CDL PS2011/16

Lecture 4
Object recognition

Reading

Essential
- Bruce, Green & Georgeson (2003), Chapter 4, pp. 80-82; Chapter 5, pp. 85-99; Chapter 6, pp. 136-140; Chapter 7, p. 201; Chapter 9, pp. 265-287
- Eysenck & Keane (2000), Chapter 4, pp. 83-96

Recommended
- Gordon, (1997), Chapter 8
- Roth and Bruce, Part 2, pp. 71-136

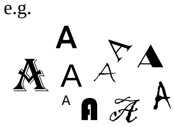
Additional
- Wade & Swanston, Chapter 6, pp. 226-235
- Pinker, Chapter 4.

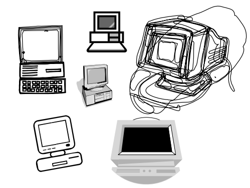
The problem of stimulus equivalence

If the information about objects available to an observer is a static two-dimensional retinal image,

then an infinite number of retinal images can correspond to a particular object

e.g.





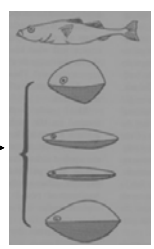
What different images of the same object have in common?

Simple mechanisms of recognition

e.g. Tinbergen (1951)
- Aggressive display by sticklebacks towards male intruders in proximity of nesting site
- male sticklebacks characterised by red belly
- key stimulus = red patch

Detailed model of stickleback is ineffective →

Presence of red patch in roughly the right position triggers response →



More sophisticated visual recognition processes require matching complex configurations with internal representations

Theories of pattern recognition
- simplified tasks motivated by applied problems in computer vision
  - -Template matching
  - - Feature analysis
  - - Structural descriptions

Template matching theories

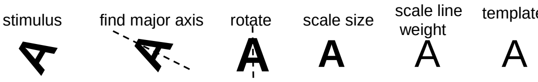
- memory templates of patterns
- stimuli are compared with templates
- if they match the template the stimulus is recognised

Problems with template theories

matching would fail even with minor differences between stimulus and template

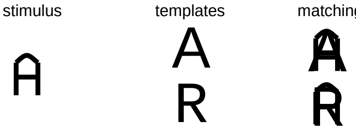
Standardisation processes are required
e.g.

stimulus | find major axis | rotate | scale size | scale line weight | template



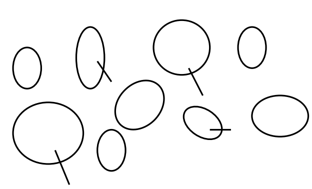
even after standardisation mismatches are likely to occur

stimulus | templates | matching



An even more serious problem with template theories is to decide which characteristics of a stimulus have to be modified in order to match it with the relevant template.

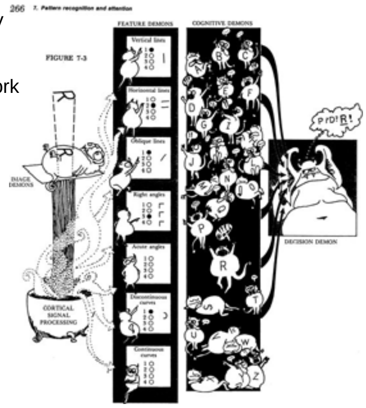
e.g. should these stimuli be matched with an O or with a Q ?



Feature analysis theory

Mostly influenced by Hubel and Wiesel's work on feature detectors in the visual cortex

e.g.
Selfridge's (1959)
"Pandemonium"



Problems with feature detection

- Recognition based on lists of features
- loss of information concerning differences between different instances of the same class of patterns

e.g.

pandemonium

pandemonium

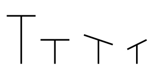
Unable to distinguish patterns with the same features in different configurations
e.g.

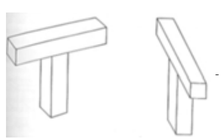
Structural descriptions

not a theory
set of representations
propositional (statements with objects and predicates)
explicit description of relationships

e.g. letter T
horizontal line, vertical line (supports, bisects)

can be abstract and robust enough to be exported to 3D





Outline of Marr's theory of vision

Complete computational theory of object recognition

Input for vision = grey level description
pattern of intensity of light in the retinal image

Series of representations to achieve object recognition

primal sketch
- raw primal sketch = symbolic representation of bars, blobs etc.
- full primal sketch = raw primal sketch + grouping procedures

21/2-D sketch = primal sketch + depth cues
- viewer centred representation

3-D model representation = three dimensional representation of object parts and their relationships
- object centred representation
- can be matched to memory templates for recognition

Modularity

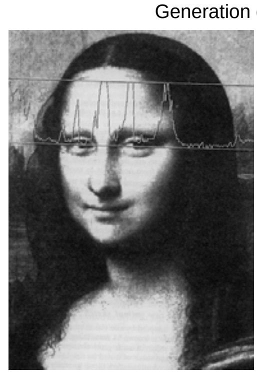
Grey level description

pixels
units of area of the image

values of light intensity in each pixel
possibly coded by receptors in the retina

3	2	3	3	3	3
2	3	2	2	2	2
1	2	3	2	3	3
5	3	2	11	5	4
7	9	24	34	36	36
32	33	37	36	36	37
36	38	36	36	36	36

Generation of raw primal sketch



zero crossings
changes in light intensity values

not all zero crossings correspond to boundaries
need for smoothing

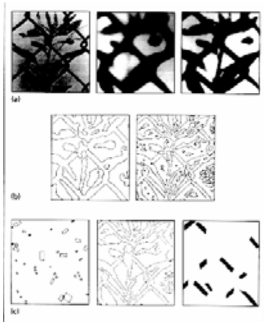


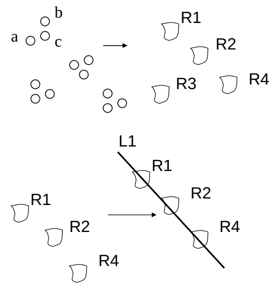
Image processed using different spatial frequency filters

Zero-crossings derived for differently filtered images

Zero crossings combined to detect blobs, edges and bars

Generation of full primal sketch

primitives of raw primal sketch + grouping

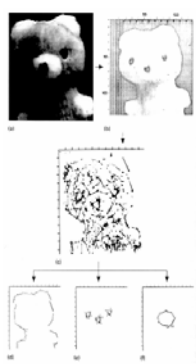


labelling
- explicit naming
  - e.g. regions = R
  - lines = L
- place tokens

Grouping e.g. by proximity, good continuation etc.

output of full primal sketch
- occluding contours
- components within contours

Summary of Marr's early stages of visual processing



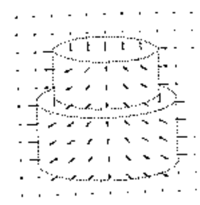
retinal image
analysis of light intensity values
grey level description

Zero crossings
identification of significant regions
raw primal sketch

Grouping
identification of occluding contours and significant components within them
full primal sketch

$2^1/_2$-D Sketch

Derived from full primal sketch using depth cues e.g. stereopsis, shading etc.



Range map
-description (relative to the viewer) of orientation and depth of surface elements
-surfaces hidden from view are not represented

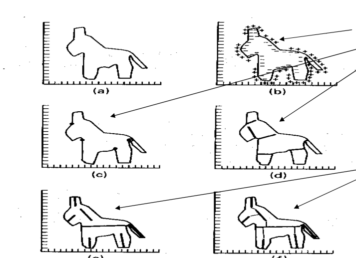
A viewer centred representation must become an object centred representation to enable recognition from different view points.

3D Representation

Representation of objects
- hierarchical arrangement of generalised cones

Generalised cones
- major axis
- cross section of variable size

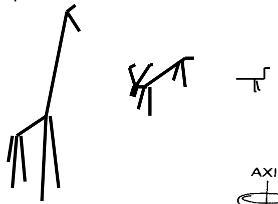
The detection of concavities (-) and convexities (+) in an image is used to identify its main component parts and their major axes



Concavities indicate likely junctions of different parts

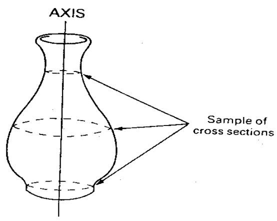
Then the major axis of each component part can be derived

Axes provide information on size and orientation of parts



Generalised cone

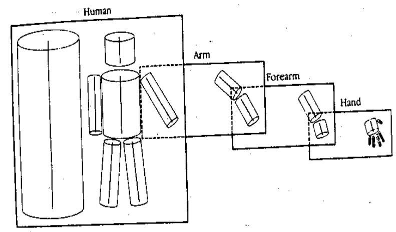
created by moving a cross-section of constant shape but variable size along an axis

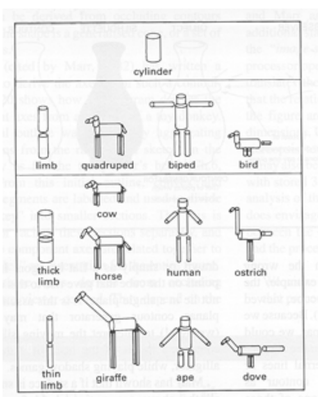


Objects as hierarchy of 3-D models

at each level representation of:

- major axis
- component axes
- information about shape of component axes





Recognition

matching of 3D description from image and catalogue in LTM

catalogue can be addressed from different levels of the hierarchy and from sub-components

General principles of Marr's approach

- Modularity
- Explicit naming
- Least commitment
- Mostly bottom up until access to memory catalogue is required for recognition

Summary

Stimulus equivalence problem

Pattern recognition
- Simple recognition mechanisms
- Template theories
- Feature theories
- Structural descriptions

Marr's approach
- primal sketch
  - raw
  - full
- 2½ D sketch
- 3D representation
  - hierarchy of generalised cones
- Object recognition
  - matching with 3D LTM catalogue