Cognitve Psychology – A Students's Handbook 1. Mai 2002

Chapter 2
Visual Perception: Basic Processes

CHAPTER OVERVIEW (SUMMARY)

1. *Perceptional organisation*: The Gestalt psychologists identified several laws, eg:
- law of similarity,
- law of proximity,
- law of good continuation,
- law of closure,
- law of common fate

determing which elements of a visual display will be grouped together. They focused on lines and shapes and tried unsuccessfully to explain visual organisation in terms of electrical field forces in the brain.

2. *Depth and size perception:*
- Monocular cues: linear perspective, aerial perspective, texture, interposition, shading, familiar size, motion parallax
- Oculomotor cues: convergence, accomodation (both very limited usefull)
- Binocular cues: stereopsis (very important) involves establishing correspondences between the information presented to the one eye and the information presented to the other eye

Information from th various depth cues is additively combined to minimise errors in depth perception.
Size constancy depends mainly on perceived distance, but familiar size is also important.

3. *Brain systems:* Anatomically distinct systems in the visual cortex which are specialised for processing different kinds of information, eg: form, motion, colour:

V3: motion processing, exspecially processing of dynamic form and obtaining 3-dimens. Structure from motion
V4: colour processing
V3,V4,IT and other: form perception
V5: motion processing

INTRODUCTION

Roth's definition of perception: " The term perception refers to the means by which information acquired from the environment via the sense organs is transformed into experiences of objects, events sounds, tastes, etc."

1. PERCEPTUAL ORGANISATION

Although the basic visual information available to us consists of simple elements such as edges, lines and blobs we perceive an organised world consisting of people and objects. We are able to work out, which parts of the visual information belong together and so to form seperate objects: *perceptual segregation*.

Gestaltist approach

The Gestaltists made the first systematic attempt to study perceptual segregation and perceptional organisation:
fundamental principle of perceptional organisation: *law of Prägnanz* "Of several geometrically possible organisations that one will actually occur, which possesses the best, simplest and most stable shape."
Other laws (subsumed under the law of Prägnanz):

° *law of proximity*: elements that are close to each other are grouped together
° *law of similarity*: elements that are similar to each other are grouped together
° *law of good continuation*: elements requiering the fewest changes or interuption are grouped together in straight or smoothly curving lines (remember: Kreuz wird als bestehend aus zwei geschwungenen Linien wahrgenommen, nicht als zwei sich im Scheitel berührende Winkel)
° *law of closure*: missing parts of a figure are filled in to compete it

And: *law of common fate*: elements, that appear to move together are grouped together
Gestalt laws are descriptive statements possessing little or no explanatory power
Doctrine of isomorphism: assumed, there are electrical "field forces" in the brain to produce the experience of a stable perceptual organisation when we look at our visual environment

Section summary

The great strength of the Gestaltist approach: indentification of several important phenomena.
The important weaknesses: no experiments to find limitations and inadequacies of their theory and no real explanation ot the phenomena they had discovered.
And: Though they emphased (betonen) the importance of the law of Prägnanz, according to which the perceptual world is organised into the simplest and best shape, they could not say which shape is the simplest and best.

Subsequent theories

Restle's (1979) way of clarifying the notion of simplicity:
studied the perception of moving dots:
complicated approach: to treat each dot as completely seperated from all the other dots and to calculate starting position, speed, direction
but what is possible: to treat the dots as belonging to groups, especially if they move together in the same direction with the same speed
greater amount of processing or computation when dots are treated seperatly than as members of groups
main finding: whatever moving dots in a display would involve the least calculation generally correponds to what is actually perceived

Julesz (1975) identified grouping processes based on basic stimulus properties such as brightness, wavelength and granularity:
extended the Gestaltists work by sudying the effects of brightness and colour:
a visual display is perceived at consisting of two regions if the average brightness in each region differs considerably
but: no perceived difference if detailed pattern of brightness in each region was different but there was only a modest difference in the average brightness
a visual display is perceived as consisting of two regions if it consists of coloured squares and the average wavelength of light in each region is clearly different
discovered *some exceptions* to the notion that average brightness or wavelength is crucial in determining wether a display is perceived as consisting of two regions:
another important factor: granularity (way in which the elements in a region are distributed)
a visual display in which the overall brightness is the same but the granularity is greater in one half than the other is perceived as consisting of two regions
Suggest: processes involved in perceptual organisation occur very early in visual perception
Proposition of the theory of *textons* (basic elements of early vision):
textons consist of elongated blobs or line segments and their termination
they coorespond directly to the representations found in the raw primal sketch
the raw primal sketch contains only basic information about a visual display -> the initial perceptual organisation does not take account of subtle (fein) variations in texture

Full primal sketch

Workers in artificial intelligence made some progress in telling us in detail about processes involved in perceptual organisation, as Marr (1976, 1982) for example:
noted that basic perceptual organisation is achieved already at the level of the two-dimensional primal sketch
the raw primal sketch contains information about light intensity changes in the form of edges, blobs, bars and terminations, it is typically ambiguous
used two general principles to design a program to achieve perceptual organisation:
principle of explicit naming: give a name or symbol to a set of grouped elements that is used to describe other sets of grouped elements all of which can form a larger grouping
principle of least commitment: ambiguities are resolved only when there is evidence as to the appropriate solution
assigned *place tokens* to small regions of the raw primal sketch , such as the position of a blob or edge or the termination of a longer blob or edge
various edge points are incorporated into a single place token
place tokens than are grouped together in various ways:
Clustering: place tokens that are close together are combined ti form higher-order place tokens
Curvilinear aggregation: place tokens that are alined in the same direction are joined to form a contour

Section summary

Why Marr's visual processing program was so successful:
the grouping principles applied to place tokens reflect what is generally the case in the real world

it does not rely (sich verlassen auf) on object knowledge or expectations but works well nevertheless

The greatest limitation of Marr's appoach:

his assumption, that grouping is based on 2-dimensional representations; in general this is the case, but grouping can also be based on 3-dimensional representations

2. DEPTH PERCEPTION

How is the 2-dimensional retina image transformad into perception of a 3-dimensional world?

Different senses of the term "depth perception":

absolute distance: distance from the observer that an object is located

relative distance: distance between two objects (used when fitting a slize of bread into a toaster)

Judgements of relative distance are often more accurate than judgements of absolute distance.

Cues to depth are often provided by movement either of the object or of the observer but here we focus on monocular cues and binocular and oculumotor cues that are available even if nothing move.

Monocular cues

Only require the use of one eye, but can be used with both eyes open as wel. Include:

linear perspective: parallel lines pointing directly away appear closer together as they recede into the distance (Schienen)

arial perspective: because of light scattering more distant objects lose contrast and appear hazy (verschwommen) -> objects with less contrast appear to be more distant

texture: textured objects (cobble-stoned road) slanting away have a texture gradient, density (Dichte) increases with distance

interposition: a nearer object hides a part of more distant object (remeber: vier Kreise, denen je ein Viertel fehlt scheinen wie von Quadrat verdeckt)

shading: a shade provides evidence that there must be a 3-dimensional object as 2-dimensional objects do not have shades

familiar size: use of retinal image size of an object to provide an accurate estimate of its distance, only possible if actual size of the object is known

motion parallax: refers to the movement of an object's image over the retina (caused by the move of the object as well as by the move of the observer's eye), motion parallax can generate depth information without all other cues

Binocular and oculomotor cues

Involve both eyes being used together, are effictive only in facilitating depth perception over short distances , include:

Oculomotor cues (less important):

convergence: eyes turn inwards to focus on an object to a greater extent close with a very close object than with one that is further away

accomodation: tickening of the lens of the eye when focusing on a close object

Binocular cue (more important):

stereopsis: stereoscopic vision depending on the disparity in the images projected on the retinas of the two eyes

Three rules to match up information from the two eyes:
Comptibility constraint: elements from the input of each eye are matched with each other only if they are campatible (eg: same colour)
Uniqueness constraint: each element in one image can match with only one element in the other image
Continuity constraint: matches between to elements are prefered where the disparities between the two images are similar to the disparities between nearby matches on the same surface.... This is the least adequate constraint.

Integrating cue information

How is the information from the various cues combined and integrated?
Additivity: all the information from different cues is added together.
Exceptions:
there can be a situation where one depth cue is overhelmed by another, which then is ignored, eg: experiment where the interposition overhelms the cue of familiar size

Size constancy

is the tendency for any given object to appear the same size whether its size in the retinal image is large or small (imagine someone walking towards you, their retinal images increases but their size seems to remain the same).

Why is it important?
we take account of an objects's apparent distance when judging its size

"size distance invariance hypothesis":
for a given size of retinal image, the perceived size of an object is proportional to its perceived distance
example: size constancy is often not shown when we look down at objects from a high building or a plane because it is hard to estaimate the distance

In sum size constancy depends on many factors including perceived distance, size familiarity, the horizon and so on.

3. BRAIN SYSTEMS

more than 30 visual areas in the cortex
in retina: great majority of M (magnocellular or largebodied) and P (parvocellular or smallbodied) cells
their axons form the optic nerve which projects on the lateral geniculate nucleus (LGN)
the LGN is organised in 6 layers:
layer 1 and 2 receive input from M ganglion cells -> Magno layers: movement detection
layers 3-6 receive input from P ganglion cells -> Parvo layers: colour, fine textures, detailed objects
ipsilateral projections (from the eye at the same side) terminate in layers 2, 3 and 5
contralateral projections (from the eye at the other side) terminate in layers 1, 4 and 6
neurons from the P layers and from the M layers mainly project to the primary visual cortex, V1

after having passed V2 there are two pathways proceeding further into the cortex:
parietal pathway: mainly concerned with movement processing ("Where"pathway for action)
temporal pathway: concerned with colour and form processing ("What"pathway for perception)
different parts of the cortex are specialised for different visual functions:
V1 and V2: colour and form, contain "pigeonholes" (Ablegefach) into which the different signals are assembled before being relayed to the specialised visual areas
V3, V3A, (V4) and IT (inferotemporal cortex): form (especially the shapes of objects in motion) but not colour
(V1 and) V4: mainly colour, many cells are also responsive to line orientation
V5: visual motion

Blindsight

Partial or total damages in V1 cause a loss of vision in part or all of the visual field, in spite of this loss of conscious vision, some of these patients can make accurate judgements and discriminations about visual stimuli. (Anm. d. Red.:) No real explanation seems to exist.

Intagration of information – the "binding" problem

Gazzaniga: "Visual perception is a devide-and-conquer strategy.":
each area provides its own limited analysis
-> processing is distributed and specialised

How is information about colour, form and motion combined? An approach:
oscillation-binding theory: neurons sometimes exhibit oscillatory activity, in which there are alternating bursts of high and low rates of firing, they could tend to oscillate in a synchronised way when they are responding to the same abject -> produce an integrated percept of the object
problem with this theory:
oscillation only seems to occur with moving stimuli! What about static stimuli?
Oscillations develop too slowly and last for too long
Tovée: *"not such a binding problem"*: there is only high visual acuity for stimuli presented to the fovea of the retina -> "tunnel vision effect, where only the visual information from the centre of the visual field is fully sampled and analysed", the different features of an object are probably integrated or combined "in a subsequent higher integrative cortical area".