Competing Frameworks in Perception

Lesson II: Perception
module 08

University of Idaho
© Steffen Werner - 2004

Views on perception

- Perception as a cascade of information processing stages
  - From sensation to percept
  - Template vs. feature models
  - David Marr’s computational theory of vision
  - Top-down vs. bottom-up processing
  - Perception as unconscious inferences
- Perception for action - what do we need to perceive to act sensibly in our environment
  - James Gibson’s theory of direct perception
  - Optical flow and how to catch fly balls
- Milner and Goodale’s theory of two visual systems

Processing stages in visual perception

- Retina
  - Retina contains rods and cones (receptors)
  - Sensitive to different wavelengths of light
  - Information is processed / sharpened immediately
- Early visual processing
  - The information is sent from the retina to LGN
  - From LGN to visual cortex
  - Different cortical areas are sensitive to different, increasingly complex “patterns”
- Late visual processing
  - Information is integrated in higher cortical areas
  - E.g., object recognition, motion perception, etc.
To recognize an object, the perceptual “input” has to be matched to a memory representation.

- Template matching assumes that...
  - a large number of specific representations are holistically stored in memory (large database)
  - the input is matched against all possible memory representations
  - the representations with the best match (overlap) constitutes the interpretation of the input

Template matching has been applied to...

- Face recognition, fingerprints, letter recognition, electronic processing of checks, bar codes, etc.

Problems with template matching

- The same object can look very different
  - Faces look different from different angles, for different facial expressions, different lighting conditions, changes in clothing
  - Even simple stimuli like letters vary dramatically in their appearance
  - A system that uses template matching must therefore include a large number of different instances for each object

- Can we avoid some of these problems?
  - Powerful normalization processes (e.g., making sure that the objects is in a standard orientation, size)
  - Careful selection of the elements making up the set of objects and templates

Examples: template matching

Color codes:
- grey = stored template
- black = input
- purple = overlap (match)
Feature-matching models

- Feature-matching
  - The visual image is processed at multiple levels to determine the presence / absence of features
  - The resulting feature list is matched to object description in memory
- Advantages over template matching
  - Features are more stable than exact appearances
  - Feature lists are compact representation of a potentially large class of different sensory inputs
  - Neural basis for feature detection has been identified (e.g., Hubel & Wiesel’s feature detectors in the cat)
  - We will encounter a specific feature model later when we talk more specifically about object recognition

Important considerations

- What are the basic features?
- Distinctiveness: Allows to discriminate between different objects (in this case letters)
- Reliability and stability: Feature should be present and detectable in most cases
- Discriminating between different letters more important than between version of the same letter

Research on letter confusability

- Which letters are often confused (P-R, Y-V, M-N)?
- Teaching pre-kindergarten children the distinctions between confusable letter pairs

A feature representation of the letter R

Visual input

Feature level

Letter level

Interpretation

“Looks like an R to me...”
Information processing stages
- (Retinal) image
- Raw primal sketch (simple features, i.e., edges, blobs, terminators)
- 2.5D sketch (identification, organization and order of surfaces)
- Hierarchical 3D model (structural description based on 3D primitives)

Three levels of analysis
- Computational level: what do we need to compute?
- Algorithmic level: how can we do it?
- Implementation: how can it be implemented?

Bottom-up theories
- Data-driven, stimulus-driven
- Perception is a process that proceeds from early sensory processes towards more and more complex processing which results in a percept / recognition
- David Marr’s theory is one example

Top-down theories
- Influence of high-level cognitive processes (expectations, knowledge) on perception
- Percepts are constructed as a compromise between the sensory data and our expectations of what we are supposed to perceive

What is interesting about these images?
easy to see if you know what to look for

The word-superiority effect

It is easier to identify a letter as part of a word than in isolation
- Faster reaction times and higher accuracy
- This effect has been the source of some of the most important theoretical models of top-down processing

| WORK | ◯◯◯◯ | __ __ K |
| K    | ◯◯◯  | K       |
| OWRK | ◯◯◯◯ | __ __ K |

The interactive activation model
Perception as hypothesis testing

- Perception is influenced by expectations
  - We see what makes most sense, e.g., yellow bananas, three-dimensional objects, etc.
  - We "perceive" the occluded legs of a table
  - We recognize "appropriate" objects faster in a scene than unusual objects
- Perception as unconscious inference
  - Knowledge of the world influences perception
  - Perception as problem solving
- Perception is influenced by statistical regularities
  - The light-from-above hypothesis

Perception and expectations

- What do you see? Why?

Direct Perception (Gibson, 1950)

- Environment contains necessary information -- No need for "perceptual problem solving"
  - A moving observer in a visual environment is presented with a rich and unambiguous set of data
  - Many visual guided behaviors can be explained by identification of perceptual "invariants"
  - Higher, cognitive processes are not necessary
- Examples
  - Radial optical flow during locomotion
  - Are we going to crash?
  - Catching fly-balls
  - Size constancy in the natural environment
Imagine you are driving down a rural highway towards an intersection and another vehicle is approaching the intersection from the right...
- How can we figure out whether this is potentially a crash situation or whether everything is fine?
- Problem: It is difficult to judge the exact time that both cars will be at the intersection

Solution
- See next slide ...

Solution
- If the angle between your car and the approaching car is decreasing, the other car will pass before you
- If the angle between your car and the approaching car is increasing, you will pass first
- What happens if the angle remains the same...?
Ever wondered how to catch a fly ball?
- For a psychologist, the question becomes: How are perception (where is the ball going?) and actions (running) coordinated?
- Given the accuracy of our perceptual system, it is nearly impossible to predict the exact location where the ball will land.

Solution
- The outfielder runs while looking at the ball.
- The outfielder tries to maintain a linear trajectory of the ball in his (her) field of view.
- If trajectory stays linear then the outfielder will be in the right position to catch it.

Solution
- The outfielder runs while looking at the ball.
- The outfielder tries to maintain a linear trajectory of the ball in his (her) field of view.
- If trajectory stays linear then the outfielder will be in the right position to catch it.

Size constancy and horizon-height ratio

Synthesizing the two approaches

- Unconscious inference
  - Semantic interpretation and high-level recognition
  - Producing conscious results
  - Necessary for ambiguous, degraded sensory input

- Direct perception
  - Linked to actions in a three-dimensional world
  - Close coupling of visual events and actions
  - Rich sensory information available

- Different foci of the two frameworks
  - Object recognition vs. action in physical world
  - Conscious vs. largely unconscious
  - Evidence for separation of two systems...?
Are visual illusions also manual illusions?

Imagine grasping the central disc in this example of the Ebbinghaus-Titchener illusion. What do you think will happen?

Milner and Goodale’s theory

- Perception for action
  - Milner and Goodale assume that there are two basic visual systems - one, conscious, for the recognition of objects, and a second, unconscious, for action

- Neuropsychological data from case studies
  - Some patients cannot consciously identify the orientation of a slit but can orient their hand appropriately when putting things through it

- Dissociation of object recognition and visually guided action in non-clinical populations
  - While the size of the central circles is misjudged in the Ebbinghaus-Titchener illusion, the hand shows less bias when grasping it (this is not undisputed)