1 ## Visual Cognition

September 11, 2008

2 ## Overview of Visual Cognition

- Visual system: mission critical
- Multivariate inputs, unitary experience
- Multiple types of vision means many types of units in a distributed visual network
- Segregating vs. parsing
- Visual experience reveals nature of underlying representations
- Theory >> data

3 ## Who's Coming to Dinner?

- Gestalt Psychology (interested in basic dimensions of perceptual organization)
- Cognitive Psychology (interested primarily in effects of cognitive operations on visual experience)
- Cognitive Neuropsychology (interested in effects of selective lesions of visual cortex on visual behavior)
- Electrophysiology (interested in single- and multiple-unit activity of cells in visual cortex)
- Cognitive science (interested in developing overall theories of visual experience and object recognition)

4 ## Sensation and Perception

- **Sensation**: Conscious outcome of sense organs and projection regions. ("I detect something", not necessarily conscious, and not necessarily meaningful)
- **Perception**: means by which information acquired from the environment via the sense organs is transformed (organized) into conscious experiences of objects, events, sounds, tastes, etc. ("I know, recognize, appreciate what I am sensing, and it means something to me")

5 ## Three Stages of Visual Processing

- Reception: absorption of physical energy by receptors
- Transduction: translation of physical energy into electrochemical activity of the nervous system
- Coding: how information is encoded and represented; relationship between aspects of the physical stimulus and resultant nervous system activity (e.g., rate, temporal patterning of neuronal activity)
  - Population coding
  - Sparse coding

6 ## Basic Perceptual Phenomena You Already Know About

- **Size and Shape Constancies**: reflect mechanisms whereby perception of a distal object stays the same despite proximal changes in the perception – implies storage of a "structural representation"
- **Depth Cues:** monocular or binocular sources of information that convey information about relative distance of objects from viewer – involve cues about the structure, movement, and proximity of objects

7

8

9 ## Milner & Goodale: Perception-Action Model

- Alternative to the Mishkin-Ungerleider model
  - Dorsal: "How do I interact with that object"?
  - Ventral: "What is that object"?

• Evidence: distinctions between object recognition and visual control of action; Balint's syndrome; optic ataxia

10 Perceptual Organization

• Two complementary issues:
  – ***Organizing*** coherent objects and events out of segregated sensory/perceptual inputs ("binding")
  – ***"Parsing"*** the perceptual world; understanding which inputs belong together and which come from separate objects
  – http://www.michaelbach.de/ot/mot_feet_lin/index.html

11 Perceptual Organization

• Gestalt Psychology
  – **Law of Prägnanz**: perceptual system organizes to the simplest and most stable shape possible from the array
  – other laws describe how disparate perceptual elements are grouped
  – although their "laws" are probably incorrect, the elementary concept of **perceptual grouping** is critical
  – Inferring processes from performance

12

13 Figure–Ground Segregation

• A type of perceptual organization in which edges are assigned to regions for purposes of shape discrimination

14

15 Is PO learned or hard-wired?

• Schematic depiction of two types of displays: (a) homogeneous displays and (b) heterogeneous displays. From Spelke et al. (1993) with permission from Elsevier.

16

17

18

19 Gestalt Theory in the 21st Century

• Basic perceptual principles have survived
• Contentions that such processes are necessarily very "early" in perception have not
  – Proximity computation occurs after depth perception
  – Segmentation is not just "bottom-up" (Vecera &Farah, 1997)

20 Illusions are fun....

But their real value is that they reveal shed light on what is stored about object structure

21

22

23

24

25

Depth Perception

Monocular depth cues – cues that provide depth information with one eye closed

- pictorial cues (e.g., linear perspective)
- aerial perspective (differing contrast)
- texture gradients (decreased detail from front to back)
- interposition
- shading (production of shadows by 3-D objects)
- familiar size as anchor
- motion parallax

26

27

28

29

30

31

32

33

Depth Perception (cont'd)

- **Binocular Cues to Depth**
  - Binocular disparity: different images to each eye as a function of object closeness; basis of stereoscopic vision
  - Binocular convergence: rotation of eyes depending on object closeness; brain receives efferent motor signals from ocular muscles and interprets distance accordingly
  - Motion parallax: differential perception of motion (speed and direction) as a function of **distance from perceiver**

34

35

How are depth cues combined?

- Additively
- Preferentially
- Multiplicatively
- Most data suggests <u>additivity</u>, though one cue may occasionally overwhelm another in certain circumstances

36

Color Vision

- Young-Helmholtz theory: three populations of cones
  - Wide receptivity to wavelenth (continuous, not categorical; i.e., not just RGB detectors)
- Opponent-Process theory: R-G, B-Y, B-W

– LGN cells show ↑ activity to some wavelengths and ↓ activity to others

- Synthesis: Two-stage theory (Trichromatic at the level of neurons; O-P at the level of fibers and pathways)

37

38 Multiple Visual Areas (Zeki)

- **V1 and V2**: responsive to color and form
- **V3 and V3A**: form and motion
- V4: color
- V5: motion

39 Color perception/cognition can be selectively impaired

- Achromatopsia: selective impairment of inferior visual cortex (V4?)
- Color anomia: a general visual-verbal disconnection (seen with other visual-verbal disturbances)
- Specific color aphasia: specific inability to name colors

40 Theories of Perception

- **“Indirect” perception**: perception is the result of interactive processes from without and from within – use of internal representations based on “top-down” and “bottom-up” processing
- **“Direct” perception** (Gibson): information from the visual world is sufficient to permit perception without the involvement of internal representations - focus in “bottom-up” processing, without the need to posit internal representations

41 “Direct” View (Gibson)

- **“Optic array”** contains all necessary visual information
- Layout of objects in space given by **texture gradients, optic flow patterns, and affordances (implied meaning of objects)**
- Perception involves **“picking up”** information through “**resonance**”

❖ Has had historical impact in restoring interest in the perceptual environment

❖ Has been criticized as being underspecified, and neglects role of knowledge in stimulus exploration

42 Data in Favor of “Indirect” View

- **Context effects in perception**: what is seen depends upon surround
- **Effects of emotion** (e.g., weapon focus)
- **Visual illusions**: idea is that general knowledge about objects is applied inappropriately to the perception of two-dimensional figures (e.g., Muller-Lyer)

43

44 Template Theories of Pattern Recognition

- **Template theories**: a miniature copy or template of each known pattern is stored in long-term memory
  - straightforward template theory
  - normalized template theory
- **Problems**
  - not adaptable
  - impose large storage requirements

45

46 Feature Theories of Pattern Recognition

- **Feature theories:** pat-terns consist of a set of specific features or attributes
- **Advantages:**
  - elementary features can combine to form multiple objects
- **Problems:**
  - context effects in perception
  - recognition can take place when features are occluded

47 Prototype Theories of Pattern Recognition

- Individual instances are not stored; what is stored is an “exemplar” or representative element of a category
- Recognition based on “distance” between perceived item and prototype
- Nature of computation still relatively unknown

48

49

50 Theories of Object Recognition

51 Object Recognition: Key Questions

- When does one object end and another begin? (grouping/segmentation)
- Viewpoint independence (perception of objects as objects, regardless of view)
- How do we know that that two things belong together, and how do we know that dogs are dogs and not cats? (categorization)

52 More Key Issues

- **The “binding problem”:** if perception of a given object is distributed throughout visual areas, how does the brain combine features to produce unitary percept?
- **Conscious vs. unconscious perception**: perception of some object or object qualities may proceed pre-consciously or without effort (a general principle)
  - Example: Threat

53

54

55 The Building Blocks of Object Recognition: Marr’s Representations

- **Primal Sketch**: 2-D representation of light intensity changes, information about edges, contours, and blobs
  - Raw: pure light intensity changes
  - Full: uses information to identify shapes
- **2 1/2-D Sketch**: depth and orientation of visible surfaces, shading, texture, motion, binocular disparity; observer-centered
- **3-D Sketch**: three dimensional description of objects independent of view

56 Marr-Hildreth Algorithm

- Attempts to account for development of primal sketch
- Idea of “blurred repres-entations”
- **“Zero-crossings”** identify edges within a visual image
- Only **‘reliable’** zero-cross-ings are kept
- **Four types of tokens**: edge-segments, bars, terminations, and blobs

57 Marr & Nishihara (1978)

- Development of 3-D sketch based on processing of more elementary shape primitives (basic primitive is a cylinder with a major axis)
- Hierarchical organization of primitives

- Concavities important in segmenting parts
- This is a completely computational, not empirical, model

58

59 Biederman's RBC Theory

- Objects built from primitives called "geons" (n = 36)
- Viewer must decide appropriate way a visual object is to be segmented into geons
- Two key components of decision:
  - locating concavity
  - deciding which edge information remains invariant across different viewing angles (invariant properties like curvature, parallelism, etc.)

60 Biederman's Recognition-by-Components Theory

- Adapted from Biederman (1987)

61 Recognition-by-Components Theory

- Biederman (1987): five invariant properties of edges
  - Curvature: points on a curve
  - Parallelism: sets of points in parallel
  - Cotermination: edges terminating at a common point
  - Symmetry: versus asymmetry
  - Collinearity: points sharing a common line

62 Biederman (1987). Participants were presented with degraded line drawings of objects. Recognition was much harder to achieve when parts of the contour containing concavity information were omitted than when other parts of the contour were deleted. This confirms the assumption that information about concavities is important for object recognition. Figures adapted from Biederman (1987).

Biederman's view is "viewpoint invariant"

63

64

65 Viewpoint-dependent and Viewpoint-invariant Theories

- Biederman (1987)
  - The ease of object recognition is *not* affected by the observer's viewpoint
- Tarr (1995), Tarr and Bülthoff (1995, 1998)
  - Changes in viewpoint reduce the speed and/or accuracy of object recognition
- Milner and Goodale (1995)
  - Dorsal pathway makes use of viewpoint-dependent information
  - Ventral pathway makes use of viewpoint-invariant information

66 Common Elements in Object Recognition Theories

- Edge coding
- Grouping or encoding into higher-order features
- Matching to a stored "structural representation"
- Access to semantic knowledge

67 ## Face Recognition

- Two general theories:
  - Neural substrate model (module)
  - Specialized processing model
- Face recognition can be selectively impaired (prosopagnosia), or can it?
- Global (configurational) vs. local (feature-based) processing

68

69

70

71

72

73

74 Examples of “Greebles”. In the top row, four different “families” are represented. For each family, two members of different “genders” are shown (e.g., Ribu is one gender and Pila is the other). The bottom row shows a new set of Greeble figures constructed on the same logic but asymmetrical in structure. Images provided courtesy of Michael J. Tarr (Brown University, Providence, RI).

75 Speed of Greeble matching as a function of stage training and difference in orientation between successive Greeble stimuli. Based on data in Gauthier and Tarr (2002). Figure shows that, although speed increases with training, Greeble identification is still viewpoint dependent

76 ## Face Recognition

- Probably involves specialized processing/visual expertise
- Dedicated neural substrates aren’t the only answer
- Hybrid model: neural substrates dedicated to configural processing

77 ## Disorders of Object Recognition – Visual Agnosia

- Apperceptive agnosia
  - Object recognition is impaired because of deficits in perceptual processing
- Associative agnosia
  - Perceptual processes are essentially intact, but object recognition is impaired partly or mainly because of difficulties in accessing relevant knowledge about objects from memory

78 Riddoch and Humphreys (2001)

A hierarchical model of object recognition and naming, specifying different component processes which, when impaired, can produce varieties of apperceptive and associative agnosia.

79

80 ## Structural Descriptions

- **Structural descriptions** consist of propositions which “describe the nature of the components of a configuration and make explicit the spatial arrangement of these parts” (Bruce & Green, 1990)
- **Perceptual representation systems** for faces, objects, etc. Evidence for separate systems: category-specific recognition defects

81

82