Vision II

Object Recognition
Overview

• Perception versus recognition
• Where does it occur?
• What problems have to be solved?
• Evaluation of models
• Implications from neuropsychology
Recap on hierarchical organisation

Top-down vs bottom up

<table>
<thead>
<tr>
<th>Model layers</th>
<th>Corresponding brain area (tentative)</th>
<th>RF sizes</th>
<th>Num. units</th>
<th>Num. subunits</th>
</tr>
</thead>
<tbody>
<tr>
<td>classifier</td>
<td>PFC</td>
<td></td>
<td>10^0</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>AIT</td>
<td>7°</td>
<td>10^2</td>
<td>6,000</td>
</tr>
<tr>
<td>C3</td>
<td>PIT - AIT</td>
<td>7°</td>
<td>10^3</td>
<td></td>
</tr>
<tr>
<td>C2b</td>
<td>PIT</td>
<td>7°</td>
<td>10^3</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>PIT</td>
<td>1.2° - 3.2°</td>
<td>10^4</td>
<td>100</td>
</tr>
<tr>
<td>S2b</td>
<td>V4 - PIT</td>
<td>0.9° - 4.4°</td>
<td>10^7</td>
<td>100</td>
</tr>
<tr>
<td>C2</td>
<td>V4</td>
<td>1.1° - 3.0°</td>
<td>10^5</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>V2 - V4</td>
<td>0.6° - 2.4°</td>
<td>10^7</td>
<td>10</td>
</tr>
<tr>
<td>C1</td>
<td>V1 - V2</td>
<td>0.4° - 1.6°</td>
<td>10^4</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>V1 - V2</td>
<td>0.2° - 1.1°</td>
<td>10^6</td>
<td></td>
</tr>
</tbody>
</table>

- Simple cells
- Complex cells
- Tuning
- Main routes
- MAX
- Bypass routes
How do we know object recognition is distinct from object perception?

- Patient GS “agnosia”
- From Gk gnostic – failure to “know”
- Colour, form, shape & motion intact
- Unable to name or recognise objects
- Faces ok, object recognition via other senses spared
What is distinct about recognition?

- Perceive unified objects/scenes
- Intimately linked with memory
- Perception triggers memories
Two cortical pathways for visual perception

M pathway

Superior longitudinal fasciculus

"Where"

Posterior parietal cortex

P pathway

V1

"What"

Inferior longitudinal fasciculus

Inferior temporal cortex

1982 Ungerleider & Mishkin propose 2 routes for visual output

Damage to this route = agnosias
Other features of the ventral pathway

- Deficits restricted to visual modality
- Posterior disrupts visual discrimination
- Anterior damage impairs visual memory
- Deficits only after bilateral lesions
“What” is bilateral, “where” more unilateral

Until CC fibres cut..
Physiological properties of parietal cells

Both P&M large receptive fields, but parietal non-selective

Single cell recordings from neuron in monkey posterior parietal cortex (PPC). The cell responds to both small and large stimulus, but the magnitude of activity is correlated with the size of the stimulus. Many exclude fovea and cover periphery.
Physiological properties of temporal cells

SCRs from neuron in monkey inferior temporal cortex (ITC). Cells here rarely respond to simple stimuli e.g. lines/spots of light but sensitive to complex objects such as hand drawings. Weak response to mitten – activity not related to general hand shape – neither responsive to comblike objects (despite similar features).
Early PET expt supports what/where in ventral/dorsal & laterality

Kanwisher et al, 1997. PET study to dissect cognitive components

Unexpectedly, both new and familiar objects activate occipito-temporal areas.
Visual Streams - what/how

Milner and Goodale 1991

- Worked with a patient, D.F. with extensive bilateral ventral-stream lesions since 1991.
- D.F. has profound visual form agnosia
- But some interesting spared abilities...

It’s the need to act on the stimulus that allows her to perform. She still doesn’t know what it is or (maybe) where it but knows how to post something into it.
Patient DF suggests processing systems make use of different sources of perceptual information. Dorsal system used to locate and guide actions.
How do we recognise objects?

- Primarily depends on analysis of shape
- However, colour/texture/motion play a role (e.g. seashore but pink elephants?)
- Key issues:
  - What are the salient features?
  - How do we cope with huge variation in views?
The visual system is capable of object constancy.

Object constancy refers to the ability to recognise an object in countless situations despite the fact that the image on the retina is vastly different for each drawing.
Sources of object variability

- Viewpoint (distance, orientation)
- Illumination (shadows)
- Partial occlusion
- Category variability (size, shape, colour)
Theories of object constancy

• **View-dependent theories**
  – Perception depends on recognising an object from specific viewpoint *(example)*
  – Numerous different representations in memory
  – But, heavy burden on perceptual memory although less so if interpolation process

• **View-invariant frame of reference**
  – Sensory input defines basic properties (short-cuts)– object’s other properties are defined with respect to these properties
  – Marr’s theory posits initial experience of an object defines major and minor axes – but some exceptions.

NB both theories propose recognition depends on breaking down a scene/object into its component parts
CUPS?
Major axis

Minor axis
How do we encode shape?

• Exploit salient sources of information
  – Key points of distinction
  – View-invariant cues
  – Perceptual rules (parallelism, symmetry)

• Recognise objects by analysing its parts
  – Collinearity
  – T-junctions
Study these 2 drawings – if you were to draw them from memory, it's very likely you would include the points of greatest curvature – the most salient sources of information that discriminate the 2 objects.
Exploiting view-invariant cues

The parallel lines intersected by the cylinder and that describe the rectangular box remain parallel across any viewing position.
The Ames room illusion: parallelism

This illusion exploits the fact that we expect in rooms, the back wall is parallel with the front wall and the ceiling a constant height. Our perception system assumes the 2 people are at the same distance and so concludes that the closer one is a giant.
Recognition by parts analysis

Collinear: discontinuous line

T-junctions imply 2 objects

Invariant properties persist even if object rotates allowing parsing. Recognition occurs by determination of defining component parts

There are 2 salient cues to recognising the fact that there are 2 objects here. (i) upper borders are collinear and form discontinuous line (ii) edges of both objects form t-junctions (none found in 2D projection of rectangle)
The Geon theory of object recognition

All objects can be described as a configuration of parts. 24 geometric volumes/ions form a perceptual alphabet. Biederman (1990)
The Geon theory of object recognition

• Problems:
  – All exemplars of a category aren’t the same in geon terms (coffee cups +/- handles)
  – We can discriminate 2 objects that have near identical geon descriptions (dog breeds)
  – OK for between object classes, poor within
Grandmother cells: do they really exist? Hierarchical encoding taken to an extreme

Geons?

View-invariant properties?
Limitations of grandmother cell hypothesis

- Electrophysiology testing limited by no. of objects and no. of cells.
- The constant state of *spontaneous firing* and *refractoriness* renders single cells susceptible to error.
- If cell died, no longer recognise object.
- Fails to explain recognition of novel objects (wouldn’t “be” like anything else).
Ensemble encoding hypothesis

- Object recognition results from activation across complex feature detectors (some cells respond to shape, others hair etc)
- Collective activation not one unit
- Better explains:
  - Confusion between 2 visually similar objects
  - Object recognition persists if one cell dies
  - Novel objects recognised by similarity to existing feature representations
Brain Injury Patterns of Agnosia and Visual Object Pattern and recognition deficits
Agnosia

- GL, 80-yr-old salesman, hit his head during a storm. Initially reported problems with seeing but he could still see; his problem was recognising the things about him.
- GL had almost normal visual acuity for his age, & could draw quite accurate copies of objects he couldn’t recognise; thus vision intact in at least some respects. Hadn’t lost his knowledge of objects either.
- Usually refers to a recognition disorder not resulting from general intellectual deterioration, language impairment, or basic sensory dysfunction. In visual agnosia the patient can still see things but can’t recognise what they are. There’s also tactile & auditory agnosia.
- Lissauer (1890) distinguished between ‘apperceptive’ & ‘associative’ agnosias, suggesting that they were the result of impairments at different stages.
Failures of object recognition

(a) Agnosia

(b) Memory loss
Change Blindness examples
Subtypes of Agnosia

• Apperceptive agnosia – failures in object recognition linked to problems in perceptual processing

• Associative agnosia – normal visual representations but unable to use this information to recognise things.
Apperceptive agnosia

Can see
Can do obstacle avoidance etc
Good acuity etc, but
Cannot recognise objects
Cannot see global structure (full shape)
Cannot match to sample
Cannot copy

Rare. Usually occurs following Right IPL damage

Copying example
Tests for Apperceptive Agnosia

(a) Unusual views test

(b) Shadows test

Graphs showing percentage correct for Left and Right hemisphere patients with Anterior and Posterior damage.
Perceptual impairments

- Patient HJA (Humphreys & Riddoch, 1987)
  - Impaired perceptual grouping mechanisms to translate intact perception of lines into more complex visual descriptions required to access stored knowledge “integrative agnosia”. (supports the idea of Gestalt-based grouping mechanisms)

- Patient EL (Harris et al, 2001)
  - Can produce names of items in usual orientation but not when principal axis obscured “object orientation agnosia” (supports the theory of object constancy)
Associative agnosia

Good acuity etc
Can see global structure (full shape)
Can match to sample*
Can copy
But still cannot recognise objects
Cannot recognise own copy

Usually occurs following
left occipitotemporal damage

*but not nonmorphologically identical representations of
the same object
(e.g. matching a line drawing with the real object).
The Matching-by-Function Test

Subject must choose the 2 objects most similar in function. In both cases the correct match is with the object on the right, despite greater physical similarity of the objects on the left.
Semantic Impairments - AB

- AB - former civil servant, referred to hospital for investigation of his deteriorating memory. Could distinguish shapes & basically able to derive an adequate object-centred representation.

- Although his perceptual abilities were intact AB was severely impaired in recognition tasks.

- Warrington (1975) suggested an impairment of semantic memory. She showed AB 40 photographs of animals/objects, presented one at a time in random order. He had to decide whether or not the photographs fitted into different categories. He was able to judge quite well whether or not the photographs were of animals, but struggled with more object specific information.

- So AB presents a pattern of intact initial, viewer-centred & object-centred representations, together with impairments on semantic tasks, suggesting selective impairment of some aspects of semantic memory.
Warrington's Model of Agnosia

Left-hemisphere systems

Lesioned $\Rightarrow$ associative agnosia

Semantic categorization

Visual analysis

Right-hemisphere systems

Lesioned $\Rightarrow$ apperceptive agnosia

Perceptual categorization

Visual analysis
Problems with Warrington’s model

• Not all associative agnosia in LH lesions
• Unilateral RH lesions can produce agnosia more associative than apperceptive
• Patients like HJA require greater specification of perceptual categorization stage
Category Specificity in Visual Object Recognition

• It is hypothesized that higher visual areas are specialised for recognising different categories of object such as animals, faces, places, words and bodies.
• Evidence in favour of this position comes from patients with dissociations of spared and impaired performance in the recognition of different classes of objects.
• Alternative theory is that different categories of stimuli require somewhat different kinds of processing (e.g. words are recognised by parts, and faces recognised holistically).
Neurones in temporal cortex

- Infero-temporal cortex (aka, inferior temporal cortex, IT) contains cells that respond to very complex visual objects
- In TE they have larger receptive fields and more complex preferences than TEO
- Famous for ‘face’ cells (like the profile-selective cell
Face recognition in humans

• In humans there is also an area that responds selectively to faces - the Fusiform Face Area (named because it lives on the Fusiform Gyrus) shown using PET and fMRI

• In humans here is also a selective deficit in recognising faces (prosopagnosia), probably caused by damage to the FFA
Prosopagnosia

Can’t identify faces - even extremely familiar ones (even themselves!)
But can identify people by other means
And can recognise a face as being a face and
Can discriminate between faces.
Other routes to recognition

- Cosmetics and hair
- Voices
- Location
- Relationships
Big Question: Is prosopagnosia special (i.e. is there a special face processing area in the brain) or just another type of agnosia

Competing theories of prosopagnosia:
1) agnosia for specific faces
2) a type of perceptual disorder involving a system specific for face perception
3) a type of amnesia syndrome in which the perception of a face is dissociated from previous memories of the face.

Inability to identify specific class of objects rather than just faces?

No case of 'pure Prosopagnosia' - i.e. no other perceptual deficit e.g. Prosopagnosia’ often co-occurs with other types of agnosia, (anatomical coincidence?)
Prosopagnosia’ can occur following damage to a variety of brain regions. + area that responds to faces not the same in all subjects.