Perception of Space

Source: Vision Science: Phisics & Phenomenology
Stephen E. Palmer

The Problem of depth ambiguity

Depth perception from 2-d images is inherently ambiguous
The Problem of depth ambiguity

[Palmer 99]
Four Stages of Visual Processing

Image of a simple scene
Image of a simple scene

[Palmer 99]

Edge Detection

[Palmer 99]
Surface-Based Representation

[Palmer 99]

Surface-Based Representation

[Palmer 99]
Object-Based Representation

Category-Based Representation
Neural pathways from the eye to visual cortex

**Visual Pathways**

- **Retina**
  - Bipolar Cells
    - Graded potentials
    - Intracellular recording
  - Ganglion Cells
    - 2D sheet parallel to retina
    - Discrete discharging
    - Extracellular recording
    - On-center cells
    - Off-center cells

LGN = Lateral Geniculate Nucleus
Possible neural wiring of cortical simple cells from LGN cells

Responses of on-center, off-surround ganglion cells

Receptive field structure of ganglion cells

[Palmer 99]
Lateral Geniculate Nucleus (LGN)

- Located in the thalamus
- 3D structure
- Receives input from both eyes
- Each cell is monocular
- Divided into 6 layers of 2D sheets of neurons

- Lower 2 layers: Magnocellular “Large”
  - Large cells
  - Sensitive to contrast
  - Not very sensitive to color
  - Large receptive field
  - Short response to retinal stimulation

- Upper 4 layers: Parvocellular “Small”
  - Insensitive to contrast
  - Highly selective to color
  - Small receptive field
  - More sustained response

Diagram of the mapping from eyes to layers of the LGN

Retinotopic mapping or topographic mapping:

Nearby regions on the retina project to adjacent regions of the LGN

[Palmer 99]
Striate Cortex
- Primary cortical region for vision
- Hubel and Wiesel (1959)
  - Simple cells
  - Complex cells
  - Hypercomplex cells

Simple cells
“Their responses to complex stimuli can be predicted from their responses to individual spots of light”

- Edges detectors (light-to-dark \ dark-to-light)
- Line detectors (bright or dark lines)

Receptive fields and response profile of vertically oriented simple cells in V1

[Palmer 99]
Complex cells (~ 75% of the cells in striate cortex)
- Nonlinearity
- Motion sensitivity
- Position insensitivity
- Spatial extension
- Receive input from multiple simple cells or from LGN

Response of a complex cell in visual cortex of a cat

Hypercomplex cells “end-stopped cells”
- Extending a line or edge beyond a certain length causes them to fire less vigorously than they do for a shorter line or edge

Response of a hypercomplex (end-stopped) cell in visual cortex of a cat to moving bars of different lengths
Striate Architecture

We know that:
- Cells in visual cortex respond selectively to a limited set of spatial features:
  - Orientation
  - Position
  - Size
  - Color
  - Motion

- Each hemisphere → Retinotopic map of half the visual field
- Retinal topography is preserved
- Distortion occurs due to cortical magnification of central areas relative to peripheral ones
  “small portion of the retina (fovea) occupies a disproportionately large area of cortex”

Computational Approach

Edge detection: luminance edge
- change in the reflectance of the surface (e.g. change of material)
- change in the amount of light falling on it (e.g. shadow)
- change in the surface orientation relative to the light source

Marr & Hildreth Zero-Crossing Algorithm

A neural model of the Marr-Hildreth zero-crossing algorithm

[Palmer 99]
- Vision is a heuristic process in which inferences are made about the most likely environmental condition that could have produced a given image.

- Most of the time in everyday circumstances we are right.

- We perceive the spatial arrangement of surfaces by determining:
  - Depth
  - Surface orientation

**Marr’s 2.5-D Sketch**

- Between the 2D properties of an image-based representation and the 3D properties of a true object-based representation.

[Palmer 99] Flowchart of depth-processing modules

---

Marr’s 2.5-D Sketch

[Palmer 99]
Sources of Depth Information

<table>
<thead>
<tr>
<th>INFORMATION SOURCE</th>
<th>Ocular/ Optical</th>
<th>Binocular/ Monocular</th>
<th>Static/ Dynamic</th>
<th>Relative/ Absolute</th>
<th>Qualitative/ Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation</td>
<td>ocular</td>
<td>monocular</td>
<td>static</td>
<td>absolute</td>
<td>quantitative</td>
</tr>
<tr>
<td>Convergence</td>
<td>ocular</td>
<td>binocular</td>
<td>static</td>
<td>absolute</td>
<td>quantitative</td>
</tr>
<tr>
<td>Binocular Disparity</td>
<td>optical</td>
<td>binocular</td>
<td>static</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Motion Parallax</td>
<td>optical</td>
<td>monocular</td>
<td>dynamic</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Texture Accretion/Deletion</td>
<td>optical</td>
<td>monocular</td>
<td>dynamic</td>
<td>relative</td>
<td>qualitative</td>
</tr>
<tr>
<td>Convergence of Parallels</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Position relative to Horizon</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Relative Size</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Familiar Size</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>absolute</td>
<td>quantitative</td>
</tr>
<tr>
<td>Texture Gradients</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>quantitative</td>
</tr>
<tr>
<td>Edge Interpretation</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>qualitative</td>
</tr>
<tr>
<td>Shading and Shadows</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>qualitative</td>
</tr>
<tr>
<td>Aerial Perspective</td>
<td>optical</td>
<td>monocular</td>
<td>static</td>
<td>relative</td>
<td>qualitative</td>
</tr>
</tbody>
</table>

Accommodation

- Ciliary muscles in the eye control the optical focus of the lens by temporarily changing its shape
- Information about the tension of the muscle = information about the distance of the focused object
- Weak source of depth perception
- Provide accurate depth information for objects at short range (beyond 6-8 feet accommodation provides little or no depth information )

Depth information from lens accommodation

- Thick Lens → Close
- Thin Lens → Far
Convergence

-Convergence - The extent to which the two eyes are turned inward to fixate an object
-Used by the visual system to determine distance to the fixated point only up to a distance of a few meters

Depth information from eye convergence

Convergence as a function of distance

Convergence Angle in degrees (a)

\[ d = \frac{c}{2 \tan \left( \frac{a}{2} \right)} \]

Distance in meters (d)

[Palmer 99]
Binocular Disparity

- **Stereopsis**: The process of perceiving the relative distance to objects based on the binocular disparity
- **Binocular disparity**: Lateral displacement in the two retinal images
- **Finger experiment**
- **Crossed and uncrossed binocular disparity**:
  
  When a point \( P \) is fixated, closer points are displaced outwardly in crossed disparity, whereas farther points are displaced inwardly in uncrossed disparity

Crossed versus uncrossed binocular disparity

---

Horopter: The set of environmental points that stimulate corresponding points on the two retinas

The horopter in the horizontal plane

The theoretical versus the empirical horopter in the vertical plane

[Palmer 99]
Effectiveness of stereoscopic vision is limited to a range of less than about 30 m.

Stereograms
Each eye is presented with an image that differs only by the appropriate lateral displacement of the same object.

Shape-first or Stereopsis first?

Flowchart for two theories of stereopsis

[Palmer 99]

Da Vinci Stereopsis

[Diagram showing various conditions for both eyes and occlusion in right and left images]
- Correspondence problem: The visual system starts with two images and has to discover which features in the left image correspond to which features in the right one.

- Random dot stereograms prove stereopsis-first

Random dot stereogram of a face

[ Gareth and Peter's SIRDS pages ]

Physiological Mechanisms

Theory: Two separate physiological systems are involved in stereoscopic depth perception

V1: Highly sensitive to small or zero disparities (binocular fusion within Panum’s area)

V2: Sensitive to large disparities in both crossed and uncrossed direction (depth perception in front and behind the horopter)
3D Displays

- Stereoscopic (aided viewing)
- Autostereoscopic (free viewing)

Stereoscopic Techniques

- Color-multiplexed displays (anaglyph)
  - The left and right-eye images are filtered with near-complementary colors (red-green, red-cyan, green-magenta)
  - Color-filter glasses are used for separation
  - Limited color rendition (binocular color mixture)

Example of an anaglyph

[Paul Bourke]
- Polarization-multiplexed displays
  - Linear or circular polarization
  - Orientation has to be good

- Time-multiplexed displays
  - The human visual system is capable of merging the constituents of a stereo pair across a time-lag of up to 50ms
  - Uses LCS (Liquid Crystal Shutter) glasses
  - Impairing cross-talk may result from the persistence of CRT phosphors

Autostereoscopic Techniques
- Holovideo: Electronic Holography
  - MIT Media Lab Spatial Imaging Group
  - Uses lasers
  - Fringe patterns
  - Currently, the scope of this approach is limited to very small and coarse monochromatic holograms
  - The pixel rate for a monochrome display with a diameter of 15cm, a depth of 20cm, and a viewing zone of 30° amounts to 2 Gigapixel/sec (at 30 Hz frame rate)

- Parallax methods
  - Lenticular sheets

[Paul Bourke](http://www.paulbourke.net/visions/3d/)
[Paul Bourke](http://www.paulbourke.net/visions/3d/)
Motion Parallax

- The images of points at different distances from the observer move at different retinal velocities as the observer’s station point changes.
- Finger experiment.
- Highly used in 2D “scrolling” video games (Game Boy Advance).
- Controlled experiments have shown that motion parallax is sufficient for depth perception when no other sources of depth information are present.
- Unlike binocular disparity, motion parallax can provide effective depth information at great distances.

Example of motion parallax

[Krantz, Hannover College]
Convergence of Parallel Lines

- All parallel lines in any plane converge to the same vanishing point

The Pyramid of Khufu

A draftsman’s drawing in two-point perspective projection

[Palmer 99]

Position Relative to the Horizon of a Surface

The long winding road

Example of position relative to the horizon

[Palmer 99]
Relative Size

An object's smaller size on your retina when it is farther away from you is called relative size.

Sculls by Gustave Caillebotte

- Heuristic used: Two otherwise identical objects have the same objective size so that their relative distances can then be determined from relative image sizes.

Example of relative size

[Palmer 99]
Familiar Size

- Most objects tend to have a characteristic size or range of sizes with which experienced perceivers are familiar

Texture Gradients

- Systematic changes in the size and shape of small texture elements that occur on many environmental surfaces
- Provide info about:
  - Depth
  - Orientation
  - Curvature

Heuristic: Objects forming the texture are similar in size

Example of natural texture gradients

[Palmer 99]
Edge Interpretation (contour)

- Specifies only ordinal depth relations
- 4 types of edge interpretation:
  - 1. Orientation edges
  - 2. Depth edges
  - 3. Illumination edges
  - 4. Reflectance edges

Four kinds of edges

- Orientation and depth edges in objects with flat surfaces are mutually exclusive
Shading Information

<table>
<thead>
<tr>
<th>Shading and Shadows</th>
<th>optical</th>
<th>monocular</th>
<th>static</th>
<th>relative</th>
<th>qualitative</th>
</tr>
</thead>
</table>

- Shading: Variations in the amount of light reflected from the surface as a result of variations in the orientation of the surface relative to a light source.

- The brain implicitly assumes that illumination comes from above.

Direction of illumination and perceived convexity

Top row looks like convex bumps \ Lower row like concave dents

Aerial Perspective (Atmospheric Perspective)

<table>
<thead>
<tr>
<th>Aerial Perspective</th>
<th>optical</th>
<th>monocular</th>
<th>static</th>
<th>relative</th>
<th>qualitative</th>
</tr>
</thead>
</table>

- Refers to certain systematic differences in the contrast and color of objects that occur when they are viewed from great distances.
- Contrast is reduced by the additional atmosphere through which they are viewed (additional particles scatter light).

Example of atmospheric perspective

[Palmer 99]
Conclusion

Depth perception is a vital element of the human perception system.
Depth perception mechanism uses a lot of heuristics and provide accurate results in most everyday life circumstances.
It is still not well understood how the different depth cues are integrated together.
We do know however from experiences that there exist some sort of hierarchy between them.

Discussion