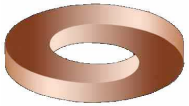
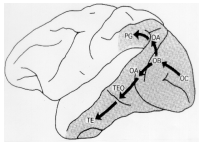


The Neuroscience of Vision III

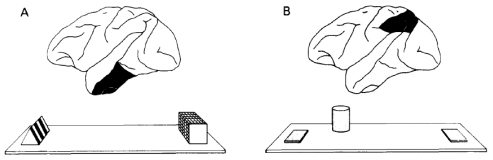
Putting the Pieces Together

- Drawing upon the apparent differences in processing in the temporal lobe (object identification) and parietal lobe (spatial processing), Ungerleider and Mishkin proposed that there are two hierarchically organized visual pathways
  - A ventral what system projecting through V4 to the temporal lobe
  - A dorsal where system projecting through MT to the parietal lobe



Different Pathways

- In support of the what/where differentiation of the two pathways, Miskin et al. focus on the types of tasks subjects with lesions in each pathway cannot perform
  - With temporal lobe damage, subject cannot recognize the object that matches or differs from a previous experienced sample
  - With parietal lobe damage, subject cannot pick object by is relation to a landmark



Clicker Question

How does Milner and Goodale's characterization of the two visual pathways differ from that of Mishkin et al.?

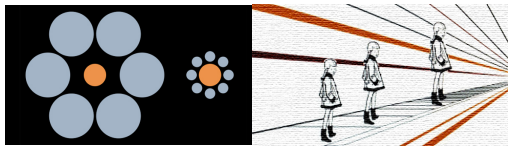
- They viewed the ventral system not as involved in recognizing objects but in identifying how valuable they are to the viewer
- They claimed that one pathway processed both object and location information while the other was actually involved in relating visual to audition
- They argued that one pathway processed familiar stimuli while the other processed novel ones
- They viewed the dorsal stream not as involved in spatial representation but in enabling visual control of actions

Recharacterizing the Two Pathways

- Instead of distinguishing what and where, Milner and Goodale proposed a distinction in terms of vision for perception and vision for action
  - Both pathways process what and where information
- Milner and Goodale identify the difference between the streams not in terms of inputs but the outputs they serve
  - Ventral: "enduring characteristics of objects and their spatial relations"
  - Dorsal: "mediate the visual control of skilled actions"
    - Crucial to implementation of action

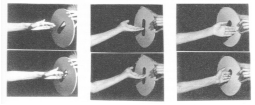
Distinguishing two Ways of Seeing

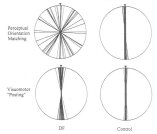
- When asked to report which yellow circle is larger, people experience the Ebbinghaus and Ponzo illusions
- But not in reaching behavior such as preparing to grip them
  - Yet effect of the illusion is found if response is delayed, requiring reliance on memory

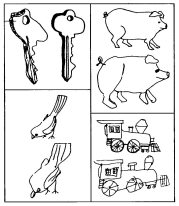


Doubled Dissociation

Optic ataxia: With damage to the parietal pathway, patients are unable to put hand through slot in correct orientation







Visual agnosia: When damage is to temporal pathway, copying is slow and slavish and patients cannot name objects

Clicker Question

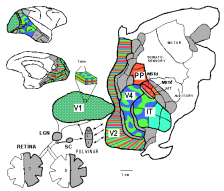
When van Essen and Gallant describe the visual system as hierarchical, to what are they referring?

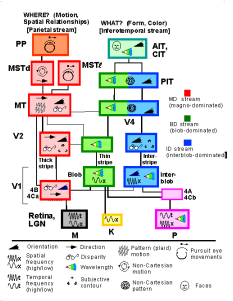
- That as you move higher in the brain areas have greater importance
- That processing proceeds though a sequence of processing areas in which areas lower in the hierarchy project to those higher
- That processing areas lower in the hierarchy are regulated by those higher in the hierarchy

Mechanism for Visual Processing

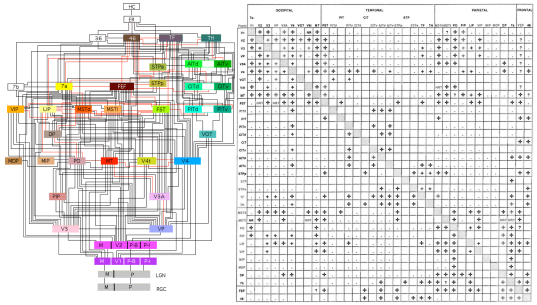
Van Essen: Schema of overall organization of visual processing

Represents the combined efforts of recording, lesion, stimulation



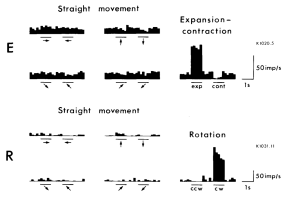


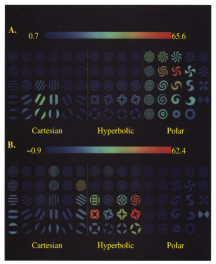
Visual System: A Hierarchical, Interconnected Network



Finding Only What You Look For

- Looking for cells that respond to non-Cartesian movements and shapes, van Essen and Gallant found them in MSTd and V4
- What else should one look for?



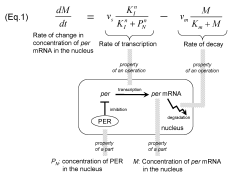


Making Sense of the Visual Mechanism

- The research strategies on which we have focused have primarily been directed at decomposing the mechanism of visual perception to identify its parts and the operations they perform
- But to understand how they carry out the task of seeing, researchers need to recompose the system
- One can recompose a mechanism in a diagram that traces the flow of activity
  - But a diagram is static and fails to reveal how the components will interact
  - Animation is helpful in showing how we think components interact, but their activity is not generated from the components
- Mathematical models can show what happens as the components each perform their operations

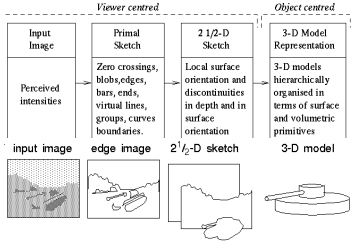
Mathematical Models

- Represent the components of the mechanism as variables
- Write difference or differential equations to specify how each component affects others
  - With parameters specifying the degree of effect
- Solve the equations analytically when possible
  - More commonly, simulate their solution by applying the equations iteratively on a computer



Marr's Algorithmic Account

- "Vision can be understood as an information processing task which converts a numerical image representation into a symbolic shape-oriented representation."

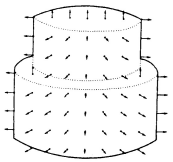


From Image to Primal Sketch

- Extract information regarding edges and intensity changes
  - Zero-crossings
  - Blobs
  - Edge segments
  - Boundaries

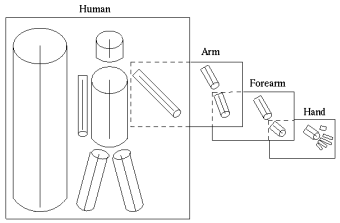
2½D Sketch

- Explicitly represent surfaces, their orientation and rough depth and the contours of discontinuities in a viewer—centered coordinate systems
  - Not a complete 3D representation, but estimate of location of objects relative to the viewer
- Infer depth from binocular disparity, texture gradients, occlusion, convergence, and relative sizes
- Typically, the 2½D sketch captures surfaces as we are aware of them



3D Representation of Objects

- Describes shapes and their spatial organization in an object-centered coordinate system
  - Modular hierarchical representation
- Corresponds more to our *understanding* than our *perceptual awareness*



Marr's Levels of Analysis

- Reacting to perceived stagnation in understanding the brain based on studying individual neurons or writing programs to describe how they work together, David Marr (1982) argued:
  - "there must exist an additional level of understanding at which the character of the information-processing tasks carried out during perception are analyzed and understood in a way that is independent of the particular mechanisms and structures that implement them in our heads.This was what was missing – the analysis of the problem as an information-processing task. Such analysis does not usurp an understanding at the other levels, of neurons or of computer programs – but is a necessary complement to them, since without it there can be no real understanding of the function of all those neurons."
- Note: these are not levels in a hierarchy of processing areas

Clicker Question

Which of the following was not one of the levels of analysis Marr claimed was required in a complete account of the visual system

- the computational level
- the representation and algorithm level
- the mathematical analysis level
- the implementational level

Marr's Levels

- Proposed that understanding had to proceed at three levels of analysis
  - **Computational theory**
    - What is the goal of the computation?
    - Why is it appropriate?
    - What is the logic of the strategy by which it can be carried out?
  - **Representation and algorithm**
    - How can this computational theory be implemented?
    - In particular, what is the representation for the input and output?
    - Why is the algorithm for the transformation?
  - **Hardware implementation**
    - How can the representation and algorithm be realized physically?

Discussion Question

Around 1900 divers found this object on a shipwreck off Antikythera (Greece) that dated to around 65 BCE. CAT scans revealed a large number of gears and subsequent reconstruction showed it to be an analog computer. How could you go about figuring out what it computed?

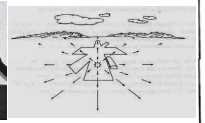
- Find the builder and ask her
- Reflect on the culture and consider what would have been important to compute
- Consider what problems in the world the object could have solved by the computer
- Look for other similar artifacts and use knowledge of what they do

The Antikytheran Computer

- CAT scanning allowed researchers to identify many of the gears of the mechanism and even to read the information on the dials on its outside
- Researchers concluded that it was an analog computer used to predict the positions of the sun and the planets in relation to the fixed stars, including eclipses
  - What was crucial to this evaluation was determining that the relations between the positions of the various dials corresponded to the relations between astronomical phenomena
- Researchers were implicitly operating at Marr's computational level—asking what relations in the environment the outputs of the mechanism corresponded to

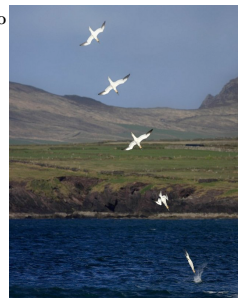
Marr vs. Gibson

- Marr contended that the process of vision was one of inference—inferring what one is seeing from minimal cues
- Gibson contended that there is rich information "in the light"
  - Specify affordances: possibilities of action
  - And where the agent is and how it is moving in the environment



Information in the Light

- How does the gannet know when to fold its wings when diving into water to catch a fish?
  - Too soon and it is a missile with no guidance system
  - Too late and it is wingless
- David Lee showed that gannets are sensitive to a simple measure--the rate of expansion of the target in their visual field
- Note: on Van Essen's account neurons in MSTd do exactly this

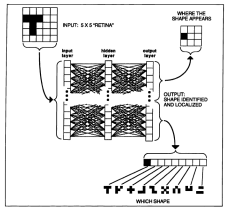


Inferential/computational (Marr) vs. Ecological Perception (Gibson)

- Marr viewed Gibson as a competitor
  - Gibson spoke of "directly perceiving" information in the light and rejected the idea of vision as making inferences/ computations
- But he also claimed that Gibson had come closest to understanding the computational level
  - By looking at the task posed in the world we can understand what the visual system must do
  - There is still a difference: whereas Marr often relied on intuitive accounts of what vision is for, Gibson relied on experiments
    - These often produced results at odds with intuition

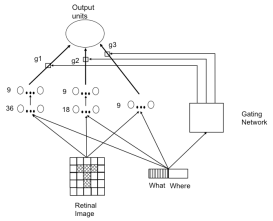
Modeling the Two Pathways

- Rueckl, Cave, and Kosslyn (1989) developed two neural network models for identifying what letter was presented on the input and in what location
  - The model that processed both what and where information in the same network (shown) performed less well than
  - The modular version in which separate networks processed what and where information
- Conclusion: The reason for separating what and where processing is computational efficiency



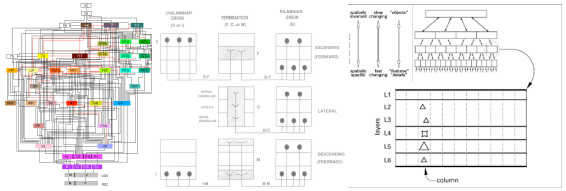
Learning to Segregate Processing Tasks

- Subsequently, Jacobs, Jordan and Barto (1991) developed a model in which the network itself learned to use different component networks for What and Where processing
  - Where information required and was better evaluated by a simpler network (without a hidden layer)
  - What information required a more complex network

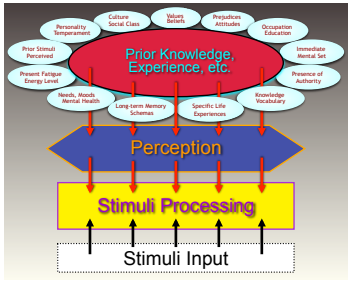


Top Down/Recurrent Connections

- Most of the connections in van Essen's diagram involve both upwards and downwards projections
  - Areas local in the hierarchy have smaller receptive fields whereas those higher cover more of the visual field
  - Fan in going up, fan out going down



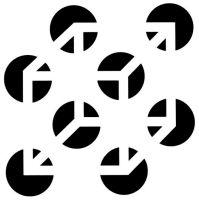
Top-Down and Bottom-Up Interact

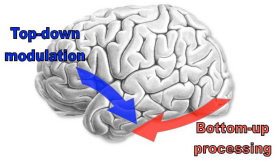


Created by Dr. Gordon Vessels 2005

Top-Down/Recurrent Connections: What Do They Do?

- The figure below can be seen either as a cube floating in front of 8 discs or as a cube seen through eight holes. Only in the first case do illusory contours appear
- The fact that how the figure appears shifts with how the individual chooses to view it illustrates top-down control





Richard Gregory on Top-Down Perception

Learning to See

- Top-down processes can change how the more basic processes operate so that one can learn to see things one hasn't seen before

Top-Down Processing: Nuisance or Useful Strategy?

- Top down processes can easily lead to misperception
  - Seeing what one wants to see rather than what is there
- So why does the brain employ top down processes?

- Helps to resolve ambiguity in sensory input
  - Instantaneous input can be interpreted in many ways
  - Previous context narrows the range of possibilities
- Provides a basis for learning
  - We are constantly predicting how the world will appear next
  - When it fails to fulfill out expectations, the network can adjust so as to predict better in the future

A further step: The brain as endogenously active

- A feedforward mechanism takes in input, produces an output, and does nothing more until a new input comes along
  - This is how researchers have often approached the human mind/brain
- But a mechanism with feedback loops can exhibit ongoing activity
  - As long as there is a supply of energy (there is no getting around the second law of thermodynamics)
  - Thus, one can have a system that is always doing something even without input
    - And whose activity may only be modulated by inputs

Discussion Question

What would you do if you showed up for an appointment and were shown to a room and asked to wait

Do absolutely nothing, since I wasn't asked to do anything

Get out my cell phone and check in on Facebook

Get out my cell phone and tweet a message

Think about different things going on in my life

Vision as Primarily Predictive

- The accounts of vision we have focused on assumed a reactive perspective
  - The visual system takes in input and register it, making the information available to other cognitive activities
- An alternative perspective is that the brain is predicting our next experiences
  - What we register from experience are the respects in which our predictions turn out false
  - We don't bother processing information that we already expected to be true
- Think of experiences in which you failed to notice things or events that you clearly saw