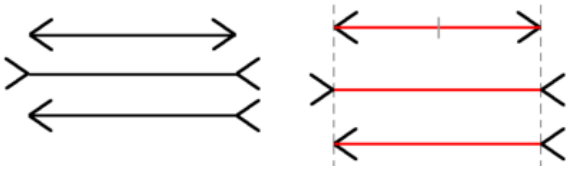
Contours

- Edges of objects
- Can recognize objects by simple line drawings
  - Lack color
  - Lack texture
  - Lack 3d shape
    - Depth cues

Illusions

- The visual system makes assumptions
- These assumptions can be exploited
  - Become illusions
- Automatic
- Knowledge does not prevent

Muller-Lyer Illusion



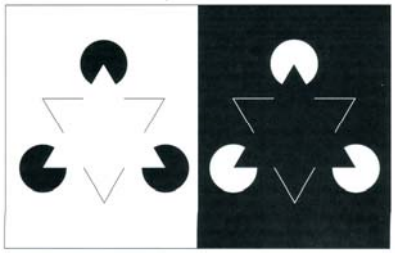
Which horizontal line is longest? Shortest?

They are the same size.

Perceived Size – shape indicates size

Muller-Lyer Illusion

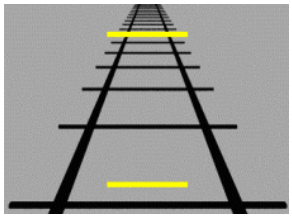
Kanisza Triangle



Filling-In

Illusory Contours/Brightness Illusion

Ponzo Illusion



Which horizontal line is longer?

Object size is based on background

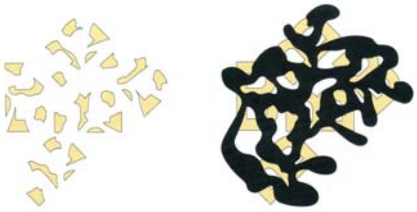
Perspective vs Framing Hypotheses

Perceived Size

Which woman is bigger?

Size Constancy – 2 objects of same retinal size at different distances must be different in real size

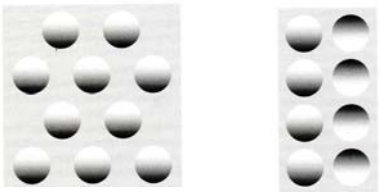
Occlusion



Filling-In

Assume contours continue under occluders

Physics in the Real World



Are the spheres concave or convex?

Assume light comes from above – e.g. sun

Genetic or experiential

Naïve Physics

What direction will the water flow?

Circular: assume contour continues
Rightward angle: assume inertia

Conclusions

- Perception is an active process that:
  - Takes the image on the retinas
  - Applies rules
    - Hard-wired
    - Developmental
    - Learned
  - To create an internal representation
- From these rules we can start to understand how the visual system works

Visual System – Multiple Cortical Areas

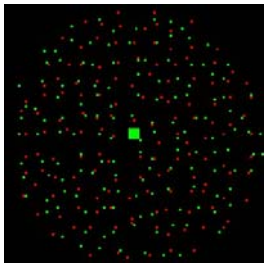
- Different areas process different features of an object
  - Motion
  - Form
  - Color
- However, we are conscious of a single unitary percept

Hierarchical processing

- Some areas build on previous stages of processing

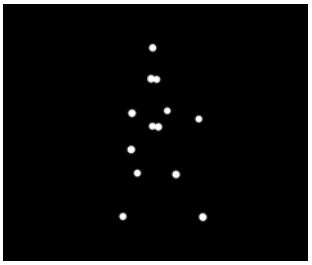
Hierarchical processing

- Some areas build on previous stages of processing



Hierarchical processing

- Some areas build on previous stages of processing



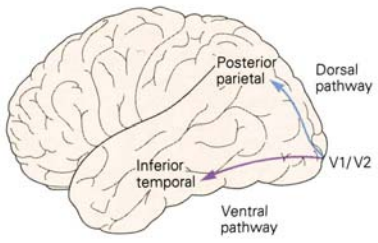
Retinotopic Maps

- Organized two-dimensional map of the visual field
- Representation of the retinal image, keeping the spatial structure
- Repeated in many visual areas
- Maps relative spatial locations
- Allows for processing in a stream

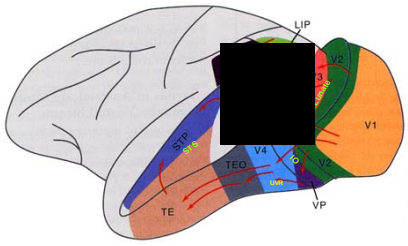
Visual Processing

- Is everything connected in one stream or are there parallel systems?
- Anatomically
  - 2 major pathways in the retina
  - Continues through LGN, V1, extrastriate areas

Two Visual Streams



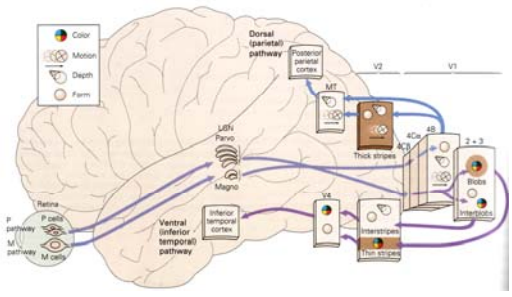
Two Visual Streams



Two Visual Streams

- Dorsal
  - V1 -> V2 -> MT -> MST -> PPC
  - Sensitive to
    - Motion
    - Depth
    - Spatial locations
  - Poor at
    - Color
    - Stationary objects
- Ventral
  - V1 -> V2 -> V4 -> TE -> TEO -> IT
  - High resolution
  - Sensitive to
    - Edges
    - Orientation
    - Contours
    - Color
    - Object recognition

Visual Pathways in detail

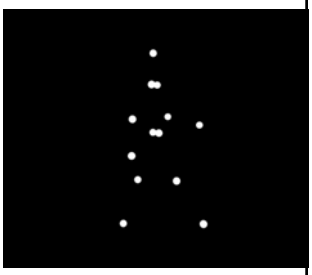


Two Visual Streams

- Dorsal vs Ventral
- What vs Where
  - Ungerleider & Mishkin
- Perception vs Action
  - Milner & Goodale

Two parallel streams?

- Are they truly distinct, parallel systems?
- Which stream is responsible for this?



The Binding Problem

- How are neurons that are selective for different features in different areas bound into an object representation?
  - How are black, smooth, cylindrical bound into a hockey puck?
  - How about black, rough, flat bound into a road?
  - How about red & white, smooth & edged, cylindrical bound into a can of Coca-cola?

The Binding Problem

- Not likely to have neurons that represent every combination – too many would be needed
- Alternatively: the information from different cell populations in different areas must be bound: dynamically associated together
- We do not yet know what the binding mechanism is

Attention

- Treisman & Julesz, independently, showed that binding needs attention
- Preattentive process
  - Rapidly scans features to distinguish figure/ground
  - Parallel processing of elementary features: color, orientation, size, direction of movement, etc

Preattentive Process

- Simple variation in a property is easily distinguished

Preattentive Process

- Complex differences are not detected

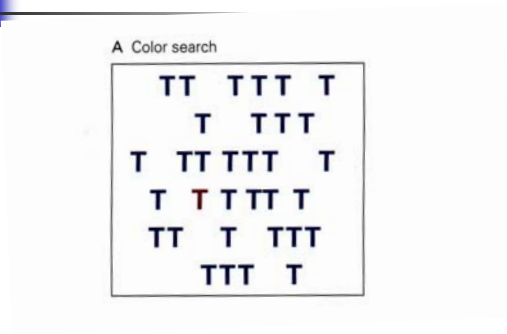
Preattentive Process

- Bottom-up processing
  - Focused on properties of individual elements
  - Does not require cognitive control

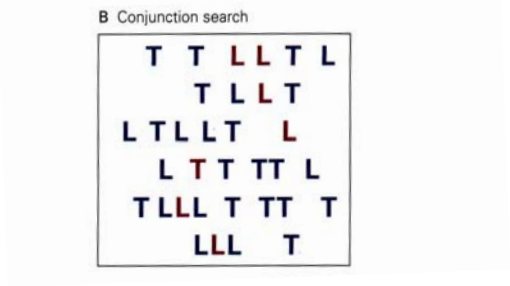
Attentive Process

- Selects the segregated features of an object
- Works serially: spatial attention
- Top-down processing:
  - Selected item is identified independently of the individual elements in the visual field

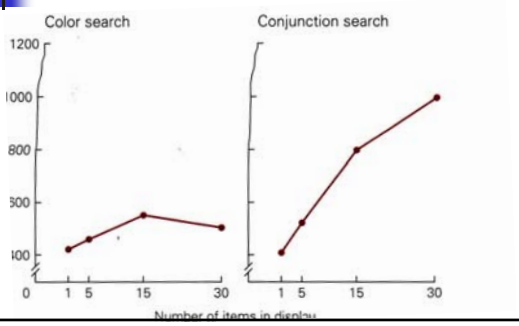
Visual Search: Preattentive vs Attentive



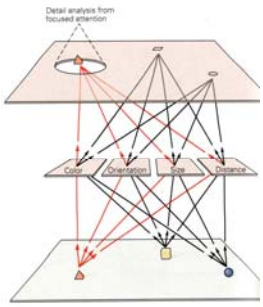
Visual Search: Preattentive vs Attentive



Visual Search: Preattentive vs Attentive



Feature Integration Theory - Treisman



Saliency Map

Feature Maps

Visual Field

Spotlight of Attention

- Not just for objects, but also spatial location

Why do we need attention?

- Limited capacity

Conclusions

- Two pathways help in understanding visual processing, but it is more complicated than that suggests
- Different features are processed in different areas, but they need to be bound
- We do not yet know the binding mechanism
- The visual system has a limited capacity
- Attention mechanisms have developed to overcome this bottleneck