The Neuroscience of Vision III

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

- 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.”

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?
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 they 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 can they be reconciled?
  - Gibson as providing insight into Marr’s computational theory—the account of the goal and appropriateness of perception
  - And perhaps showing that some things don’t have to be computed (e.g., a complete, internal representation of the world and our place in it) since the world is always there to give further information as needed

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**

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**

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/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.

---

**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 our expectations, the network can adjust so as to predict better in the future