Suggested reading:

- Read JCA (2005) Early computational processing in binocular vision and depth perception. Progress in Biophysics and Molecular Biology 87: 77-108

Depth Perception

Contents:

- The problem of depth perception
- Depth cues
- Panum’s fusional area
- Binocular disparity
- Random dot stereograms
- The correspondence problem
- Early models
- Disparity sensitive cells in V1
- Disparity tuning curve
- Binocular energy model
- The role of higher visual areas
The problem of depth perception

Our 3 dimensional environment gets reduced to two-dimensional images on the two retinae.

The “lost” dimension is commonly called depth.

As depth is crucial for almost all of our interaction with our environment and we perceive the 2D images as three dimensional the brain must try to estimate the depth from the retinal images as good as it can with the help of different cues and mechanisms.

Depth cues - monocular

Accommodation
(monocular, ocular)

Texture Accretion/Deletion
(monocular, optical)

Motion Parallax
(monocular, optical)

Convergence of Parallels
(monocular, optical)
**Depth cues - monocular**

- **Position relative to Horizon** (monocular, optical)
- **Familiar Size** (monocular, optical)
- **Relative Size** (monocular, optical)
- **Texture Gradients** (monocular, optical)

**Depth cues - monocular**

- **Edge Interpretation** (monocular, optical)
- **Shading and Shadows** (monocular, optical)
- **Aerial Perspective** (monocular, optical)
**Depth cues – binocular**

**Convergence**  
(binocular, ocular)

**Binocular Disparity**  
(binocular, optical)

As the eyes are horizontally displaced we see the world from two slightly different viewpoints.

The visual fields of the two eyes overlap in the central region of vision but points that are not on the so-called horopter fall onto different retinal positions.

This lateral displacement which is relative to the fixation point is called **binocular disparity**.

**Panum's fusional area**

If we see different views of the world with both eyes why don't we experience double vision?

Points in an area near the horopter (called **Panum's fusional area**) are fused perceptually into a single experienced image.
Panum's fusional area

You can easily check this perceptual fact for yourself

Hold one forefinger in front of you and the other forefinger 30cm behind it. If you focus on one finger you will see the other finger twice because it is not within the limits of Panum's fusional region.

Binocular disparity

The point that we fixate on (here P) by definition has no binocular disparity as it falls directly onto the fovea at the same retinal position in both eyes.

Fixating a point P the image of points like C that are closer to the viewer than P are displaced outwardly on the two retinae in a so-called crossed disparity, whereas farther points as F are displaced inwardly in a so-called uncrossed disparity.

The direction and amount of binocular disparity therefore specify relative depth information in regard to the horopter.
Binocular disparity

Example of the use of binocular disparity for the experience of depth - Anaglyphs

Nothing new, discovered in the 1850s by Joseph D’Almeida and Louis Du Hauron, William Friese-Green created the first 3D anaglyphic motion pictures in 1889.

Random dot stereograms

Until about 40 years ago, a viable theory was that the visual system operates on the two retinal images separately, performing scene segmentation and object recognition, and then finally, for each object in the scene, compares where its images fall in the two retinas.

However, in the 1960s, Bela Julesz drew attention to the fact that stereopsis works perfectly well with random dot stereograms.
Random dot stereograms

1. two identical images with random dots
2. cut out a shape that you want to appear in a different depth plane and slide it to the right in one image – the amount will determine the depth
3. fill the empty hole that has appeared through the sliding with random dots
4. put the original image left to the newly created one – that’s it!

Random dot stereograms show that binocular disparity is such a strong depth cue that it does not need any monocular cue for depth perception.

The correspondence problem

Something is missing – the correspondence problem!

Until now we have talked about how the direction and amount of disparity between corresponding image features in the two eyes can encode depth information but the problem we have excluded until now is how to determine which features in one retinal image correspond to which features in the other?

Computer Vision:
match contents of image parts using global cost functions.
The brains solution to correspondence problem

Constraints
1. **Color constancy**: corresponding points have the same color

2. **Surface opacity**: most surfaces in the world are opaque so only the nearest can be seen → *If A10 is a correct match, then B10, C10, D10 etc. cannot be correct.*

3. **Surface continuity**: surfaces in the world tend to be locally continuous in depth → *excitatory connections between close matches*

Note! The constraints are
Early models

First Marr-Poggio Algorithm – dynamic behavior of the network

![Diagram showing dynamic behavior](image1)

Disparity sensitive cells in V1

In the human visual system the inputs of the two eyes converge upon single cells in V1.

Position disparity

Phase disparity

![Diagrams showing position and phase disparity](image2)

Recent studies indicate that most disparity sensitive cells are hybrid, showing both position and phase disparity (Ohzawa, Freeman 1997/1999 – Livingstone, Tsao 1999 – Prince, Cumming, Parker 2002).
Disparity tuning curve

Three representative retinal stimulus configurations correspond to a point in space that is nearer than, at, or further away from the fixation point.

Disparity tuning curve

Idealized tuning curves after characterization by Poggio & Fischer (1977)

Disparity tuning: tuned-excitatory, near, tuned-inhibitory, and far cells.

In each graph, the horizontal axis is disparity and the vertical axis is response rate.

The curves are Gabor functions (the product of a cosine and a Gaussian); the angle represents the phase of the cosine component.
Binocular Energy model

Ohzawa et al., 1990; Qian, 1994, Read and Cumming 2007

A complex cell $C_x$ receives input from binocular simple cells (BS). Receptive fields of two BS that differ in their phase by 90° get combined to a complex cell $C_x$.

$$BS = E = (L + R)^2$$

$$C_x = BS + BS$$

Binocular Energy model

With the Gabor filter

$$\rho(x, y; \phi) = \cos(2\pi f x' - \phi) \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right),$$

with standard deviation

$$\sigma = \sqrt{\frac{\ln 2}{2\pi f}} \times \frac{2^{1.5} + 1}{2^{1.5} - 1}$$

and

$$x' = x \cos \theta + y \sin \theta; \quad y' = y \cos \theta - x \sin \theta$$

Image value at position $(x, y)$

$$L = \int_{-\infty}^{+\infty} dx dy I_L(x, y) \rho_L(x, y),$$

$$R$$ equivalent

Receptive field of left eye simple cell

$$\rho_L(x, y) = \rho(x + \Delta x_{\text{pref}}/2, y; \phi + \Delta \phi_{\text{pref}}/2);$$

$$\rho_R(x, y) = \rho(x - \Delta x_{\text{pref}}/2, y; \phi - \Delta \phi_{\text{pref}}/2).$$
Solution to the correspondence problem

V1 neurons are sensitive to local matches, not to global solutions of the correspondence problem.

Anti-correlated stereograms

With the discovery of V1 disparity sensitive cells the question arose if the correspondence problem was solved in V1 already or in a higher brain region.

One approach to test for this were anti-correlated stereograms. These consist of normal stereograms but the brightness information in one of the images is reversed in comparison to the other image. Because of this there is no global-match solution to the stereo correspondence problem in anti-correlated stereograms.

Nevertheless the disparity sensitive cells in V1 respond to the local false matches in anti-correlated stereograms which is a strong indication that the correspondence problem does not get solved in V1.

The role of higher visual areas


V5/MT: Microstimulation can bias the monkey’s judgement in a far-near stereo task that relies on absolute disparity (DeAngelis et al. 1998), but no strong support for other aspects of stereovision.

V4: sensitive to the orientation of disparity defined planes (Hinkle & Connor, 2005).
Vergence Control

Additional reading:

References: