



Cognitive Psychology

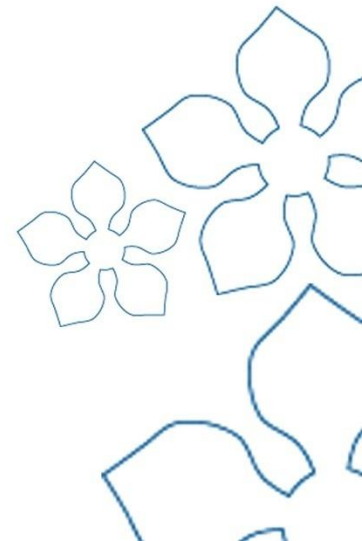
M5X00908 Brain, Behaviour, and Cognition (BBC)

Lecture 3: Perception

Dr Kate Williams



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David



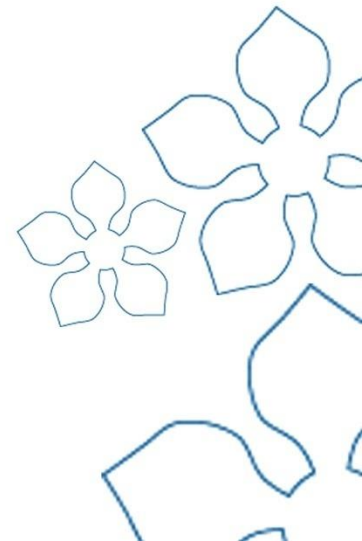


Essential Reading

- Eysenck, M. W., & Keane, M. T.
Cognitive psychology: Chapters 2, 3, 4



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

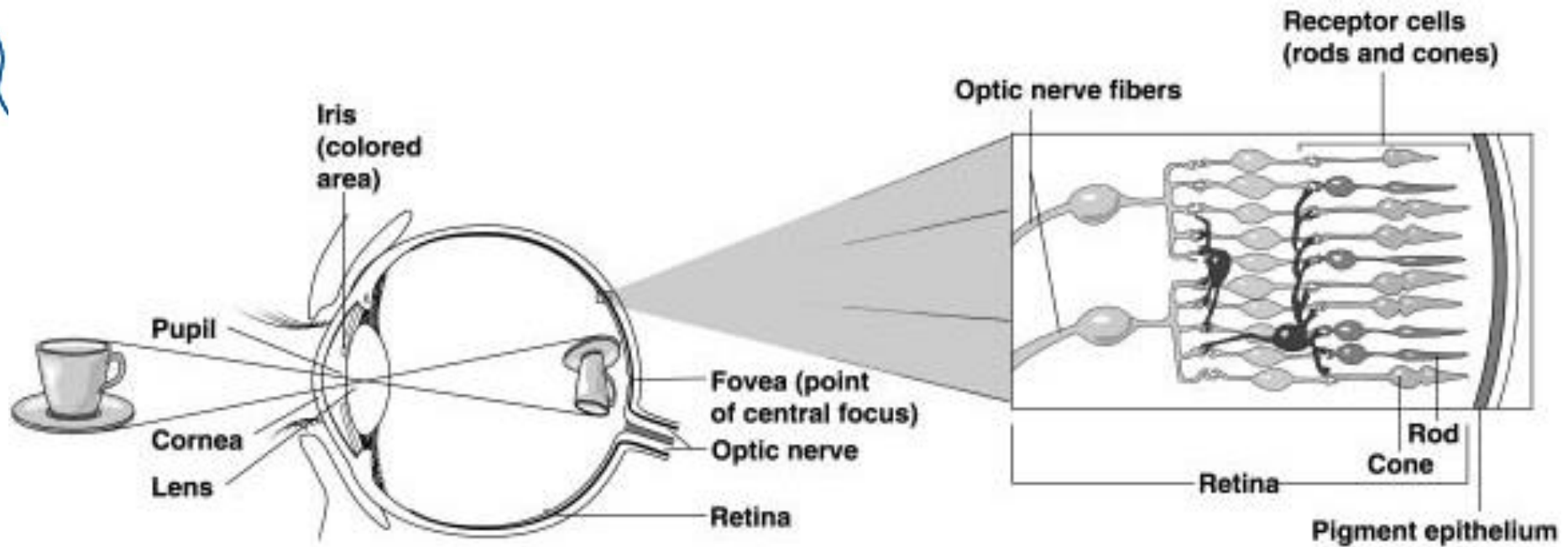


What is Perception?

“Acquisition & processing of sensory information in order to see objects in the world...” Sekuler & Blake (2002), p. 621.

- Information received through senses
- “Raw” material for perception
- Data driven processing
- Attend to selective aspects
- Perception:
 - Connects selected data to higher cognitive processes (thinking)
 - Ultimately, what we ‘perceive’ as reality

Visual Perception: From Eye to Cortex



PRIFYSGOL CYMRU

Y Drindod Dewi Sant

UNIVERSITY OF WALES

Trinity Saint David



Problems of visual perception

- Information received by sense organs is relatively impoverished:
 - Stimulation of the retina is incomplete and uneven
 - The visual scene is two-dimensional
 - Doesn't incorporate depth
 - Depth is required for distance judgements
 - The onward visual signal to the brain is not uniform in relation to the visual scene:
 - The large gap (blind spot)
 - An uneven distribution of information about colour (cones) and luminance (rods).
 - How do signals sent to brain turn in to perceptual experience.
 - Sensory inputs are somehow converted into perceptions of the objects around us.
- 

Overcoming these problems

- Top-down theories (concept driven) – the mind **interprets** the signals it receives in a rich, elaborative way (Gregory, 1970)
- Ecological theories – the visual system has evolved to derive what it needs **from the environment** (bottom-up/ data driven) (Gibson, 1966)
- Computational theories – deriving complex structure from basic input requires both processes (Marr)



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

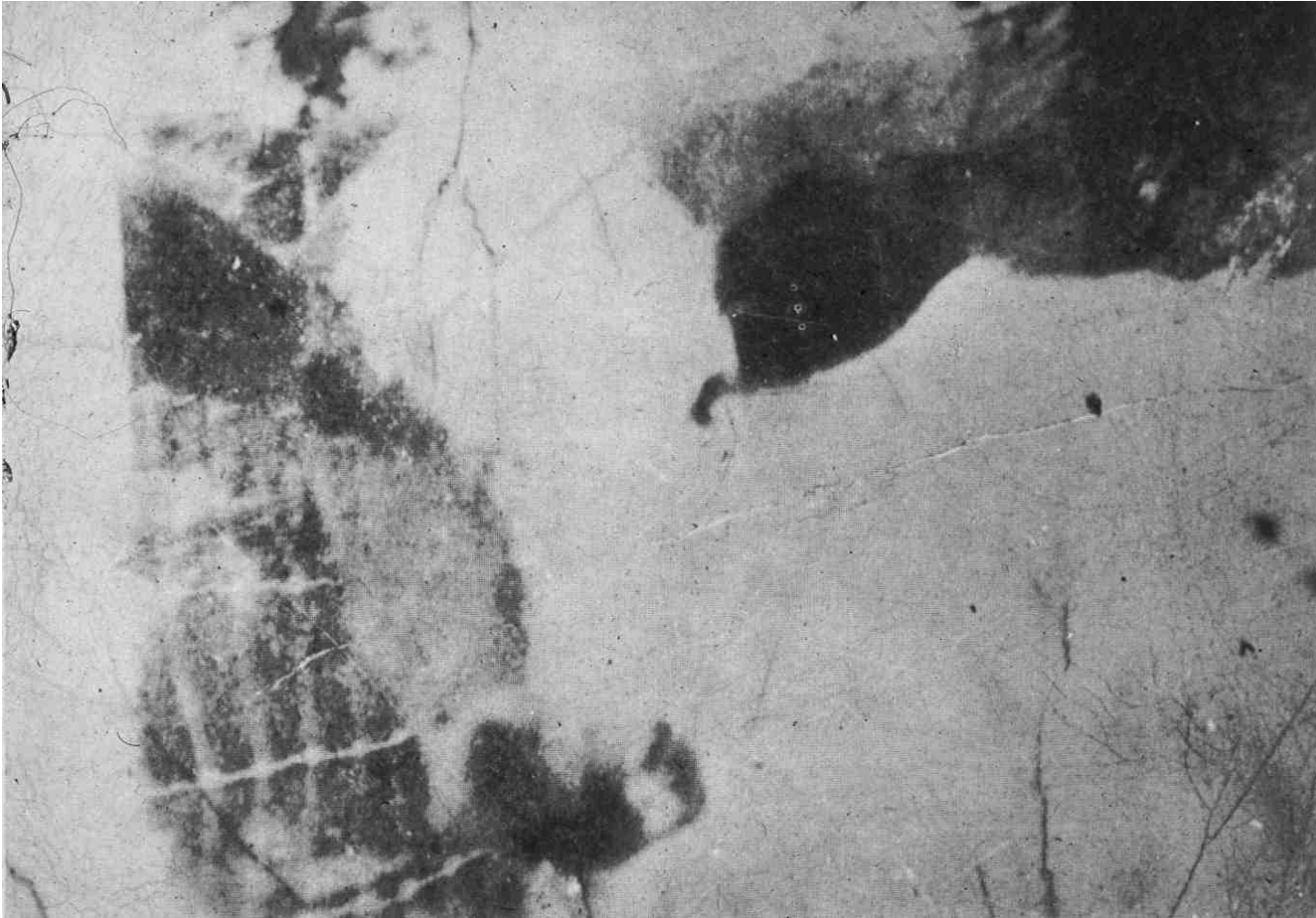
Top-down (Indirect) perception

- Sensory input is chaotic, unstable, and distorted
- Perceiver generates predictions (hypotheses) about the sensory information
- Perception is therefore constructed
 - **High level concepts** (knowledge of the world) or representations shape **low level percepts**.
- Visual illusions as support – ‘errors’ show that our hypotheses can be inaccurate (Gregory, 1963).
 - For example?



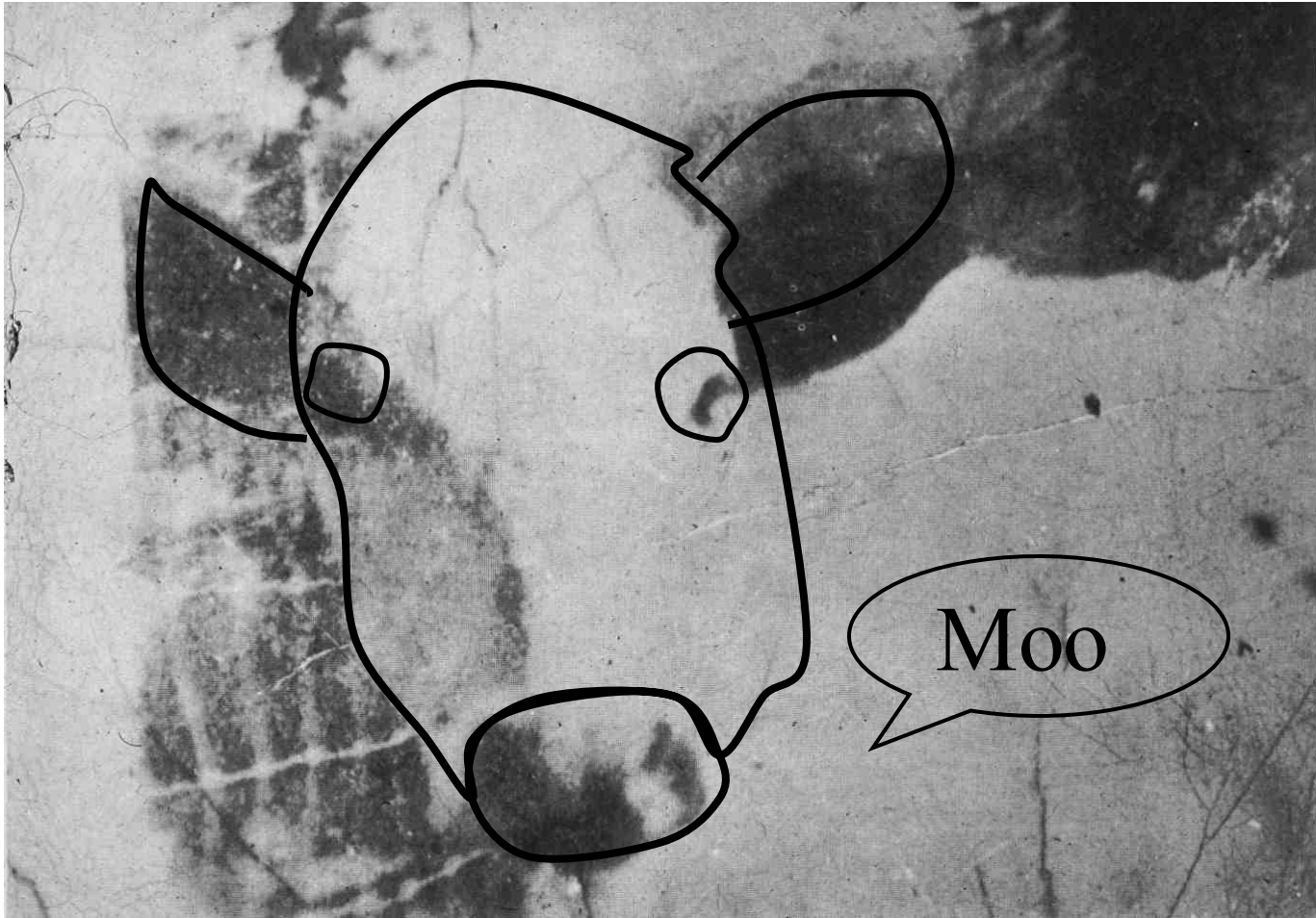
PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

What can you see?

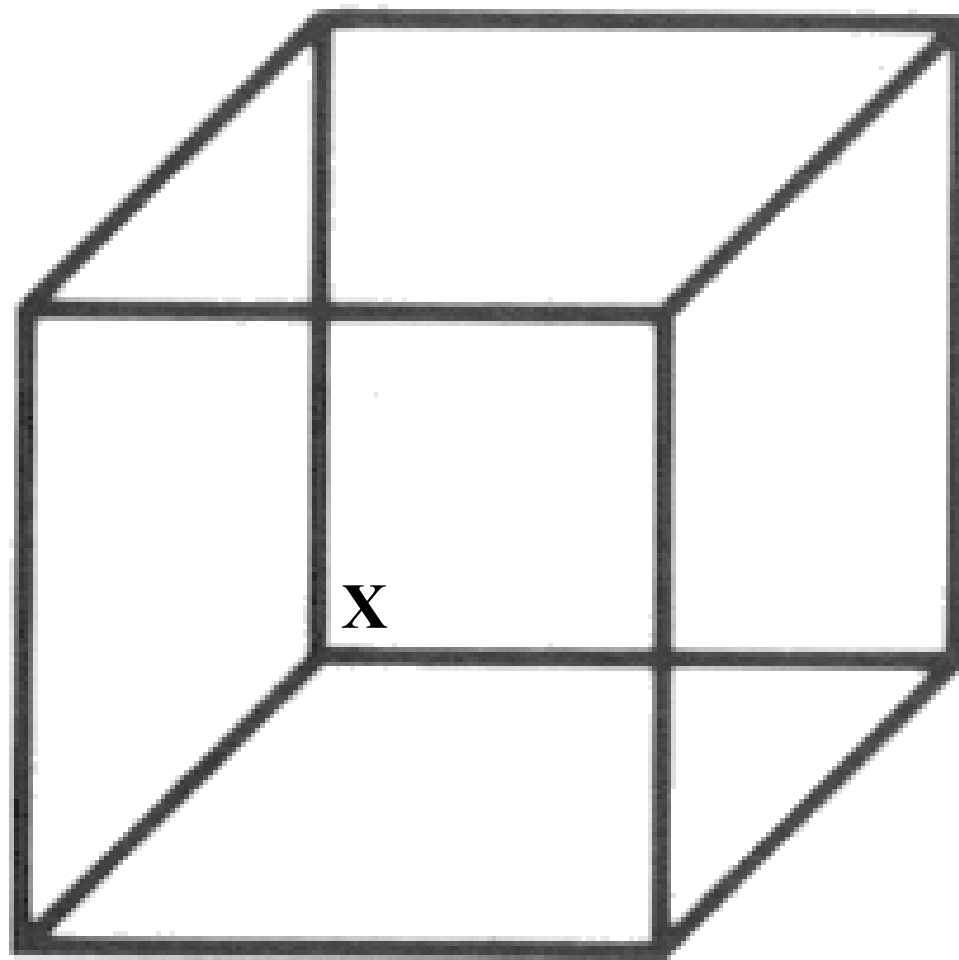


PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

What can you see? Dallenbach (1951)



- Fact that some saw a cow shows how good we are at seeing familiar things in series of blobs
- Fact that some could not see a cow shows this process is not perfect - some see roads, birds etc – they are all hypotheses which in this case were not the intention of the creator of the image



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

- Necker cube shows that we can derive two mutually exclusive percepts from a single visual input.
 - I.e., one sensation but two possible objects.
- Gregory argues Bottom-Up theories cannot explain this.
- Bottom-Up theories suggest we should always perceive the same thing from the same input.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Bottom-up (Ecological approach)

- According to Gibson (1950), Gregory's view of bottom-up processing is incorrect.
- Our visual system has **evolved** to detect what is important to us to survive/operate well.
- There is no need for inferences - all we need is present **in the visual array**.
- Ecological theory
 - Perception has evolved that support survival of species
 - Involves interpreting information as you move
 - You cannot therefore separate perception from movement and context

Bottom-up (Ecological approach)

- Gibson explained perception in terms of how it relates to properties of the world around us:
 - Perception of **movement and change** within our visual field provide crucial information about the world and our place within it (all from an egocentric perspective).
 - Distance
 - Speed
 - Location
 - **Surfaces** are important in deriving this information

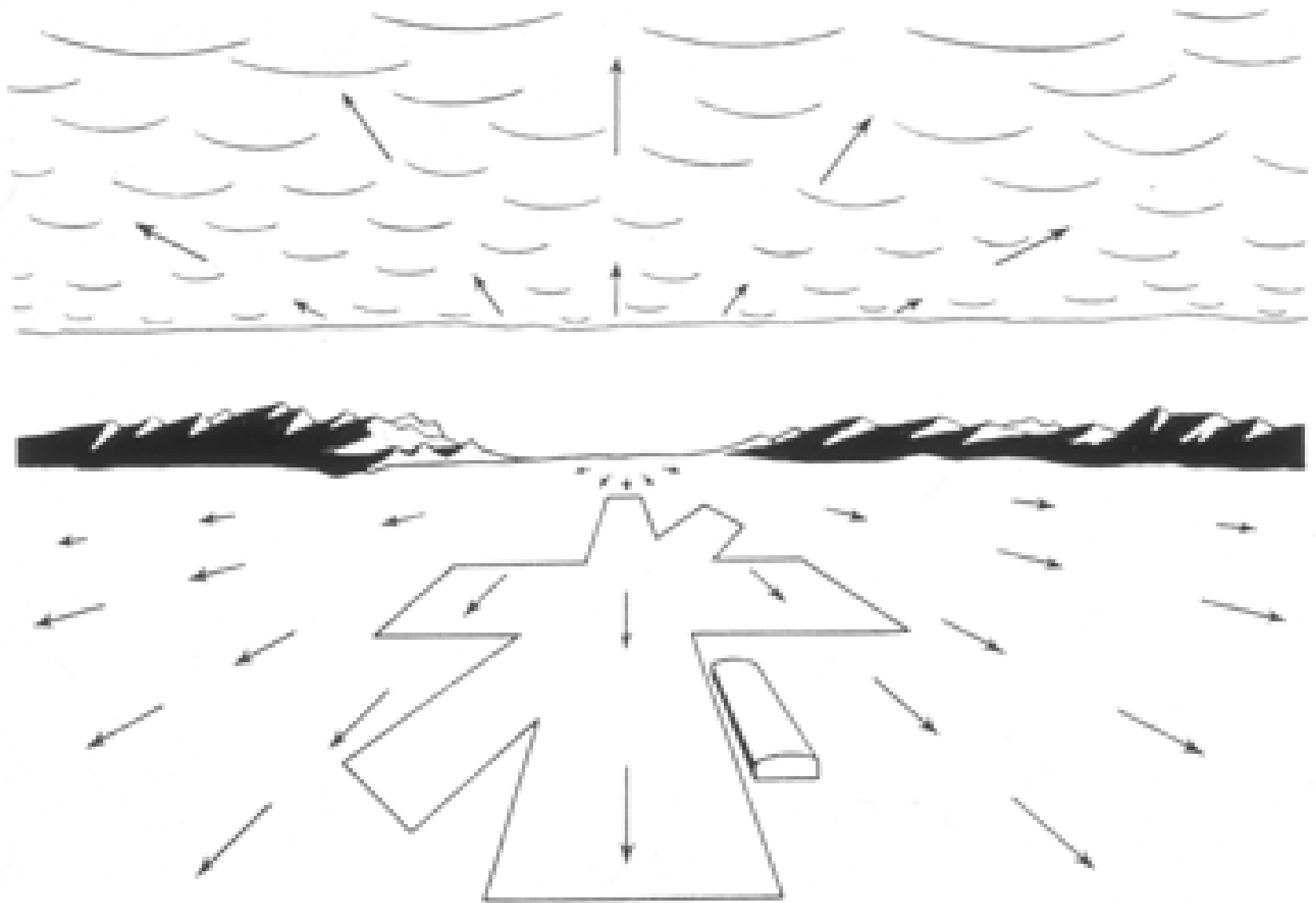


PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Movement

- Patterns of light reaching the retina consist of information about texture and gradients of objects.
- ‘Optic flow’ is one of our main sources of visual information (Gibson).
- When we move, our visual environment **seems** to flow past at different rates (apparent movement)
 - The closer a thing is to us, the faster it seems to move past
 - e.g., when on a train vs. viewing a train.
- Shifts in flow indicate direction changes
 - Inflow and outflow

Optic flow

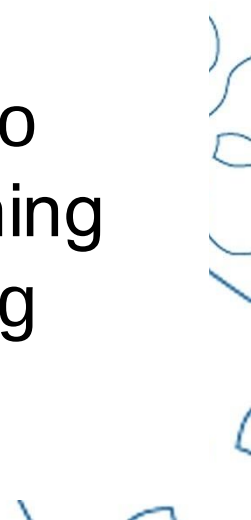


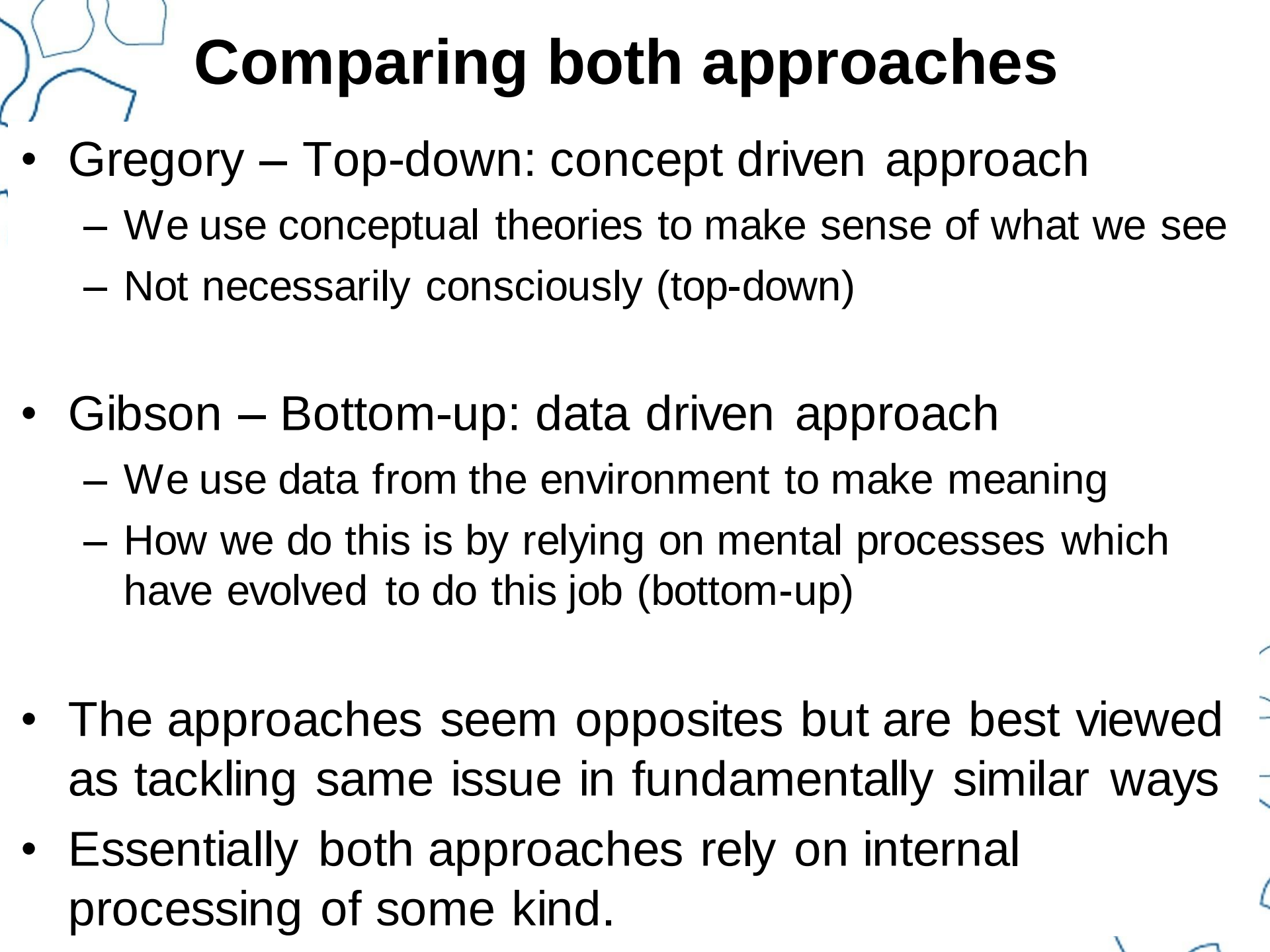
Optic flow

- Diagram shows outflow from plane landing strip
 - When moving towards an object the scene flows out from the point of reference
- Opposite is inflow where we move away from a point (move backwards)
 - The arrows flow the opposite way
- In real life, scenes are not as sparse as this – flow is achieved by objects moving forwards or backwards in our line of vision relative to our movement
- This is environmental info (data) not hypotheses.
 - They provide invaluable data to us about direction, speed, and distance of objects from us from an egocentric perspective



Bottom-up (Ecological approach)

- Optic flow demonstrates how rich environmental information is
 - Experiments involving illusions tend to ignore the complexity of the natural world – ‘air theory’
 - With such a rich array of environmental data we do not need to make hypotheses about what we are ‘perceiving’
 - Instead our brains have evolved mechanisms to pick up on this rich data and compute it’s meaning without us needing to guess what we are seeing
- 



Comparing both approaches

- Gregory – Top-down: concept driven approach
 - We use conceptual theories to make sense of what we see
 - Not necessarily consciously (top-down)
- Gibson – Bottom-up: data driven approach
 - We use data from the environment to make meaning
 - How we do this is by relying on mental processes which have evolved to do this job (bottom-up)
- The approaches seem opposites but are best viewed as tackling same issue in fundamentally similar ways
- Essentially both approaches rely on internal processing of some kind.

Marr: Computational Approach

- Four key aspects to Marr's theory:
 - Multidisciplinary - Links neurophysiology with psychology and computational AI
 - Based on an information processing view
 - Describes perception in terms of a series of computations on sensory data
 - Suggests high level visual experience is the product of these computations



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Marr: Computational Approach

Perception operates at different levels of description:

Computational level – function/purpose

Algorithmic level- detailed processes

Hardware level- what makes up the object (e.g. Brain)

He wanted to relate the function (purpose) of perception to the process of perceiving

To do this he considered both top down and bottom up approaches, evolutionary (innate) approach, as well as neurophysiological evidence



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Marr's Stages of Visual Processing

- Process of perception involves producing series of increasingly detailed representations of the environment (layers)
- Four distinct stages of processing:
 - Raw Primal Sketch
 - Complete Primal Sketch
 - $2\frac{1}{2}$ D Sketch
 - Full 3D Representation

Raw Primal Sketch

- 1st stage involves extracting information about edges of objects and basic shades of light and dark.
- Evidence from neurophysiology
 - Single cell recording of nerve cells show different orientations of line and shades of light ‘fire’ different cells in the visual cortex of animals and humans
 - (Hubel and Weisel)



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

How is this information organised to resemble the object?



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

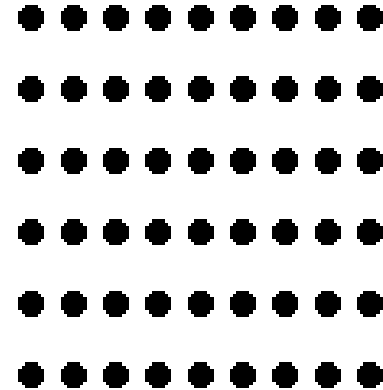
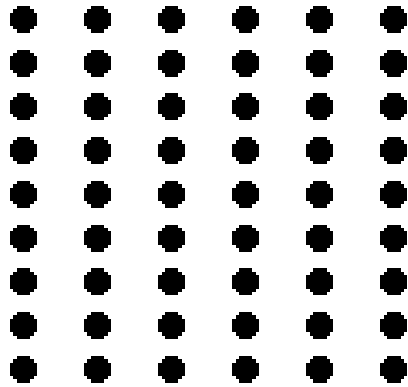
Marr: Complete Primal Sketch

- **Grouping** of surfaces and common areas.
- From **Gestalt** Psychology (1920s)
- Described different ways in which we can group objects (rules in the mind):
 - Proximity
 - Similarity
 - Continuation
 - Closure



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Proximity

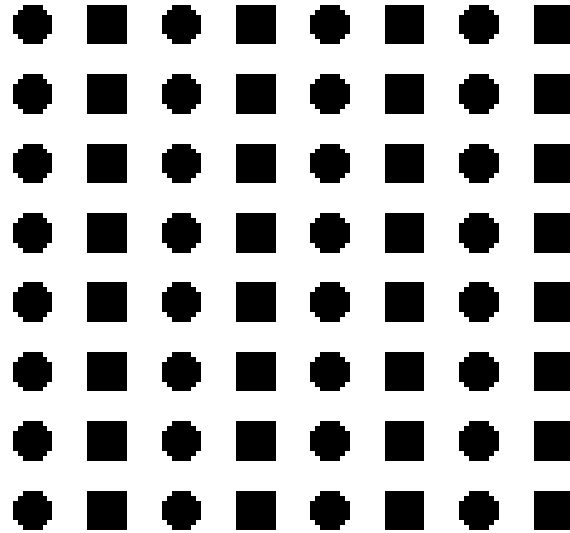


Things that are close together *seem* part of the same group – but they are just a series of dots.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Similarity

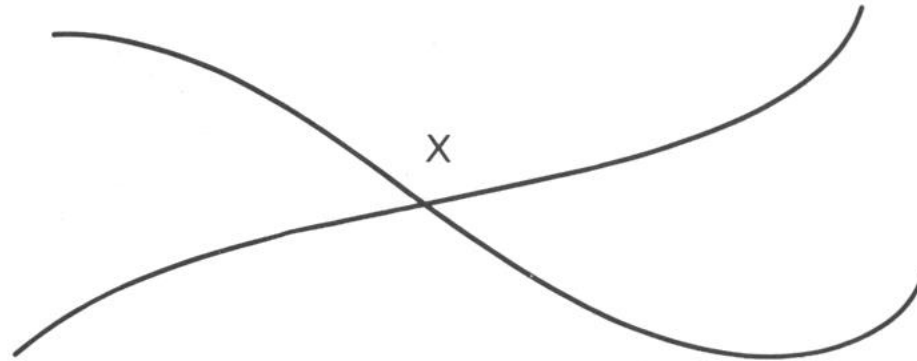


Things that are similar *seem* part of the same group.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Continuation

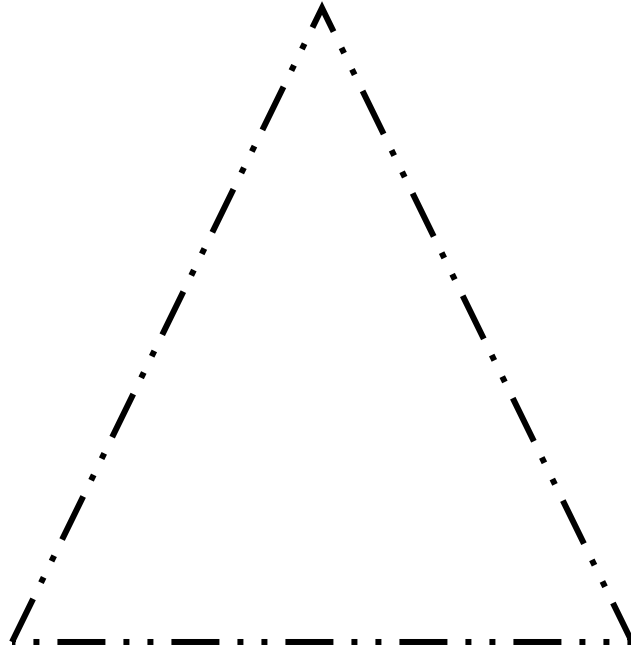


We see this as two crossing lines, rather than two ‘V’ shapes.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Closure



Images *seem* to form closed rather than open patterns.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Marr: 2½D Sketch

- All of the previous grouping principles have been used to organise the raw primal sketch
- Next we need some spatial information - the 2½D Sketch
- **2½D** to emphasise that this stage does NOT produce a full 3D representation.
- It produces an **estimate of the spatial locations** of objects (or parts) in relation to the viewer.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

2¹/₂D Sketch: Depth Cues

- Binocular disparity:
 - Each eye receives a slightly different image of the world
 - Disparity used to compute relative distance
- As well as monocular cues:
 - Texture gradients
 - Occlusion
 - Convergence
 - Relative Size
- Monocular cues – are basically Gibson's ecological approach – mind evolved to compute these

Depth: Textural Gradient



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Surfaces appear to have a finer texture as they recede into the distance.

Depth: Convergence / Perspective

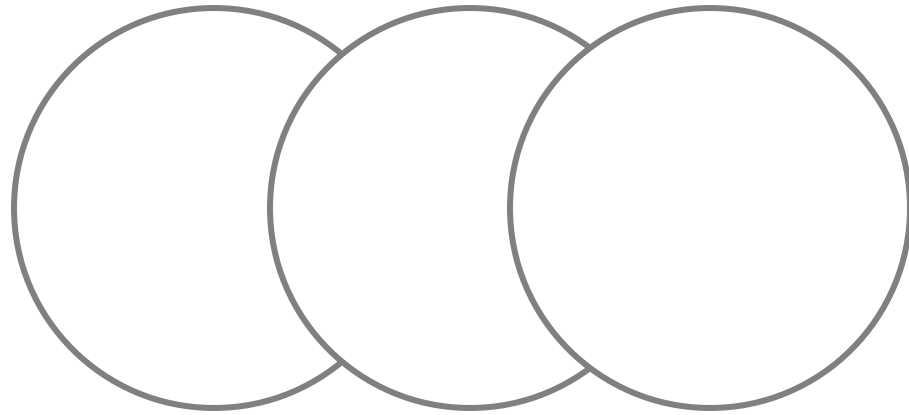


Parallel lines
appear to draw
closer together
as they go
further into the
distance.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Depth: Occlusion/ superimposition

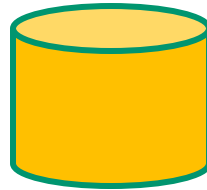


Proximal objects appear to cover over distal objects.

- These evolved abilities add texture and depth to the object
- But the 2.5 sketch does not contain information about the objects which are occluded - Are they half moons or circles?
- This is because the 2.5 D sketch is observer-centred.

Marr: 3D Model Representation

- The final stage of Marr's theory.
- A full 3D description of our spatial environment involving the identification of the volumetric structure of objects and materials in our visual scene.



- This allows us to work out the 3D environment from a non-egocentric point-of-view – a top-down process



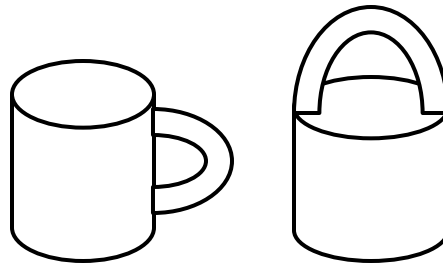
Biederman: Recognition-by-components

Developed and extended Marr's approach.

Assumes objects consist of basic geometric components known as 'geons' (geometric ions) – e.g., blocks, cylinders.

36 different types of geons make up all objects –

- e.g., an arc connected to the side of a cylinder represents a cup, where as those same 2 geons but with the arc connected to the top of the cylinder represents a bucket.



PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Biederman: Recognition-by-components

Once components of geons and their relationships are established, the image can be matched against stored object representations.

We then recognize an object based on whichever stored representation fits best with the information from the visual version of the object.

In the RBC model edge extraction is the first step in object recognition.



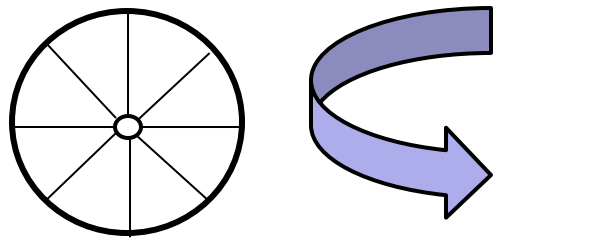
PRIFYSGOL CYMRU
Y Drindod Dewi Sant
UNIVERSITY OF WALES
Trinity Saint David

Viewpoint invariance vs. viewpoint dependent

Some edges are considered invariant – meaning an object can be recognized as well from many different viewing angles.

In the RBC model this relies on non-accidental properties

But: Imagine a bicycle wheel....what does it look like?



- Biedermans theory applies to easy categorical discriminations whereas object recognition is typically viewer-centered.
- The theory is also quite inflexible and tends to focus on perception of shape when we know other factors (e.g., colour and texture) of objects help us to identify them.

Summary

- Sensation is incomplete information
- How do we make sense of incomplete information?
 - Hypothesise the nature of the stimulus based on prior knowledge: top down concepts (Gregory)
 - Evolved to make use of basic elements from the environment (Gibson)
 - Use information and knowledge to compute and understand a more complex visual scene
 - Marr's theory represents the first detailed computational theory of perception
 - Biederman proposes a different version/extension of Marr's approach