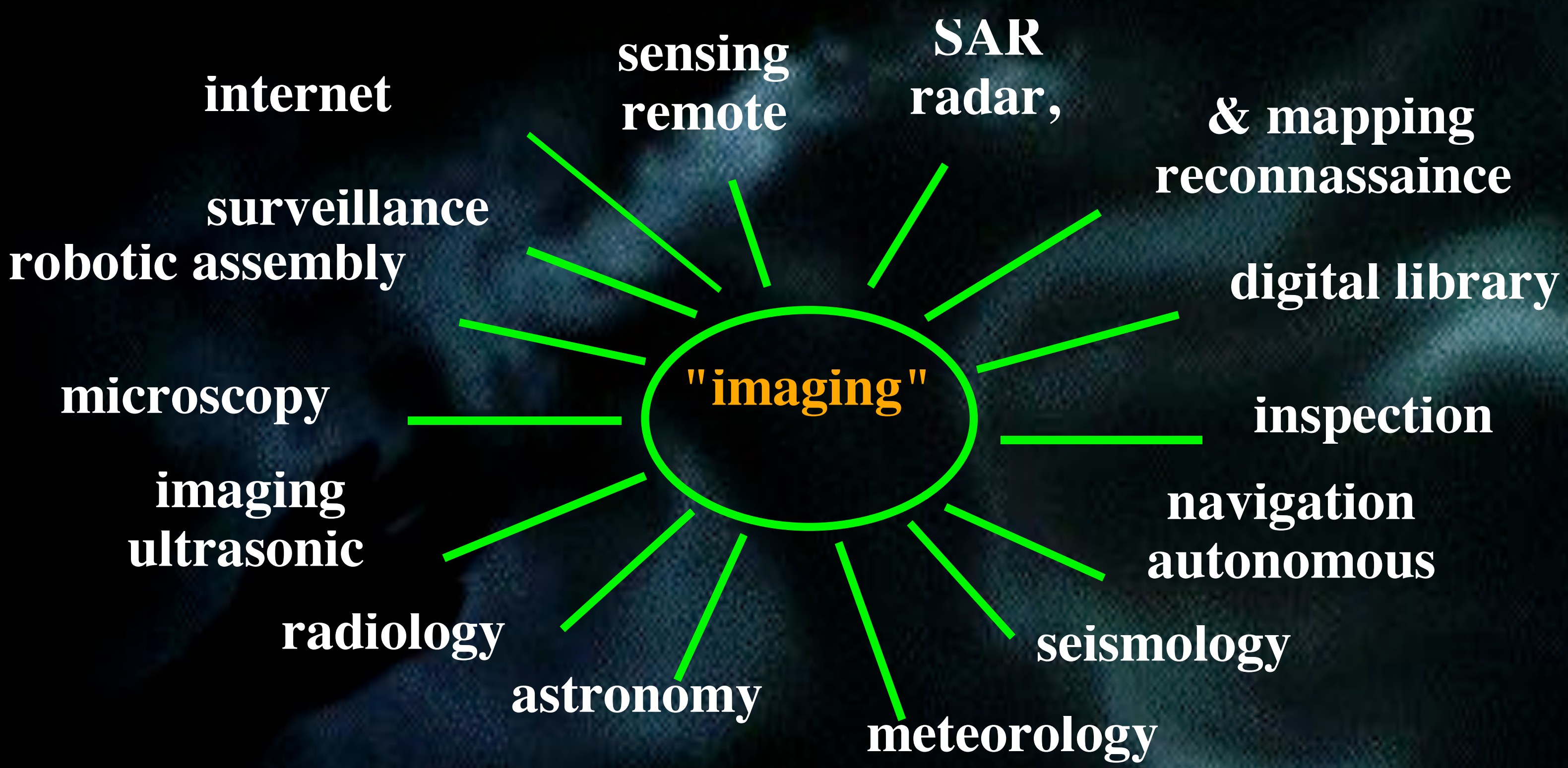
- Digital Image Processing
- (Digital Video Processing)

comp 471 / cart 498c
computer graphics:
real-time video
Monday 11 Sep 06

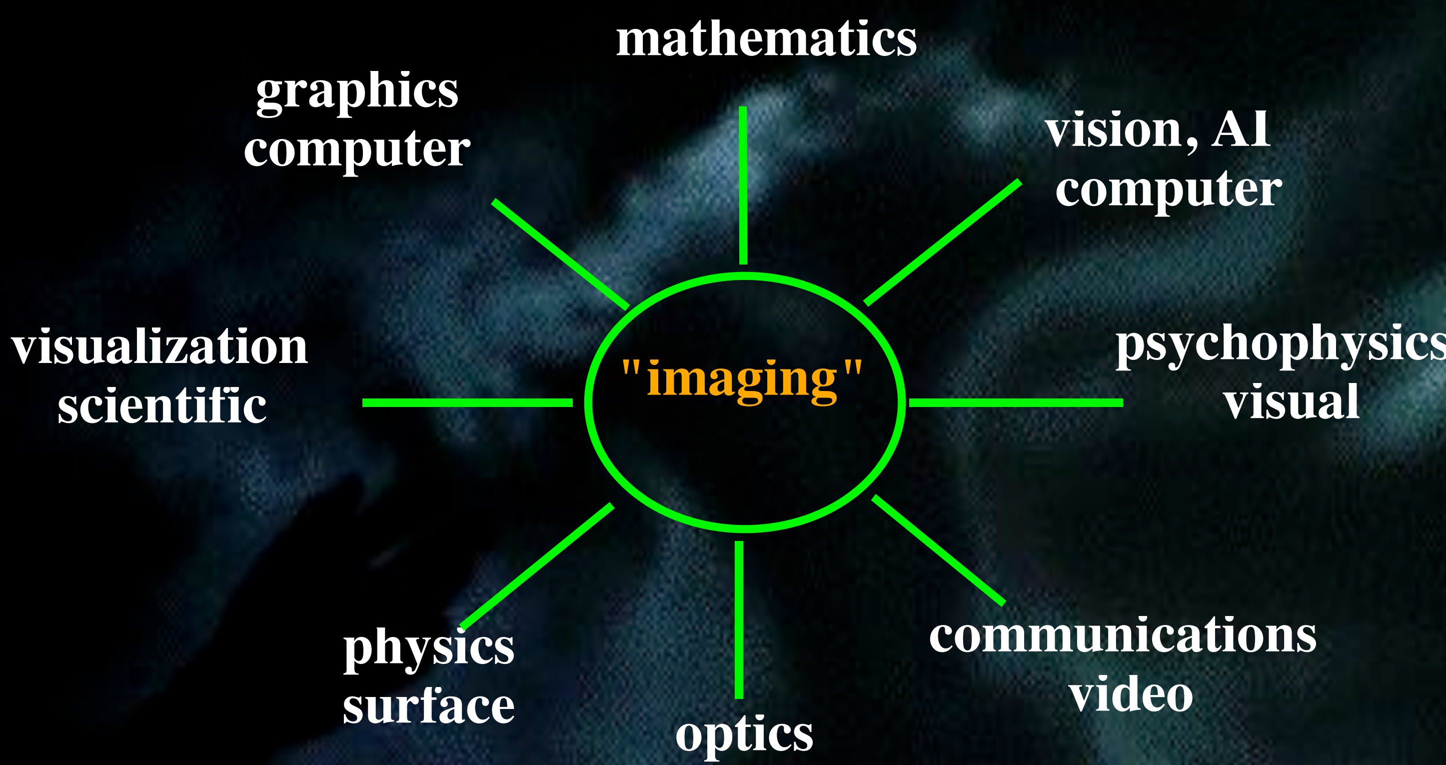
digital imaging

Digital Image Processing
Digital Video Processing

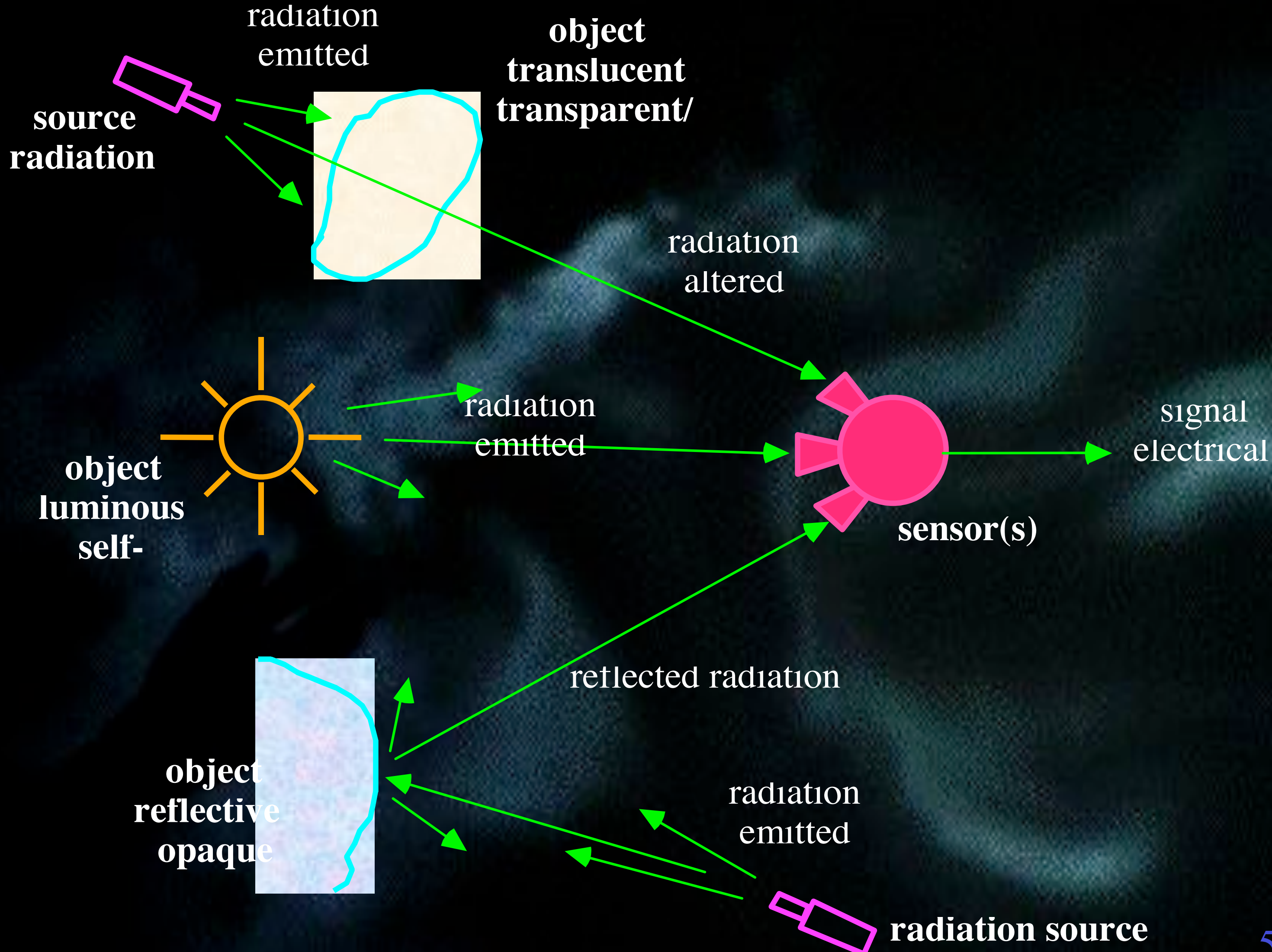
other applications of DIP/DVP



A Multidisciplinary Science



Three Types of Images



Type #1: Reflection Images

- Image information is **surface** information: how an object **reflects/absorbs** radiation
  - **Optical** (visual, photographic)
  - **Radar**
  - **Ultrasound, sonar** (non-EM)
  - **Electron microscopy**

Type #2: Emission Images

- Image information is **internal** information: how an object **creates** radiation
  - Thermal, infrared (FLIR)
  - Astronomy (stars, nebulae, etc.)
  - Nuclear (particle emission, e.g., **MRI**)

Type #3: Absorption Images

- Image information is **internal** information: how an object **modifies/absorbs** radiation

  - X-Rays in many applications

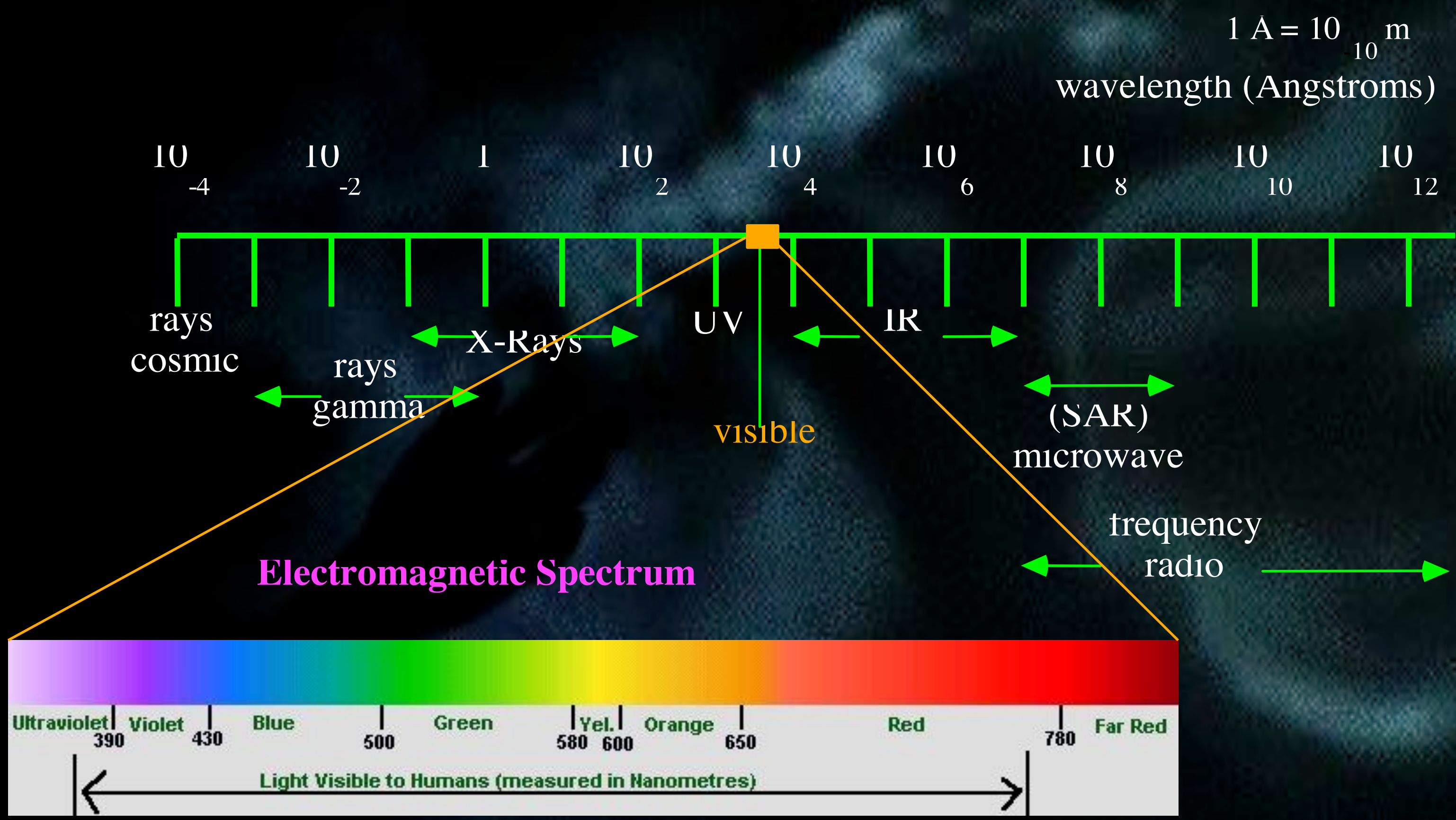
  - **Brightfield optical microscopy**

  - Tomography (**CAT**, PET) in medicine

  - “**Vibro-Seis**” in geophysical prospecting

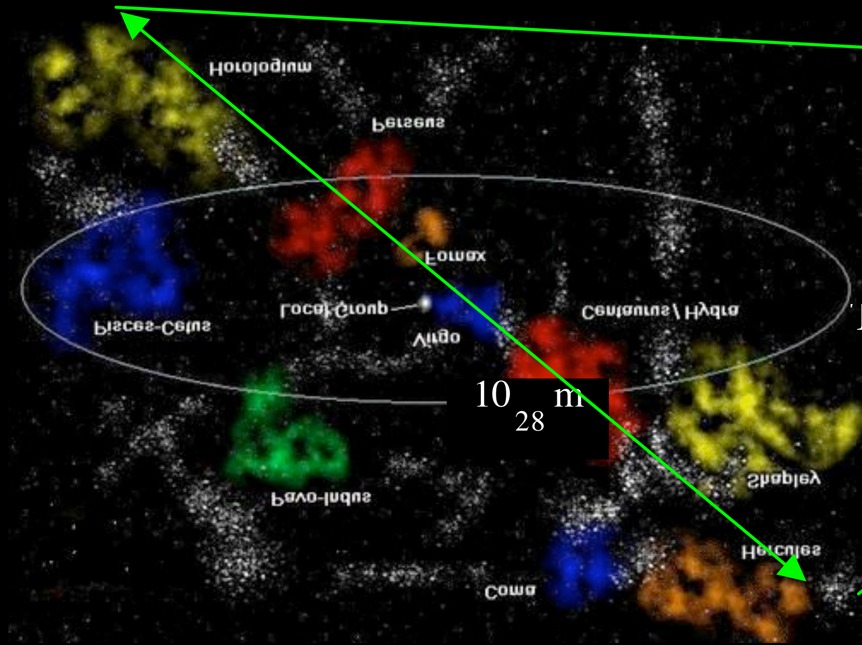
Electromagnetic Radiation

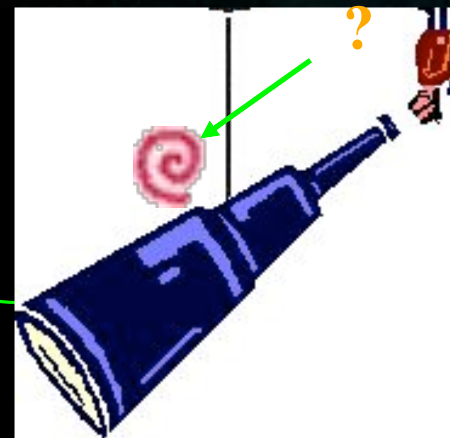
All this is used by "imagers"…



Scales of Imaging

From the **gigantic**…

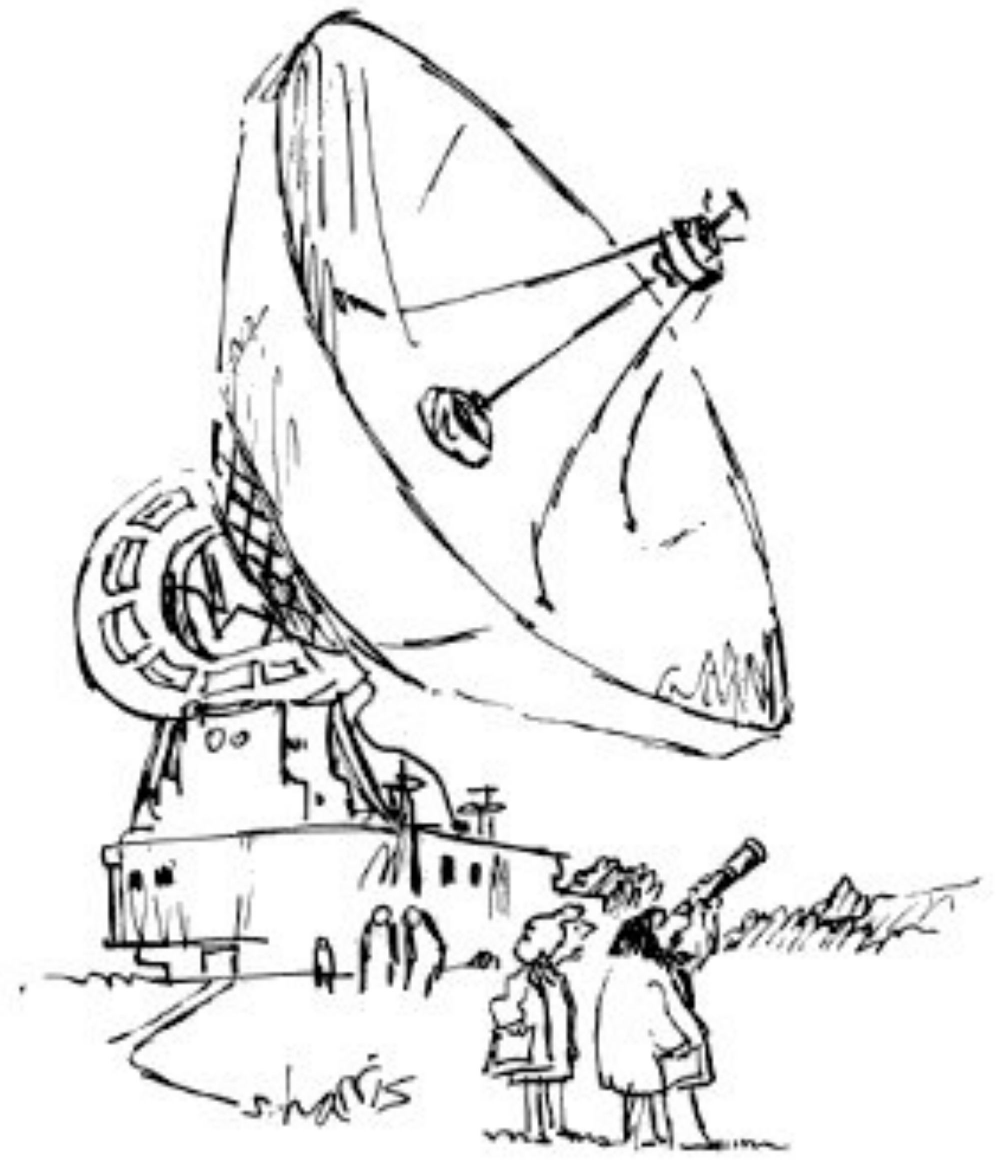




(of galaxies)

The Great Wall

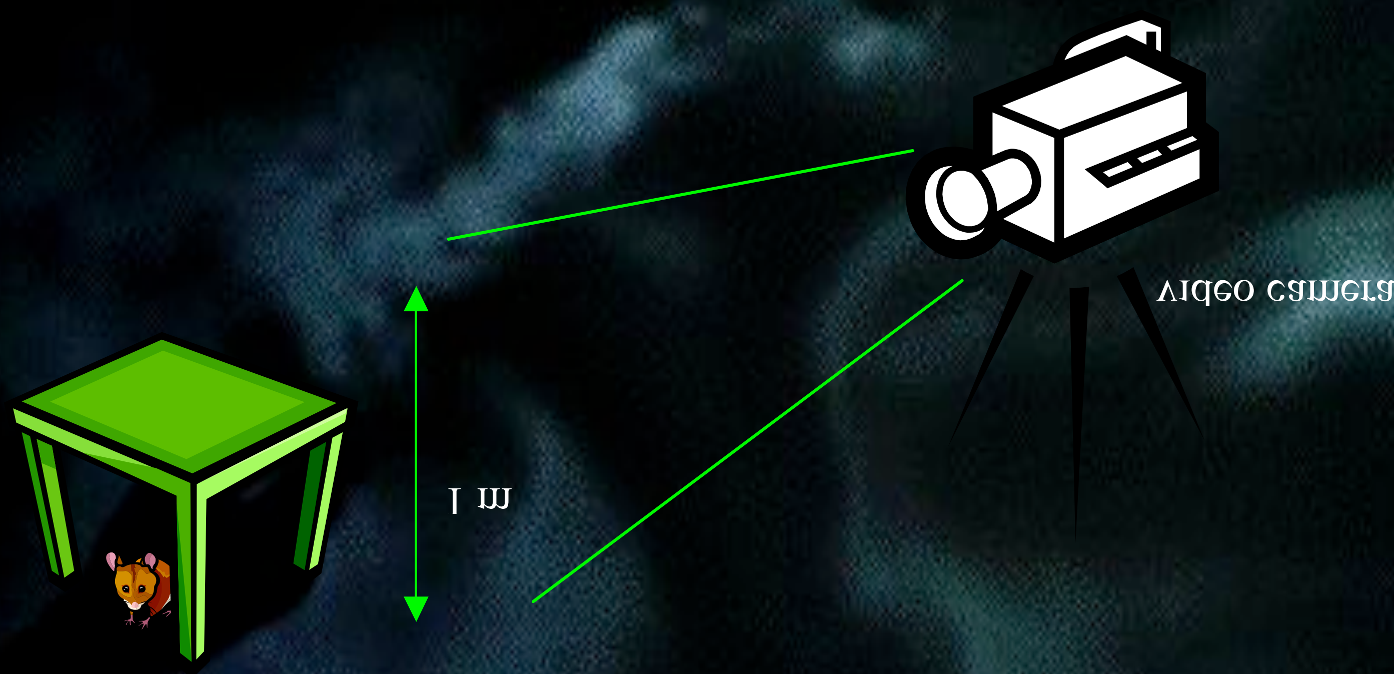
10^{28} m



"Just checking."

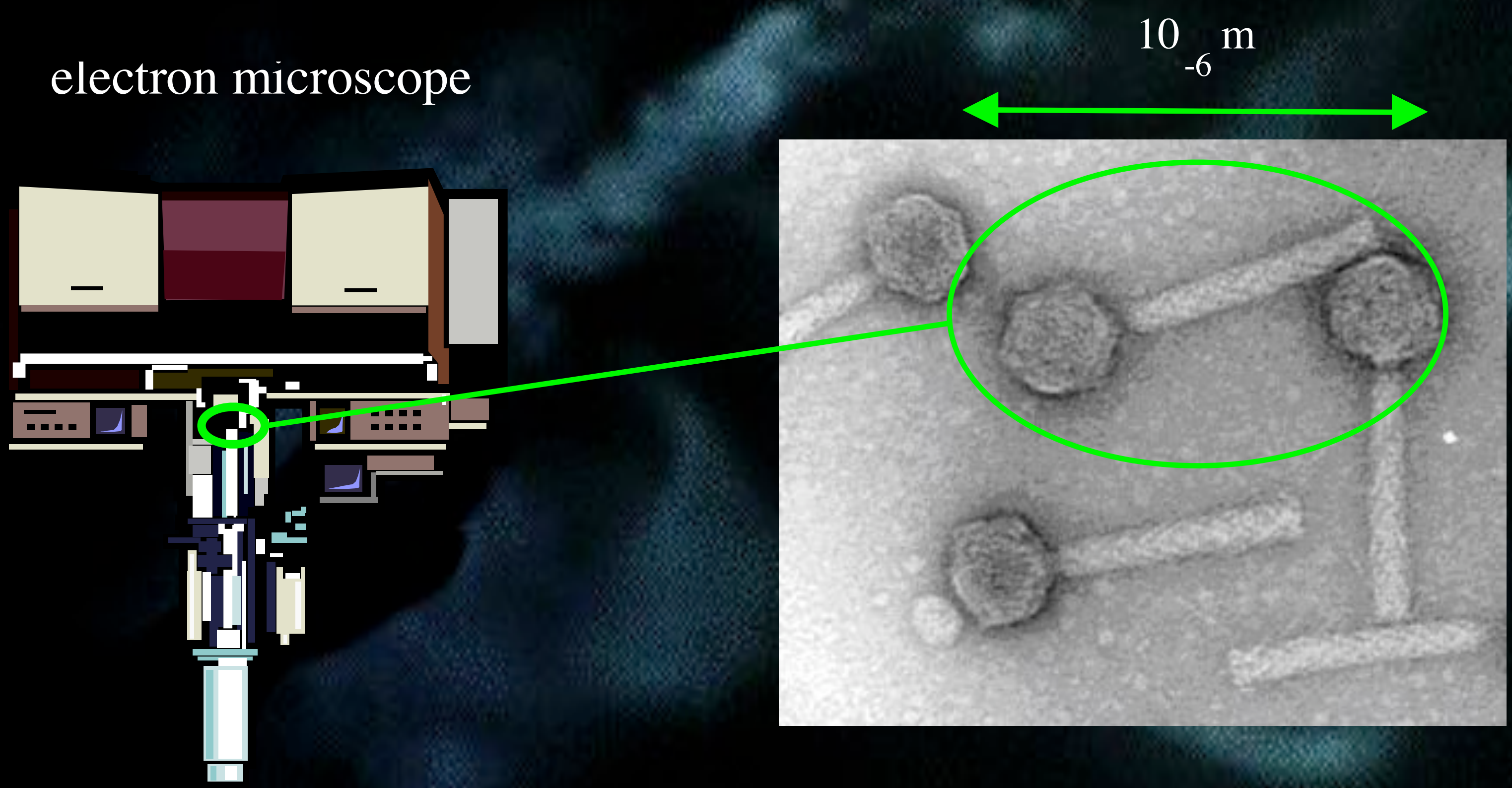
Scales of Imaging

…to the **everyday** …



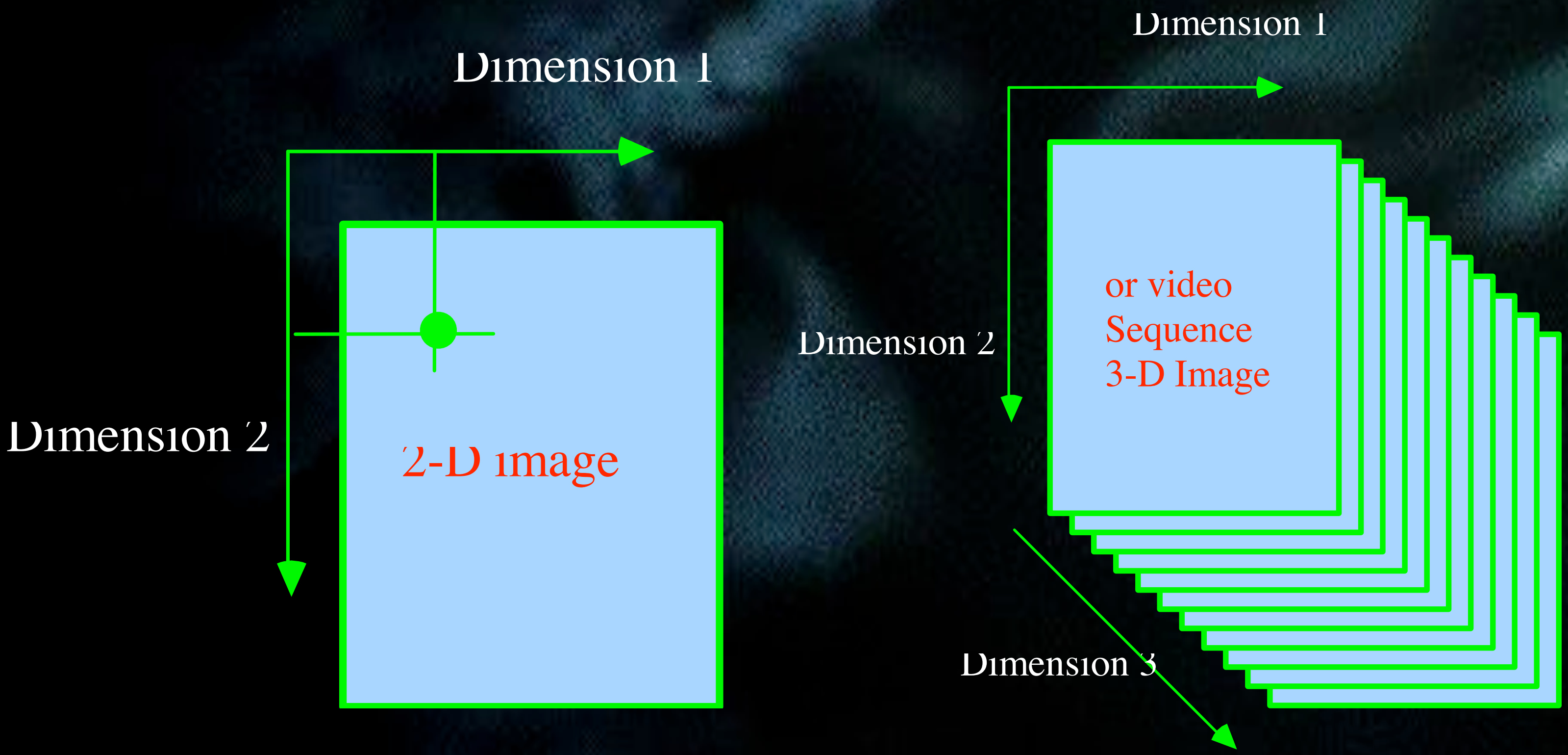
Scales of Imaging

…to the **tiny**.



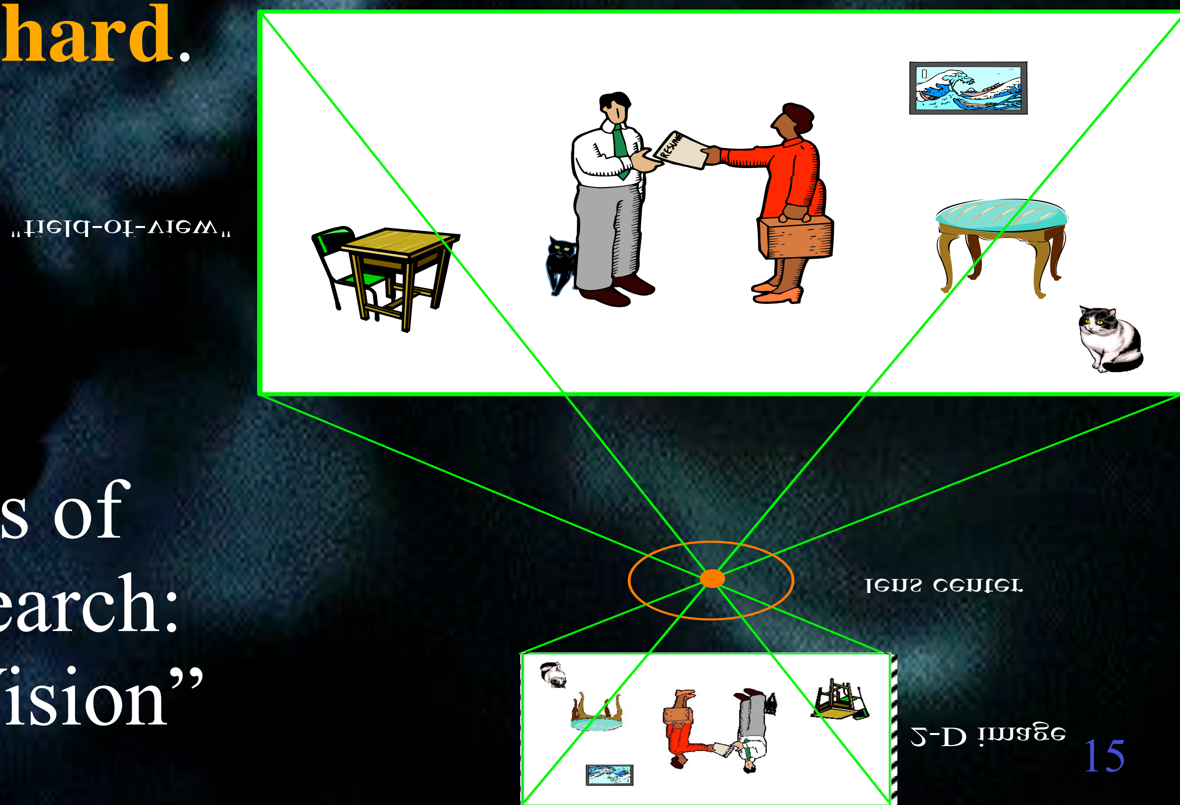
Dimensionality of Images

- Images and videos are **multi-dimensional** (≥ 2 dimensions) signals.



<u>3D-to-2D Projection</u>

- Image projection is a **reduction of dimension** (3D-to-2D): 3-D info is **lost**. Getting this info back is **very hard**.



- It is a topic of many years of intensive research: "Computer Vision"

"The image is not the object"

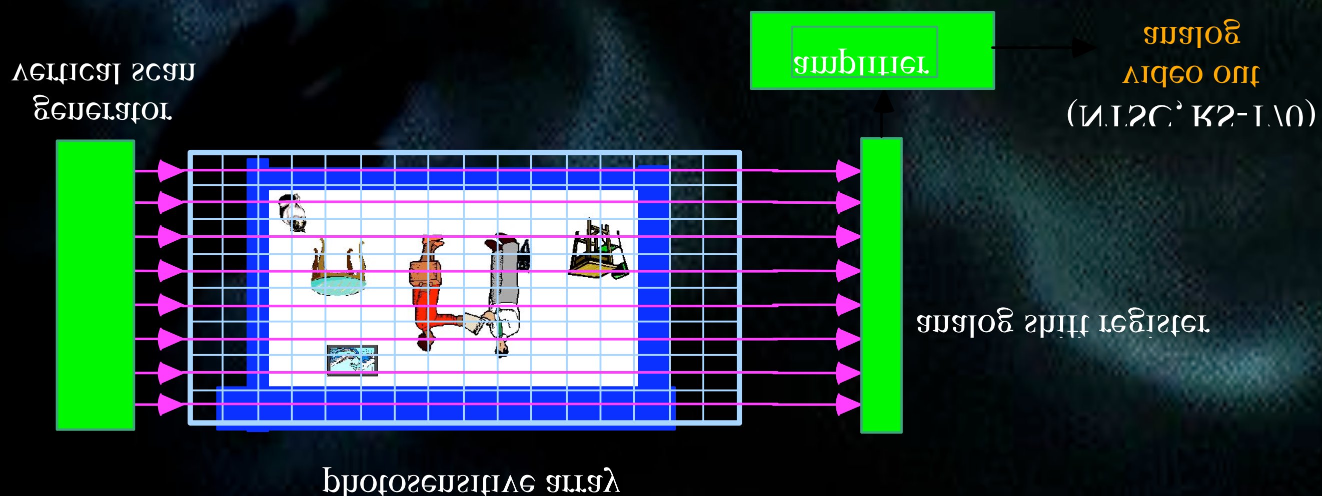
Rene Magritte (1898-1967)

Vision is a RELATION:
R(object, subject, ambient)
R(__?__, subject, ambient)
R(object,__?__, ambient)

digital image

CCD Image Sensing

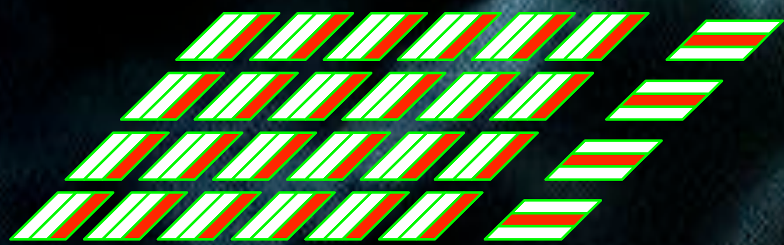
- Modern digital cameras sense **2-D images** charge-coupled device (CCD) sensor arrays.
- The output is typically a line-by-line (raster) analog signal:



CCD Image Creation

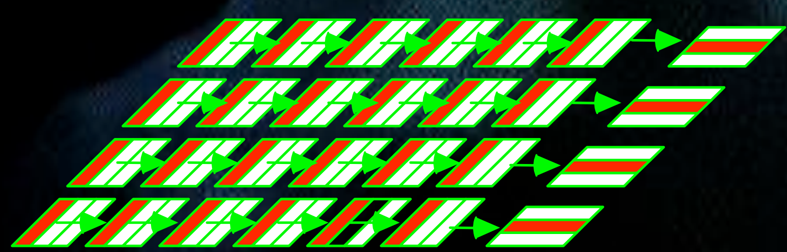
- Each CCD array cell has three "potential wells." At some instant, the middle "well" has a charge applied to it.

- Each **photon** strike creates an **electron.** The # of electrons created is proportional to the # of photons.
- At each clock the electrons are shifted twice by shifting the charges on the wells.



shift register

- At the second shift the electrons at the end sensor are shifted into the shift register



- The electrons are then shifted into an amplifier outputting a current with **voltage potential** proportional to the # of electrons

- The amplifier output is a line-by-line video **analog** waveform of standard format, e.g.

  NTSC: 525 lines/frame, 30 frames/sec

- For computer processing, the analog image must undergo **A/D Conversion**.

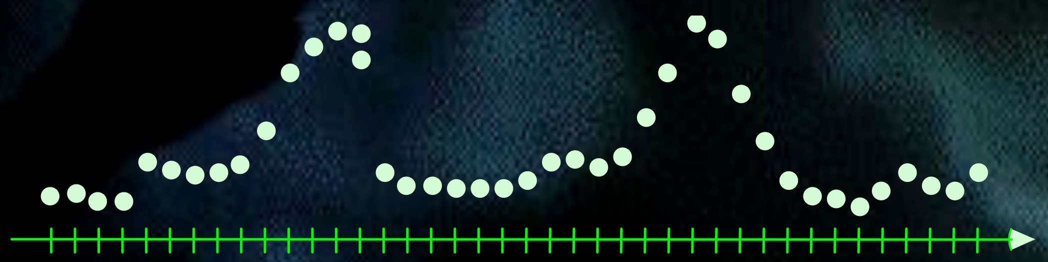
A/D Conversion

- Consists of **sampling** and **quantization**.
- **Sampling** is the process of creating a signal that is defined only at **discrete points**, from one that is continuously defined.
- **Quantization** is the process of converting each sample into a **finite** digital representation.
- Analog vs Digital Video
  - IEEE 1394 = Firewire
  - cable length limitation

Sampling

- Each video **raster** is converted from a **continuous voltage waveform** into a sequence of **voltage samples**:

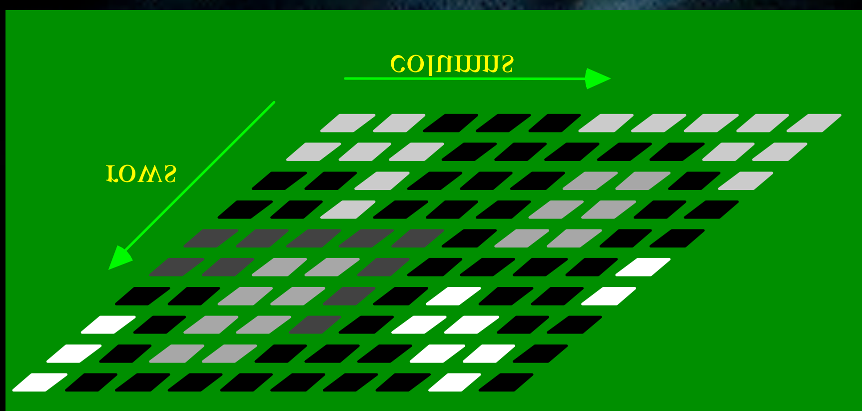
continuous electrical signal from one scanline



sampled electrical signal from one scanline

Sampled Image

- A **sampled image** is an array of numbers (row, column) representing image intensities



depiction of 10 x 10 image array

- Each of these **picture elements** is called a **pixel**.

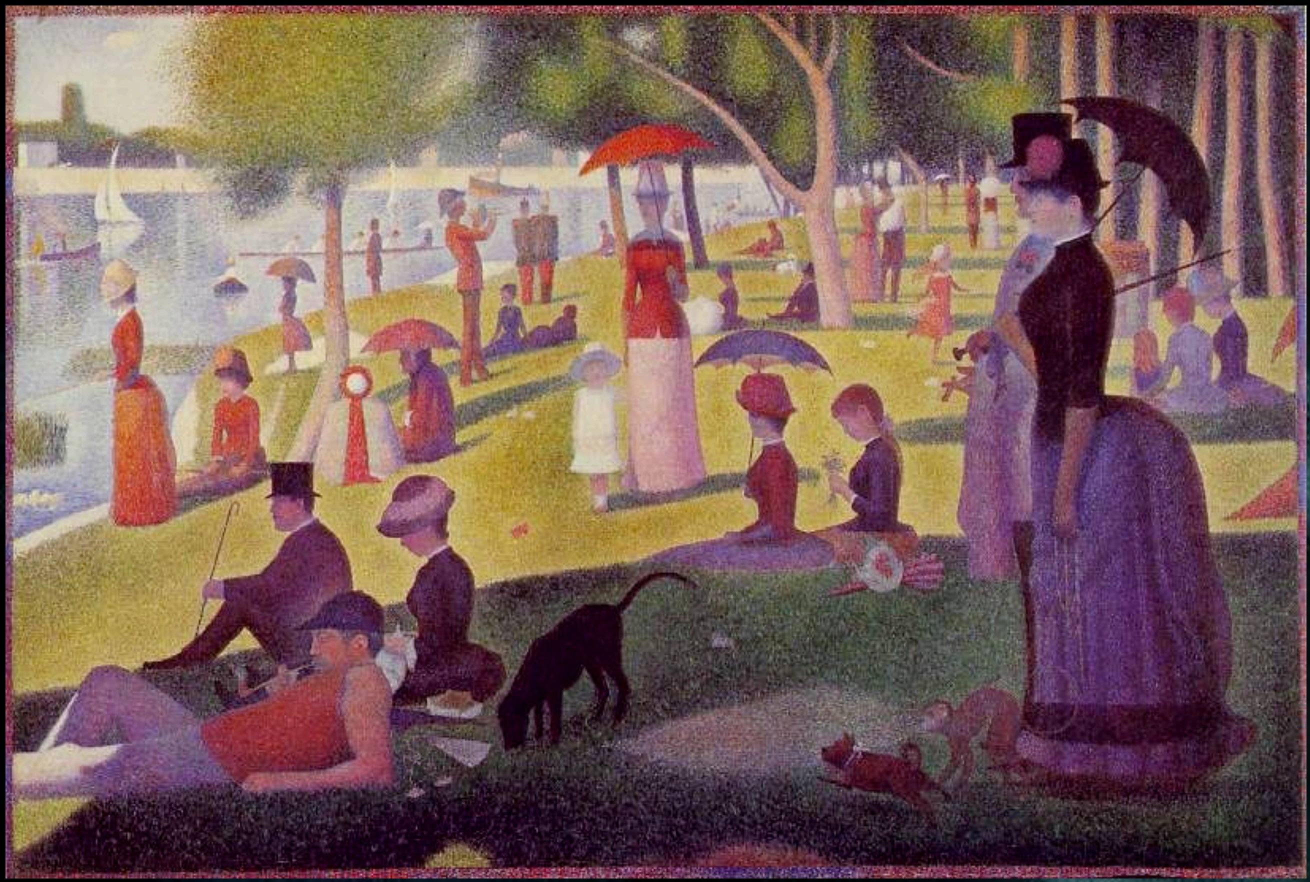
Sampled Image

- The image array is rectangular (N x M) with dimensions $N = 2^P$ and $M = 2^Q$ (why?)
- Examples: square images
- P=Q=7 128 x 128 (2^{14} ≈ 16,000 pixels)
- P=Q= 8 256 x 256 (2^{16} ≈ 65,500 pixels)
- P=Q= 9 **512 x 512** (2^{18} ≈ 262,000 pixels)
- P=Q=10 1024x1024 (2^{20} ≈ 1,000,000 pixels)

Sampling Effects

- It is essential that the image be sampled **sufficiently densely**; else the image quality will be severely degraded.
- Can be expressed via the Sampling Theorem) but the **visual effects** are most important (*make your own example!*)
- With sufficient samples, the image **appears continuous**…..

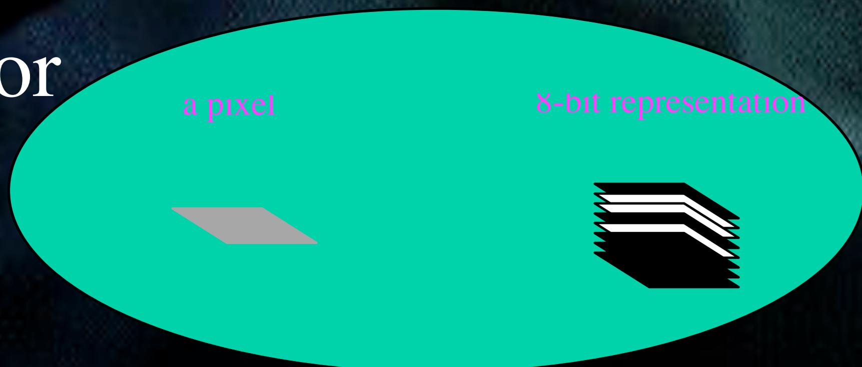
Sampling in Art



Seurat - *La Grande Jatte* – Pointillist work took 2 years to create

Quantization

- Each **gray level** is quantized: assigned an integer indexed from 0 to K-1.
- Typically there $K = 2^B$ possible gray levels.
- Each pixel is represented by B bits, where usually $\mathbf{1 \leq B \leq 8}$.
- 24bit Color

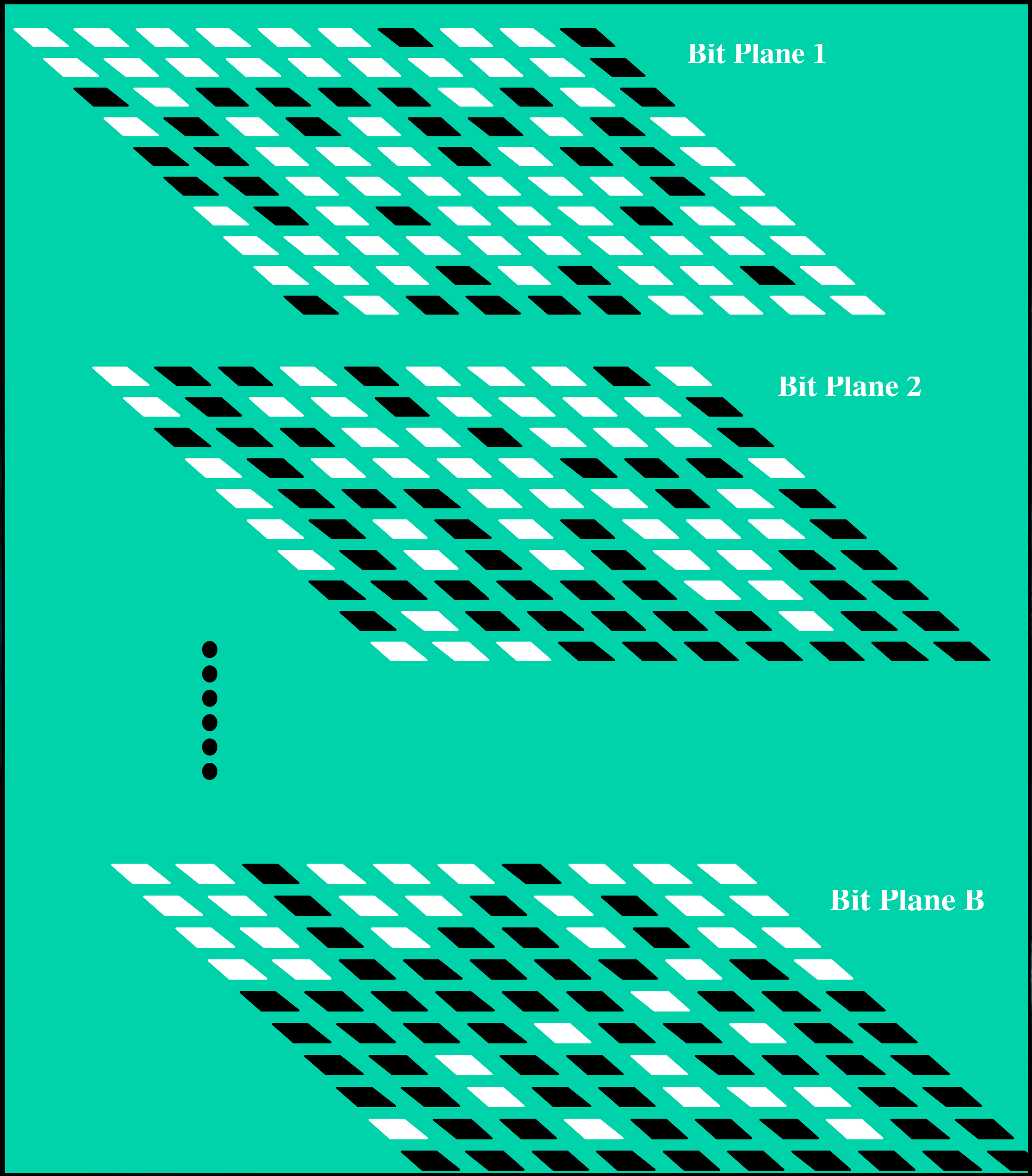


Quantization

- The pixel intensities or gray levels must be quantized **sufficiently densely** so that excessive information is not lost.

- This is **hard** to express mathematically, but again, quantization effects are **visually obvious** (*make your own example!*)

Image as a Set of Bit Planes



>> Image Notation <<

- Denote an **image matrix**

$$\mathbf{I} = [I(i, j);\ 0 \le i \le N\text{-}1,\ 0 \le j \le M\text{-}1\ ]$$

where

(i, j) = (row, column)

I(i, j) = image value at (i, j)

$$\mathbf{I} = \begin{bmatrix} I(0,0) & I(0,1) & \cdots & I(0,M-1) \\ I(1,0) & I(1,1) & \cdots & I(1,M-1) \\ \vdots & \vdots & & \vdots \\ I(N-1,0) & I(N-1,1) & \cdots & I(N-1,M-1) \end{bmatrix}$$

or I(**n**), where **n** = vector (i,j) in Z x Z

Common Image Formats

- JPEG (Joint Photographic Experts Group) images are compressed with loss – see Module 7. All digital cameras today have the option to save images in JPEG format. File extension: image.jpg
- TIFF (Tagged Image File Format) images can be lossless (LZW compressed) or compressed with loss. Widely used in the printing industry and supported by many image processing programs. File extension: image.tif
- GIF (Graphic Interchange Format) an old but still-common format, limited to 256 colors. Lossless and lossy (LZW) formats. File extension: image.gif
- PNG (Portable Network Graphics) is the successor to GIF. Supports true color (16 million colors). Somewhat new - not yet widely supported. File extension: image.png
- BMP (bit mapped) format is used internally by Microsoft Windows. Not compressed. Widely accepted. File extension: image.jbmp

The Image/Video Data Explosion

- Total **storage** required for **one digital image** with 2^P x 2^Q pixels spatial resolution and B bits / pixel gray-level resolution is

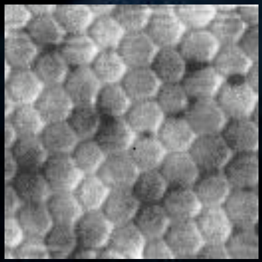
  B x 2^{P+Q} bits.
- Usually **B=8** and often **P=Q=9**. A common image size is then **¼ megabyte**.
- Five years ago this was **a lot**.

The Image/Video Data Explosion

- Storing **1 second** of a gray-level movie (TV rate = 30 images / sec) requires 7.5 Mbytes.
- A 2-hour gray-level video (8x512x512x30) requires 27,000 megabyte or **27 gigabytes of storage** at nowhere near theatre quality. That's a lot **today**.
- DIP/DVP includes ways to **compress** digital images and videos (*not this class*).

Sampling Tesselations

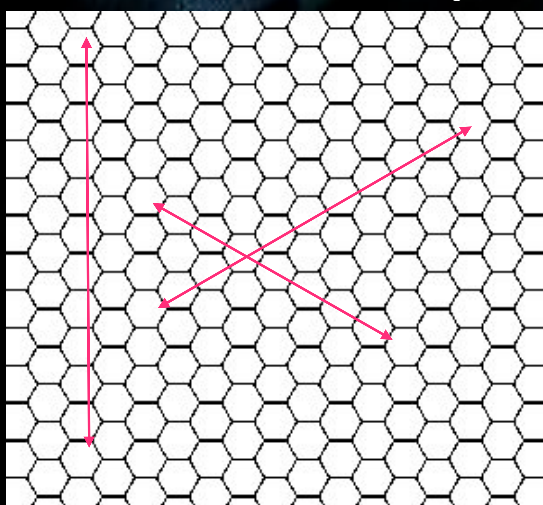
- Digital image processing systems almost always use **Cartesian (row, column) sampling** of images.
- **Simplicity of indexing** in *(procedural)* algorithms.
- Worth noting: the retina of the eye uses a hex sampling - packs pixels **more tightly**:

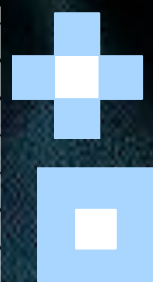


cone cells in the human fovea

Hexagonal Sampling

- Hex images can also be indexed by row-column, though the axes are not orthogonal.
- Hex sampling eliminates ambiguity in “connectivity”





4-connectivity

8-connectivity

What are the neighbors of a pixel in Cartesian coordinates?

Unambiguous hex neighbors.

Kepler Sphere Packing Problem (1611)

$$\frac{\pi}{\sqrt{18}}$$

Sir Walter Raleigh, how to pack the most cannonballs in a given volume

Kepler conjectured in 1611

Hexagonal Face-centered cubic lattice

Thomas Hales, University of Michigan

Hexagonally sampled image
(with exaggerated pixels)

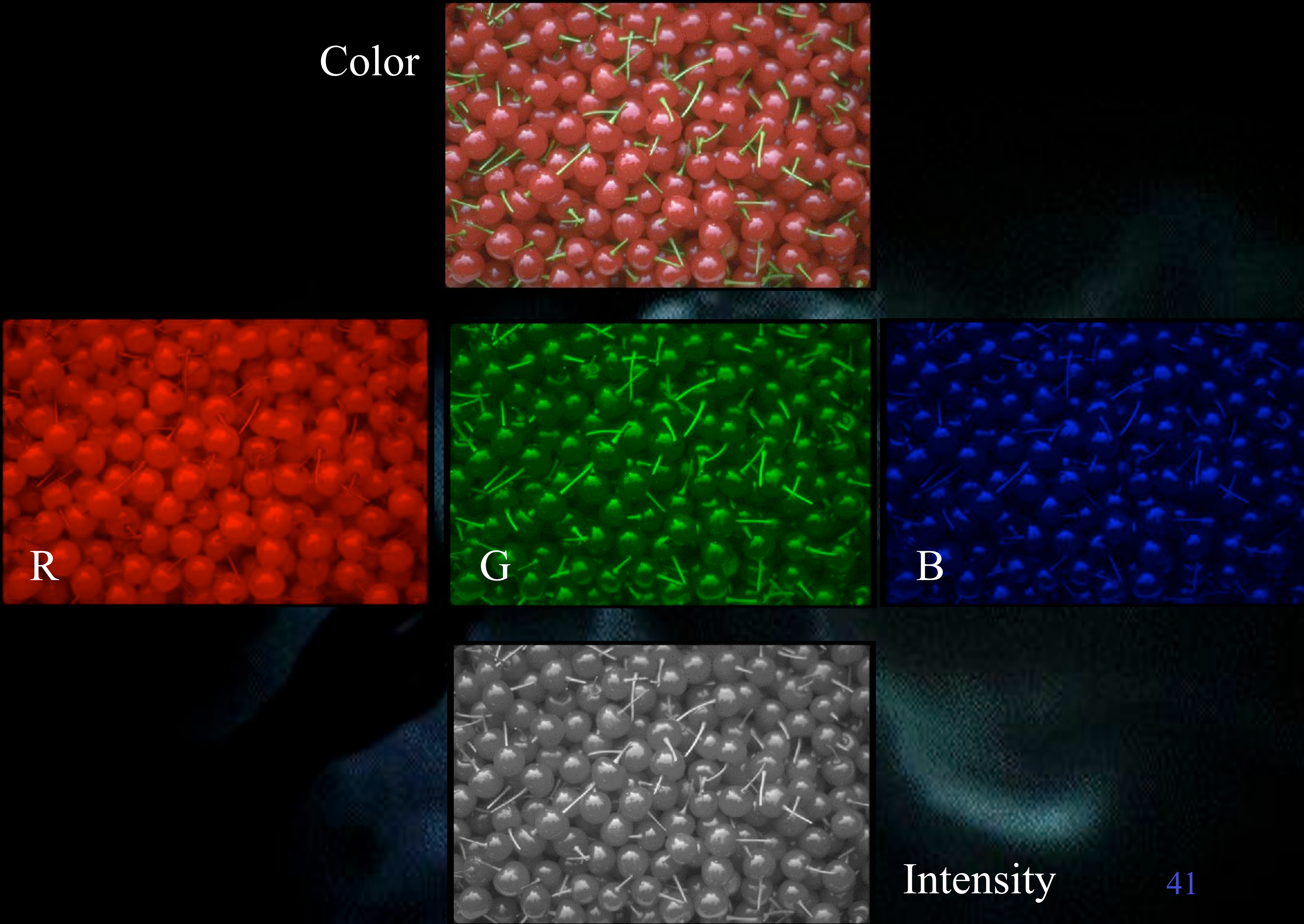
What About Color?

- Color is an important aspect of images.
- A **color image** is a **vector-valued** signal. At each pixel, the image has three values: Red, Green, and **Blue**.
- Usually expressed as three images: the Red, Green and Blue images: **RGB representation**.
- Although color is important, we will nearly always process the **intensity image** **I = R + G + B**.
- Most **color algorithms** process **R**, **G**, **B** components separately like gray-scale images then add the results.
- There are other color representations, e.g. HSB, CMYK (why also dim 3 ?).

Color is Important!

… in many ways… …although we can function without it

The Boating Party - Renoir

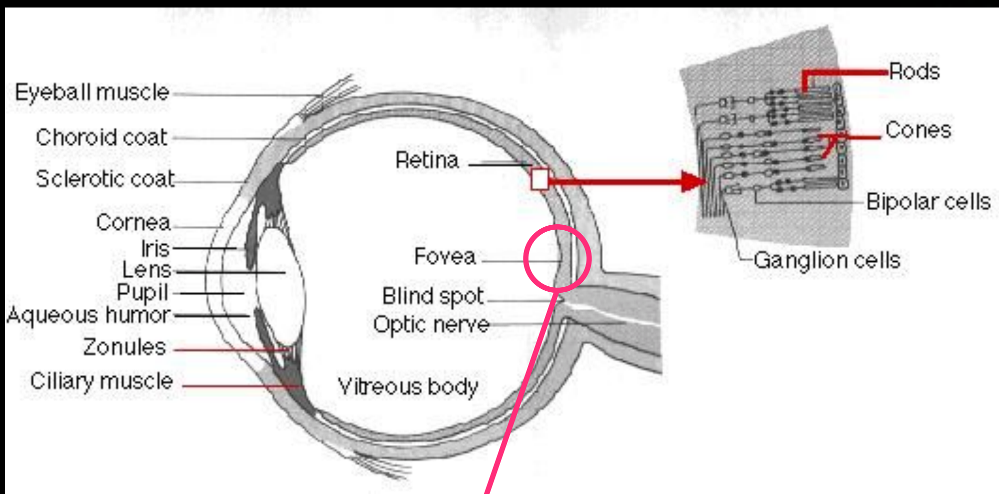

Color
R
G
B
Intensity

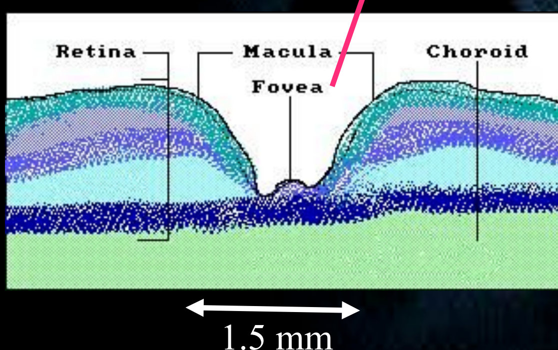
human vision

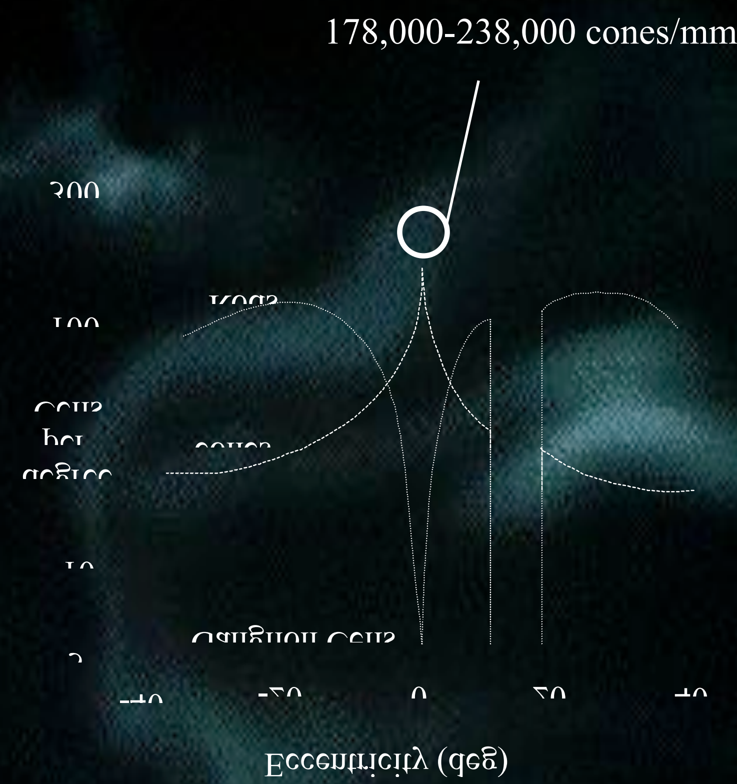
A Bit About Visual Perception

- In most cases, the **intended receiver** of the result of image/video processing or communications algorithms is the **human eye**.
- A fair amount is known about the eye:
  - the neurons (rods, cones) sample and quantize
  - the retinal ganglion and cortical cells linearly filter

The Eye - Structure







• Notice that image sampling at the retina is **highly nonuniform**!

Eye Movement

- The eyes **move constantly, t**o place/keep the fovea on places of interest.
- There are five major types of eye movement:

  - **saccadic** (attentional)

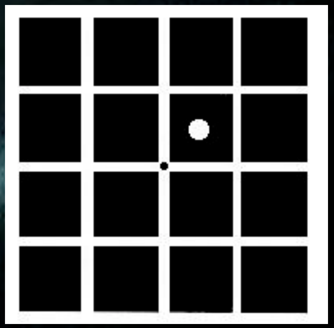
  - **pursuit** (smooth tracking)

  - **vestibular** (head movement compensating)

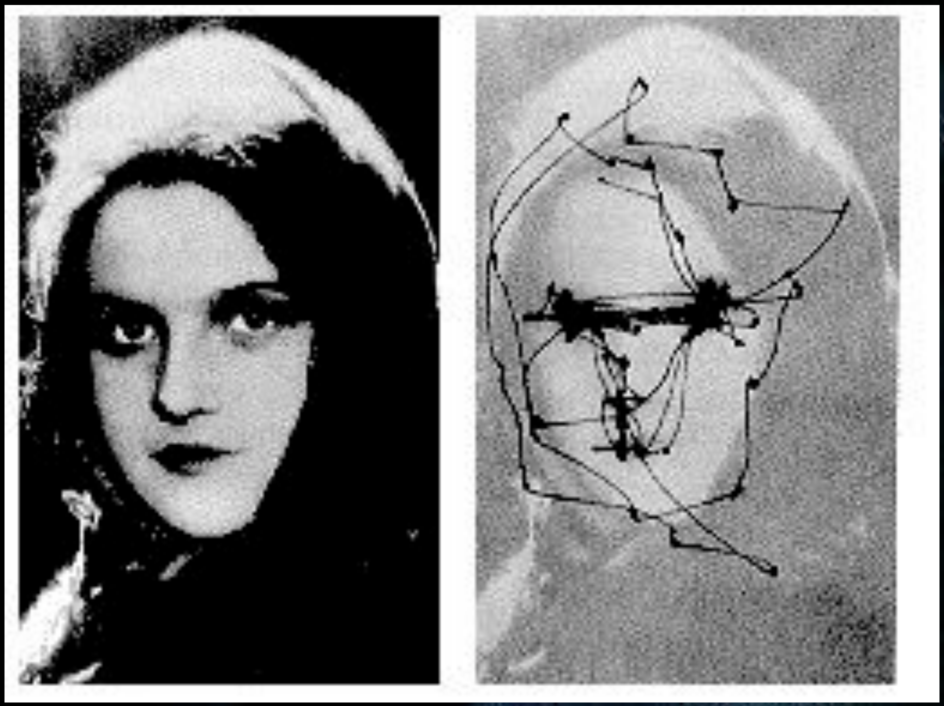
  - **microsaccadic** (tiny; image persistency)

  - **vergence** (stereoscopic)

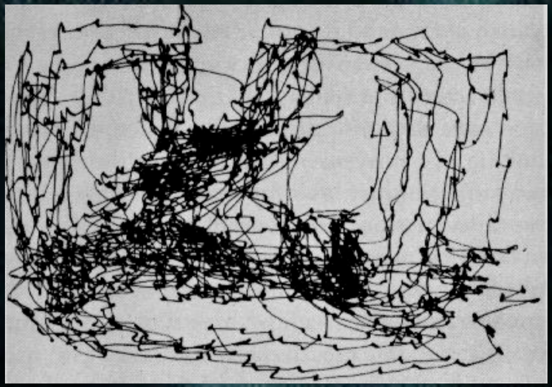
To demonstrate microsaccades, first fixate the center of the white dot for 10 sec, then fixate the small black dot. Small displacememts of the afterimage are then obvious -- the slow drifting movements as well as the corrective microsaccades.



Saccades and Fixations



Highly contextual

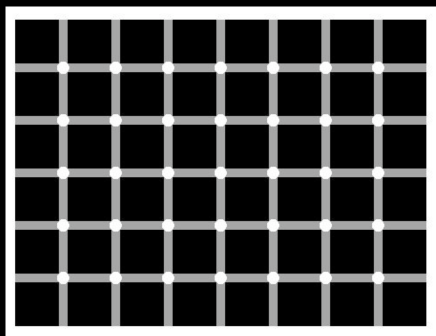


Less contextual

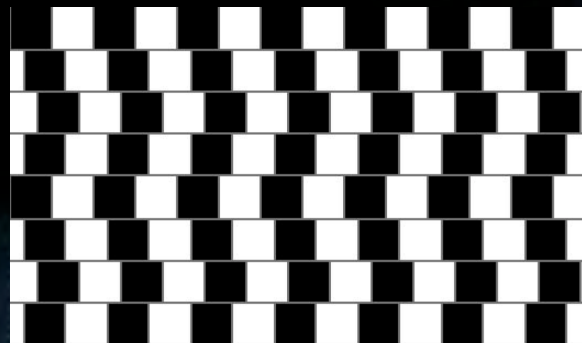
phenomenology of vision

- Constancy of scene is a construction!
- Object is a construction!

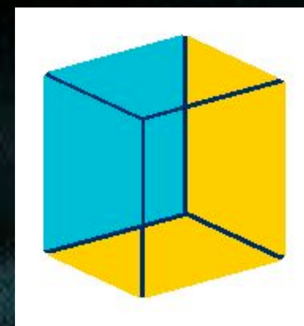
Visual ~~Illusions~~ Constructions



Find the black dot

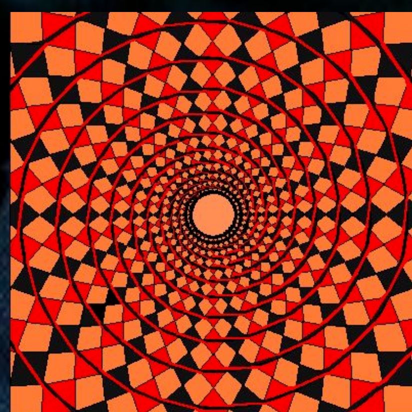


Which lines are straight?

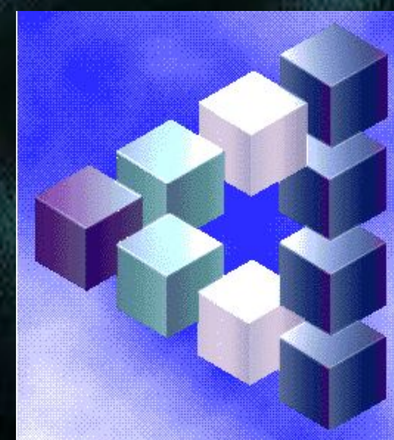


Which face is blue?

The Mars “face”

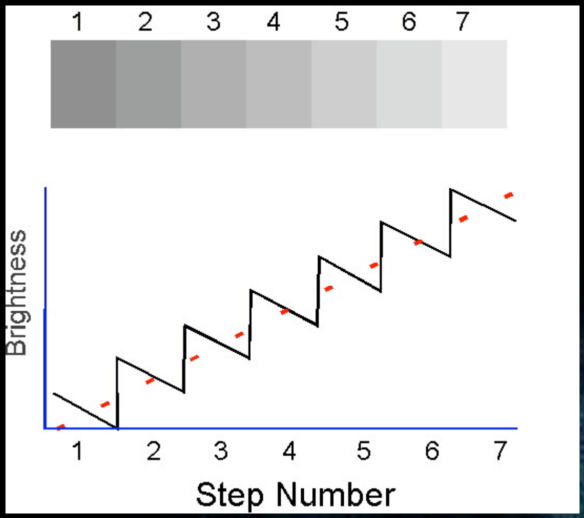


Spiral?

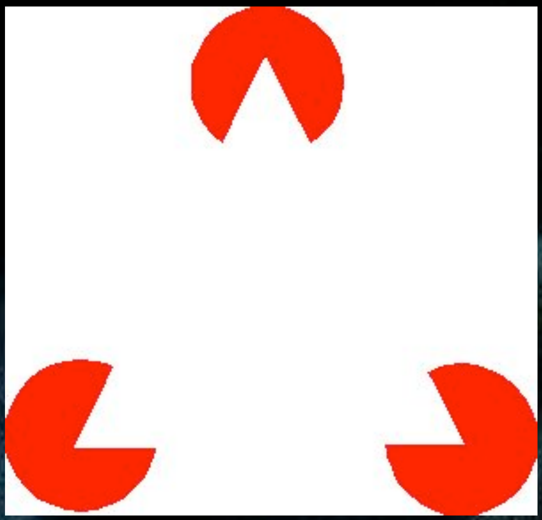


Triangle?

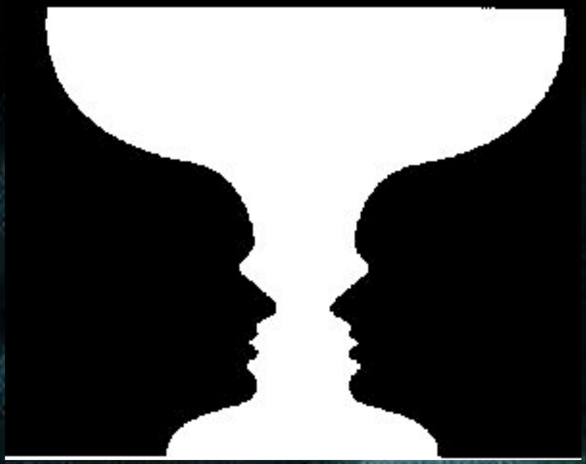
More Visual Constructions



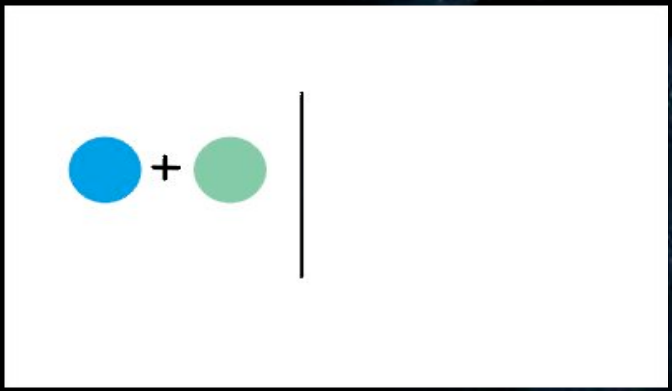
Mach Bands



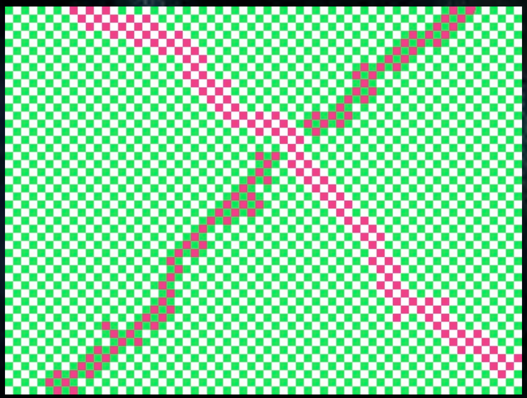
Kanisza Triangle



Reversible Image



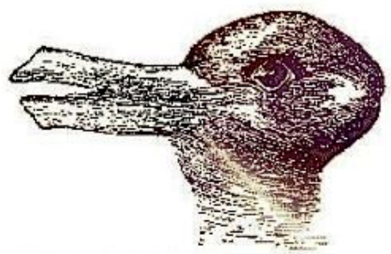
Afterimages



How many colors?

Say each color, not the words…

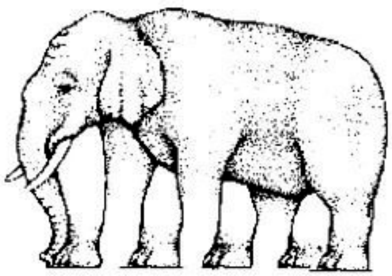
Even More Visual Constructions



A Rabbit…. Or A Duck?
hint: the duck is looking left, the rabbit is looking right

Man Playing Horn… Or Woman Sillhouette?
(hint: woman’s right eye is the black speck in front of horn handle)

Woman In Vanity… Or Skull?
hint: move farther a bit from the screen and blink to see the skull or the woman (looking at the mirror)

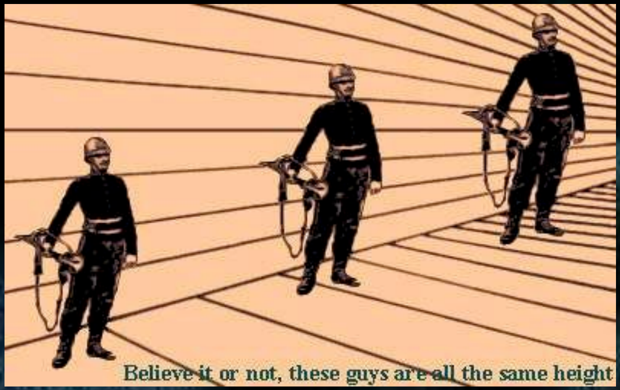


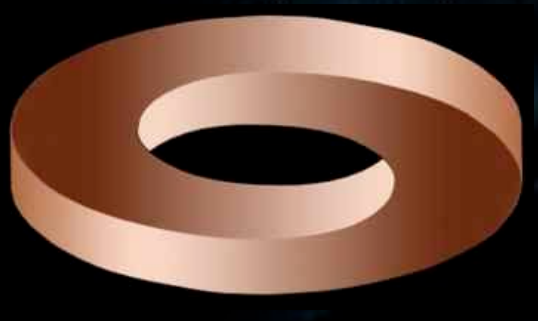
How many legs does this elephant have?

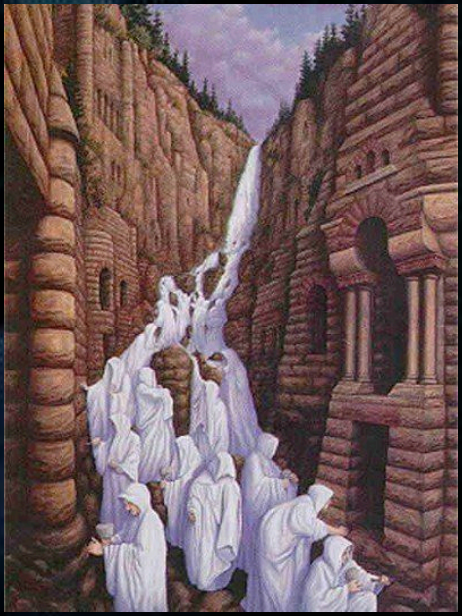
Old Woman…Or Young Girl?
hint: The old woman's nose is the young girls chin.

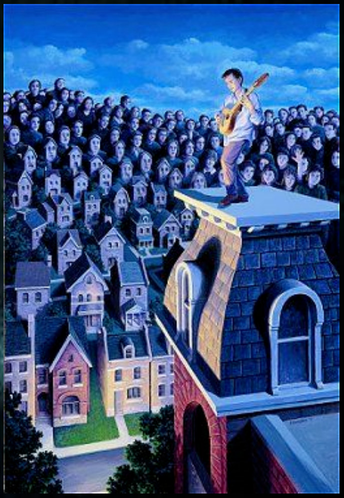
"Illusions" involving object shapes

Yet More Visual Constructions

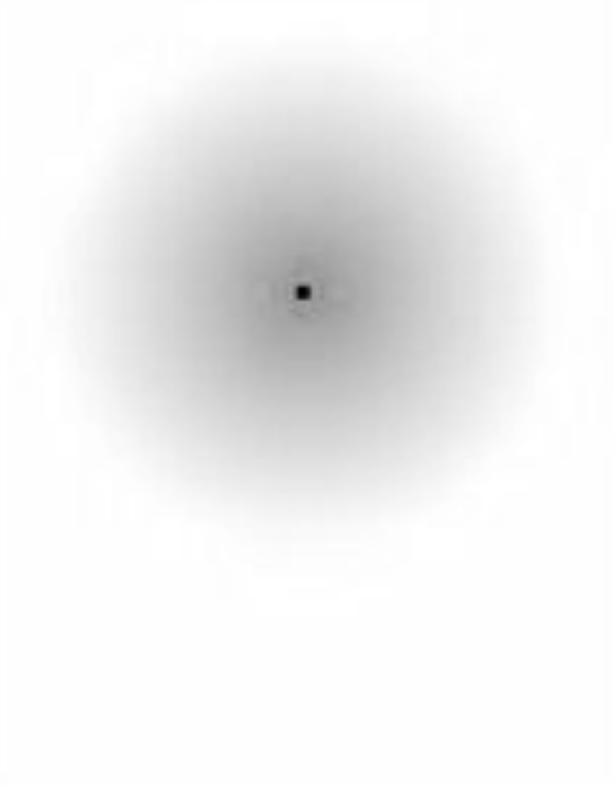


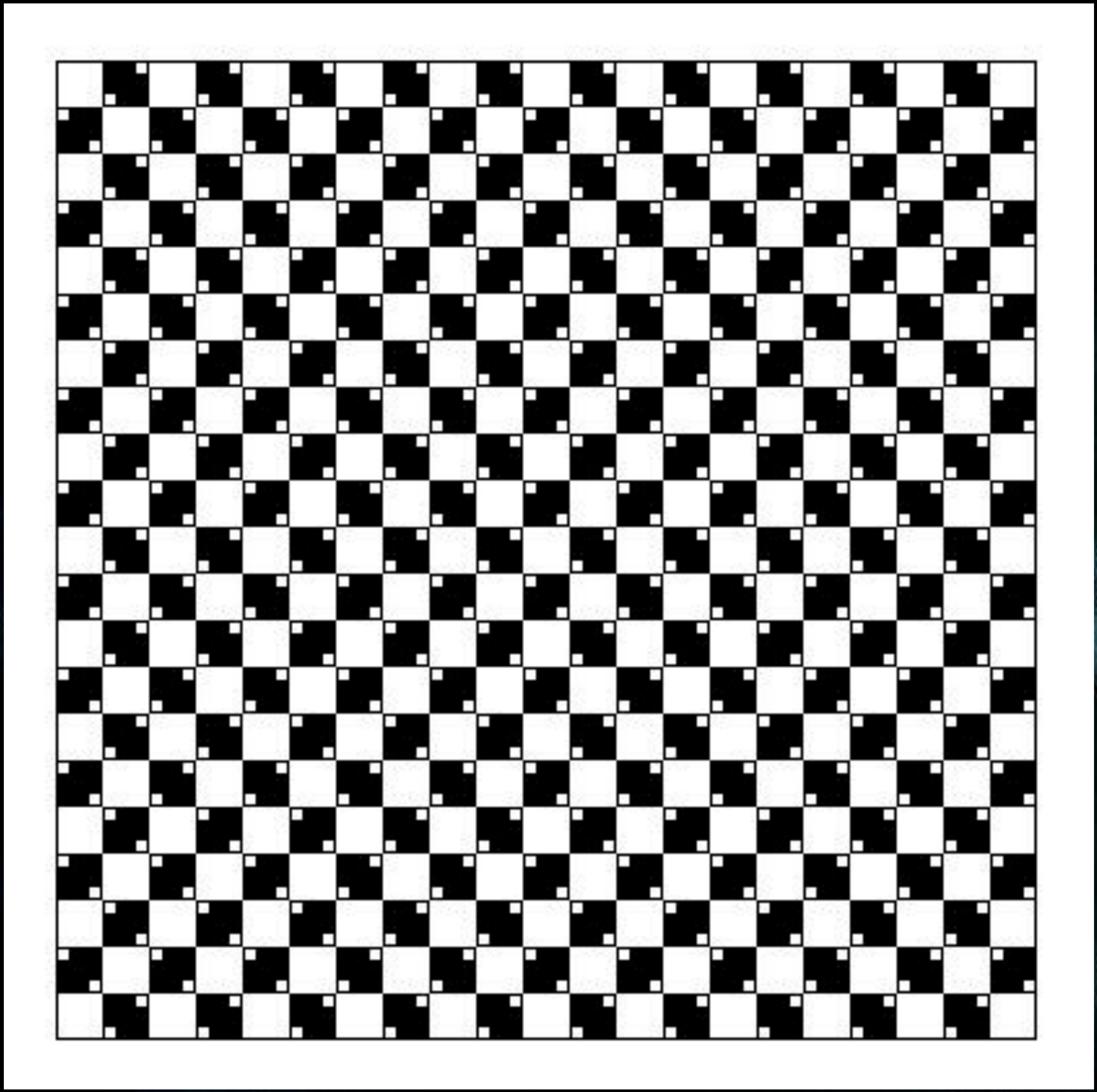




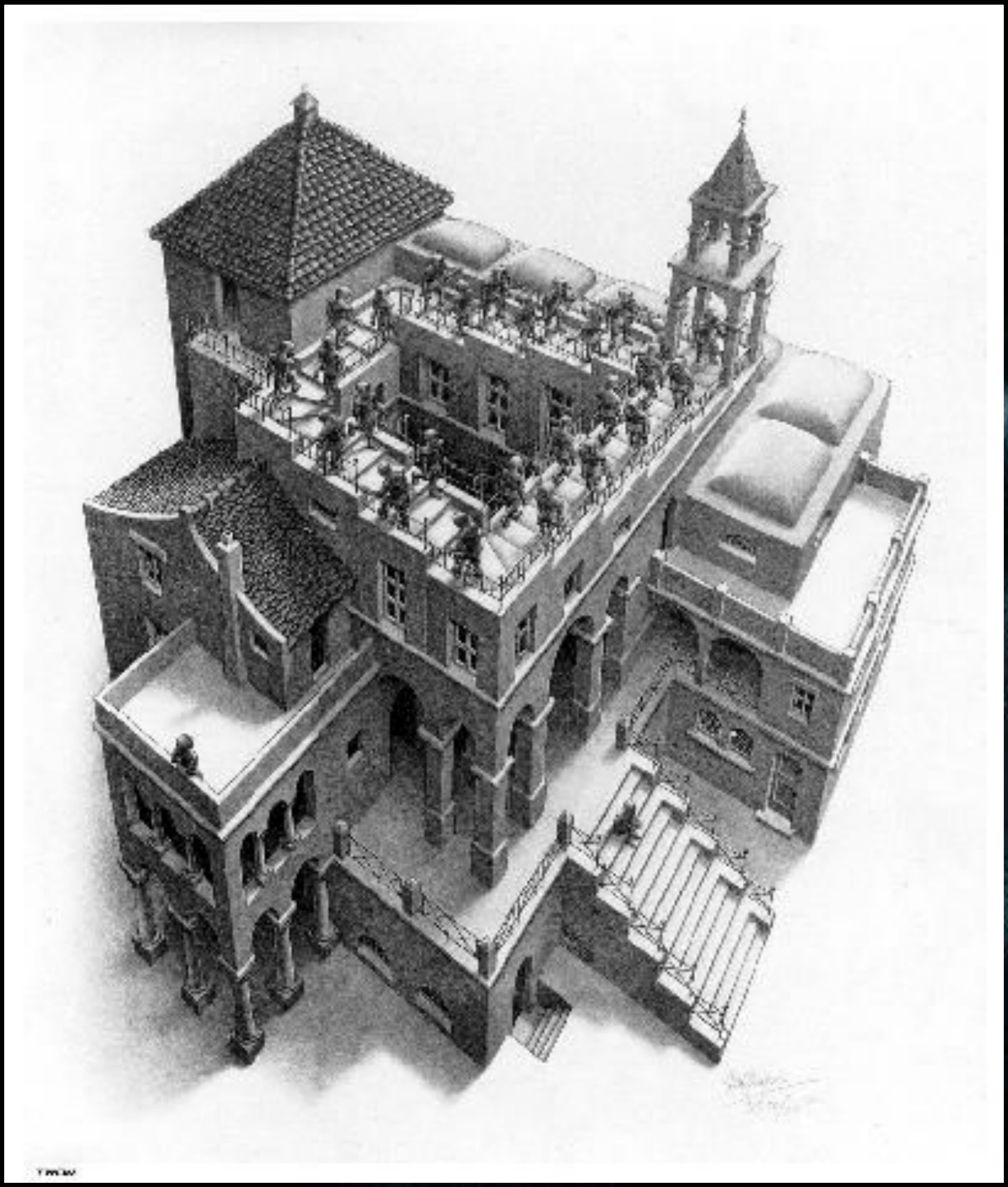


Keep staring at the black dot. After a while the gray haze around it will appear to shrink.





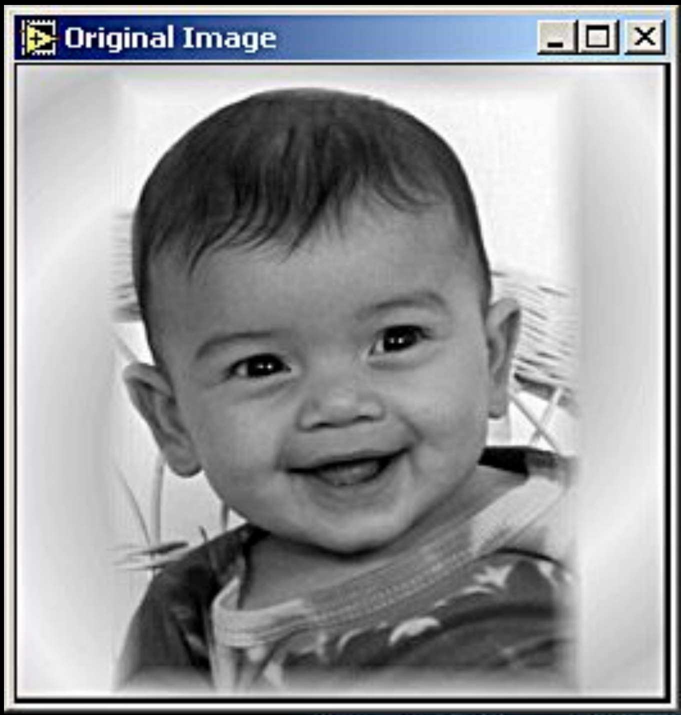
Yes, perfectly straight lines…



Ascending and Descending

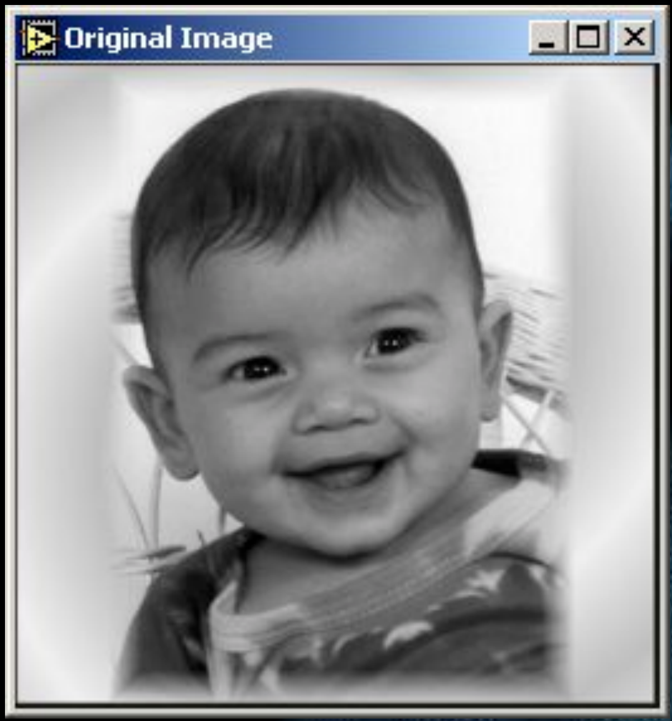
M.C. Escher

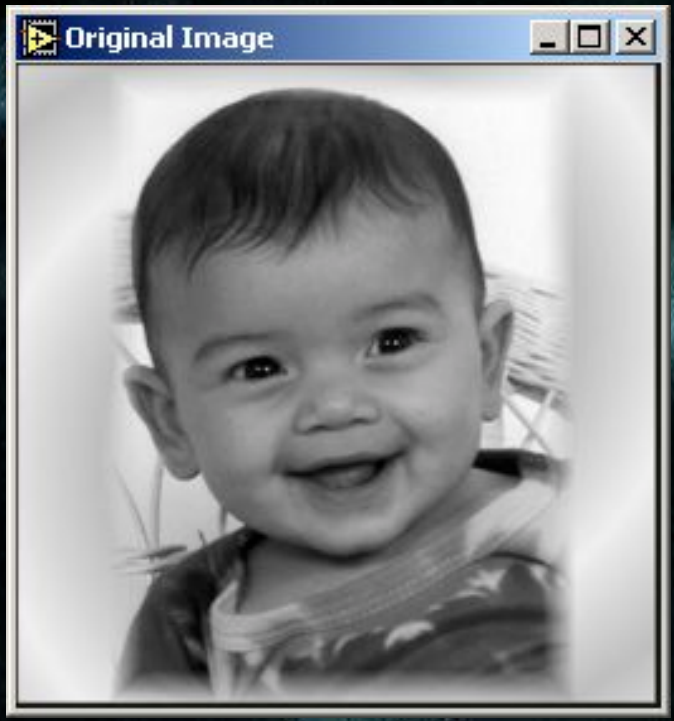
An Unusual Visual Aftereffect



Stare at the dot for ten seconds…..

An Unusual Visual Aftereffect





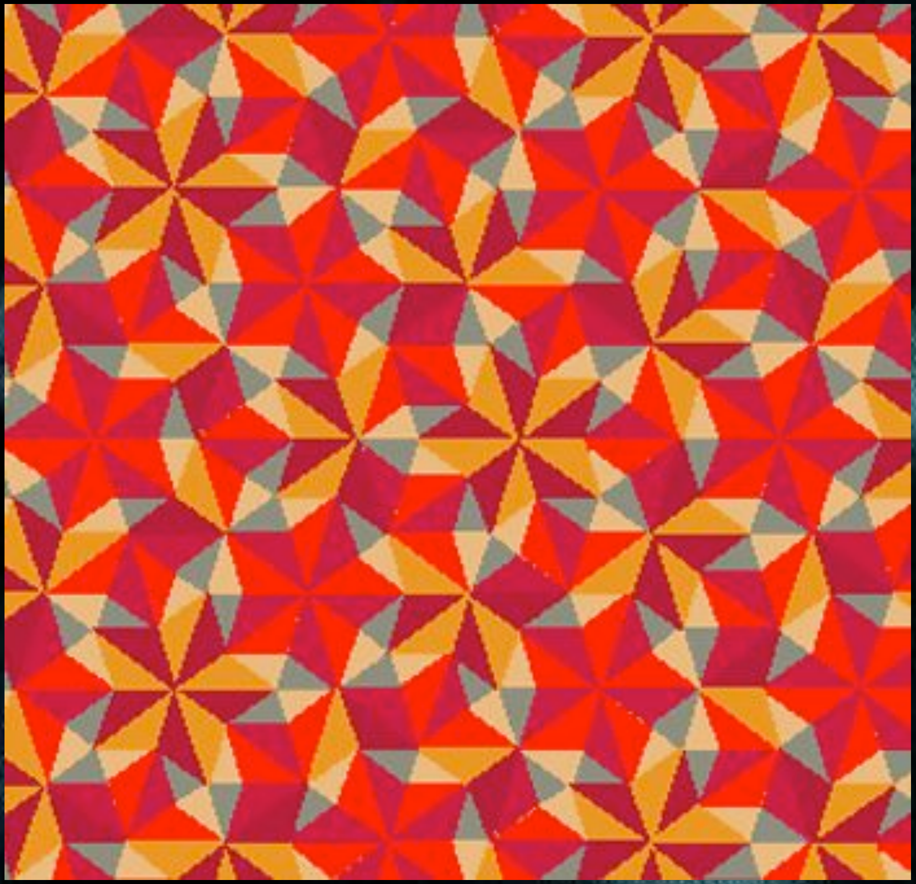
Which image is blurred?

Which Face Is Angry?

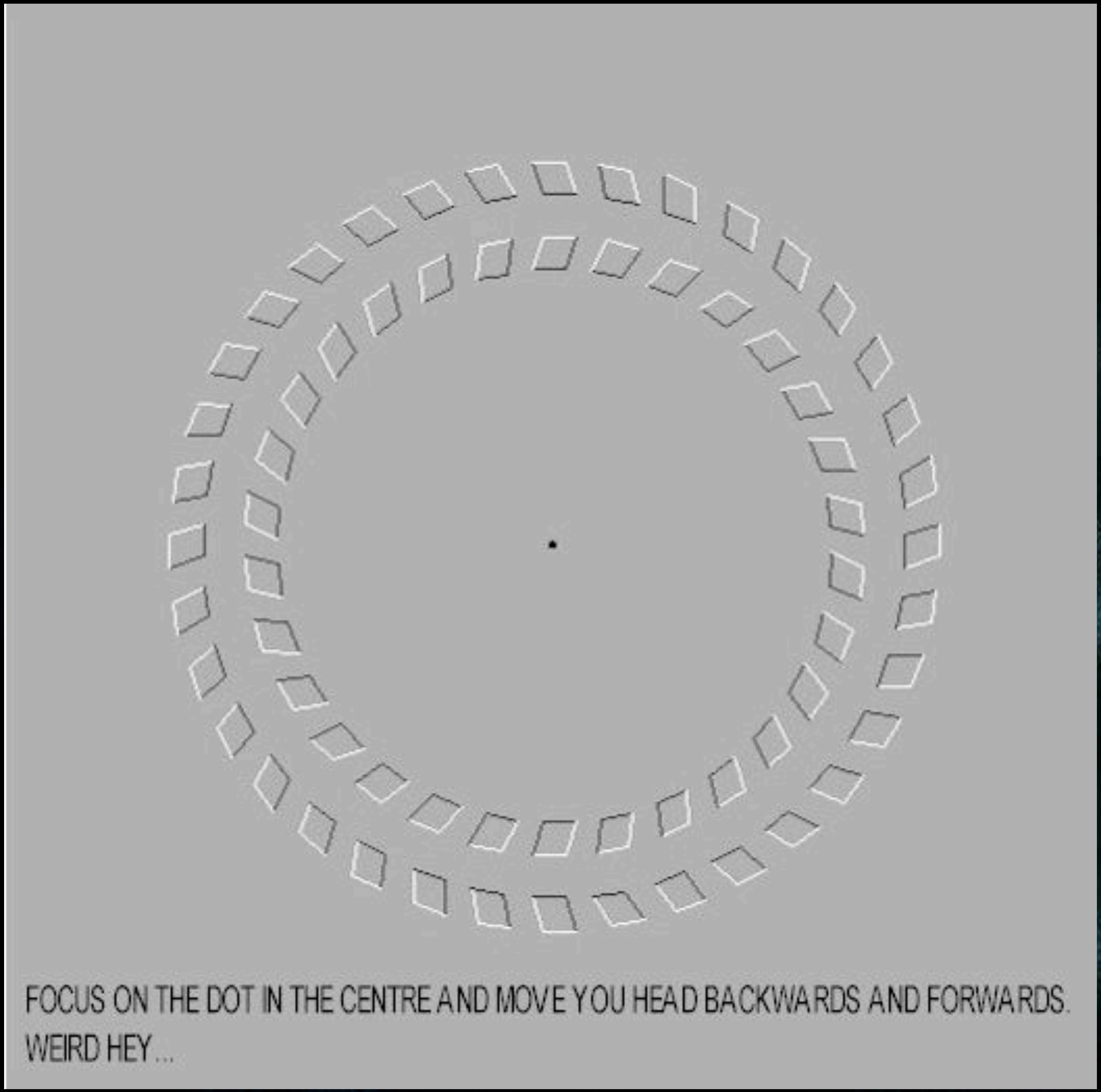
(try blurring them)

Rotating Spiral

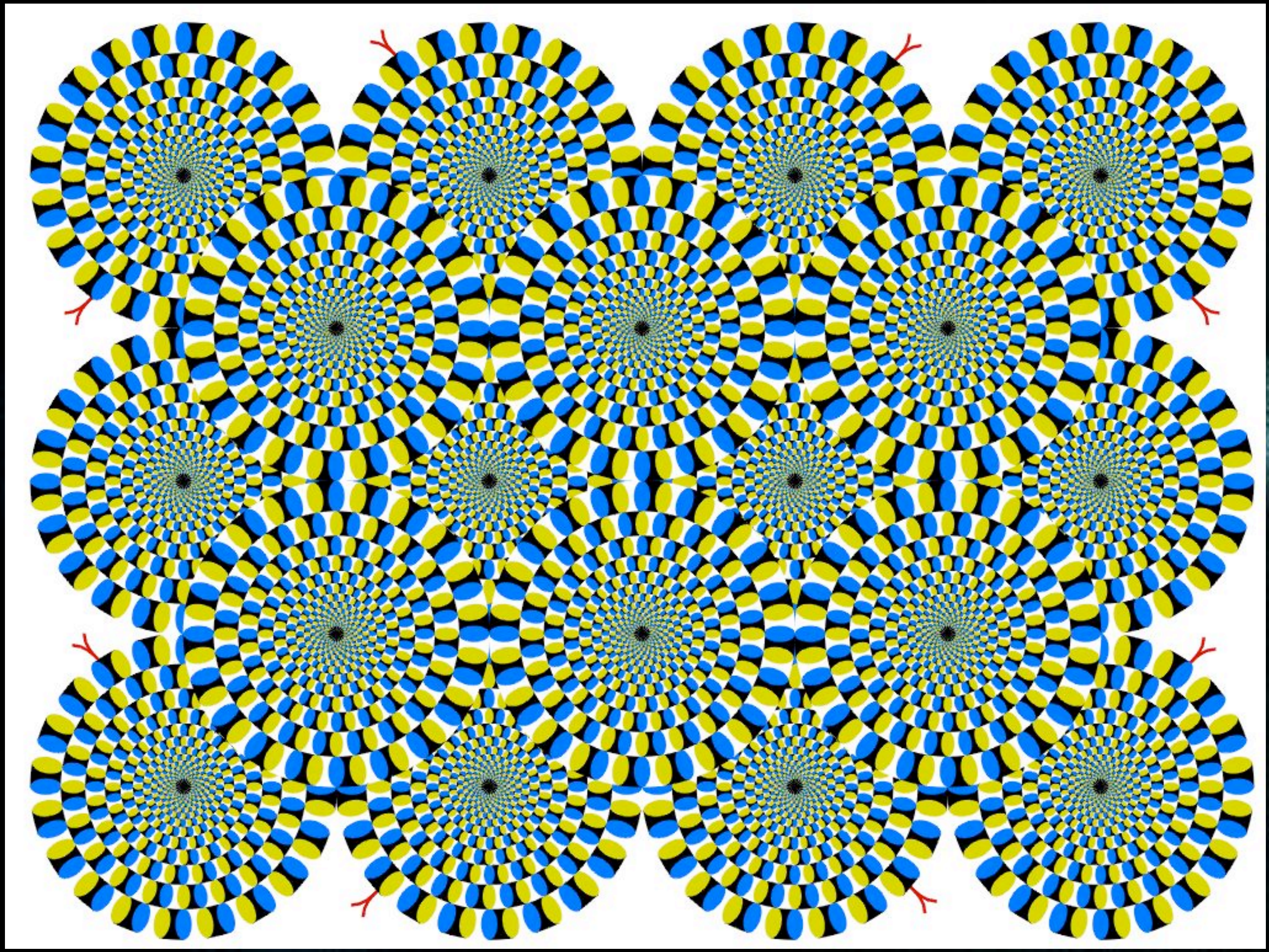
Watch this!



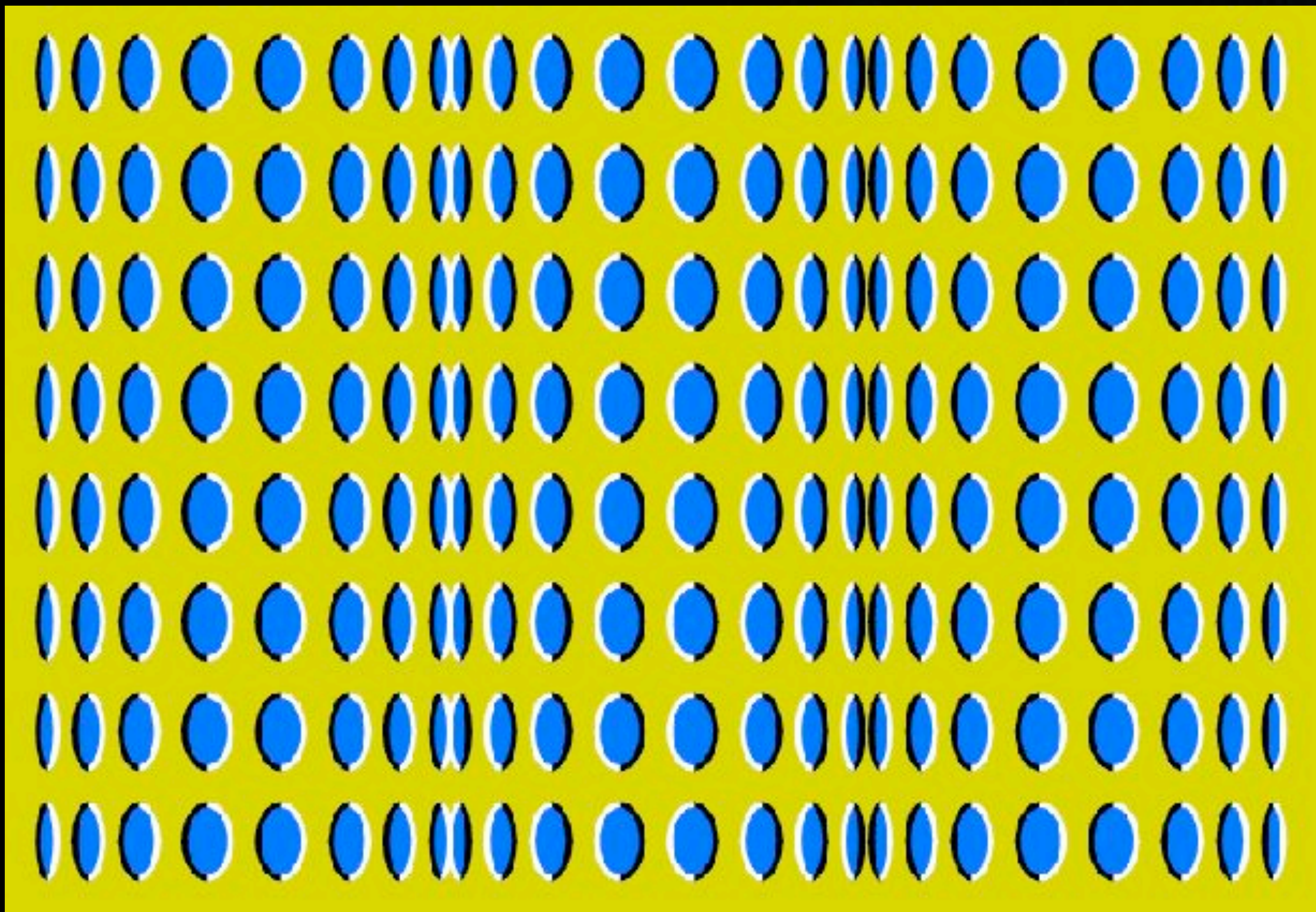
… then stare at this…

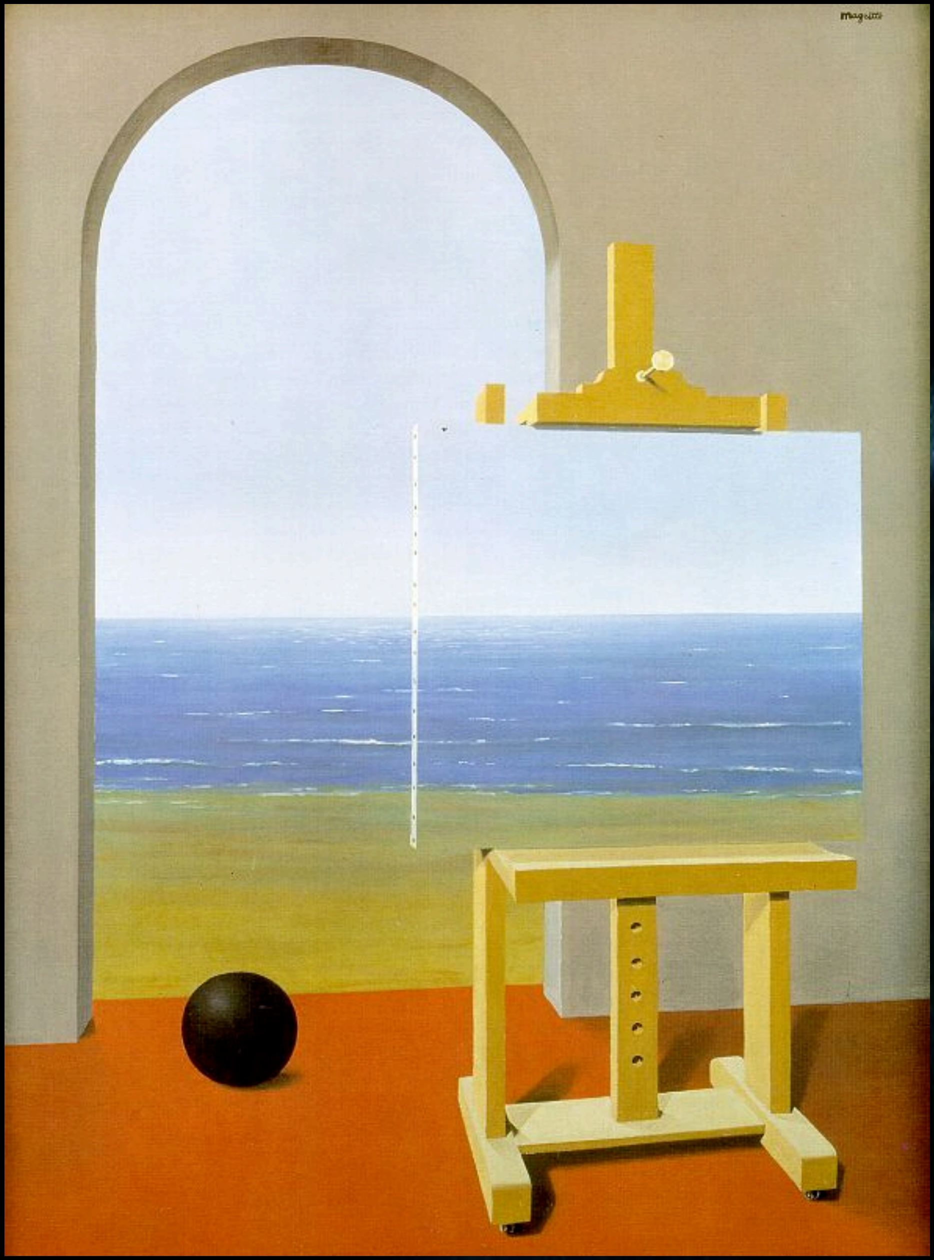

FOCUS ON THE DOT IN THE CENTRE AND MOVE YOU HEAD BACKWARDS AND FORWARDS.
WEIRD HEY...

You Thought That Was Bad…



…. And How About This





“The image is not the object”

Rene Magritte (1898-1967)

Vision is a RELATION:
R(object, subject, ambient)
R(_?_, subject, ambient)
R(object,_?_, ambient)

Wednesday

- Video Art
- Video as Structured Light
  - Installation
  - Performance
- Max / Jitter