Pattern Recognition and Perception

Outline:

1. Pattern Recognition
    - definition
    - simple models
        - template matching
        - feature models
    - the segmentation problem
    - categorical perception
    - context effects

Visual Perception

2. Classic Perceptual Psychology
3. Gestalt Psychology
4. Gibson
5. Hubel & Wiesel 1962
6. Marr's Theory of Vision (an integration)
7. Pattern Recognition (again)

1. PATTERN RECOGNITION

Simple Models

Template Matching

problems:

translation, size, orientation, variation

Feature Models

behavioral evidence

(similarity = features in common?)

confusion errors

biological evidence

stabilized retinal images

Pritchard 1961

categorical perception (see below)

Complexities involved in Pattern Recognition

Segementation Problem
- speech- lack of spaces
  coarticulation of phonemes

Categorization of Primitives?
- speech- categorical perception
  manipulate onset of voicing

Context Effects
- (top down effects)
- (conceptually driven processes)

word superiority effect: Wheeler 1970
phoneme restoration effect: Warren 1970

faces: Palmer, 1975
scene organization: Biederman et al. 1973

how can top down & bottom effects interact?

VISUAL PERCEPTION

2. Classical Perceptual Psychology
(aka as "Constructivism")

background:
British Empiricists - 19 cent. Structural Psych.
e.g., Berkeley- learn to see depth

began with:
19th cent. Helmholtz
"unconscious inference"

what they were concerned with:
inadequacy of the retinal image
how could perceptual experience arise from it?
what information does perception use?

what they studied:
illusions
tachistoscopic studies
psychophysics:
relationship between physical & psychologica

"cues" for perception
binocular disparity, stereopsis
shading
motion parallax

3. Gestalt Psychology ~1910-1950

Studied: Perception and Problem Solving

Principles:

- Whole is Greater than Sum of Parts
- Isomorphism
- Organizational Principles
  - proximity
  - similarity
  - good continuation
  - closure and good form

But no real computational theory or physiology

4. Gibson: the ecological approach to perception

historical notes:

- example of WWII motivated research
- began as discovery of new optical "cues"
  - texture gradient 1950 -
  - higher-order invariances
    - e.g., optical flow pattern

Gibson

- deemphasize processing & inferencing
  - (Constructivism)
- emphasize information "in the light"
  - direct pick up from the environment
  - "affordances"

- terms for this approach:
  - direct perception
  - ecological perception

- derivative of Gestalt approaches

Points:

- incoming information is complete
  - organism doesn't "add" anything to sensory stimulus
- nature of "processing" on that information is
  - not unconscious inferencing but rather "resonance"
    - pick up of information in the light
    - does not apply meaning or understanding
- the role of learning is not learning of concepts
  - just tuning to information present in environment

lead to a reformulation of "ecological" (vs abstract) optics:

- information in the array
  - texture gradients
  - lighting gradients
  - moving objects
  - moving observer:
    - optical flow pattern

5. Hubel & Wiesel 1962

hierarchy of cell types
- receptive fields
  - center surrounds (on/off cells)
  - edge detectors
  - complex cells
    - moving edges
    - angles
    - lines in different locations

can this be taken as a complete model of P.R.?

6. Marr's Theory of Vision

an integration of psychology, physiology, computer science

an accounting for, and integration of, preceeding work

* need for complex information processing
  vs.
    template matching
    simple feature matching

* how do Classical (constructivist) "cues" etc. play a role?

* how can we account for Gibson's "direct perception", use of invariants etc.?

* how can we account for Gestalt insights?

* how can we do this in a computational model?

* hierarchical and modular organization
  (cutting problems at appropriate joints)

  (this is a complex problem, different things are going on at different levels)

The Problem of Vision:

light is source of information re physical objects but how direct?

image- 2d array: I(x,y)
austerity of image

General Outline: Qualitatively Different Levels

1. Low Level Processes
   close to physiological model
   analyzes structures in the retinal image per se (2D)

2. Intermediate Processes
   analyzes viewer oriented surfaces in world (polar coordinates or "2.5D")

3. High Level Processes
   determine shapes in 3D world
   determine/recognize objects in 3D world

1. Low Level Processes

* physics determines correlation between
    retinal image and world: constrains image
    so that it is informative
    [Gibson]

* but, as constructivists argue, need to
    consider how retinal image is processed

* but retinal image processing is not final stage,
    more like sensory transduction to next level,
    Gibson's term "resonance" is apt at this stage

How does it work?

How to extract information from 2d array?

What to analyze? image- 2d array:
intensity array? I(x,y)
intensity gradients? "slope" at each point
acceleration of change?
2nd derivative operators
brighter, darker, brighter
(or vice versa)
center surround cells

Intensity array:

(sharp edge)

5 5 5 5 1 1 1 5 5 5 5

5 5 5 5 1 1 1 5 5 5 5

5 5 5 5 1 1 1 5 5 5 5

5 5 5 5 1 1 1 5 5 5 5

5 5 5 5 1 1 1 5 5 5 5

(soft edge)

5 5 5 4 3 2 1 2 3 4 5 5 5

5 5 5 4 3 2 1 2 3 4 5 5 5

5 5 5 4 3 2 1 2 3 4 5 5 5

5 5 5 4 3 2 1 2 3 4 5 5 5

5 5 5 4 3 2 1 2 3 4 5 5 5

Intensity gradient array:

-	+

(linear) gradient detector

/ 0 0 0 -4 0 0 4 0 0 0

/ 0 0 0 -4 0 0 4 0 0 0

/ 0 0 0 -4 0 0 4 0 0 0

/ 0 0 0 -4 0 0 4 0 0 0

/ 0 0 0 -4 0 0 4 0 0 0

/ 0 0 -1 -1 -1 -1 +1 +1 +1 +1 0 0

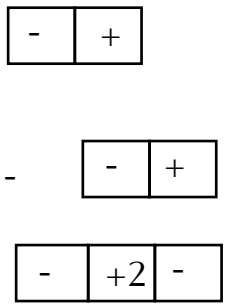
/ 0 0 -1 -1 -1 -1 +1 +1 +1 +1 0 0

/ 0 0 -1 -1 -1 -1 +1 +1 +1 +1 0 0

/ 0 0 -1 -1 -1 -1 +1 +1 +1 +1 0 0

/ 0 0 -1 -1 -1 -1 +1 +1 +1 +1 0 0

2nd. derivative (acceleration) array:



(linear) 2nd. derivative (acceleration) detector

// 0 0 4 -4 0 -4 4 0 0 /
// 0 0 4 -4 0 -4 4 0 0 /
// 0 0 4 -4 0 -4 4 0 0 /
// 0 0 4 -4 0 -4 4 0 0 /
// 0 0 4 -4 0 -4 4 0 0 /

// 0 1 0 0 0 -2 0 0 0 1 0 /
// 0 1 0 0 0 -2 0 0 0 1 0 /
// 0 1 0 0 0 -2 0 0 0 1 0 /
// 0 1 0 0 0 -2 0 0 0 1 0 /
// 0 1 0 0 0 -2 0 0 0 1 0 /

(intensity gradients repeated here for comparison)

5 5 5 5 1 1 1 5 5 5 5
5 5 5 5 1 1 1 5 5 5 5
5 5 5 5 1 1 1 5 5 5 5
5 5 5 5 1 1 1 5 5 5 5
5 5 5 5 1 1 1 5 5 5 5

5 5 5 4 3 2 1 2 3 4 5 5 5
5 5 5 4 3 2 1 2 3 4 5 5 5
5 5 5 4 3 2 1 2 3 4 5 5 5
5 5 5 4 3 2 1 2 3 4 5 5 5
5 5 5 4 3 2 1 2 3 4 5 5 5

more powerful detectors (as opposed to linear detectors)

any orientation:
circular contrast detectors

degrees of variation of gradients:
variety of sizes

feed into line detectors etc.

examples of physiological and psychological evidence

for channels of different sizes:

adaptation technique

physiology for primal sketch

organization of center-surround cells
and higher order cells

2. Intermediate Processes

Intrinsic Images, 2.5D sketch

viewer oriented surfaces in world,
(rough, not segmented)

processes take primal sketch,
work on independent information

stereopsis
motion
surface texture
shading
visual motion

example: physiological evidence for motion detection

fatigue, adaptation

3. High Level Processes

determine 3D model of world
determine objects in visual field

problem: getting from:
intermediate: viewer centered representation
to
high-level: canonical, abstract, indexed

problems of pattern recognition

decomposition to parts

representation of basic shapes

representation/recognition of complex shapes

Pattern / Object Recogniton

1. Segmentation of complex objects

Hoffman & Richards 1984
segmentation by lines of maximum concavity

2. Primitive Shapes

generalized cylinders
cross section
change of cross section over
path of axis through 3D

Biederman 1987, geons
(subset of generalized cylenders)

3. Complex Shapes:

Frame or Object like representation
hierarchical structures
subcomponents: primitives & constructs
connected at particular points