Vision: computational aspects

Richard Shillcock
Today’s goals

To explore some of the computational aspects of mapping from the visual world to the brain.
Today’s readings


The visual pathways
Magno- and parvocellular pathways

- Layers 1 and 2: magnocellular
- Layers 3-6: parvocellular
  - Layers 2, 3, and 5: ipsilateral eye
  - Layers 1, 4, and 6: contralateral eye
Topographic mapping

Stained V1 in the mouse, showing the areas that were activated by the visual stimulus. (Note also the cortical magnification of the fovea.)
Systematicity: Penfield’s homunculus
The higher visual areas become increasingly attuned to bigger receptive fields, with bilateral inputs (see, e.g., Tootell et al., 1998), and less clear retinotopic mapping.

Does the brain ever throw away information?
Systematicity is pervasive in the brain, most clearly nearer the sensorium. It is a way of importing relationships and larger-scale representation into the brain “for free”.
Vision as serving a purpose

The frog and its visual world
Ecological realism vs abstract stimuli

Pond-like backgrounds versus white backgrounds (Lettvin, Maturana, McCulloch, and Pitts).

There’s never a “null” context.

The risk of researching a technique; the assumptions become incorporated into the science in an invisible way.
The frog’s brain

- olfactory tract
- olfactory lobe
- cerebral hemisphere
- pineal gland
- diencephalon
- optic lobe
- cerebellum
- fourth ventricle
- medulla oblongata
- optic chiasma
- hypothalamus
- hypophysis

Dorsal View

Ventral View
The frog’s retina
What the frog’s eye tells the frog’s brain

Lettvin, Maturana, McCulloch & Pitts (1959)

Five types of ganglion cells – each a “feature detector”.
Each is interested in an aspect of the environment.
Contrast detectors (light/dark in a small area).
Convexity detectors (small, dark and moving).
Changing contrast detectors (moving edges).
Dimming detectors (dimming from edge or centre of the visual field).
Dark detectors (overall light intensity)
The frog’s visual world

Moving

B

Stationary

C
Convexity detection
The frog's visual world

Perception, 1981, volume 10, pages 421–422

The frog ganglion cell: not a feature detector and not a monkey cortical cell

Donald C Hood
Department of Psychology, Columbia University, New York, NY 10027, USA
James Gordon
The Rockefeller University, New York, NY 10021, USA
Received 3 March 1981

There are two fundamental misconceptions about frog ganglion cells that have been perpetuated in most recent textbooks. A large number of texts written for courses on perception, S and P, and physiological psychology use the frog ganglion cell as an example, par excellence, of a feature detector. Many go further and compare the processing done in the frog retina with that carried out in the cat and monkey cortex. The truth is, the properties of the frog's ganglion cells have never fitted the generally accepted definitions of a feature detector. Further, their response properties have recently been shown to be similar to those of other vertebrate ganglion cells.

As generally used the term 'feature detector' refers to a cell that is most sensitive to specific stimulus configurations (the 'features') of the environment. For example,
Looking further into amphibian vision

"Together, our results indicate that the salamander retina uses a population code in which every point in visual space is represented by multiple neurons with subtly different visual sensitivities" Segev et al. (2006).

We see a history of progress through reinterpretation. New analyses subsume earlier ones and introduce new concepts.
Vision as object recognition

The task becomes one of identifying invariant features.
Vision as search

The task becomes one of looking for objects.
There is clear agreement on the parts that (some) things contain.
Recognition as parsing

Is something like *syntax* going on? – Combining invariances in a rule-governed way.
Marr’s approach

Is it correct to try and “start simple and work up from there”? 

<table>
<thead>
<tr>
<th>Viewer centred</th>
<th>Object centred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Image</td>
<td>3-D Model Representation</td>
</tr>
<tr>
<td>Perceived intensities</td>
<td>3-D models hierarchically organised in terms of surface and volumetric primitives</td>
</tr>
<tr>
<td>Primal Sketch</td>
<td>Local surface orientation and discontinuities in depth and in surface orientation</td>
</tr>
<tr>
<td>Zero crossings, blobs, edges, bars, ends, virtual lines, groups, curves boundaries.</td>
<td></td>
</tr>
</tbody>
</table>

23/30
Marr’s approach

Building bottom-up versus building top-down.

Depth Reconstruction

- Image Features
  - Stereo
  - Motion
  - Shading
  - Texture
  - Shape from X

- 2 1/2-D Sketch
- Recognition (matching in 3-D domain)
- 3-D Object Models

Knowledge-based Vision

- Image Features
- Perceptual organization
- 2-D Perceptual Groupings
- Viewpoint solving and verification
- Search and probabilistic matching
- 3-D Object Models

24/30
Edges versus surfaces
Biederman & Ju (1988)

Naming and verification tasks showed no difference between photographs and cartoons.
A 19-month old boy had previously only learned to name toys and other objects. He was given line drawings of known objects. There was no evidence of learning being required.
Is simplicity really the answer?
Complexity and situatedness

- movement parallax
- binocularity
- light source
Challenges

Understand the relationship between objects, activity and whole scenes.

Decide how much we wish to base artificial systems on human cognition.

Appreciate the relationship between “clever” syntax-like solutions and “dumb” brute-force solutions.

(Never be satisfied with one-or-the-other binary choices.)
References


