

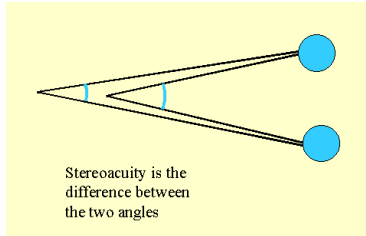
## Human stereo acuity



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

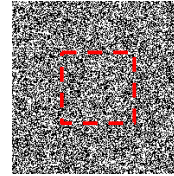
Stereoacuity is only a few seconds of visual angle

difference in depth  $\approx 0.01$  cm at a viewing distance of 30 cm

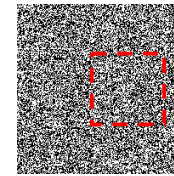


1

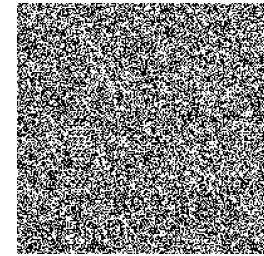
## Random-dot stereograms



left



right



- Bela Julesz, 1971
- stereo system can function independently
- we can match “simple” features
- highlight the **ambiguity** of the matching process

2

## Properties of human stereo processing

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



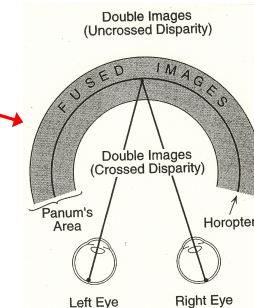
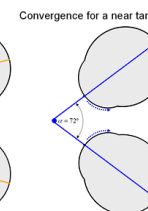
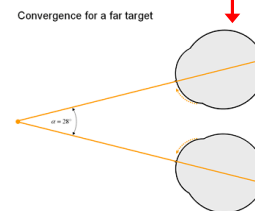
For example, a left stereo image cannot be *fused* with a negative of the right image

3

## Properties of human stereo processing

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

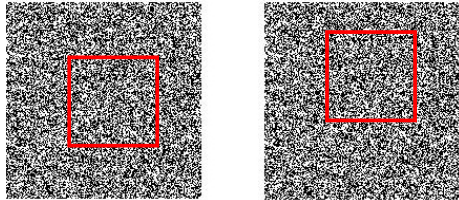
*Vergence eye movements* are needed to fuse objects over a larger range of depths



4

## Properties of human stereo vision

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

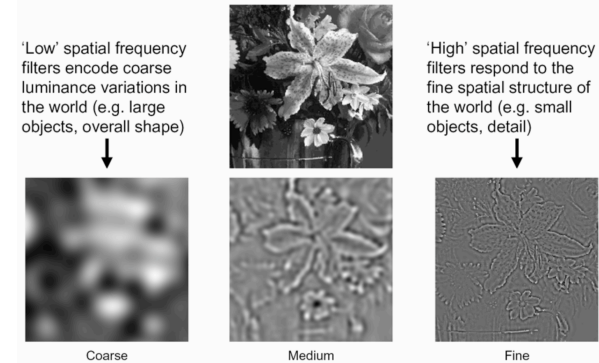


Vertical eye movements are needed to handle large vertical disparities

5

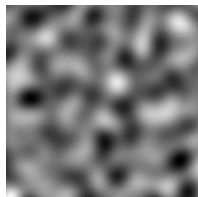
## Spatial frequency channels

In the early stages of visual processing, the image is analyzed at *multiple spatial scale* that play a critical role in stereo vision

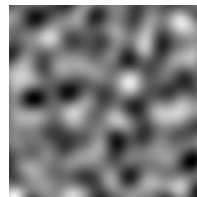


6

## Multi-scale stereo processing



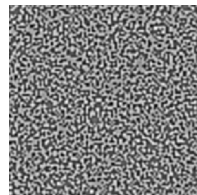
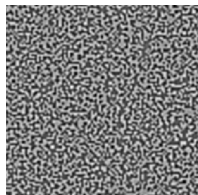
left



right

**Add together:**  
(like a hybrid image)

- coarse scale images with coherent stereo disparities
- fine scale images with uncorrelated "noise"

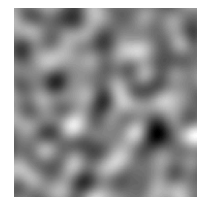


**Outcome:**

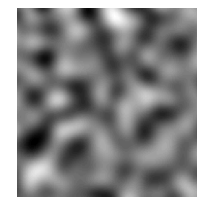
- can still fuse images
- tolerate large range of stereo disparity
- get a rough sense of depth

7

## Multi-scale stereo processing



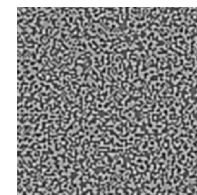
left



right

**Add together:**  
(like a hybrid image)

- coarse scale images with uncorrelated "noise"
- fine scale images with coherent stereo disparities



**Outcome:**

- can still fuse images
- only tolerate small range of stereo disparity
- more vergence eye movements

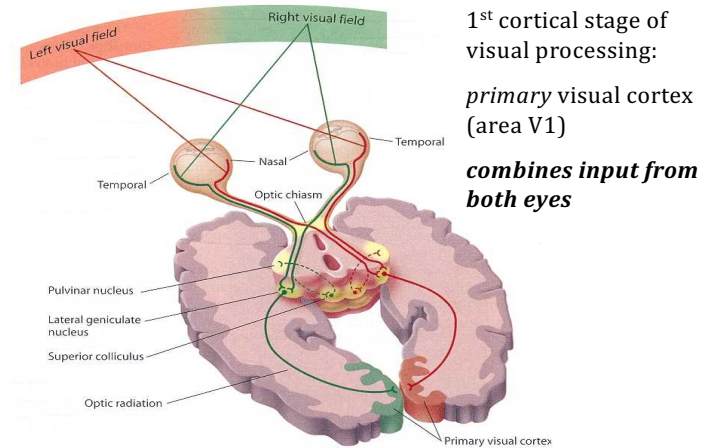
8

## Some key points from perception...

- Image features used for matching:  
~simple, precise locations, similar between left/right images
- At a single fixation, match features over a limited range of horizontal & vertical disparity
- Eye movements used to match features over larger range of horizontal & vertical disparity
- **Stereo matching is performed at multiple scales**
  - stereo information at different scales can be processed independently
  - information at coarser scales can be "fused" over a larger range of stereo disparity
  - information at coarser scales can trigger vergence eye movements that narrow the range of stereo disparity in the region of view

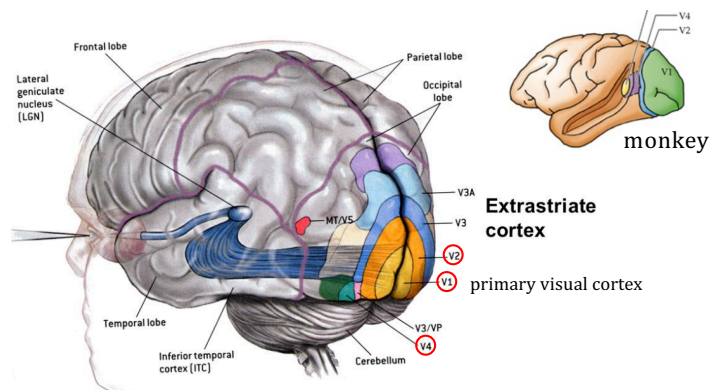
9

## Projection from the retina



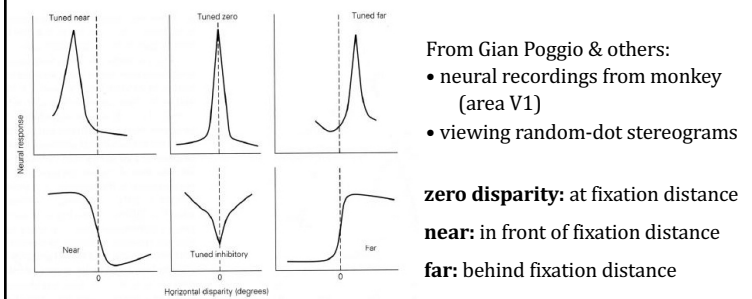
10

## Neural processing of stereo disparity



11

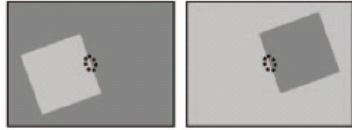
## Neural mechanisms for stereo processing



- (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!!

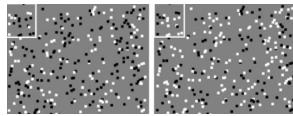
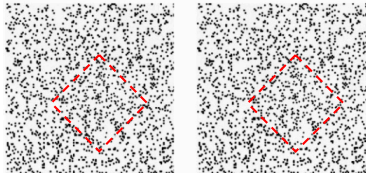
12

## 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