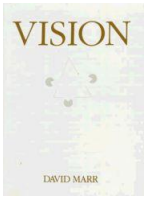
Binocular Stereo Vision

Properties of human stereo vision

Marr-Poggio-Grimson
multi-resolution stereo algorithm



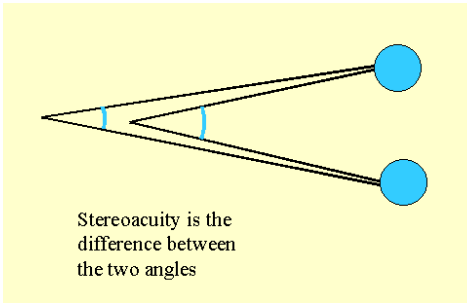
CS332 Visual Processing
Department of Computer Science
Wellesley College

Properties of human stereo processing

Use features for stereo matching whose position and disparity can be measured *very precisely*

Stereoacuity is only a few seconds of visual angle

difference in depth ≈ 0.01 cm
at a viewing distance of 30 cm



1-2

Properties of human stereo processing

Matching features must appear *similar* in the left and right images

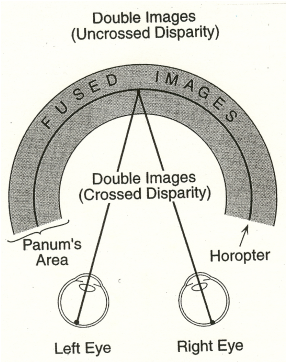
For example, a left stereo image cannot be *fused* with a negative of the right image…

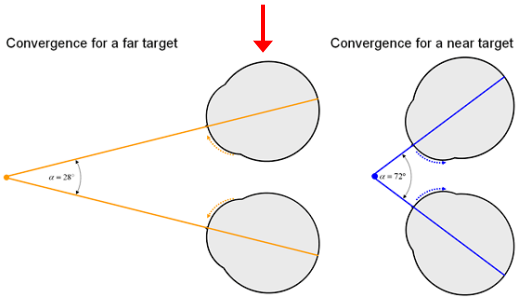
1-3

Properties of human stereo processing

Only "fuse" objects within a limited range of depth around the fixation distance

Vergence eye movements are needed to fuse objects over a larger range of depths

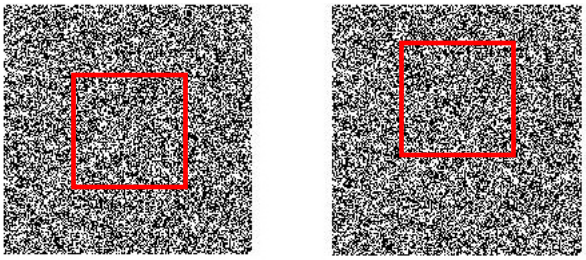




1-4

Properties of human stereo vision

Human visual system can only tolerate small amounts of *vertical disparity* at a single eye position



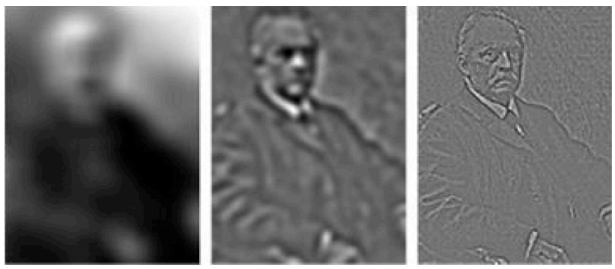
Vertical eye movements are needed to handle large vertical disparities

1-5

Properties of human stereo processing

Hermann von Helmholtz

In the early stages of visual processing, the image is analyzed at *multiple spatial scales* ...

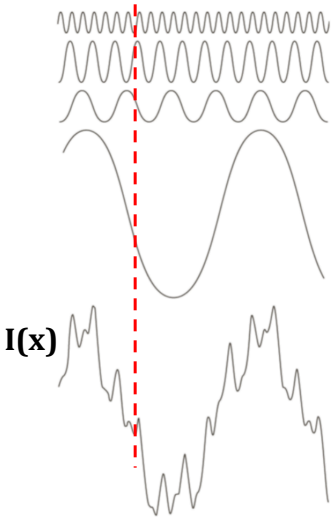


coarse medium fine

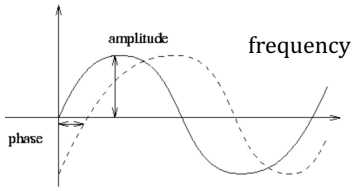
... that play an important role in the solution to the stereo correspondence problem

1-6

Spatial frequency decomposition

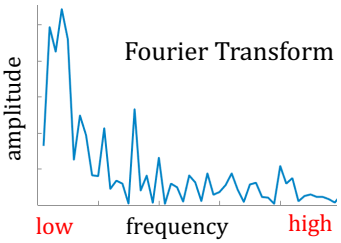


Any real signal, such as I(x), can be described as the sum of sinusoidal waves of different frequency, amplitude, and phase



frequency e.g. cycles/deg

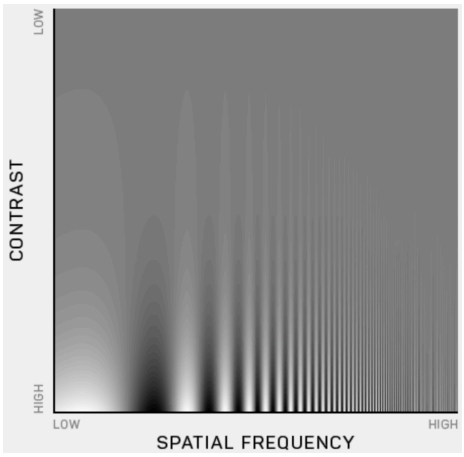
Two ways to describe 1D or 2D signals

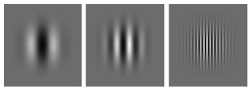


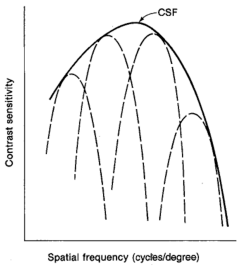
Fourier Transform

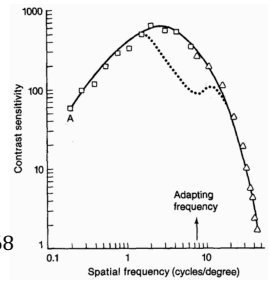
1-7

"Spatial frequency channels" in human vision









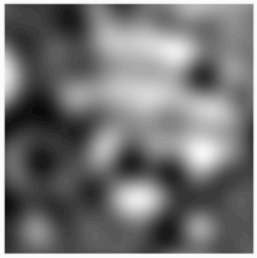
Campbell & Robson, 1968

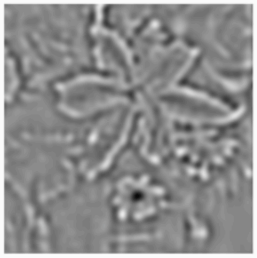
1-8

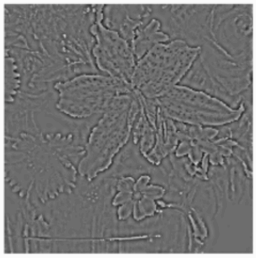
Spatial frequency channels

'Low' spatial frequency filters encode coarse luminance variations in the world (e.g. large objects, overall shape)

'High' spatial frequency filters respond to the fine spatial structure of the world (e.g. small objects, detail)







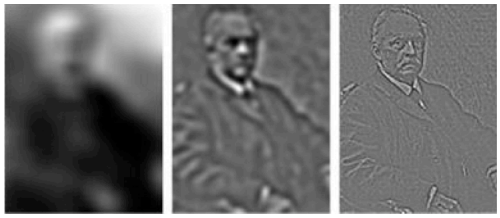
Coarse Medium Fine

1-9

Properties of human stereo processing

Hermann von Helmholtz

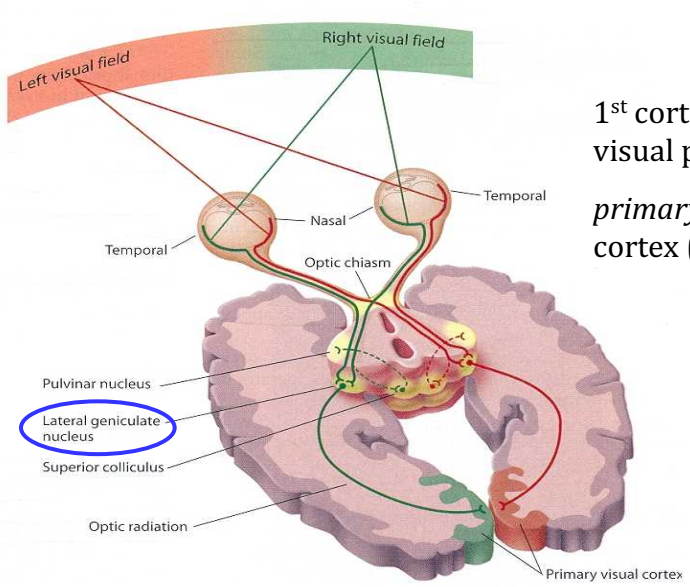
In the early stages of visual processing, the image is analyzed at *multiple spatial scales*...



- Stereo information at different scales can be processed independently
- Stereo information at coarser scales can be "fused" over a larger range of stereo disparity
- Stereo information at coarser scales can trigger vergence eye movements that narrow the range of stereo disparity present in the images

1-10

Projection from the retina

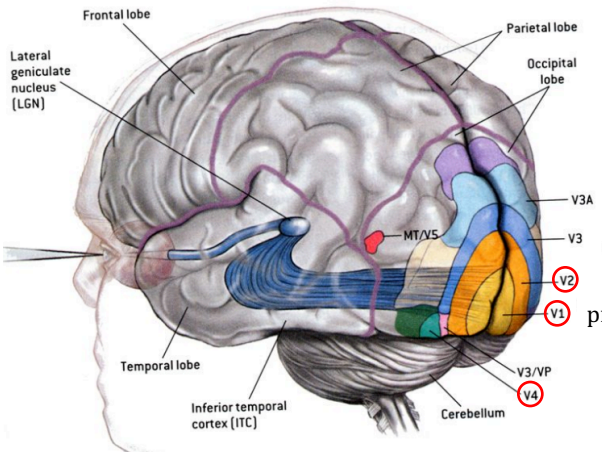


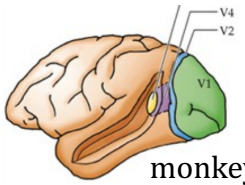
1st cortical stage of visual processing:

primary visual cortex (area V1)

1-11

Neural processing of stereo disparity





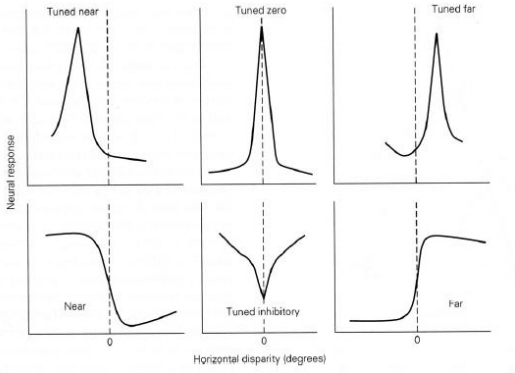
monkey

Extrastriate cortex

primary visual cortex

1-12

Neural mechanisms for stereo processing



From G. Poggio & others:

- neural recordings from monkey (area V1)
- viewing random-dot stereograms

zero disparity: at fixation distance

near: in front of fixation distance

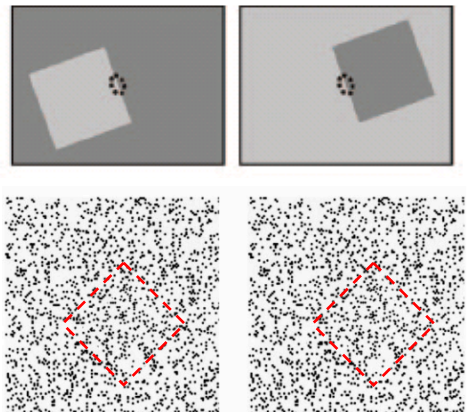
far: behind fixation distance

- (some) simple & complex cells in **area V1** are *selective for stereo disparity*
- neurons with large receptive fields are selective for a larger range of disparity

... but the stereo correspondence problem is ***not solved*** in V1!!

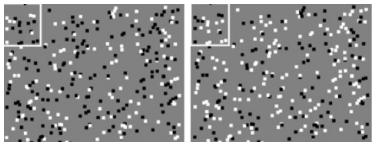
1-13

Selectivity for *stereo boundaries* in V2



Von der Heydt & colleagues:

Some V2 cells are selective for the orientation, contrast, and *side of border ownership* of an edge ... for edges defined by luminance *or stereo disparity*



"anti-correlated" stereogram

Later, in area V4, neural responses to stereo disparity appear to correspond more closely to perceived depth

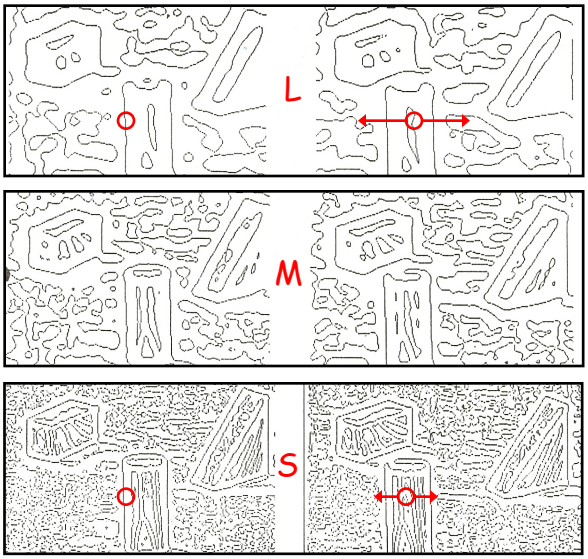
1-14

In summary, some key points…

- Image features used for matching:
  ~simple, precise locations, similar between left/right images
- At single fixation, match features over a limited range of horizontal & vertical disparity
- Eye movements used to match features over larger range of disparity
- Stereo matching performed at multiple scales
  ~independently, disparity range depends on scale
- Neurons selective for different ranges of stereo disparity, multiple processing stages V1 → V2 → V4

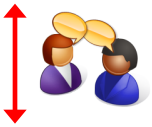
1-15

Matching features for the MPG stereo algorithm



zero-crossings of image convolutions with $\nabla^2 G$ operators of different size

rough disparities over large range



accurate disparities over small range

1-16