Image Processing and Computer Vision

An Introduction for Computer Scientists

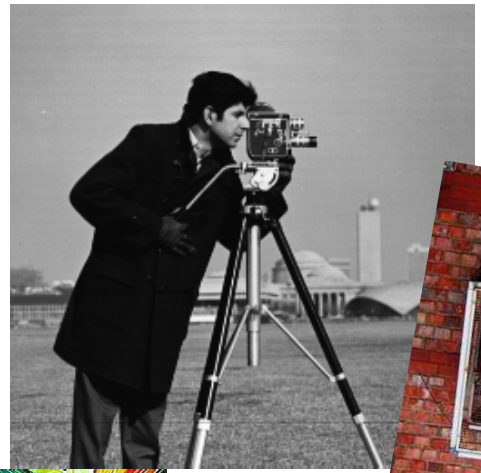
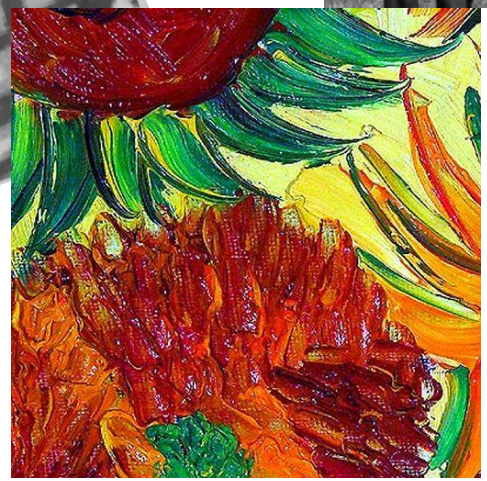
Rein van den Boomgaard Leo Dorst

University of Amsterdam

Outline

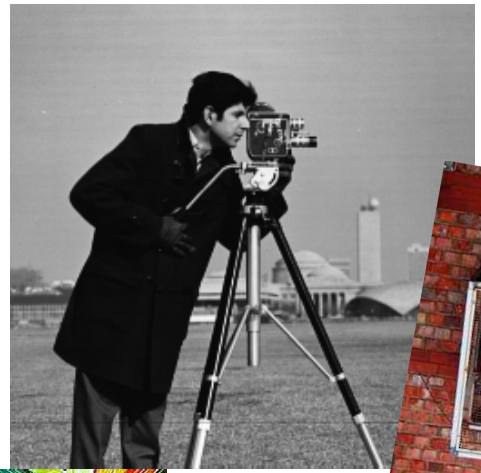
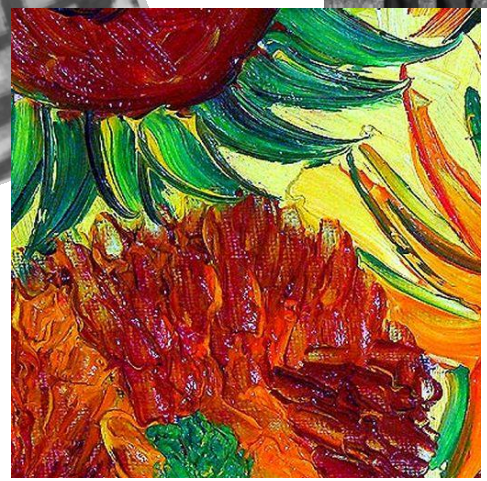
Images

- Images as we see them:


Images

- Images as we see them:

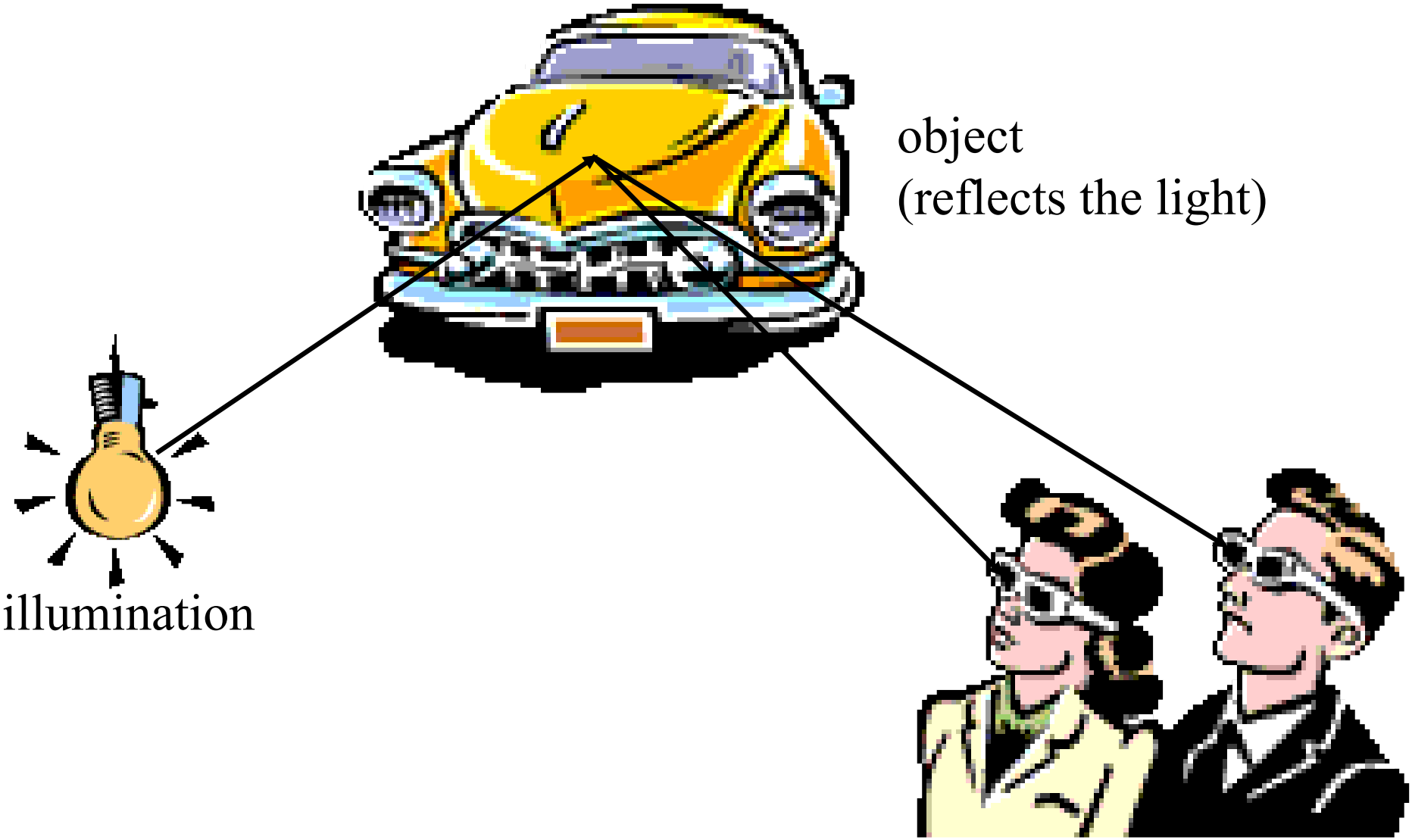





- but what are they?

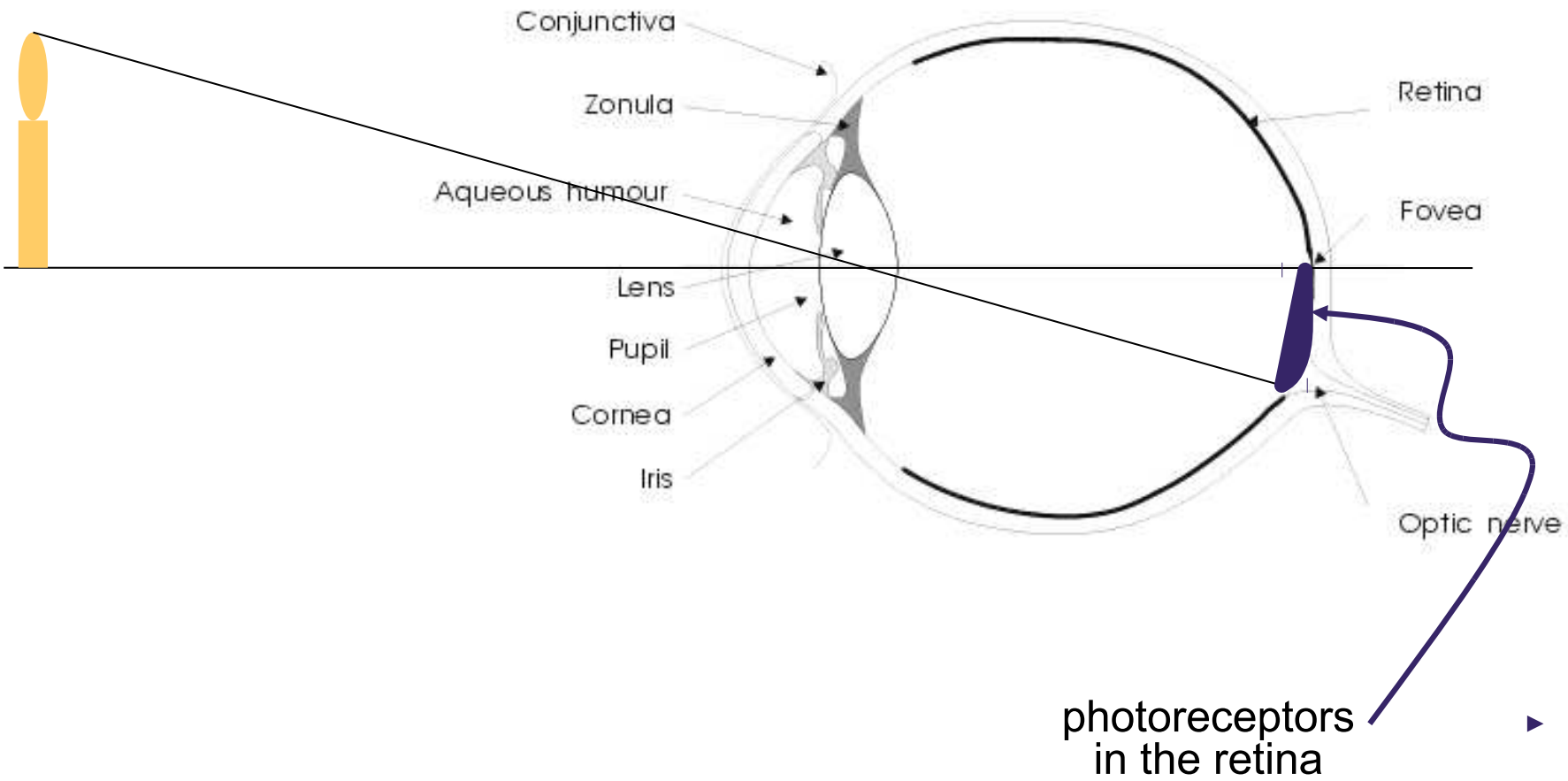
What is an image?

- Physiology, Psychology
- Physics
- Computer Science
- Artificial Intelligence

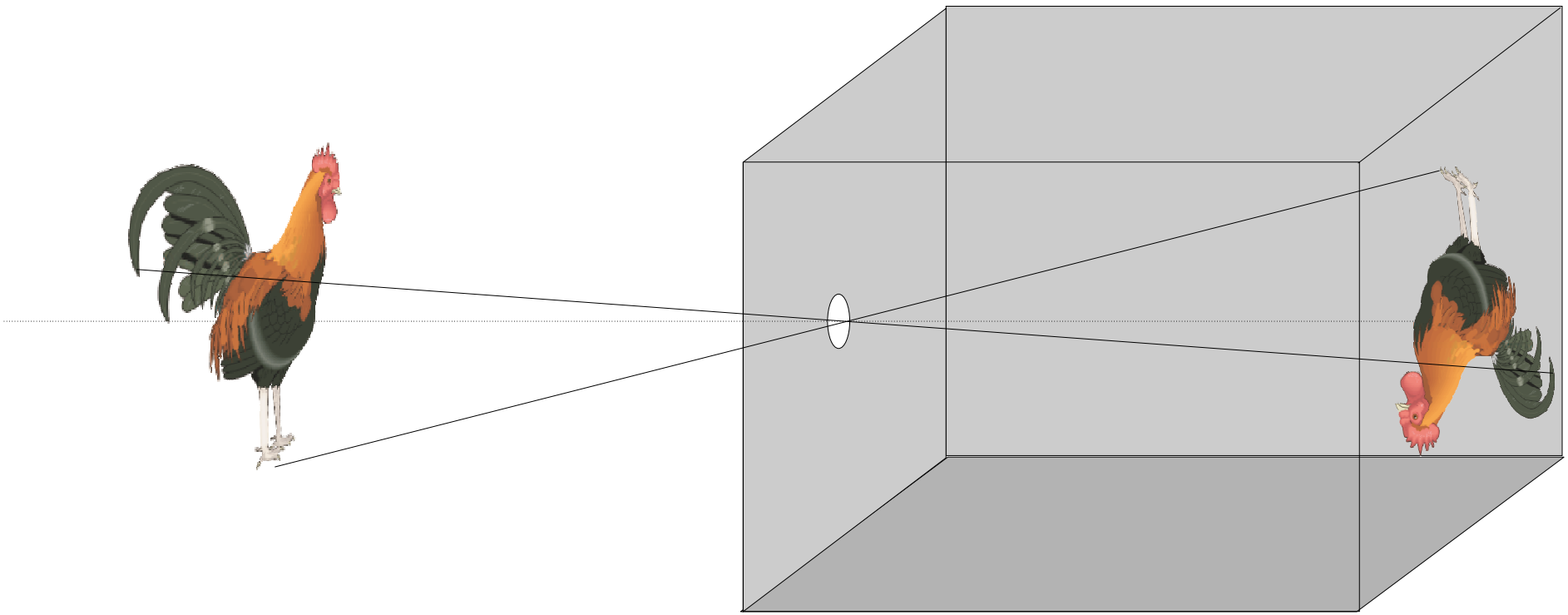
Optical Imaging



The Human Eye

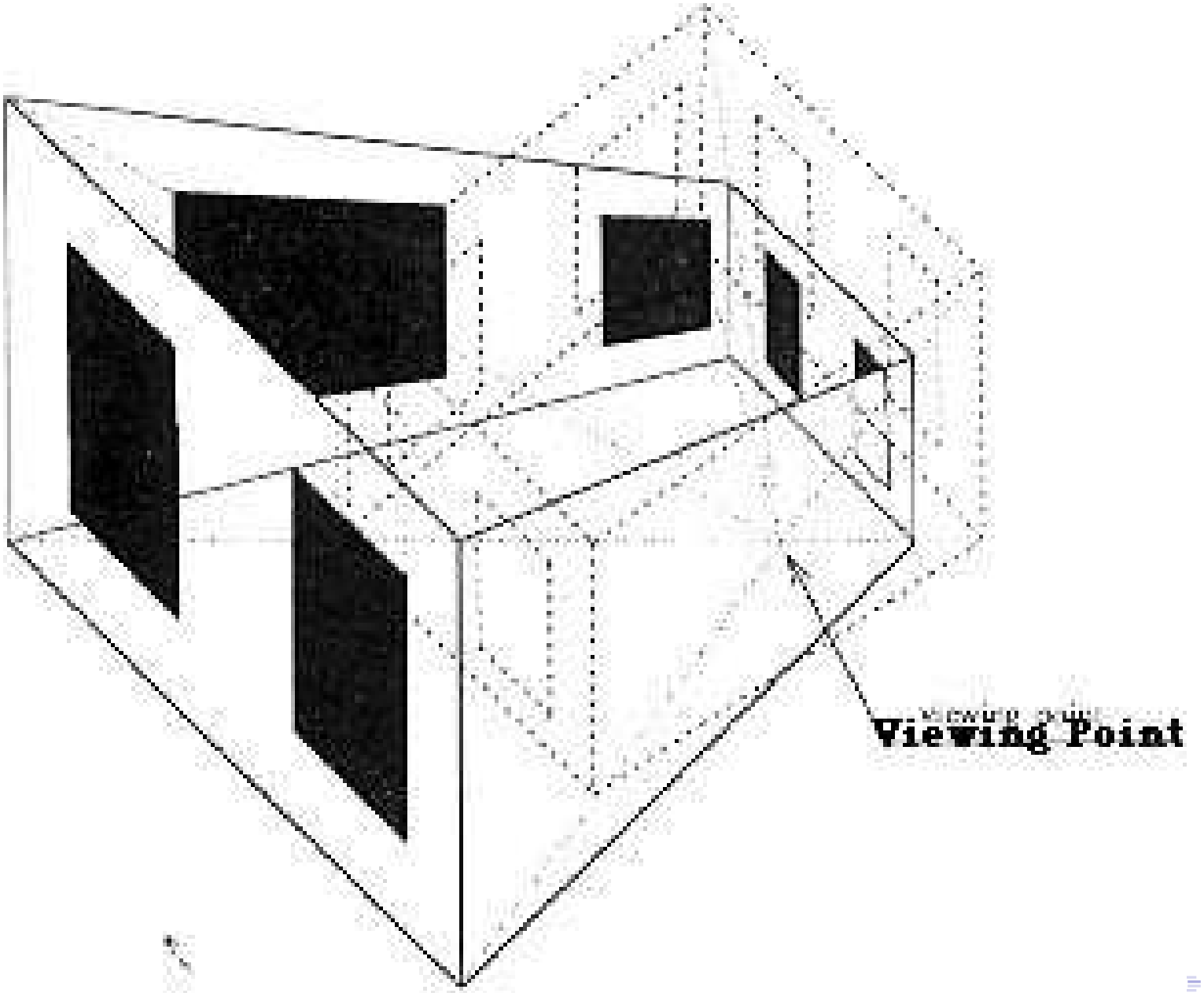


The Pinhole Camera



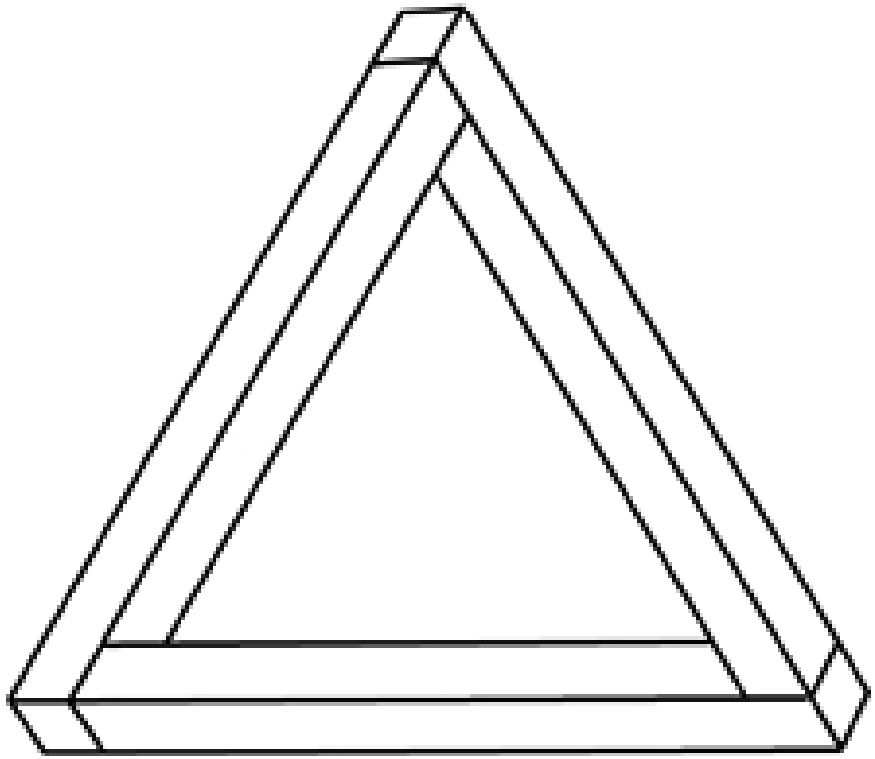
Ames Room

Ames Room



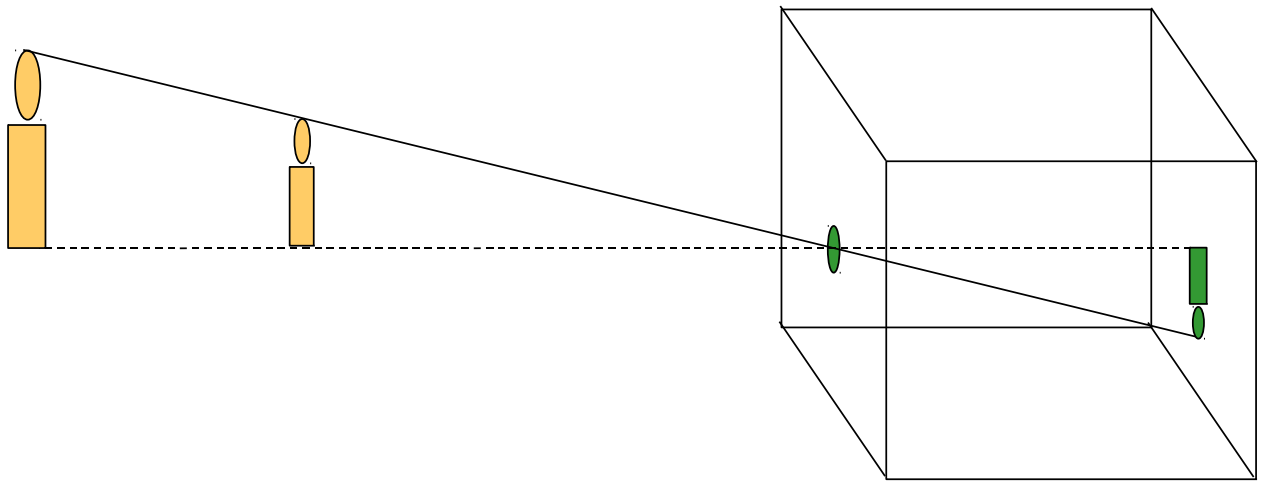
Emma's Room

Impossible Triangle



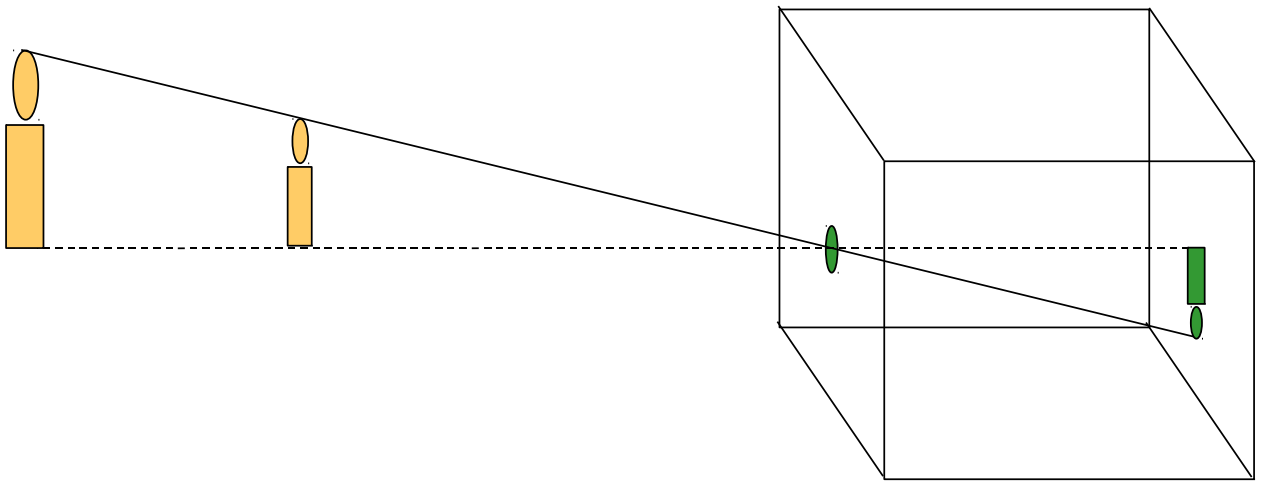
Impossible Triangle

Depth



- with one eye we can see no depth
- but then, how can we survive in a 3D world?
- can we teach a computer the same?

Depth

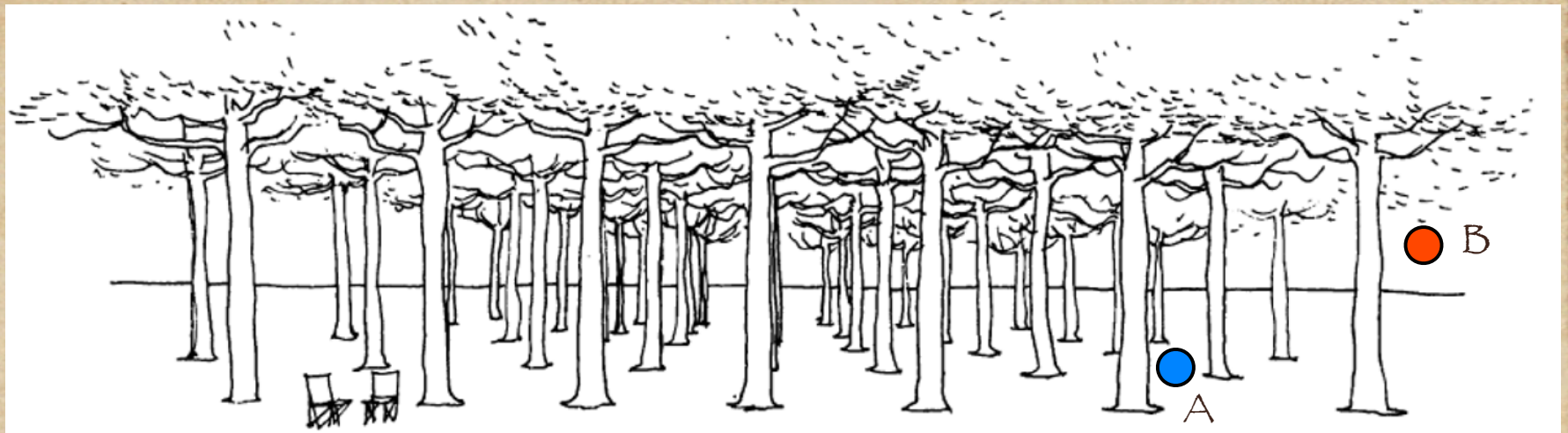


- with one eye we can measure no depth
- but then, how can we survive in a 3D world?
- can we teach a computer the same?

Depth

- we can ‘see’ depth with one eye,
- we can even ‘see’ depth in pictures,
- there are local cues for depth perception
- and our brain combines them using our familiarity of the 3D shapes surrounding us.

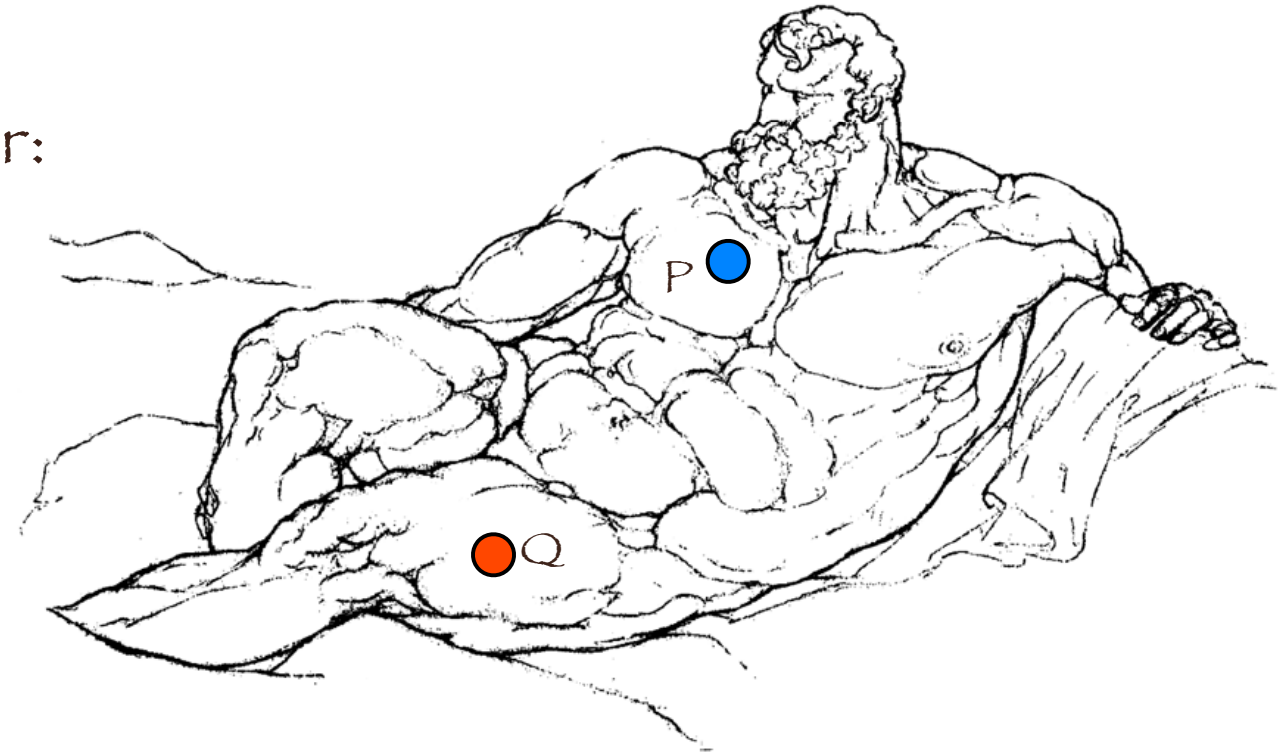
next some slides from lectures by J.J. Koenderink

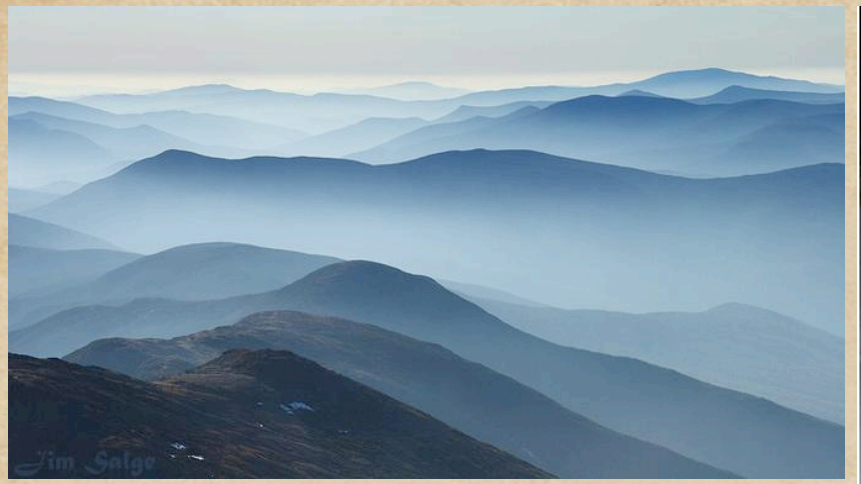


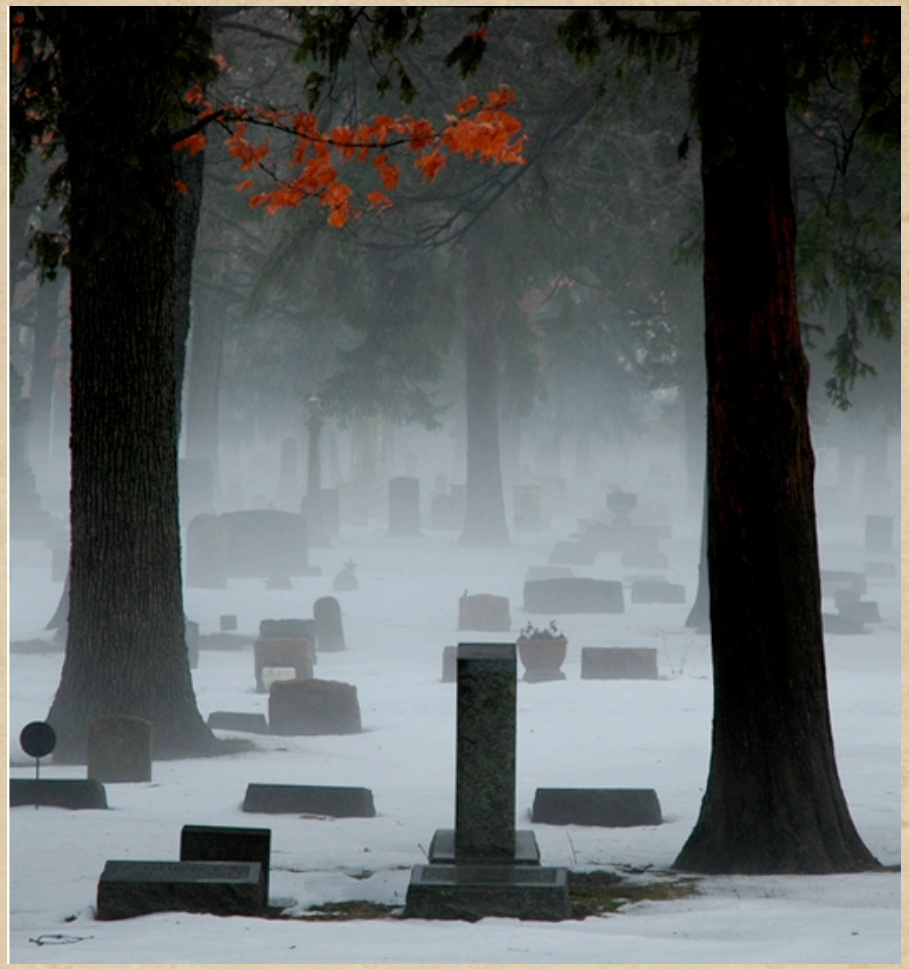
which is nearer:

P or Q?

A or B?

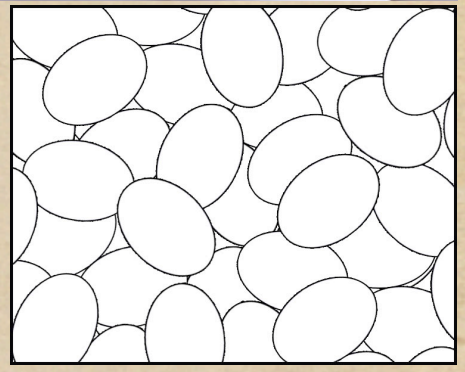




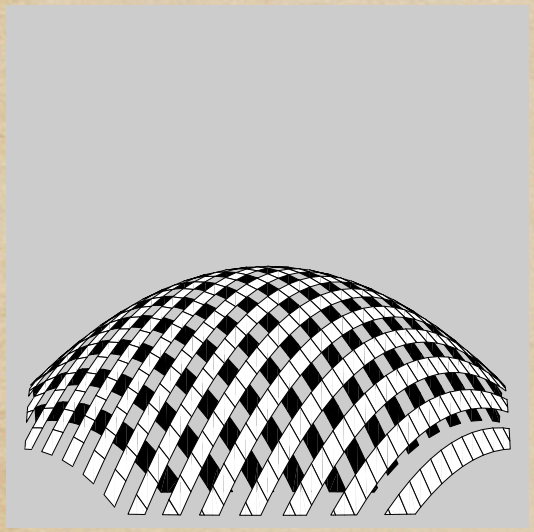


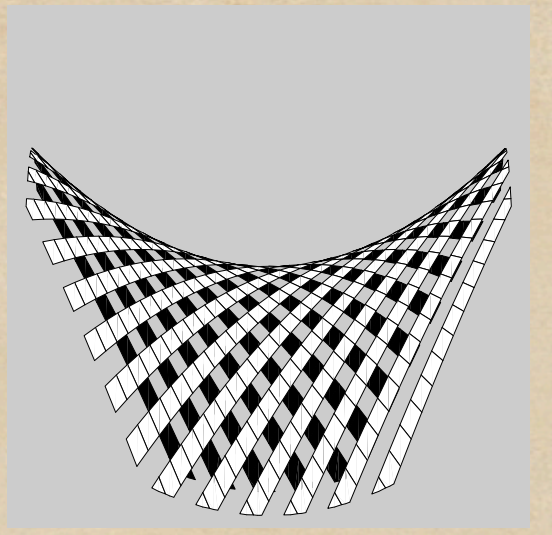
"Atmospheric perspective"

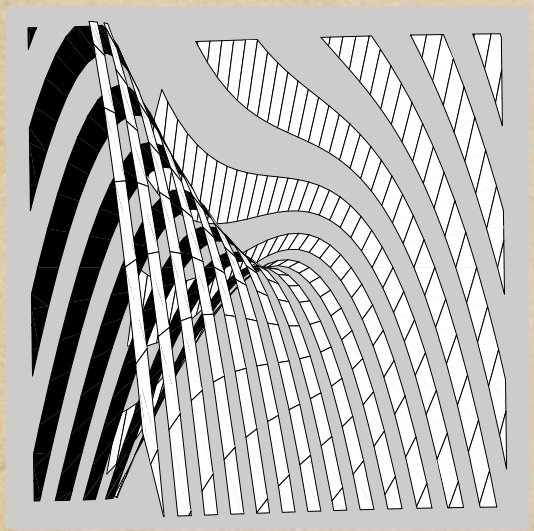
(contrast decreases exponentially with distance)

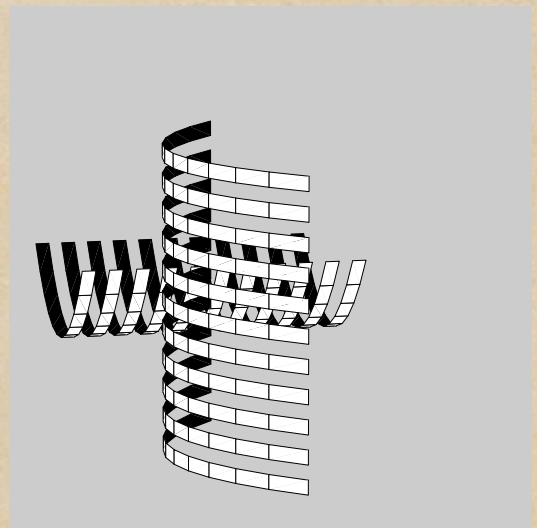


Occlusion forces an impression of depth order. However, the order relation is limited to few (2 or 3) depth layers.

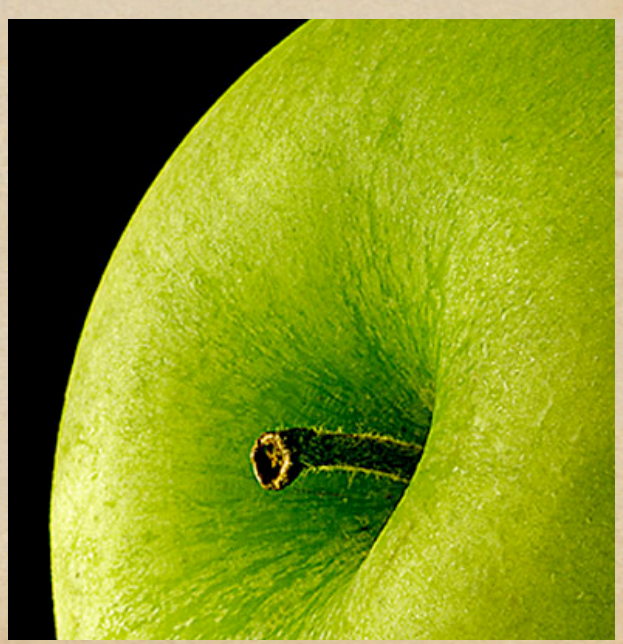


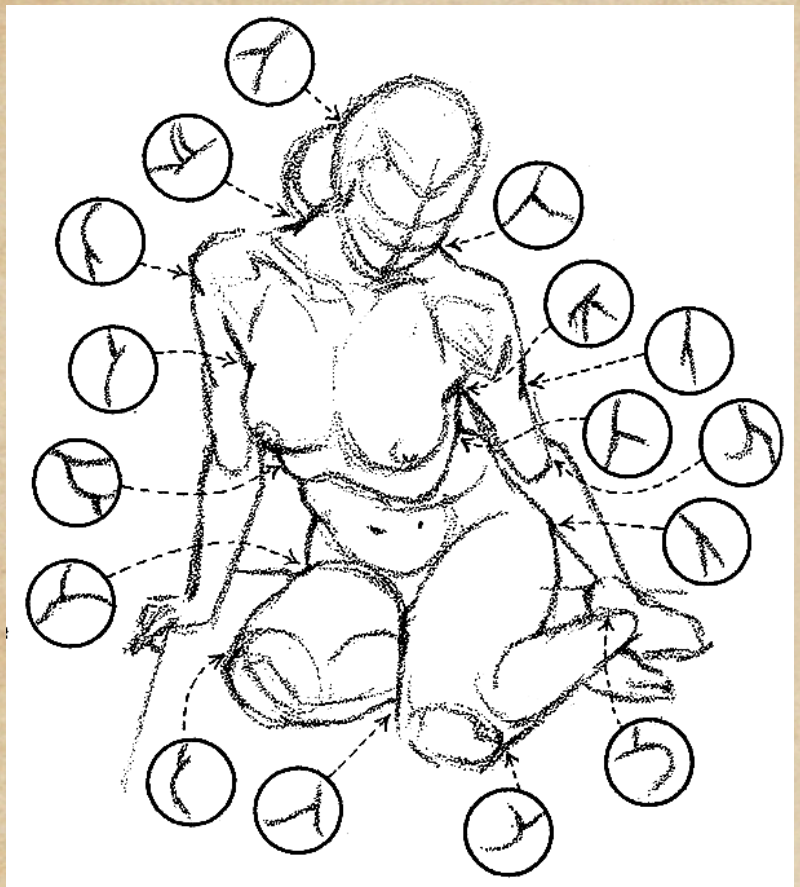


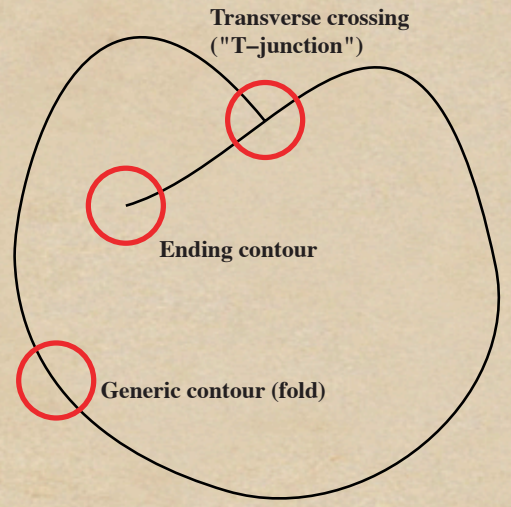


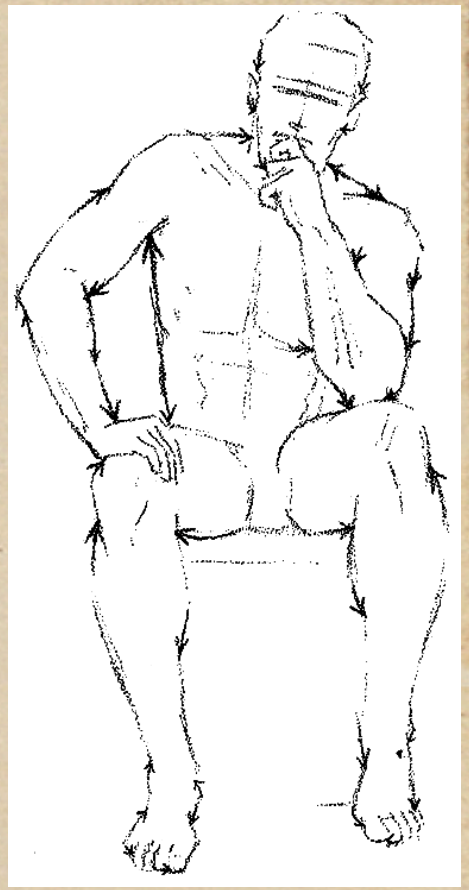


"generic singularities"

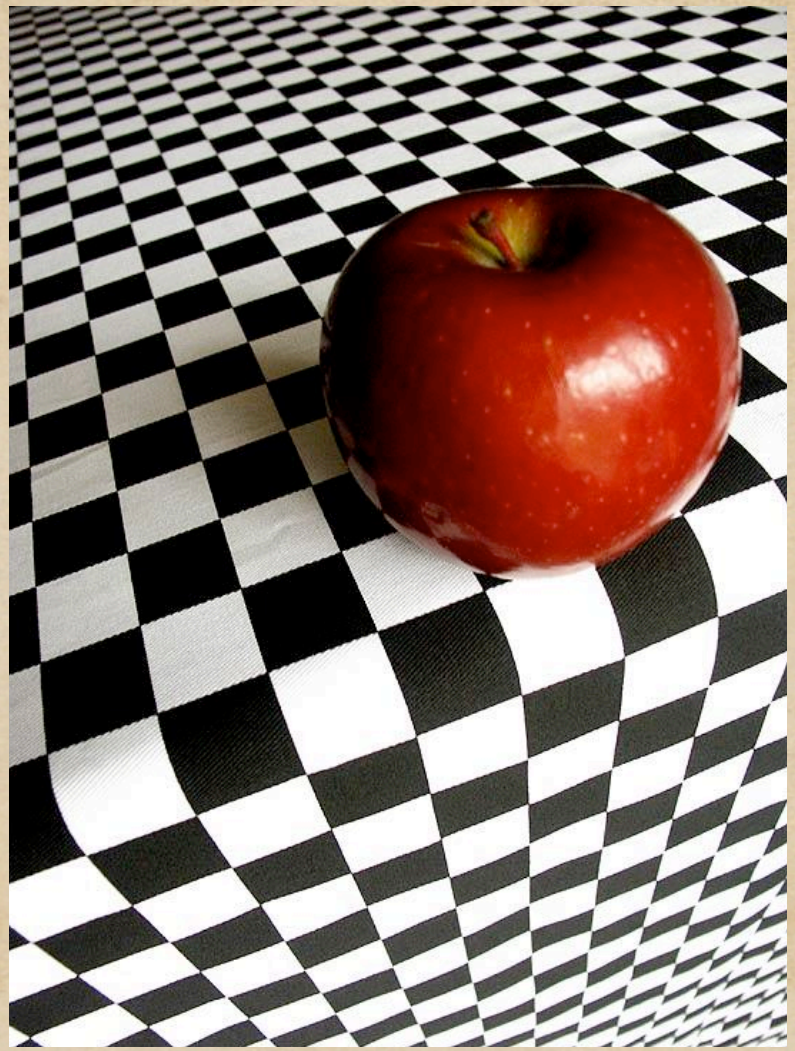








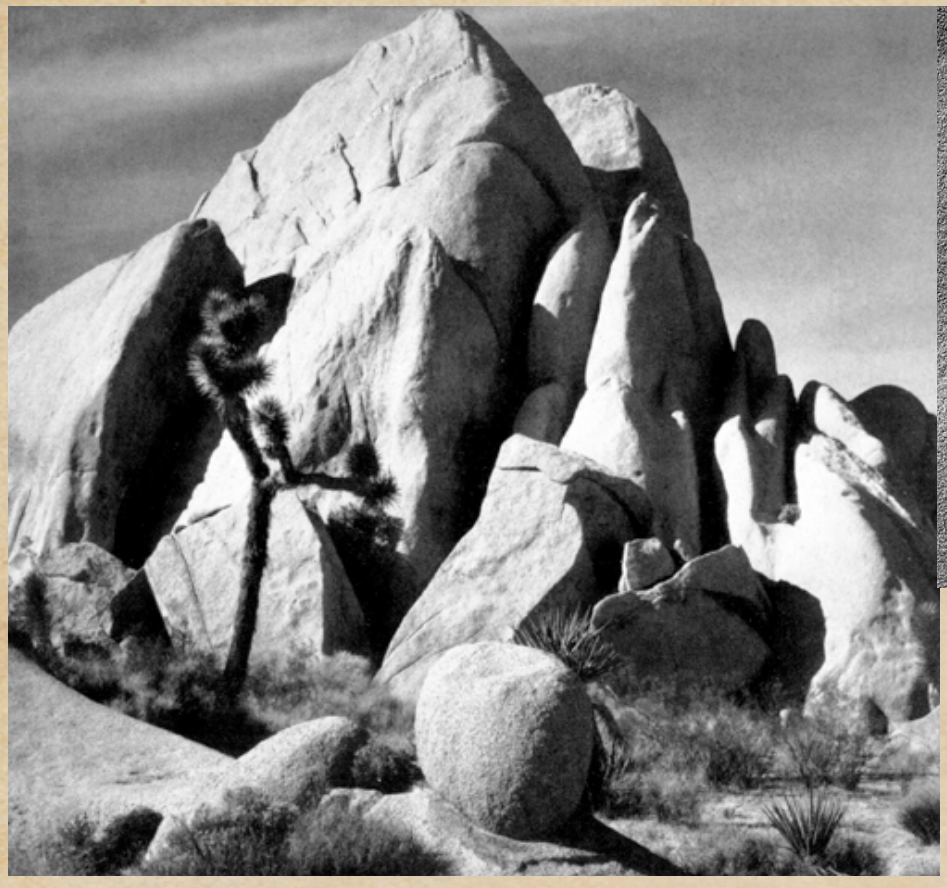
The visual contour is a twisted space curve. It yields a cue to depth order & local shape (surface curvature).

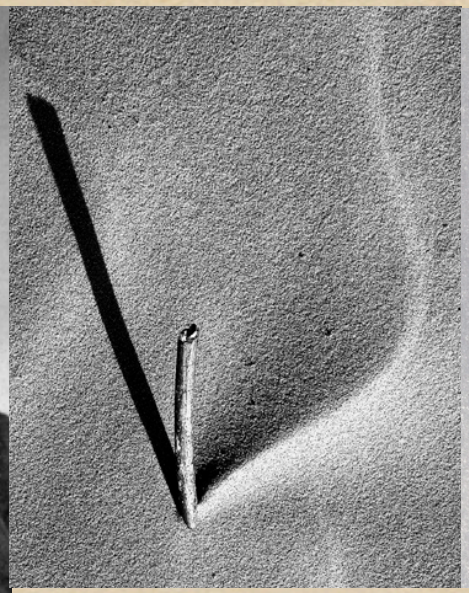


The "texture gradient" cue presupposes a statistical regularity of the world.

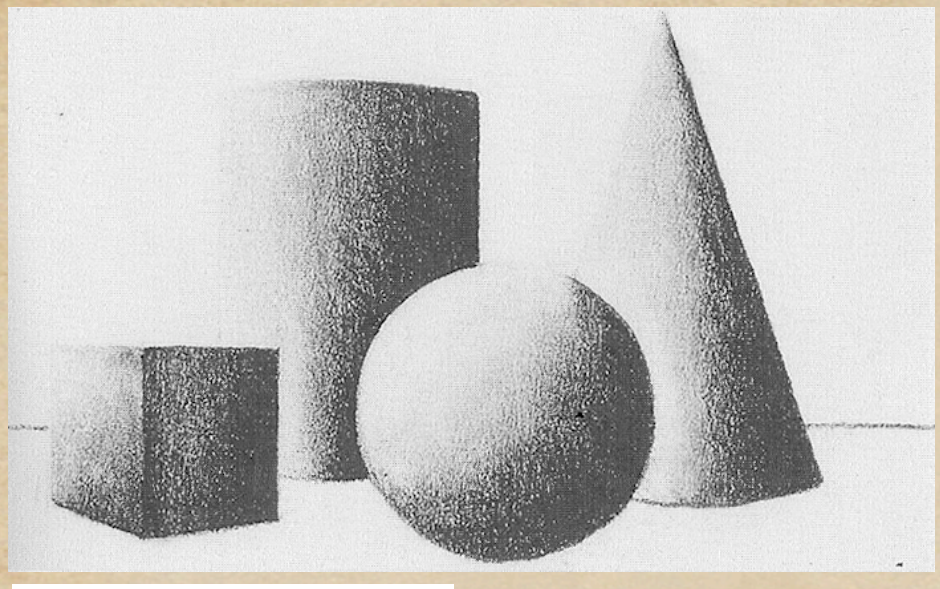
Even in the ideal case (wallpaper texture) the texture gradient cue is only weak.

Drop shadows nail objects to the ground plane

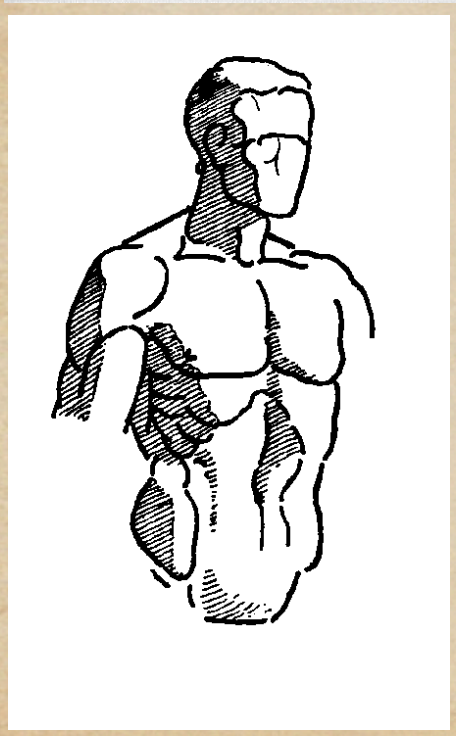




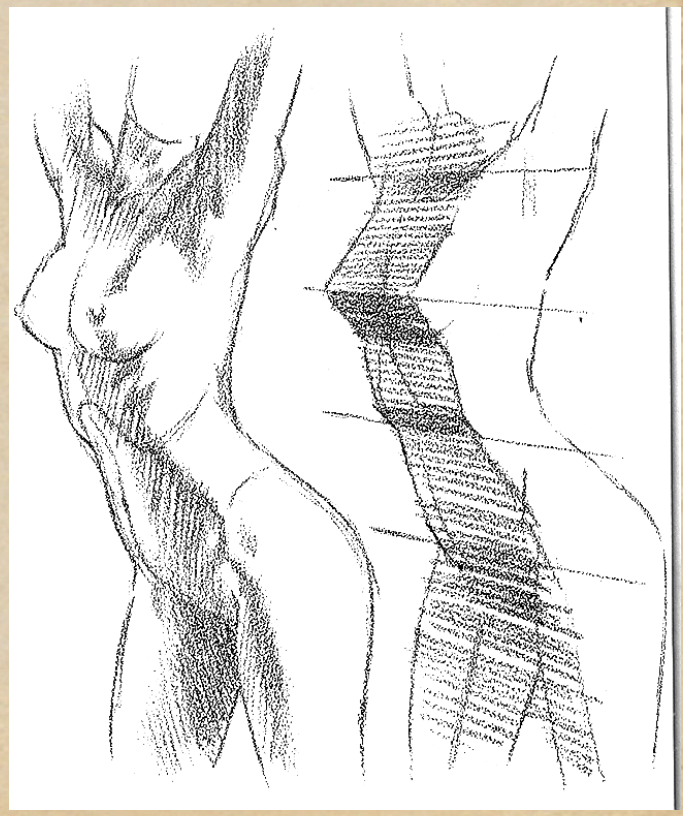
Shading indicates the spatial attitude of surface elements with respect to the direction of the local light field.



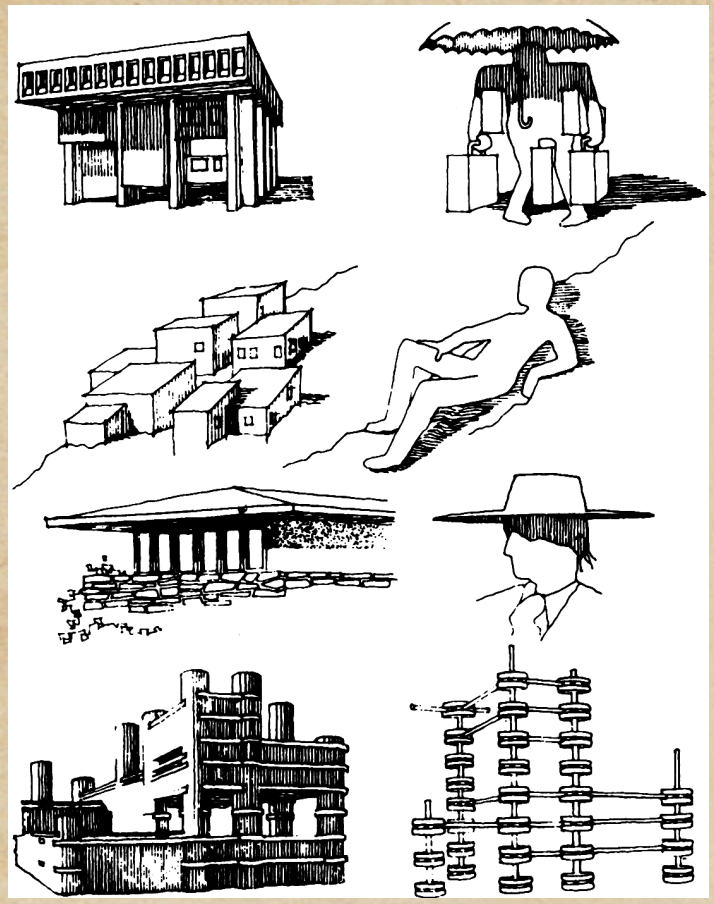
In many cases the light field is uniform, relating objects that are spatially far apart.

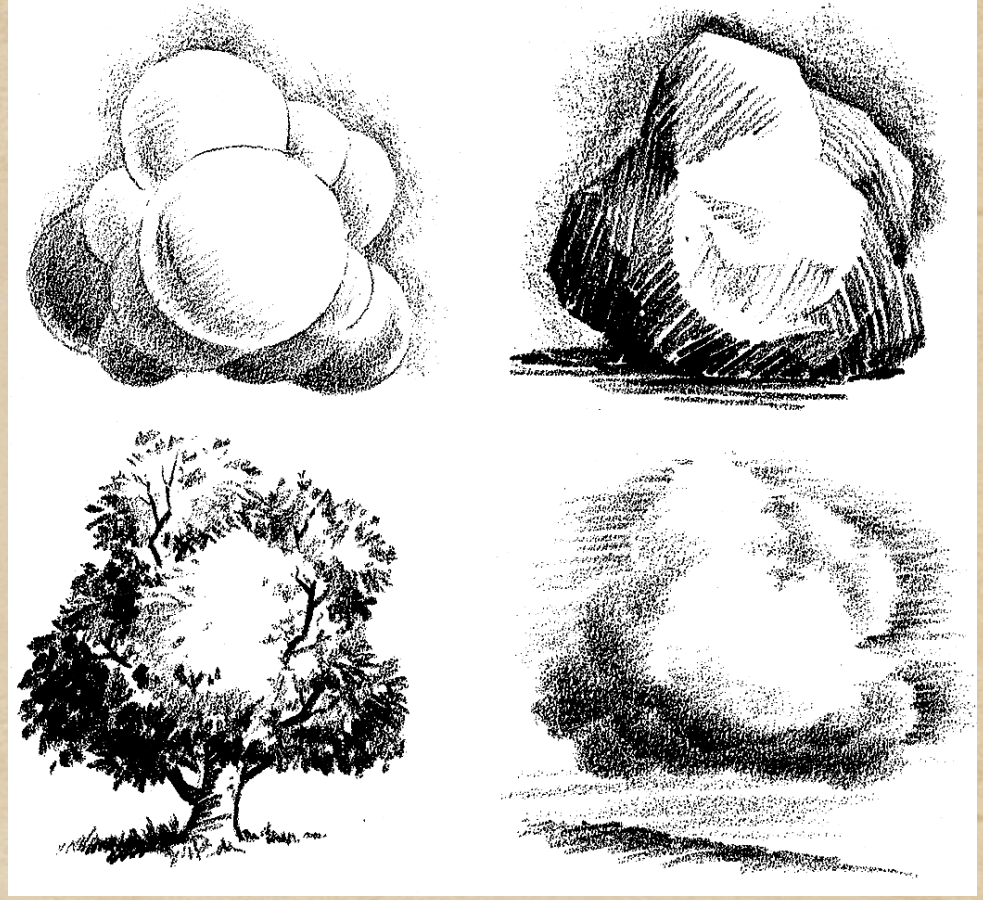


Only attitude variations in the light direction are revealed



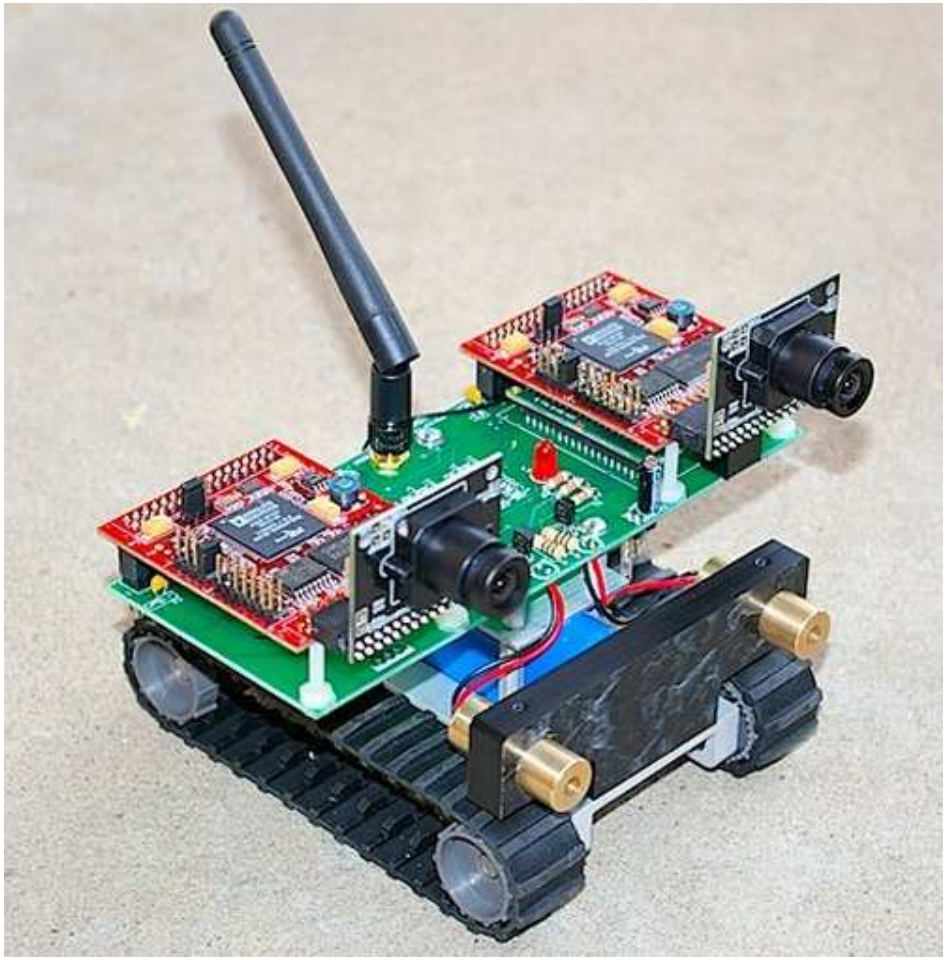
The “shading cue” is far more complicated than many textbooks would make you believe.





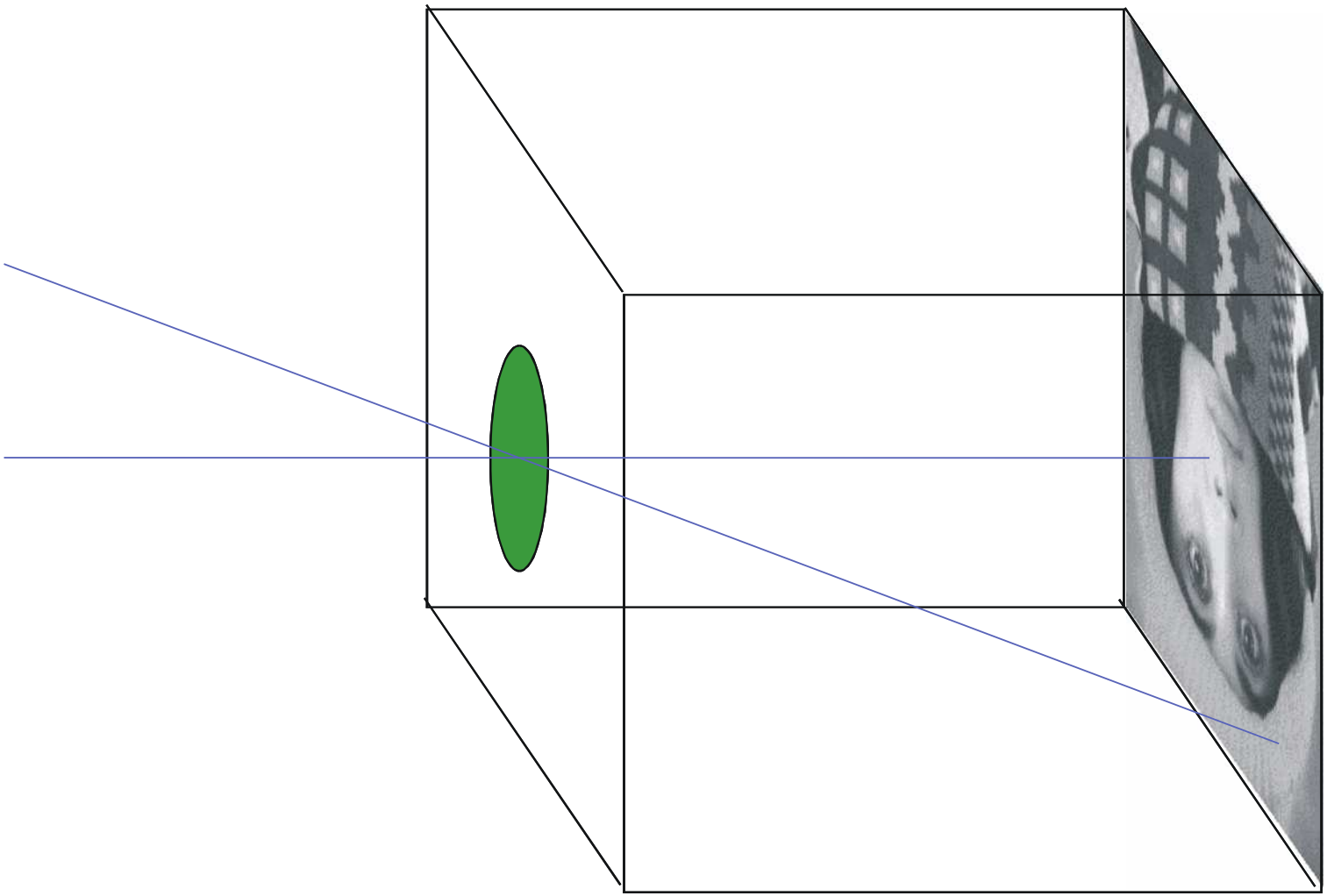
Most of the cues don't rely at all on the "meaning" in a narrow sense. "Visual objects" are not trees or houses.

Measuring Depth

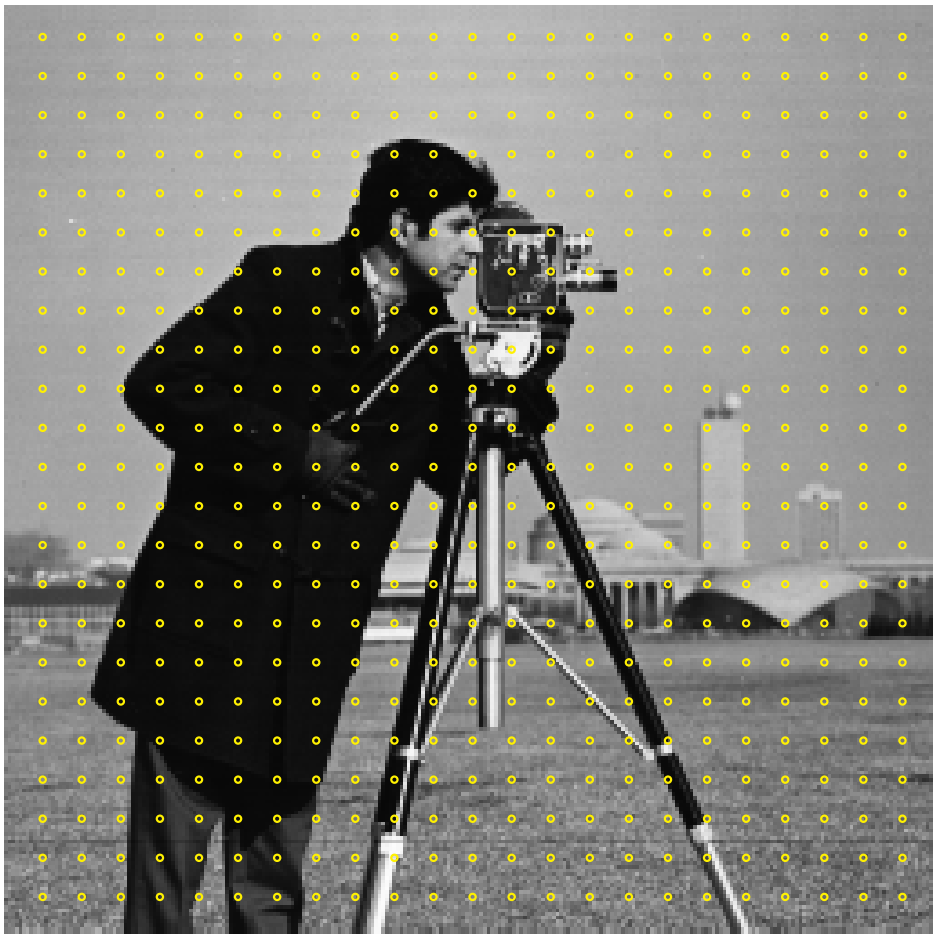


with two eyes we can measure depth

Visual Observations

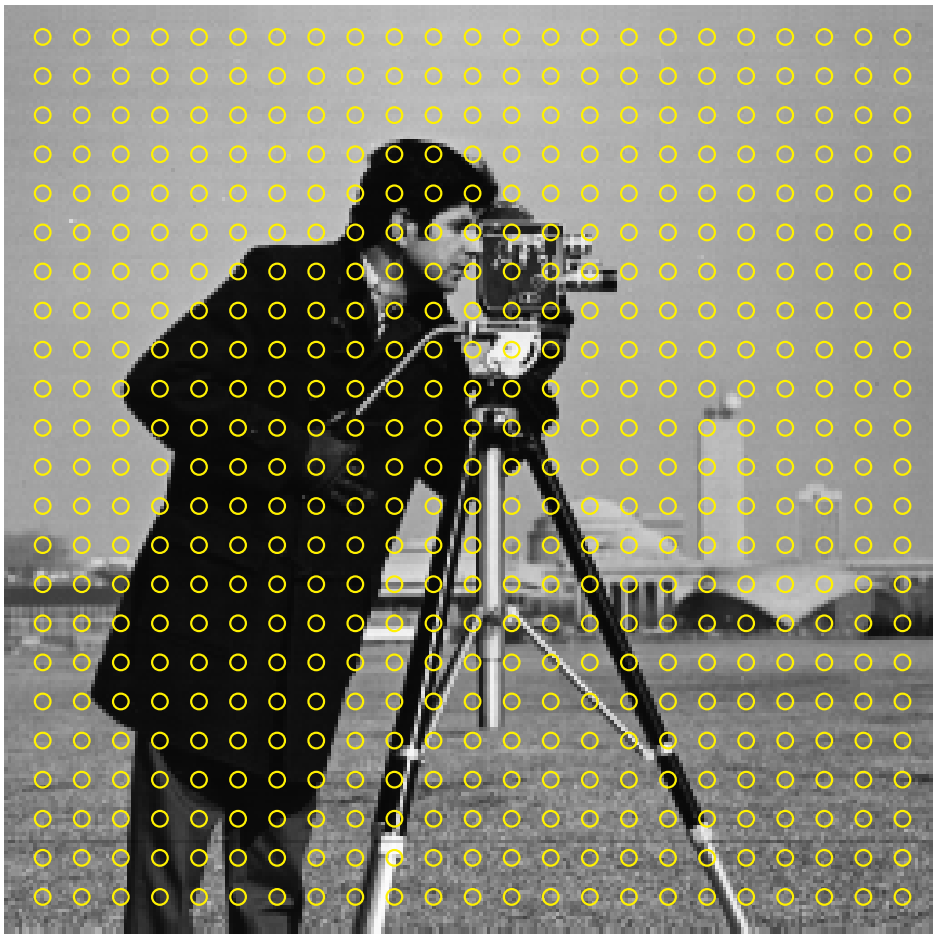


Sampling an Image



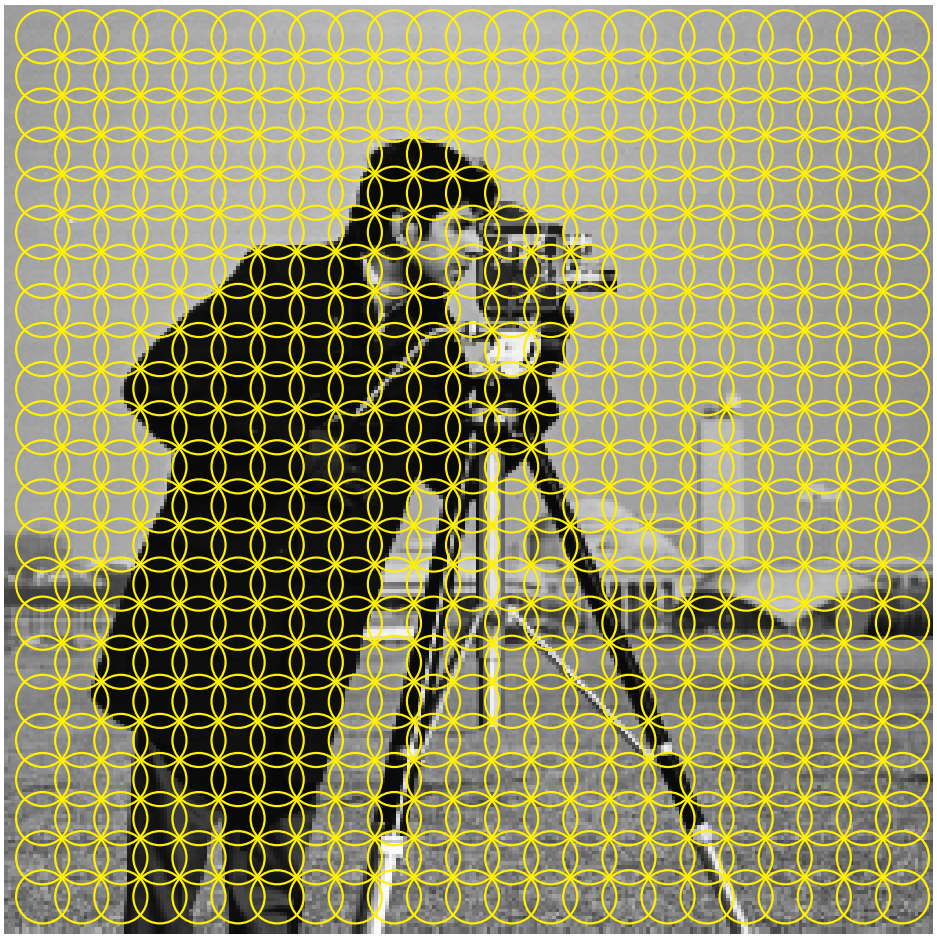
- we can only observe an image in a finite number of points

Sampling an Image



- we can only observe an image in a finite number of points
- with a probe of finite size

Sampling an Image



- we can only observe an image in a finite number of points
- with a probe of finite size
- at an appropriate scale

From Light to Pixels

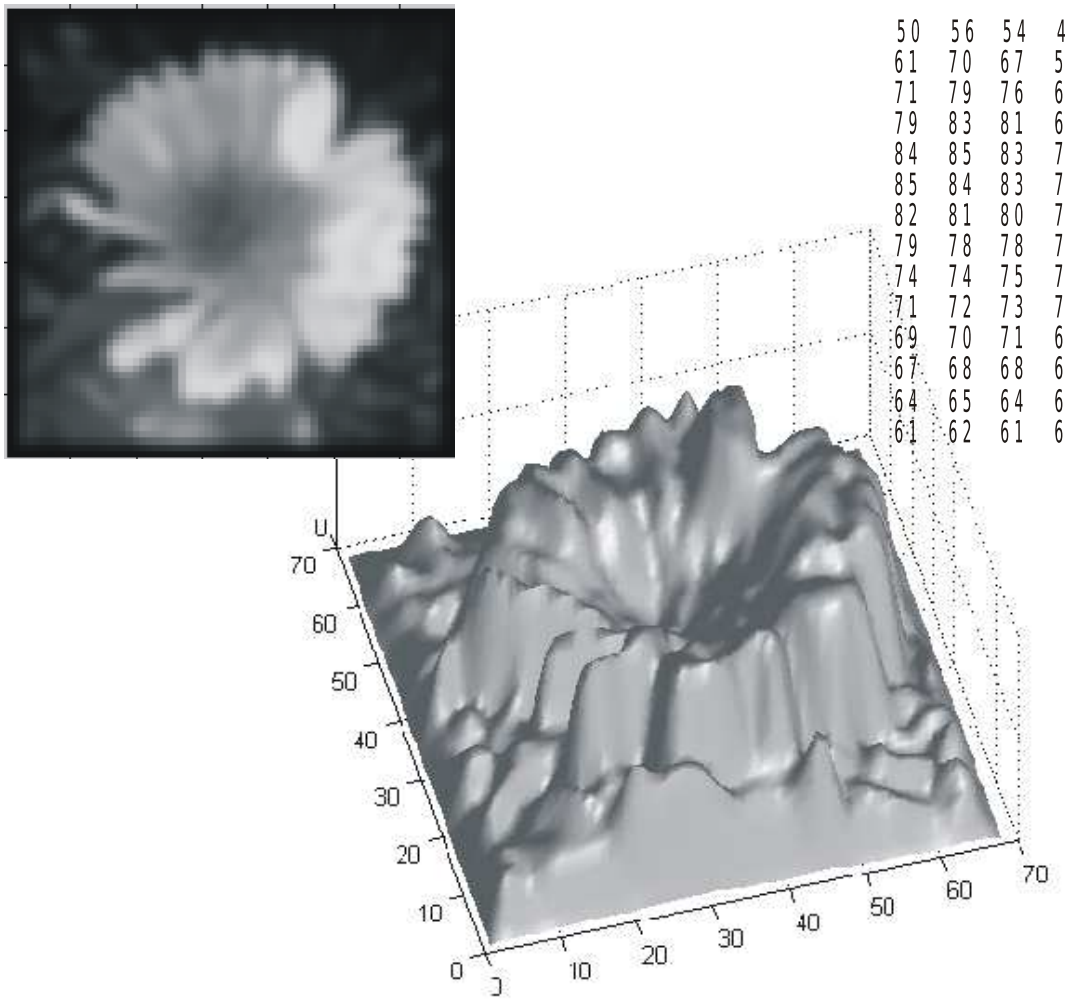
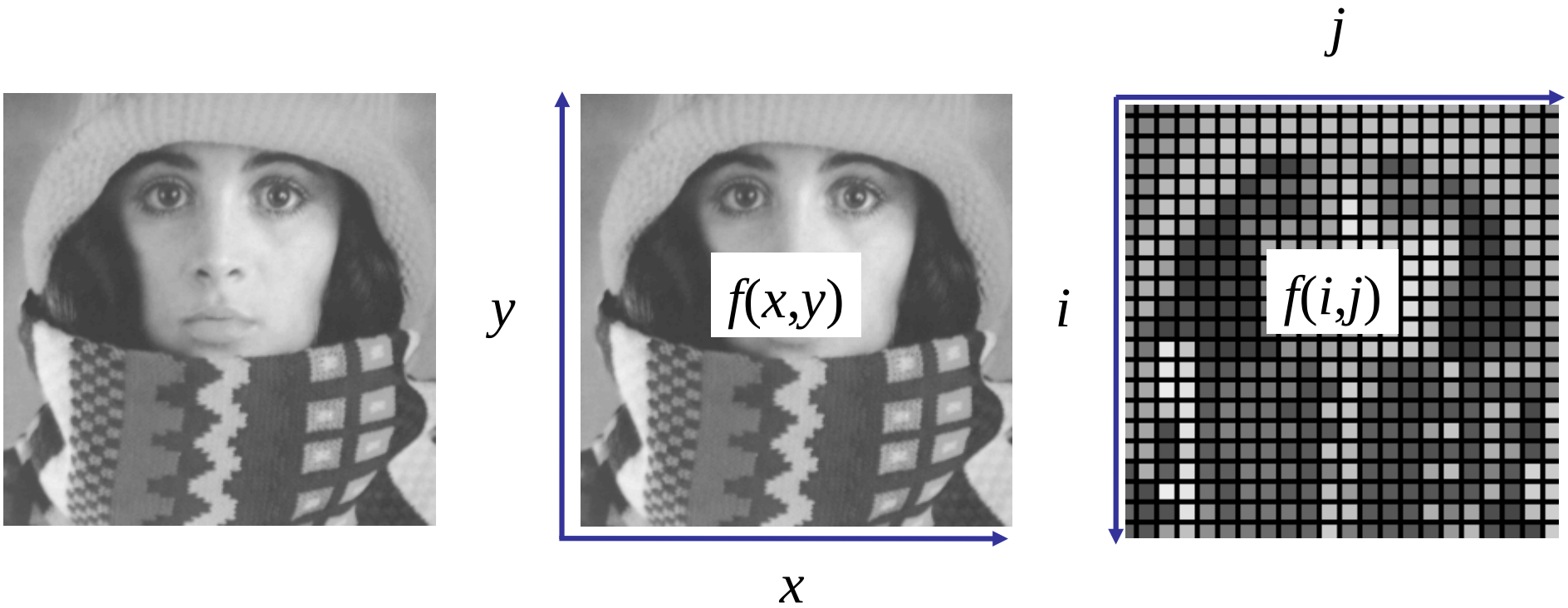
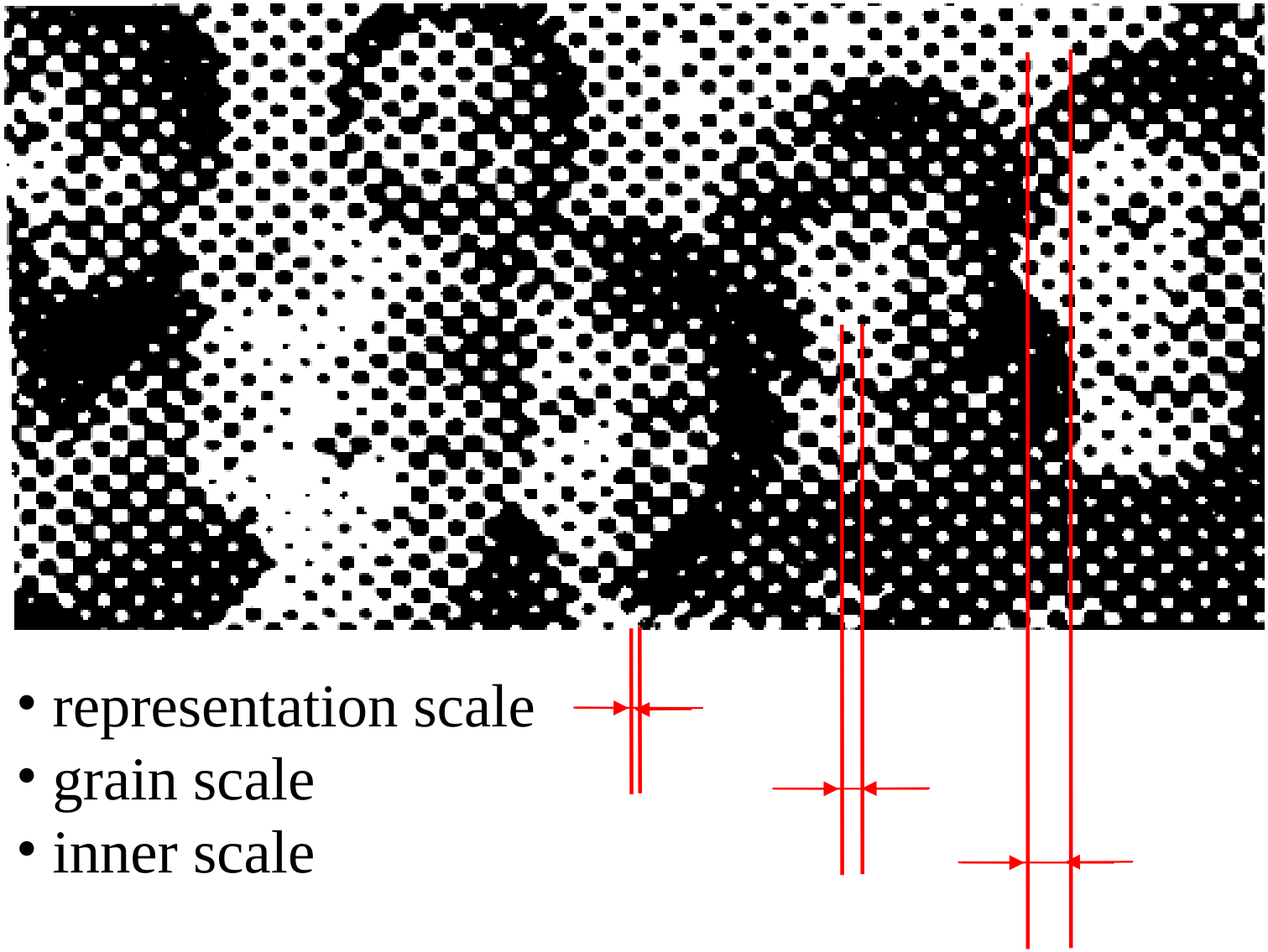


Image Representation



Scale

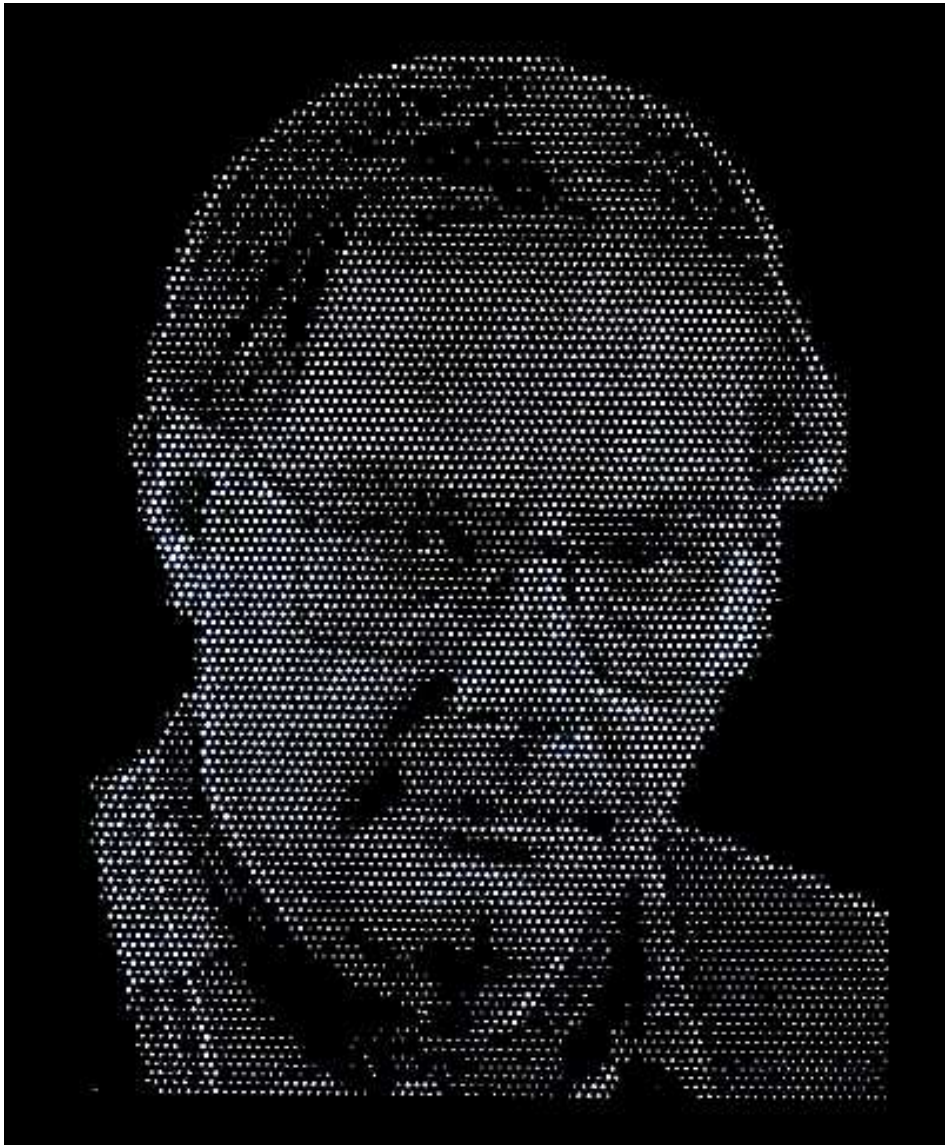


- representation scale
- grain scale
- inner scale

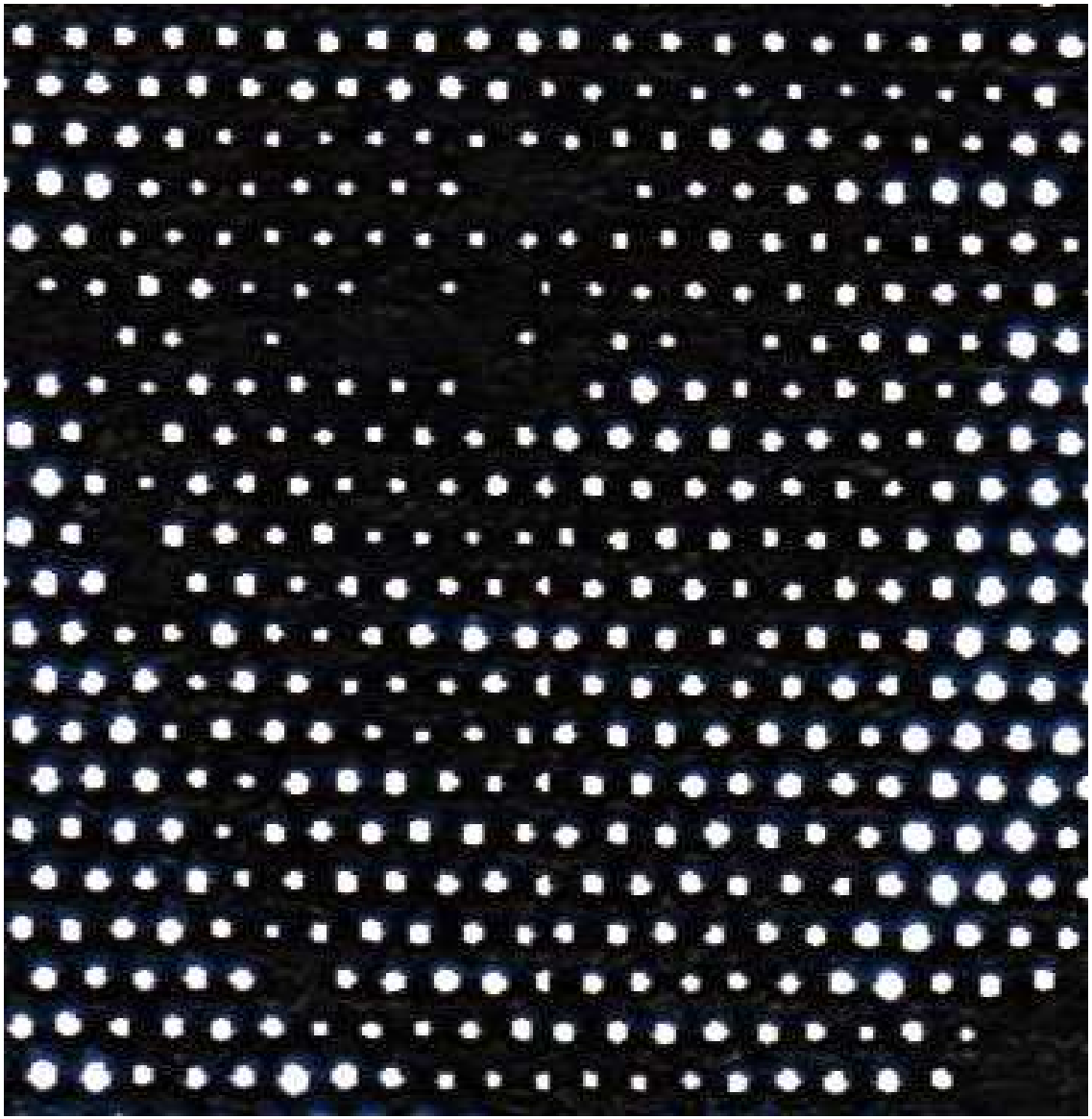
Scale

- representation scale
- grain scale
- inner scale

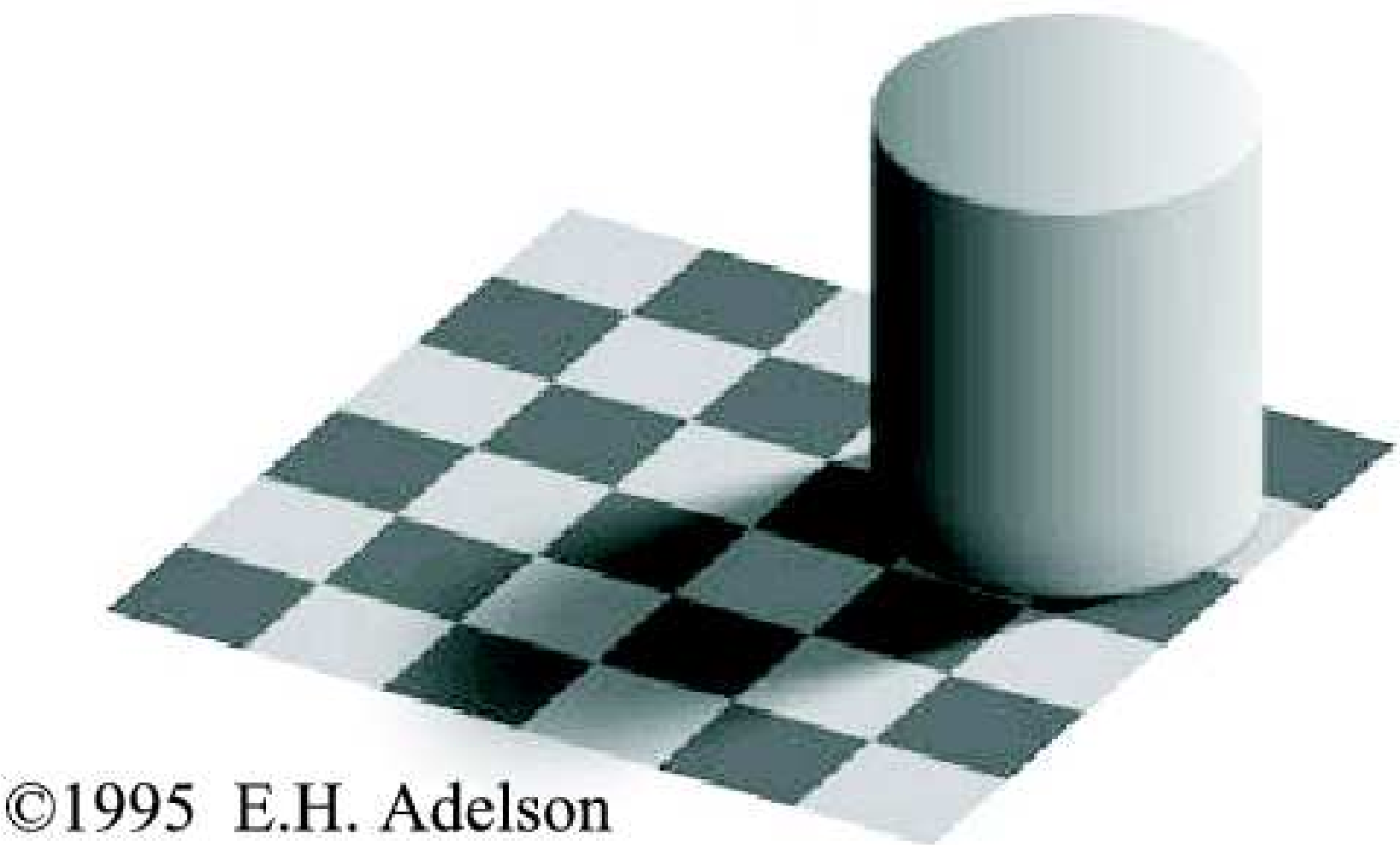
Scale



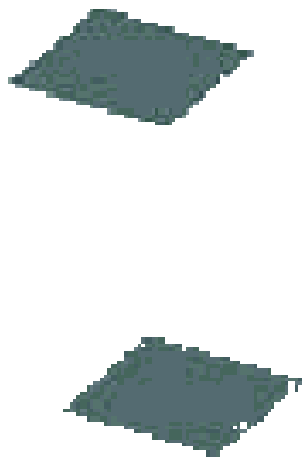
Scale



Point Observation



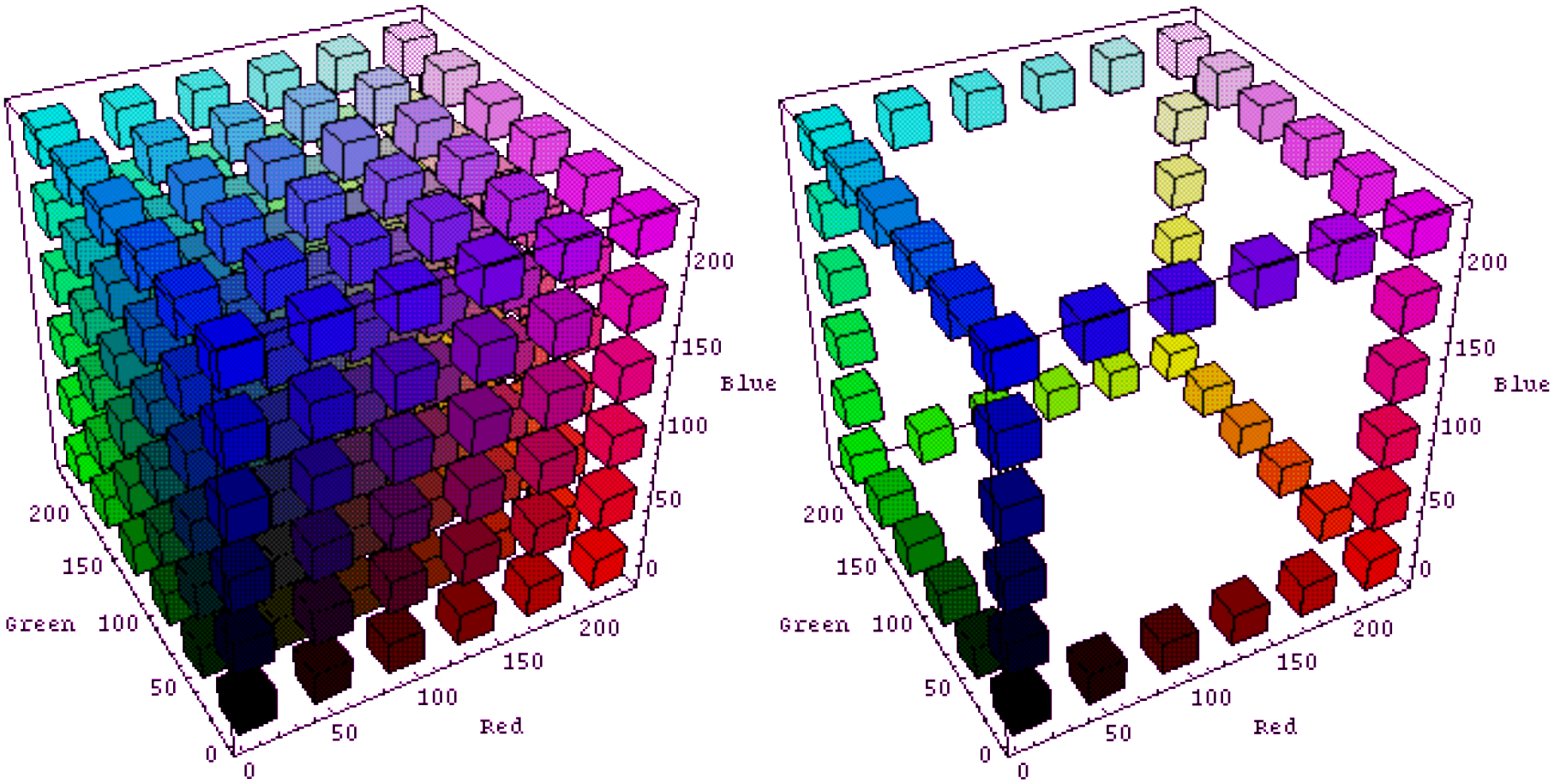
Point Observation



More Information per Sample

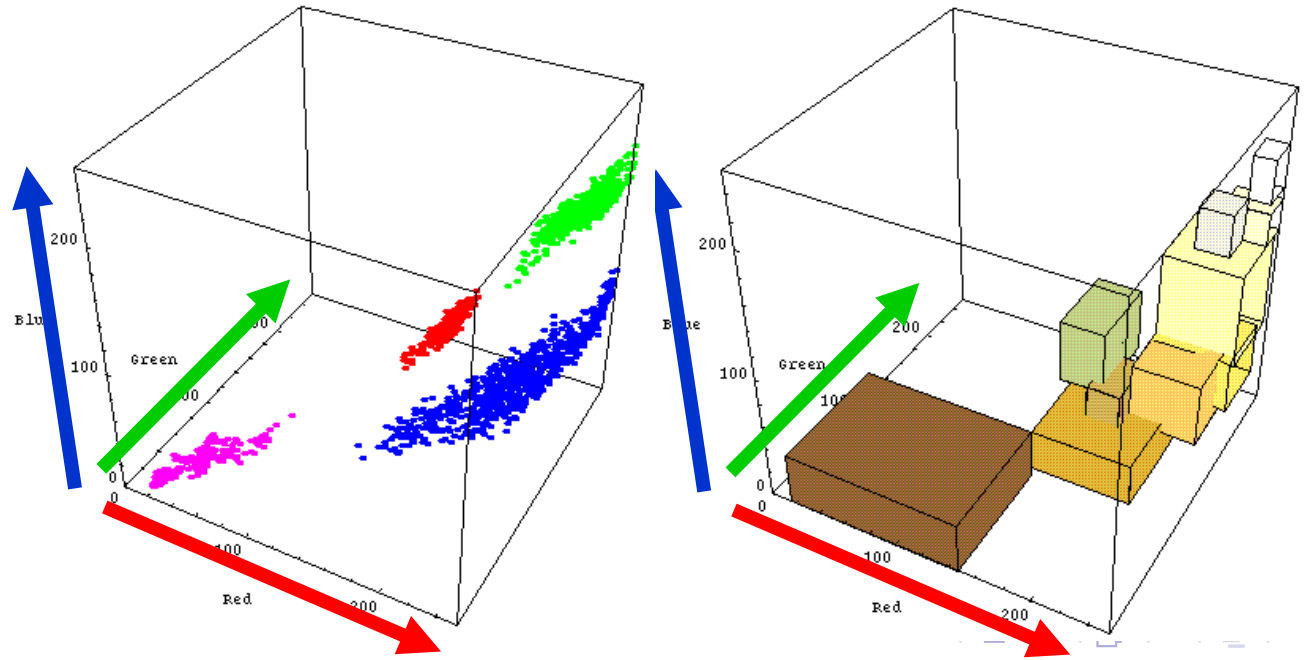
More Information per Sample

Color

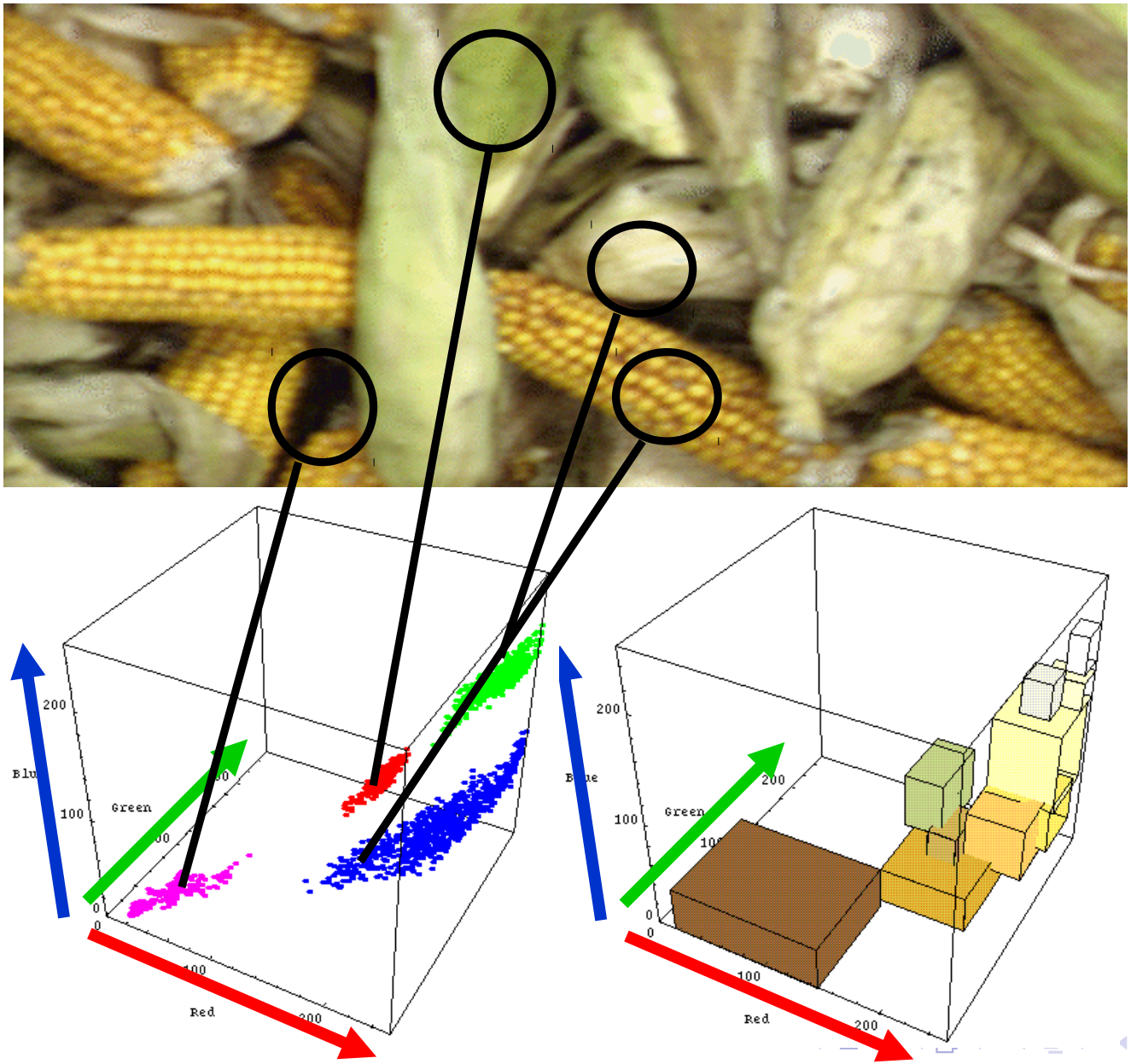


Finding Waldo

Color Histograms



Color Histograms



Visual Perception Experiment

Watch closely !!!

Visual Perception Experiment

0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 L 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 D 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0

Visual Perception Experiment

What did you see ?

Visual Perception Experiment

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	L	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	D	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

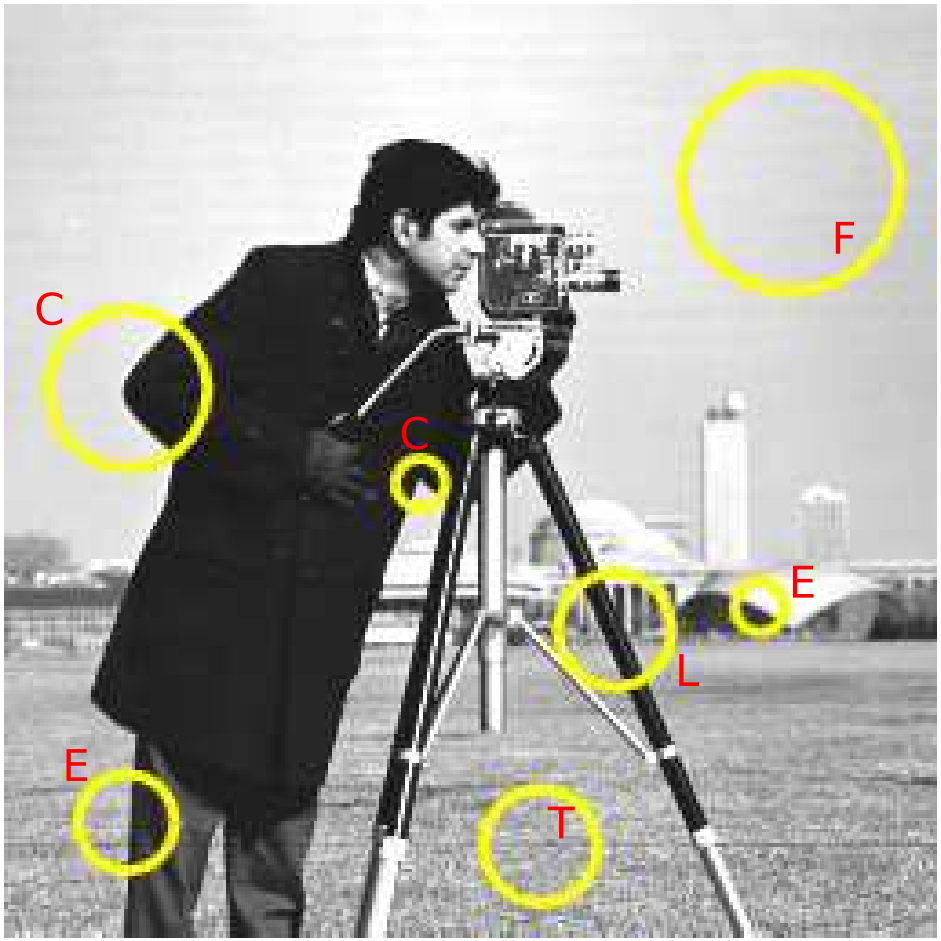
- The ‘L’ among all ‘0’s is much easier to spot then the ‘D’, but more surprisingly the time it takes to spot the ‘L’ is independent of the number of ‘0’s surrounding it.

Visual Perception Experiment

0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 L 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 D 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0

- The 'L' among all '0's is much easier to spot then the 'D', but more surprisingly the time it takes to spot the 'L' is independent of the number of '0's surrounding it.
- The visual difference between 'L's and '0's seems to be hardwired into our visual brain.
- The visual brain thus seems repsonsive to *local structure* in images.

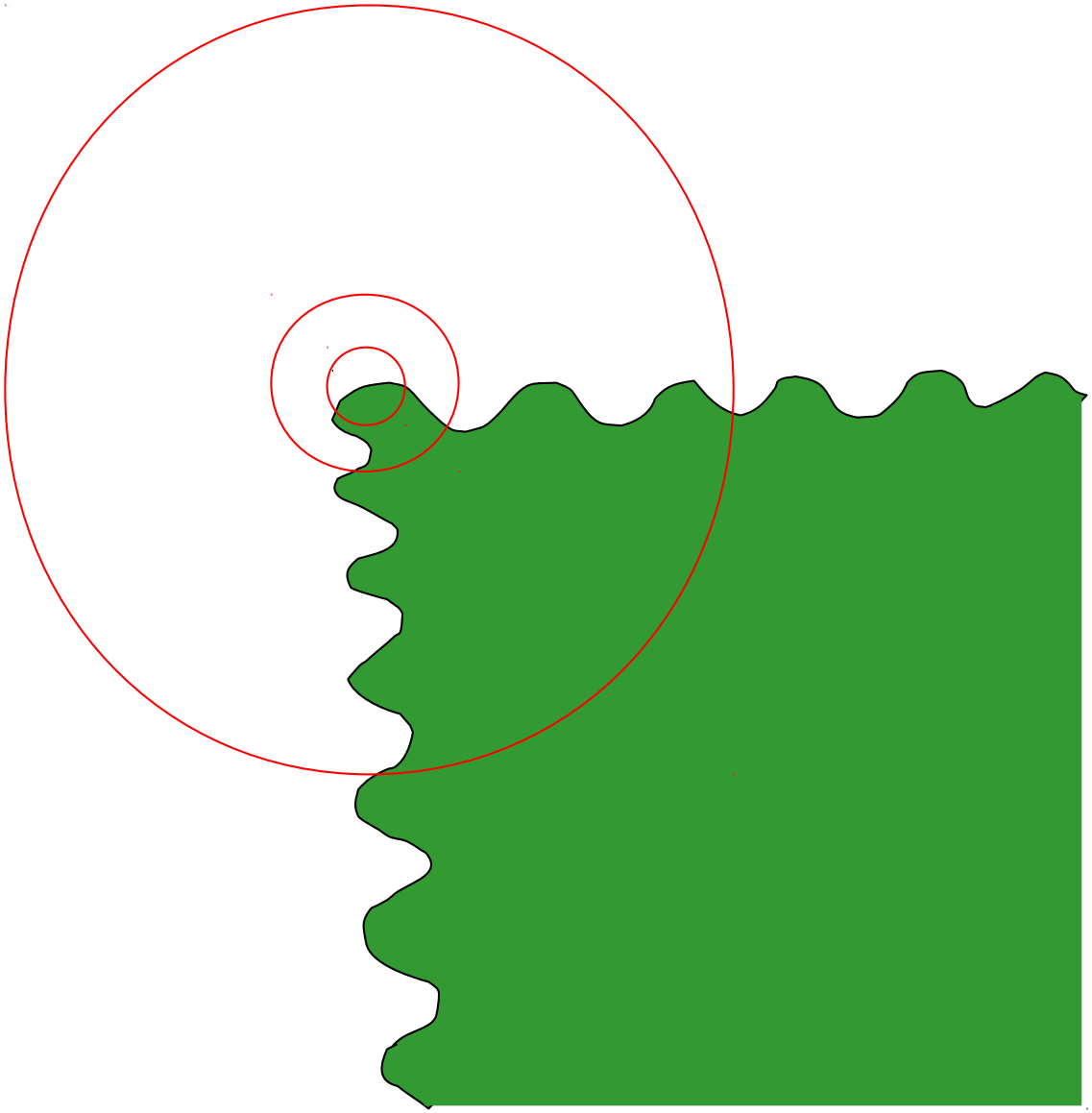
Local Image Structure



Local image structure can be characterized as image patches with

- *constant* grey value (F)
- *straight edge* (E)
- *corners* (C)
- *lines* (L)
- *T-junctions*
- *texture* (T)
- and many more...

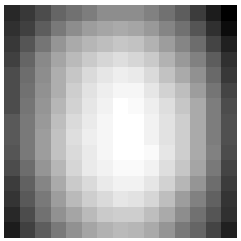
Scale Dependence

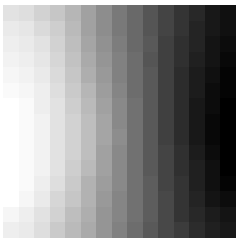


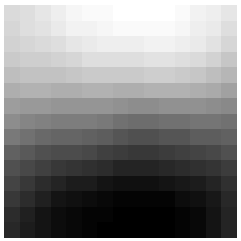
Local Structure

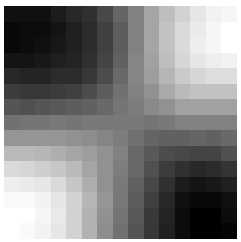
- Statistical experiments to learn which local details are the most important in natural images.
- Physiological experiments to determine for which local details the human eye is sensitive.
- Theoretical derivation from basic principles

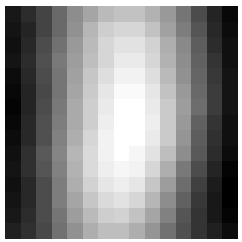
Learning Local Structure

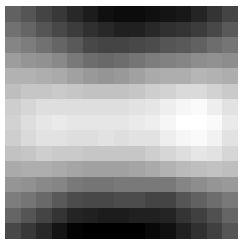






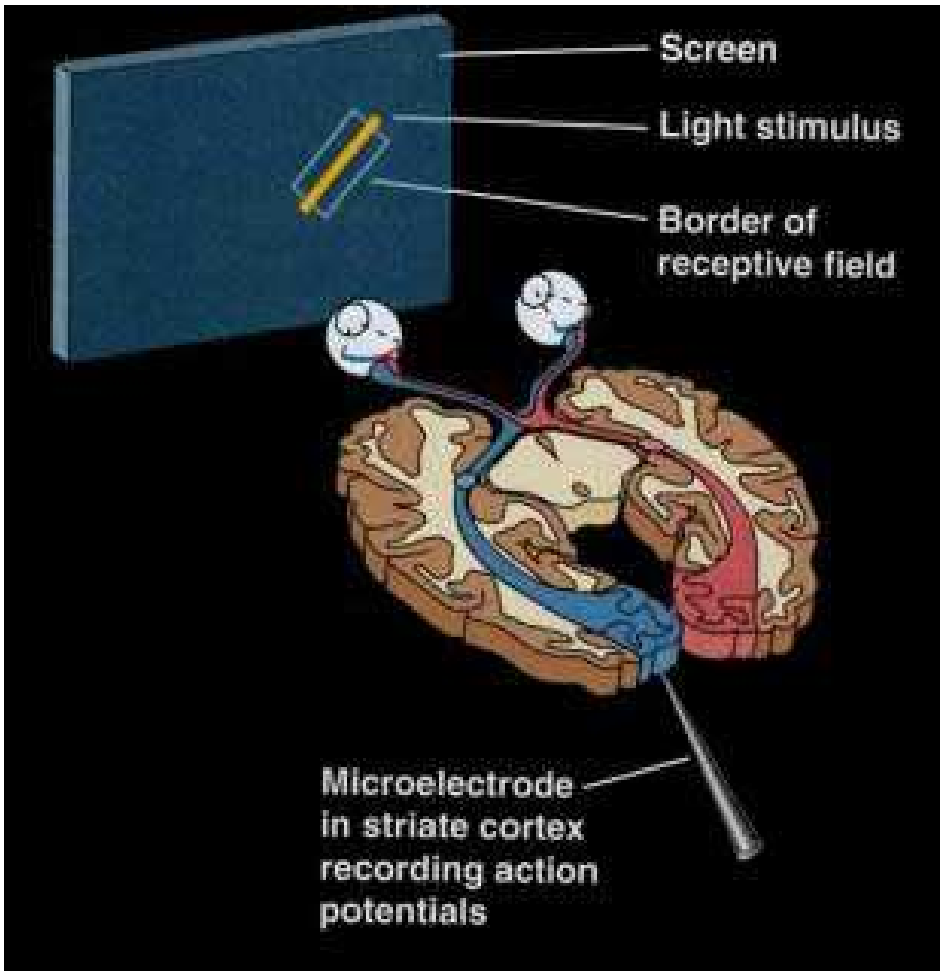






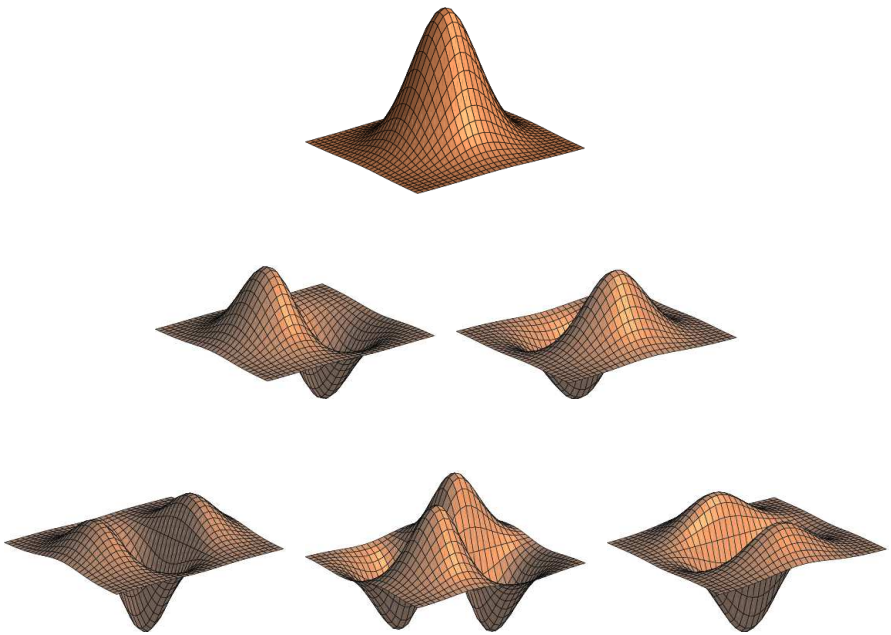
A local image patch of $N \times M$ pixels is a vector in NM-dimensional space. A principal component analysis finds the basis in image space such that any patch can be described as the linear combination of only a few basis vectors in the PCA basis.

Local Structure Detection in the Brain



Hubel and Wiesel measured the response of neurons in the visual cortex in-vivo. Their measurements indicate that the human visual system is hardwired to recognize specific details. These details show great resemblance with the local details as those learned with PCA. The human visual system thus seems to have adapted itself to the visual stimuli it is likely to see.

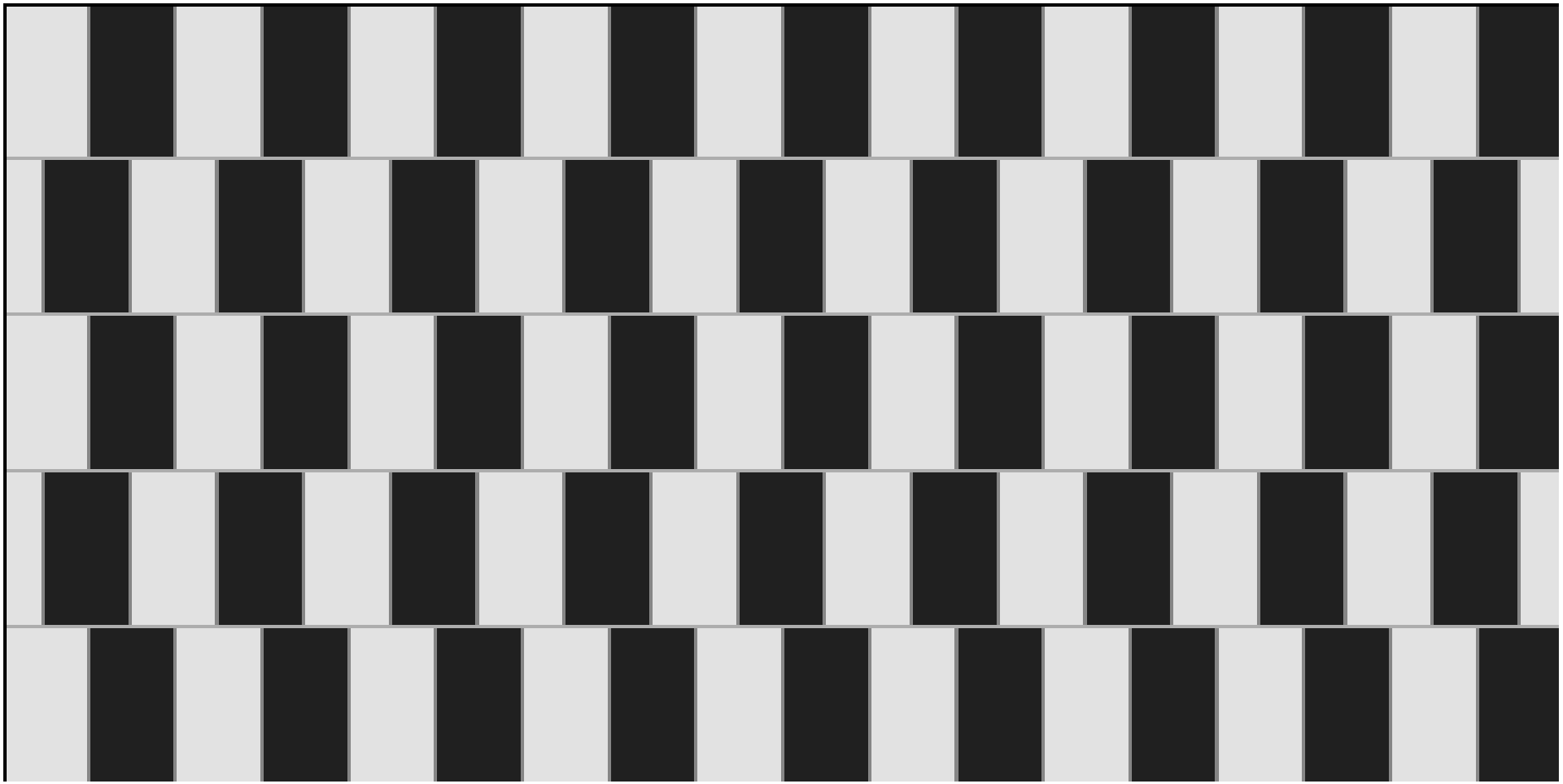
Local Structure Detection in the Brain



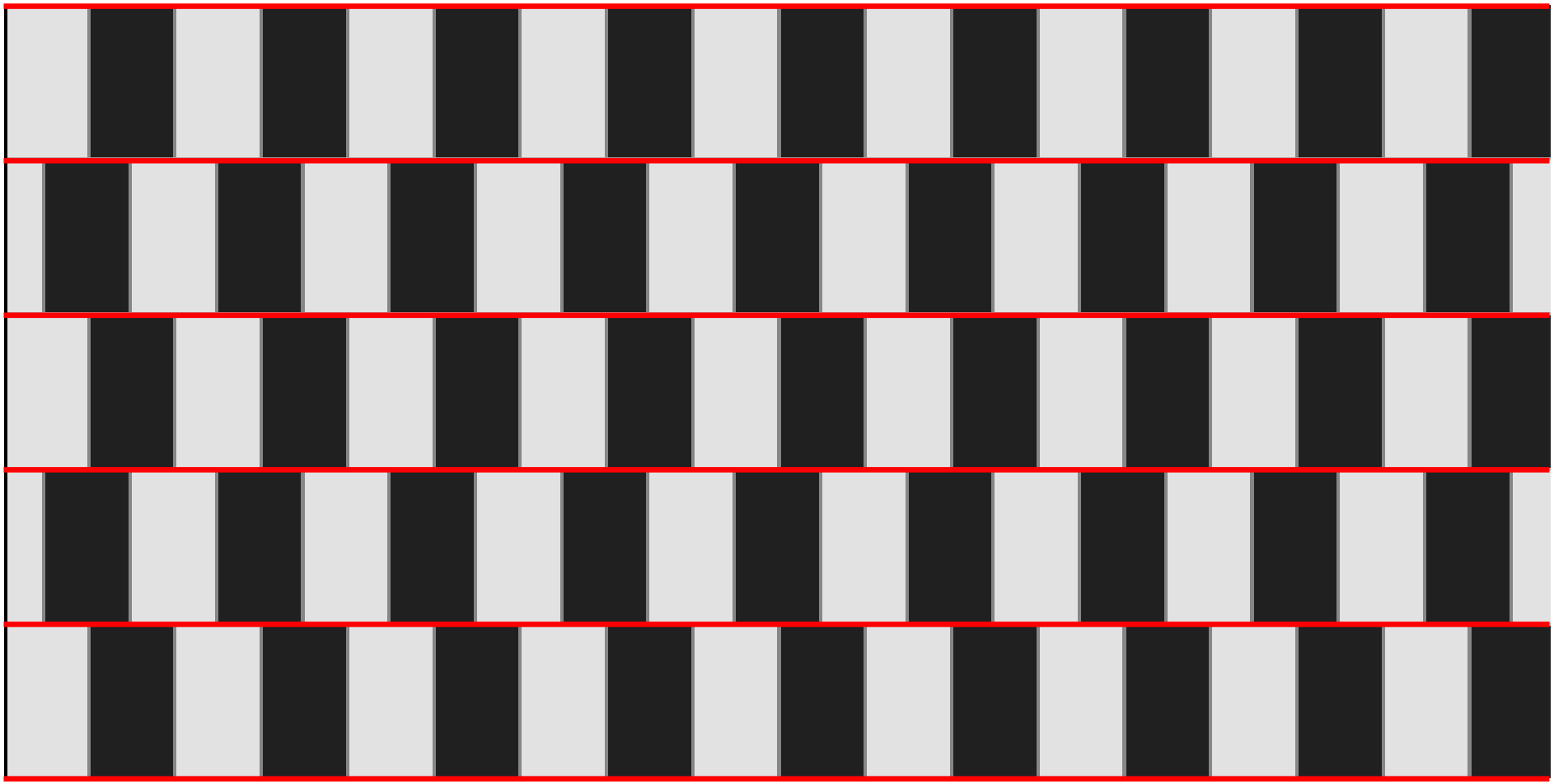
J. Koenderink (Utrecht University) showed that local image details as detected by the human brain also resemble the details that follow from a mathematical analysis based on basic (symmetry and causality) principles.

In these lecture notes a simplified version of this mathematical model is presented. The starting point is to use *image derivatives* to capture the notion of structure.

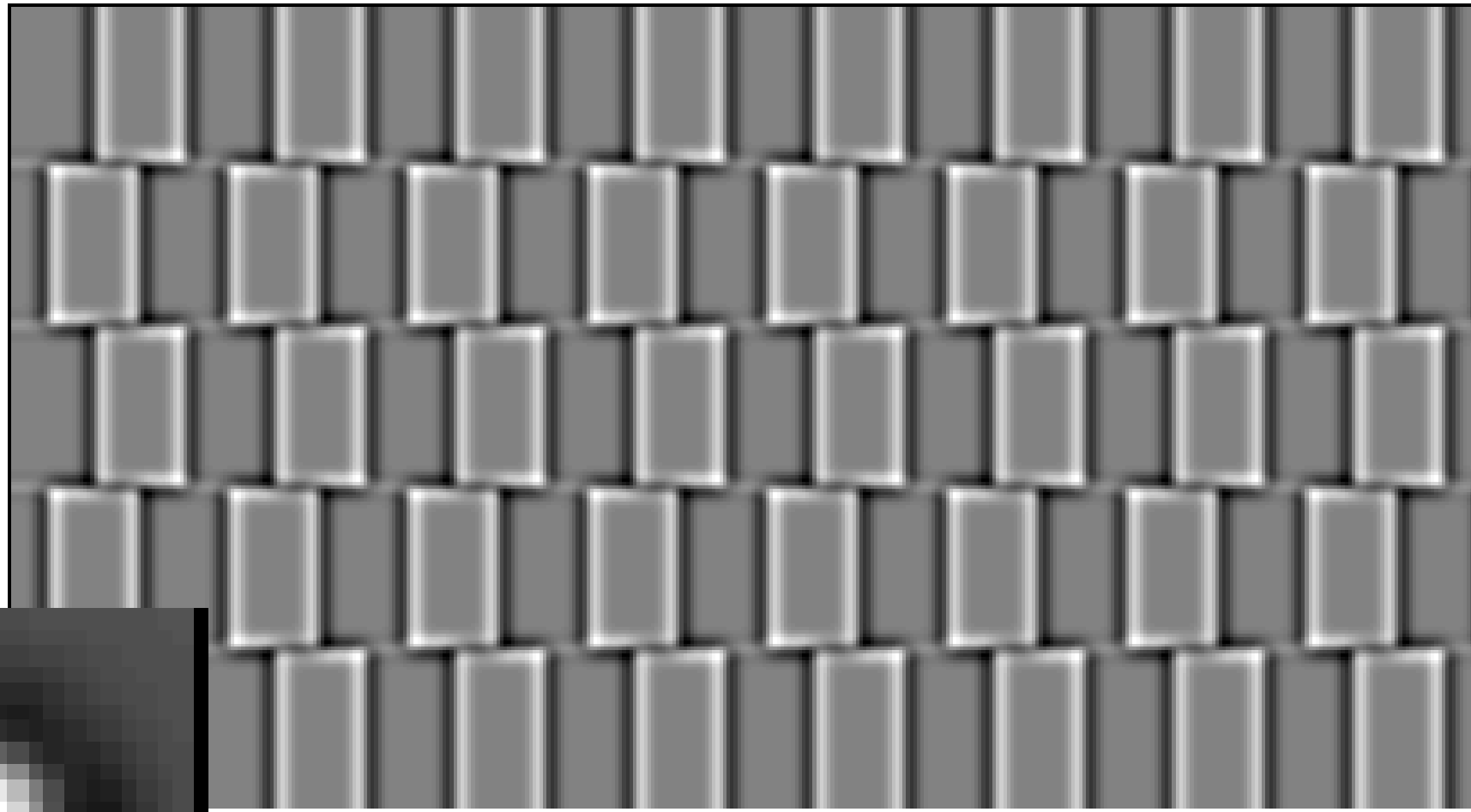
Local Structure in the Brain

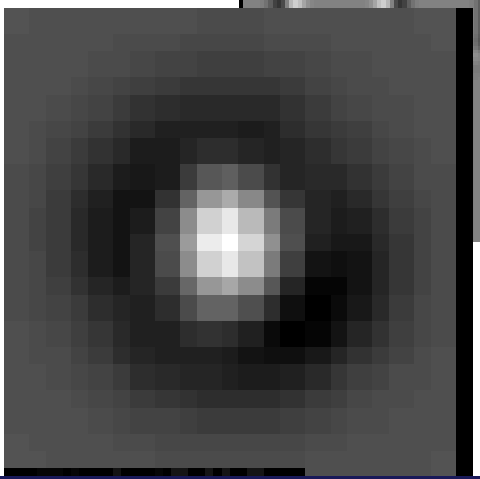


Local Structure in the Brain

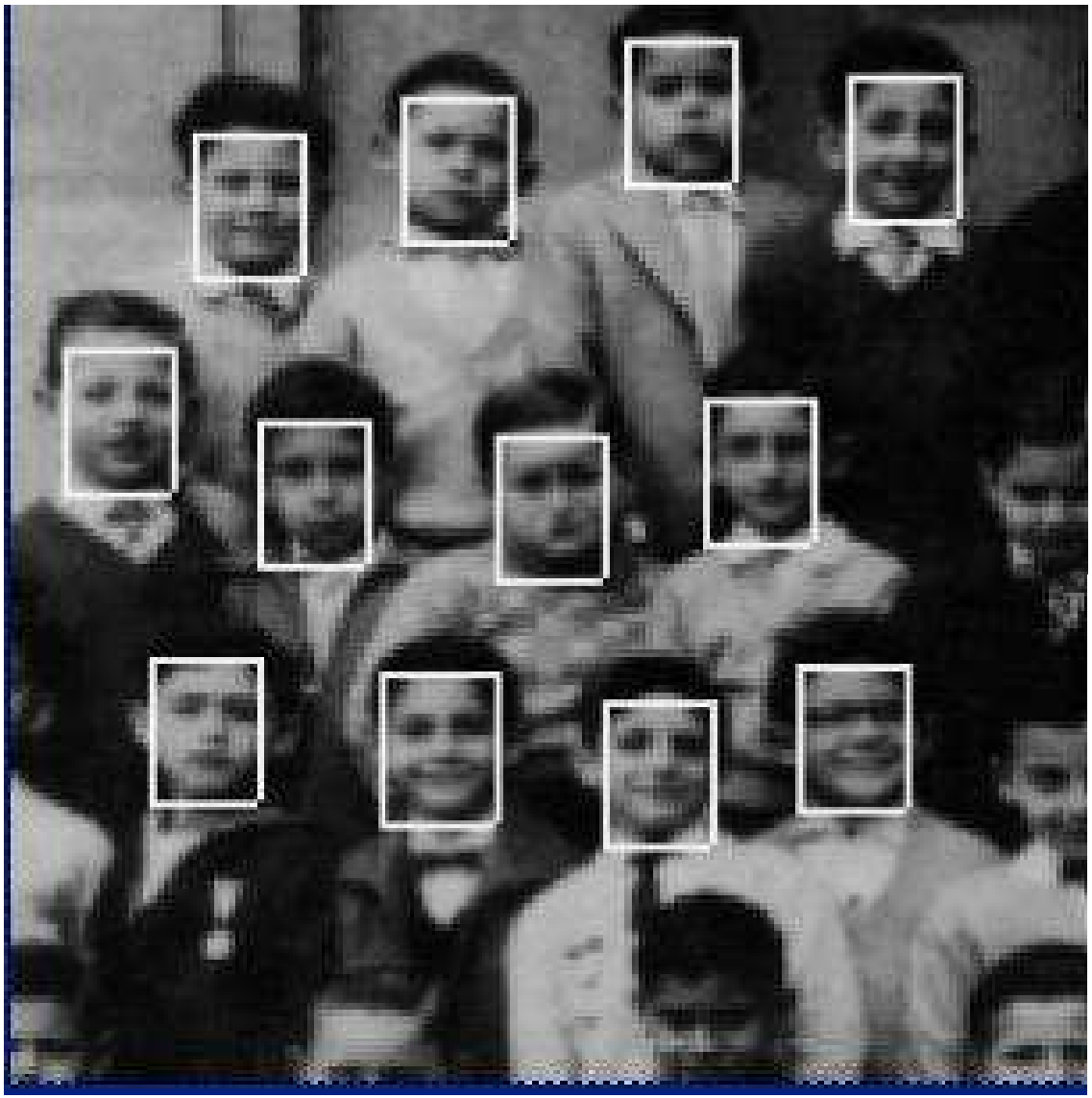


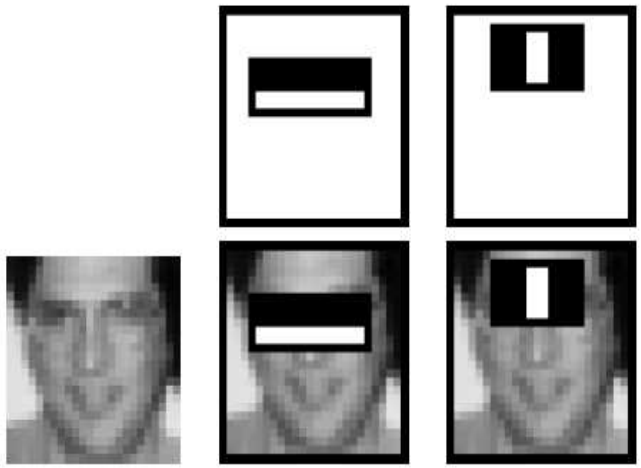
Local Structure in the Brain



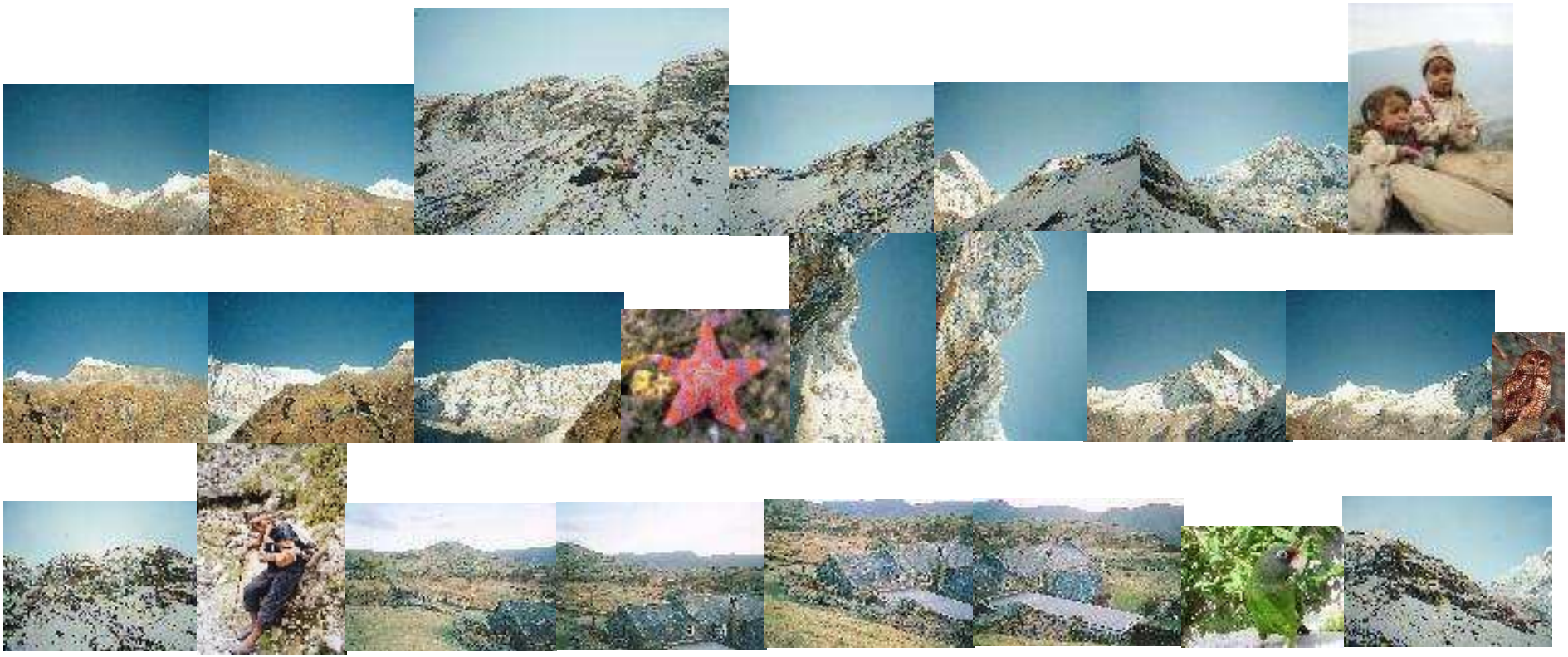


Application: Face Detection





Application: Image Stitching

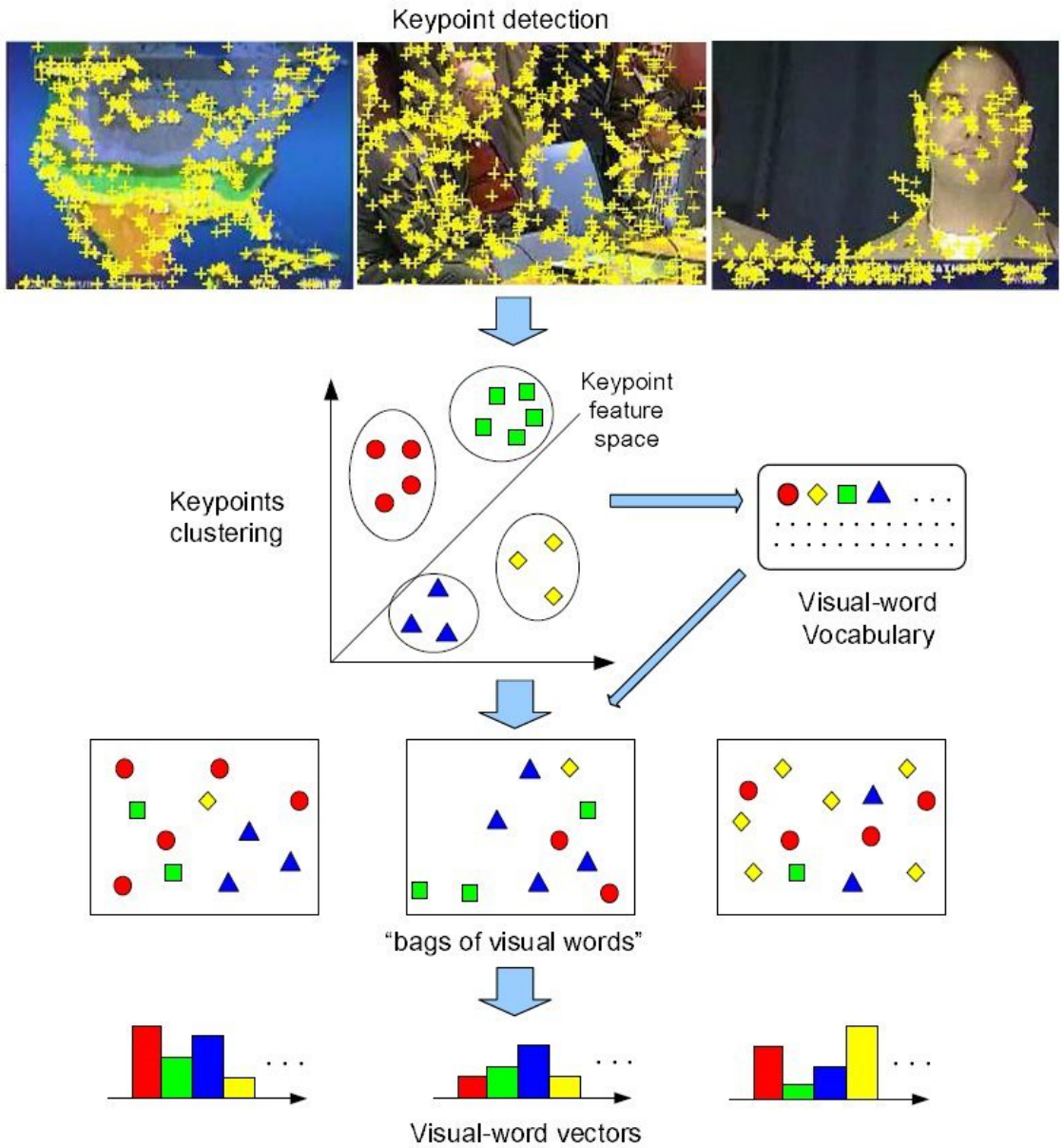


Application: Image Stitching

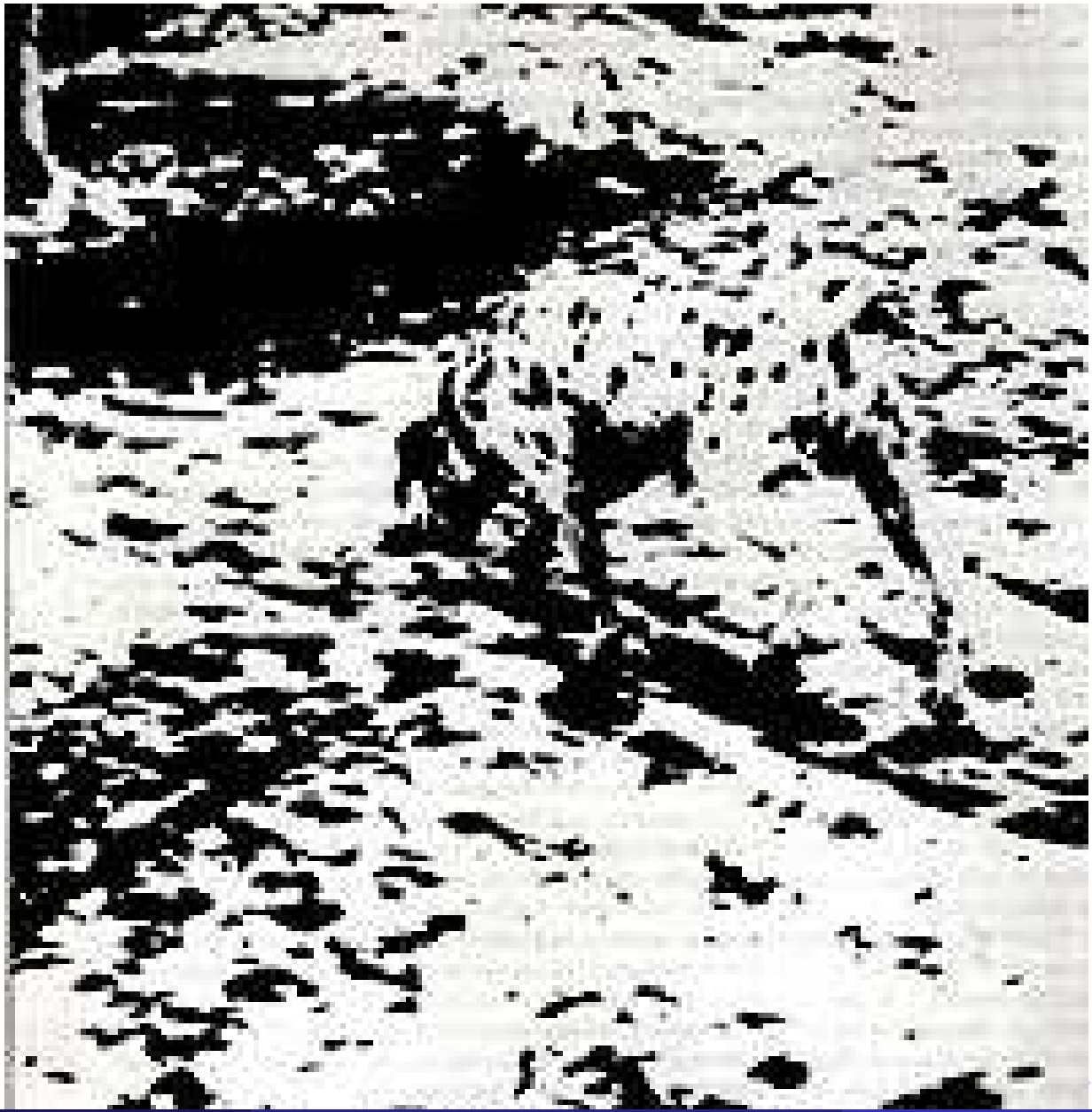
Application: Background Estimation and Subtraction

Application: Tracking Objects in Video

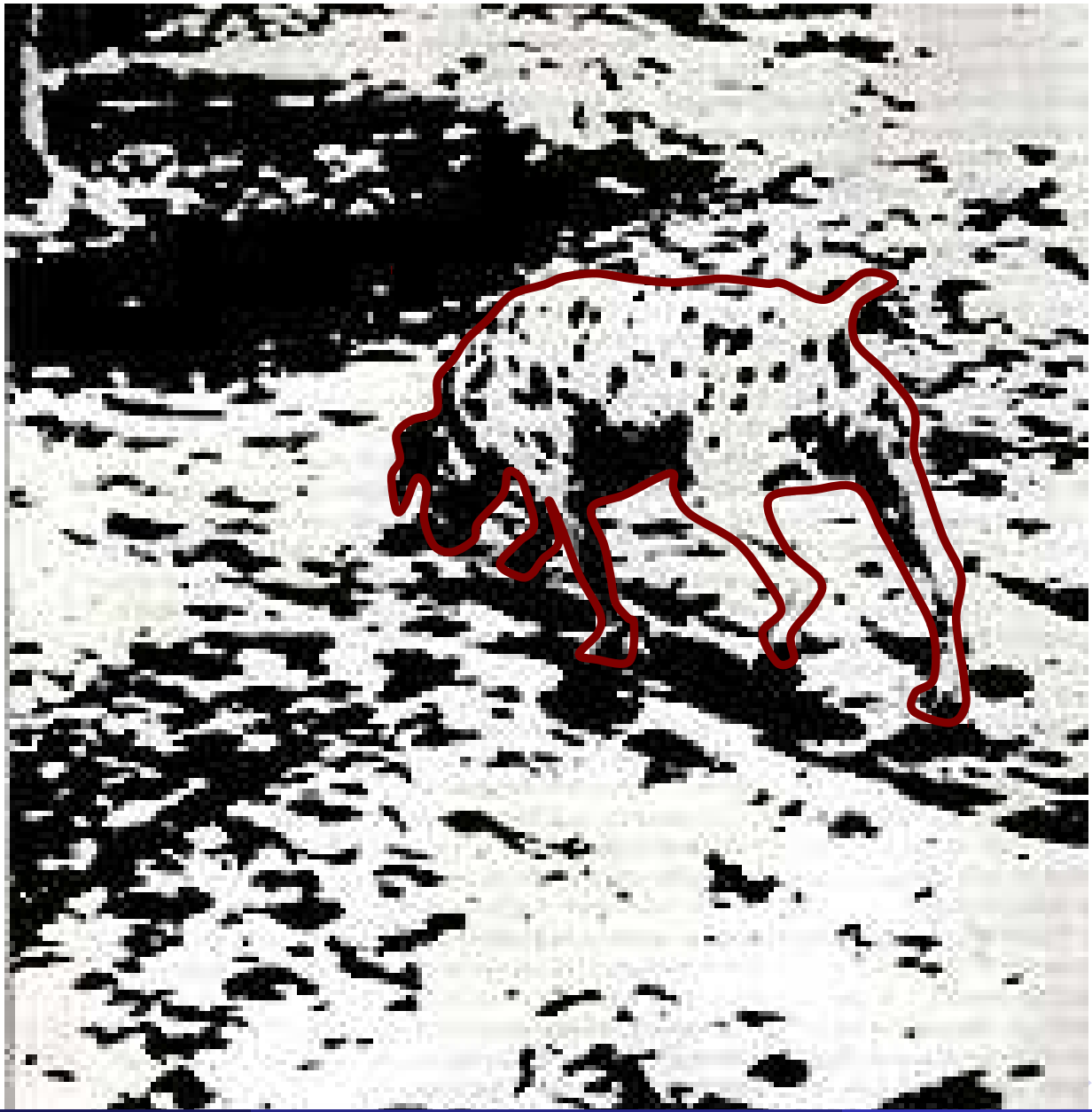
Application: Bag-of-Words



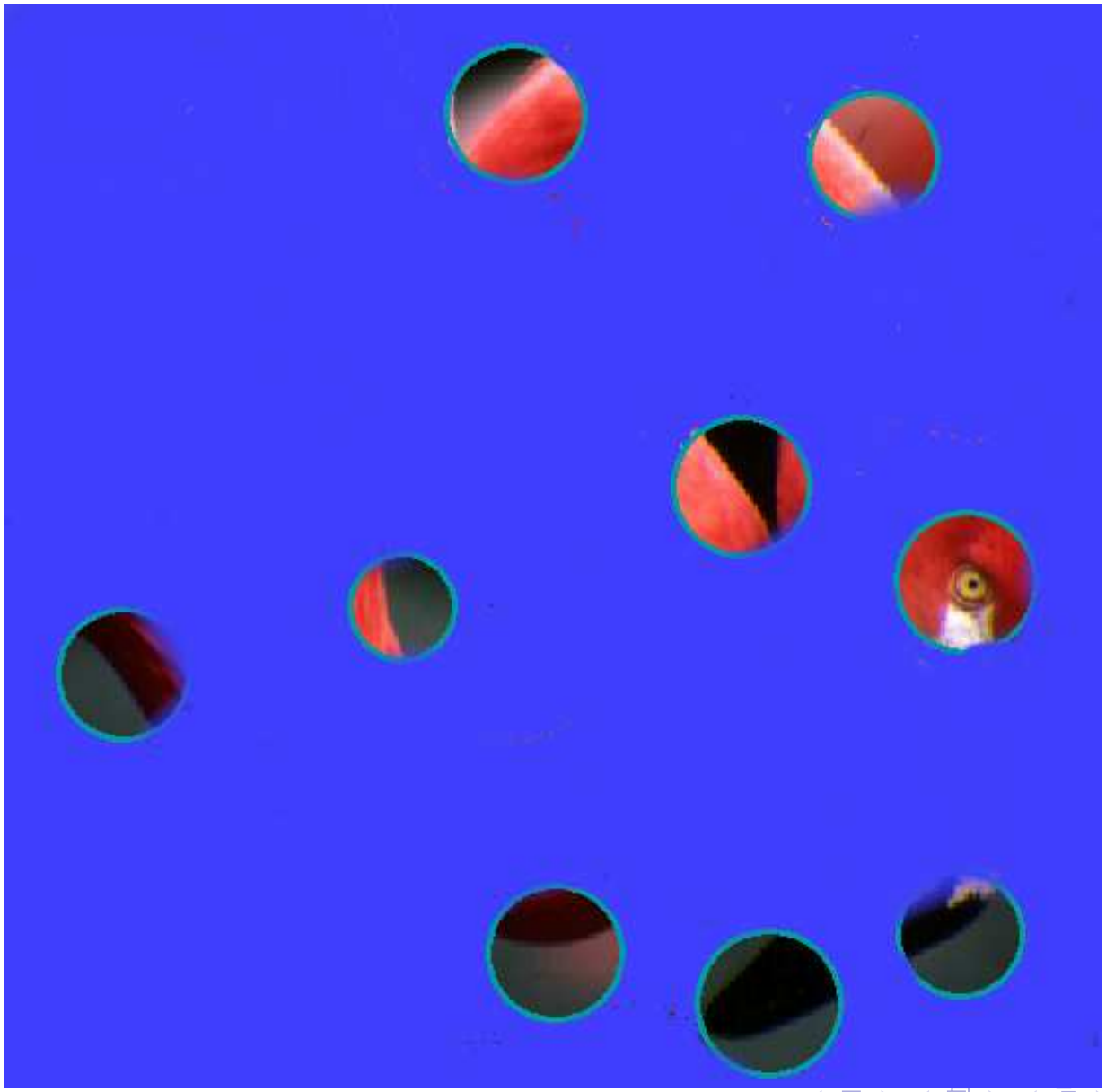
Local Cues are not enough



Local Cues are not enough

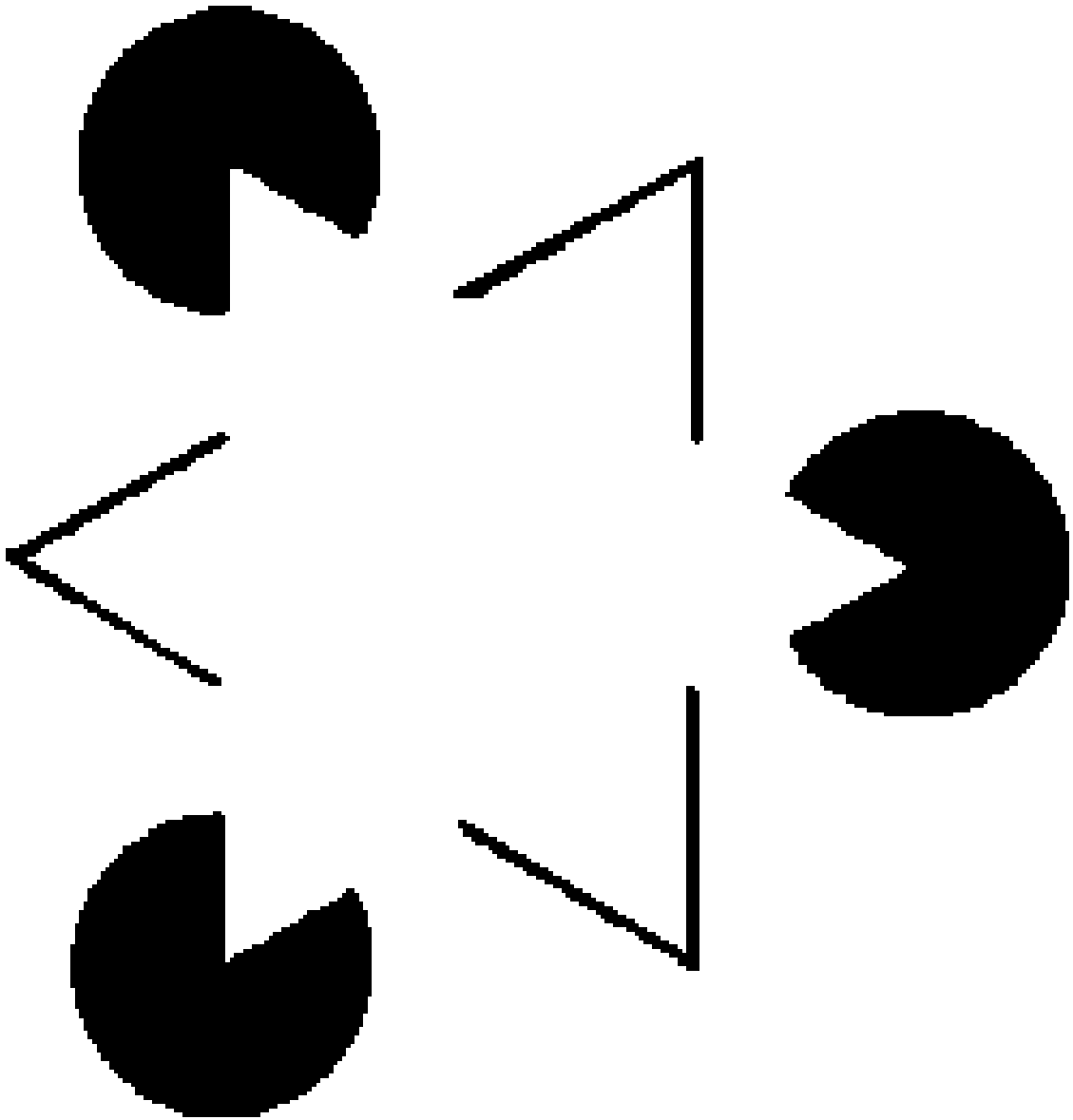


Local Cues are not enough

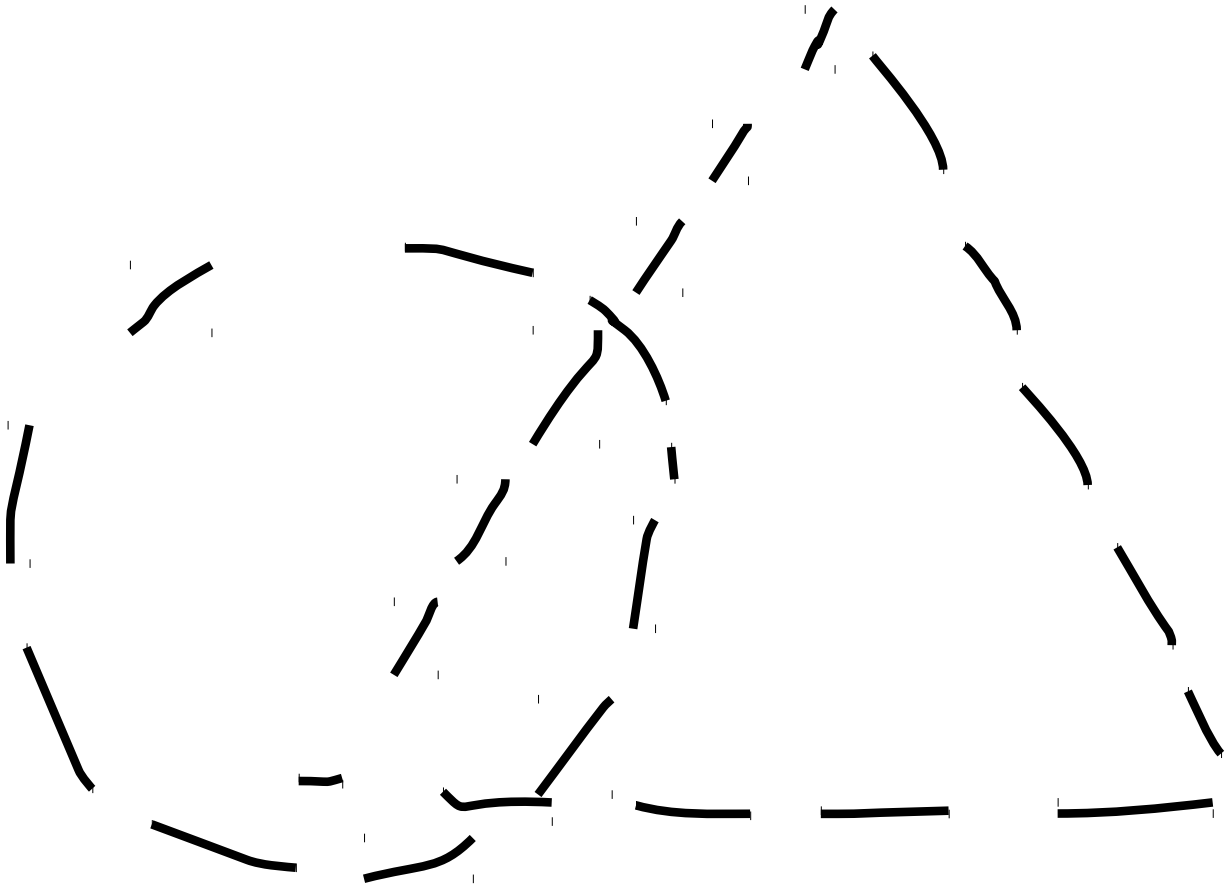


Local Cues are not enough

Perceptual Grouping



Perceptual Grouping



Segmentation

Articulated Objects