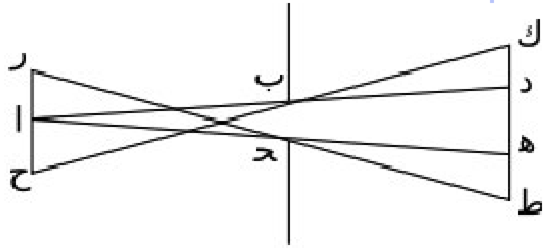
Computer Vision

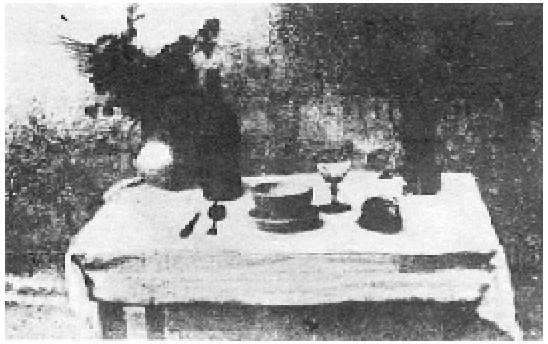
Introduction

Historical context

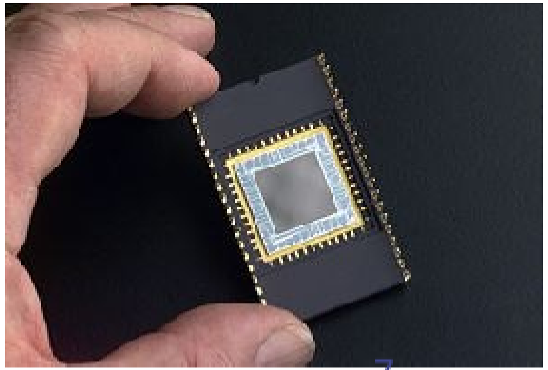
- **Pinhole model:** Mozi (470-390 BCE), Aristotle (384-322 BCE)
- **Principles of optics (including lenses):** Alhacen (965-1039 CE)
- **Camera obscura:** Leonardo da Vinci (1452-1519), Johann Zahn (1631-1707)
- **First photo:** Joseph Nicephore Niepce (1822)
- **Daguerréotypes** (1839)
- **Photographic film** (Eastman, 1889)
- **Cinema** (Lumière Brothers, 1895)
- **Color Photography** (Lumière Brothers, 1908)
- **Television** (Baird, Farnsworth, Zworykin, 1920s)
- **First consumer camera with CCD:** Sony Mavica (1981)
- **First fully digital camera:** Kodak DCS100 (1990)



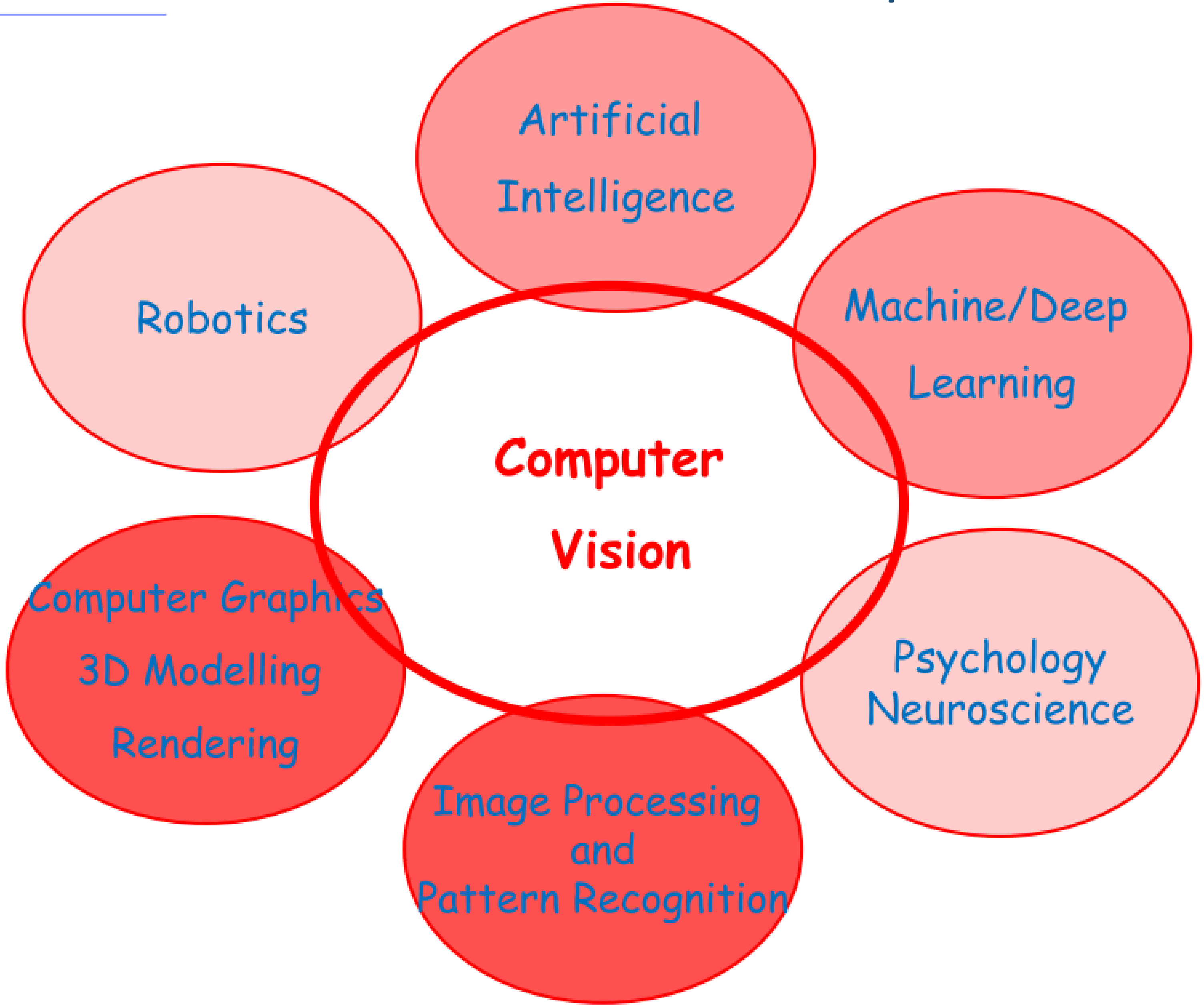
Alhacen's notes



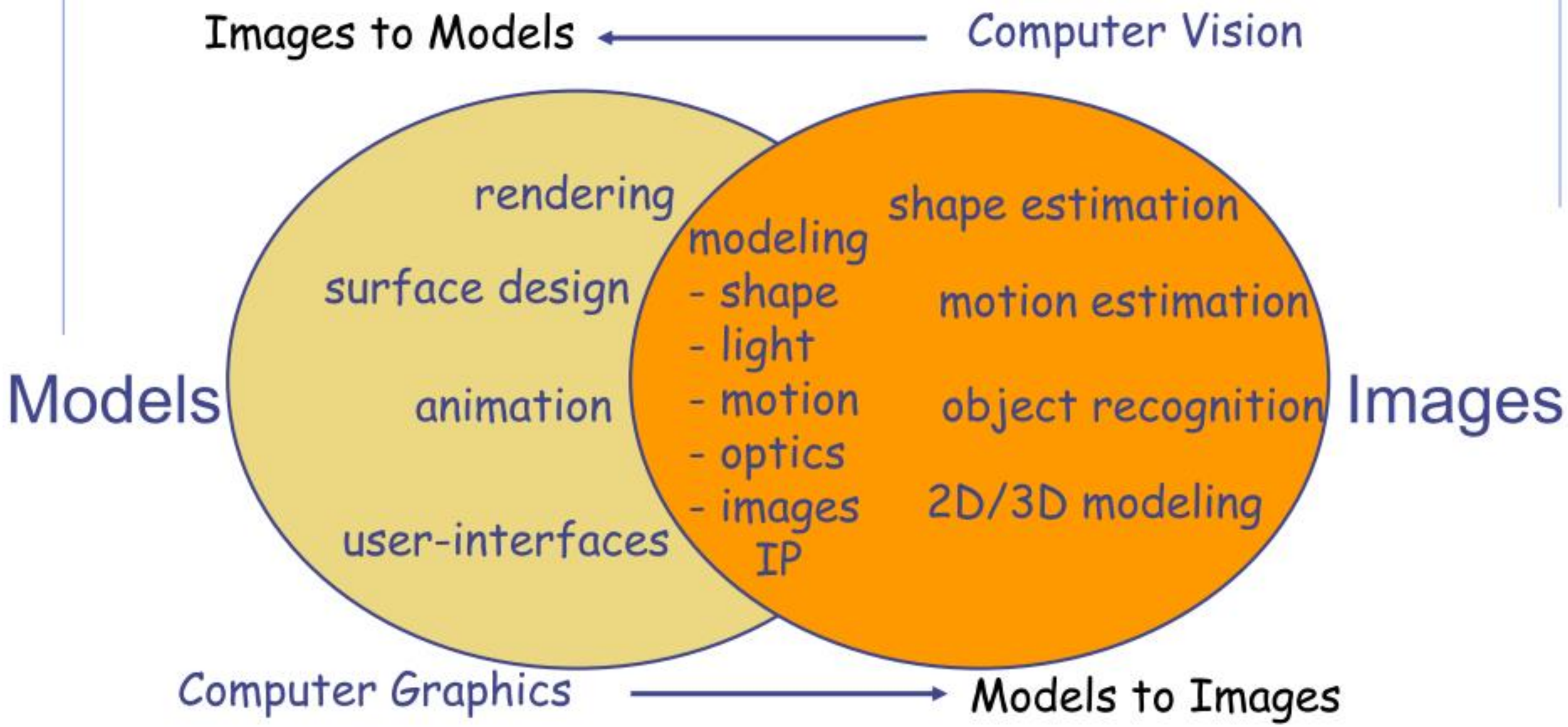
Niepce, "La Table Servie," 1822



Connections to other disciplines

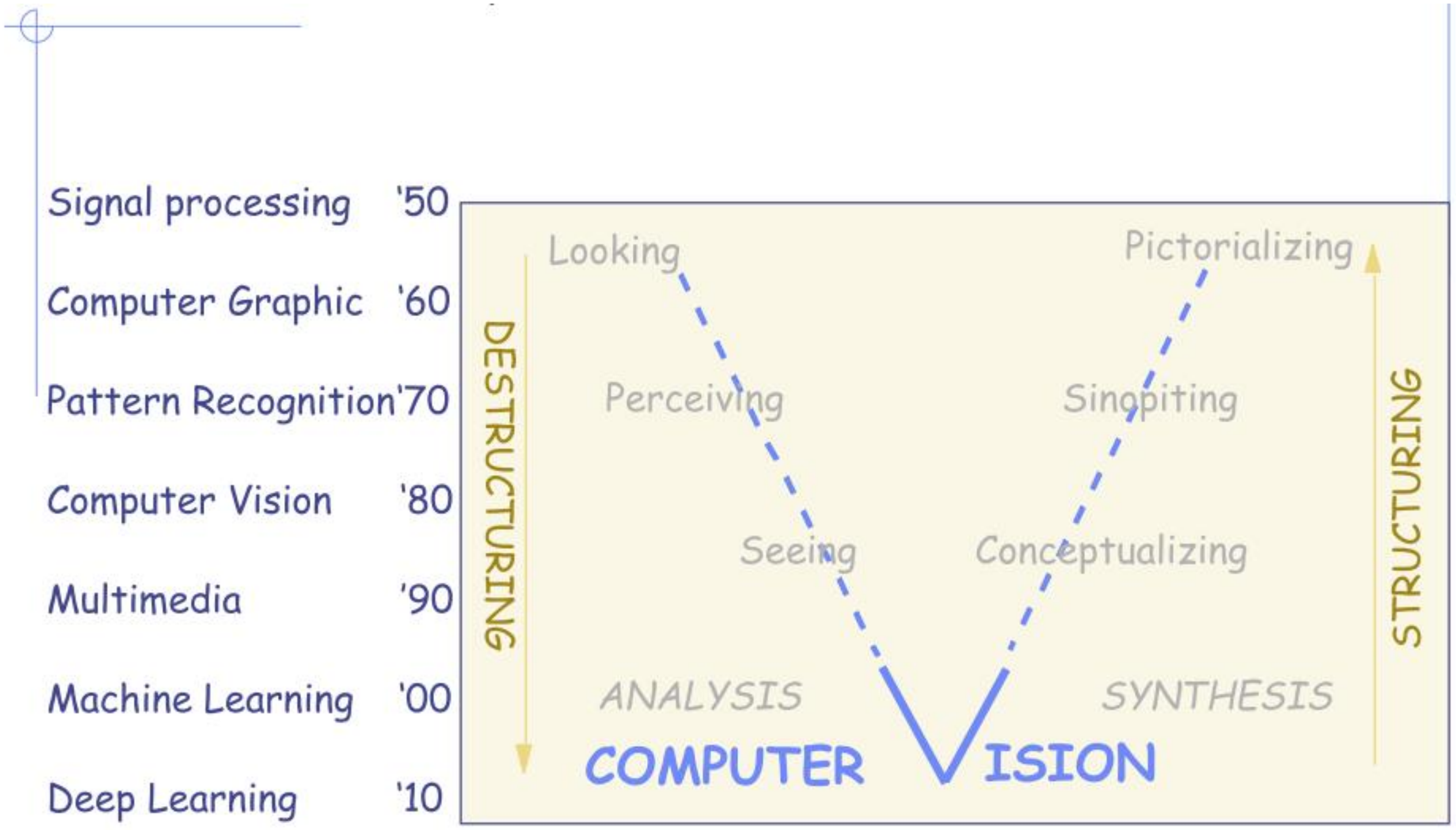


Vision and Graphics

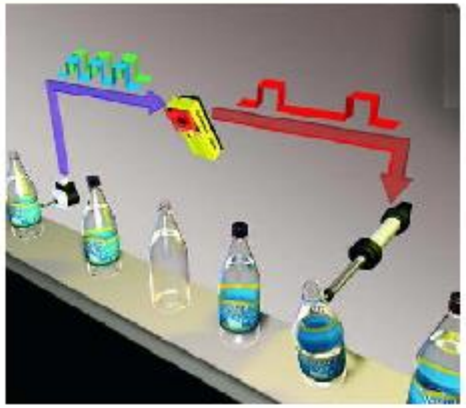


Two inverse problems: analysis and synthesis

Dual aspect of vision: analysis and synthesis

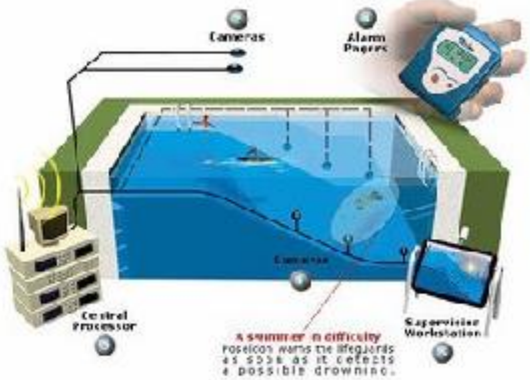


Applications of computer vision



Factory inspection

Reading license plates, checks, ZIP codes



Monitoring for safety (Poseidon)

Surveillance

Autonomous driving, robot navigation

Driver assistance
(collision warning, lane departure

Applications of computer vision

Assistive technologies

Entertainment

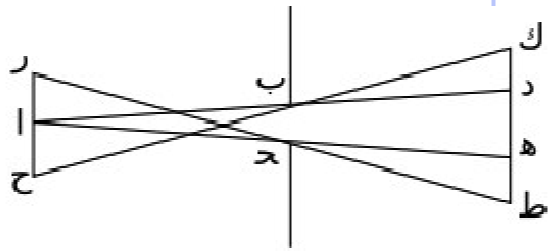
Movie special effects

Digital cameras (face detection for setting focus, exposure)

Visual search (MSR Lincoln)

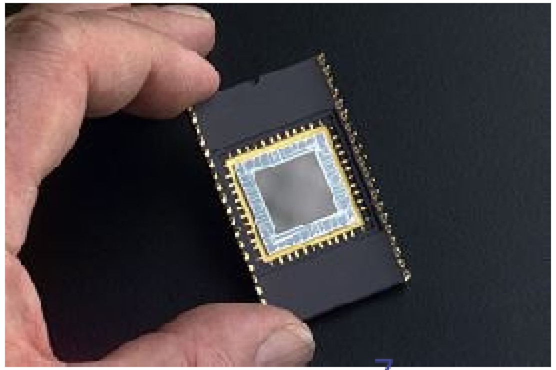
Historical context

- **Pinhole model:** Mozi (470-390 BCE), Aristotle (384-322 BCE)
- **Principles of optics (including lenses):** Alhacen (965-1039 CE)
- **Camera obscura:** Leonardo da Vinci (1452-1519), Johann Zahn (1631-1707)
- **First photo:** Joseph Nicephore Niepce (1822)
- **Daguerréotypes** (1839)
- **Photographic film** (Eastman, 1889)
- **Cinema** (Lumière Brothers, 1895)
- **Color Photography** (Lumière Brothers, 1908)
- **Television** (Baird, Farnsworth, Zworykin, 1920s)
- **First consumer camera with CCD:** Sony Mavica (1981)
- **First fully digital camera:** Kodak DCS100 (1990)

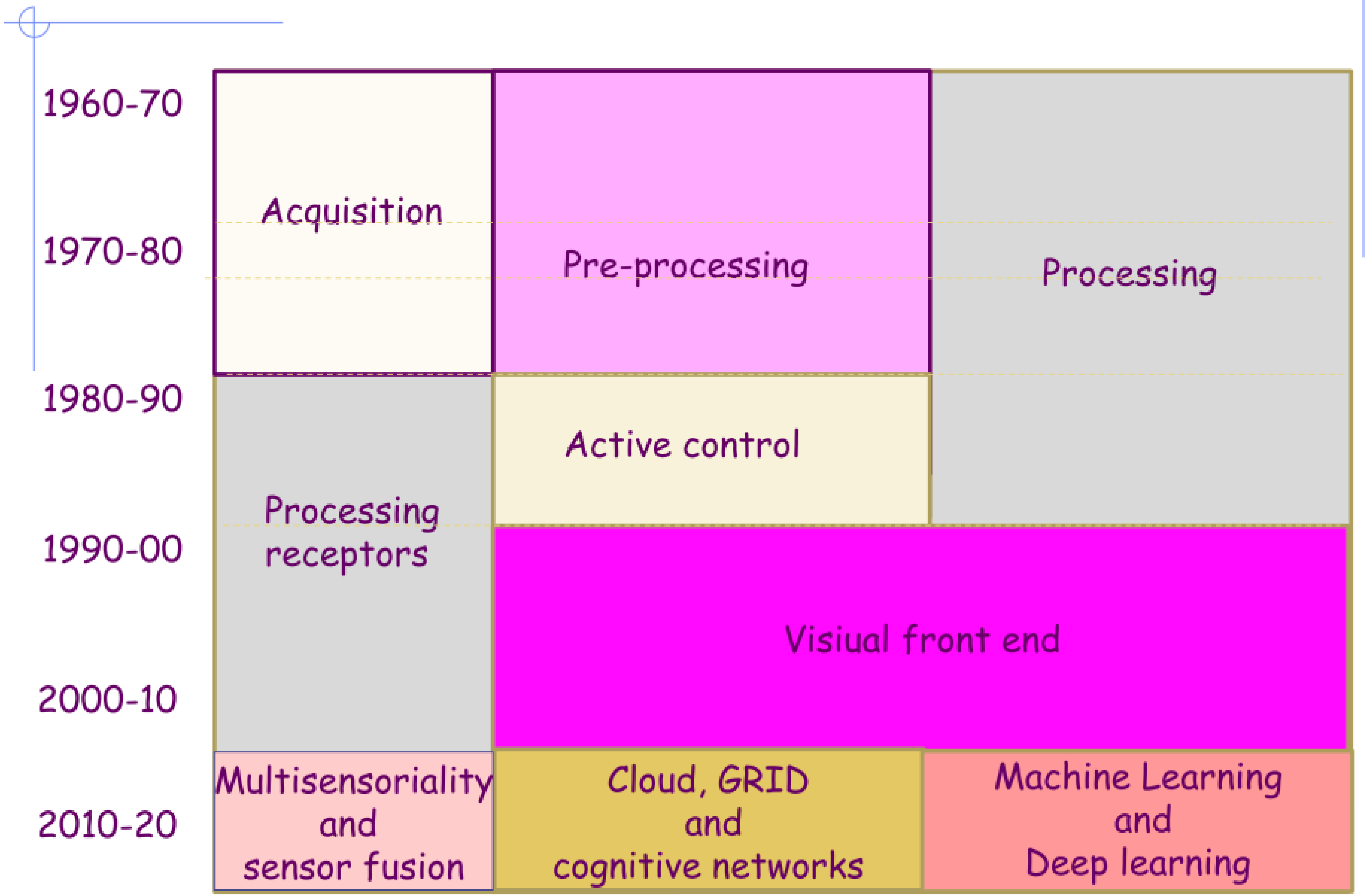


Alhacen's notes

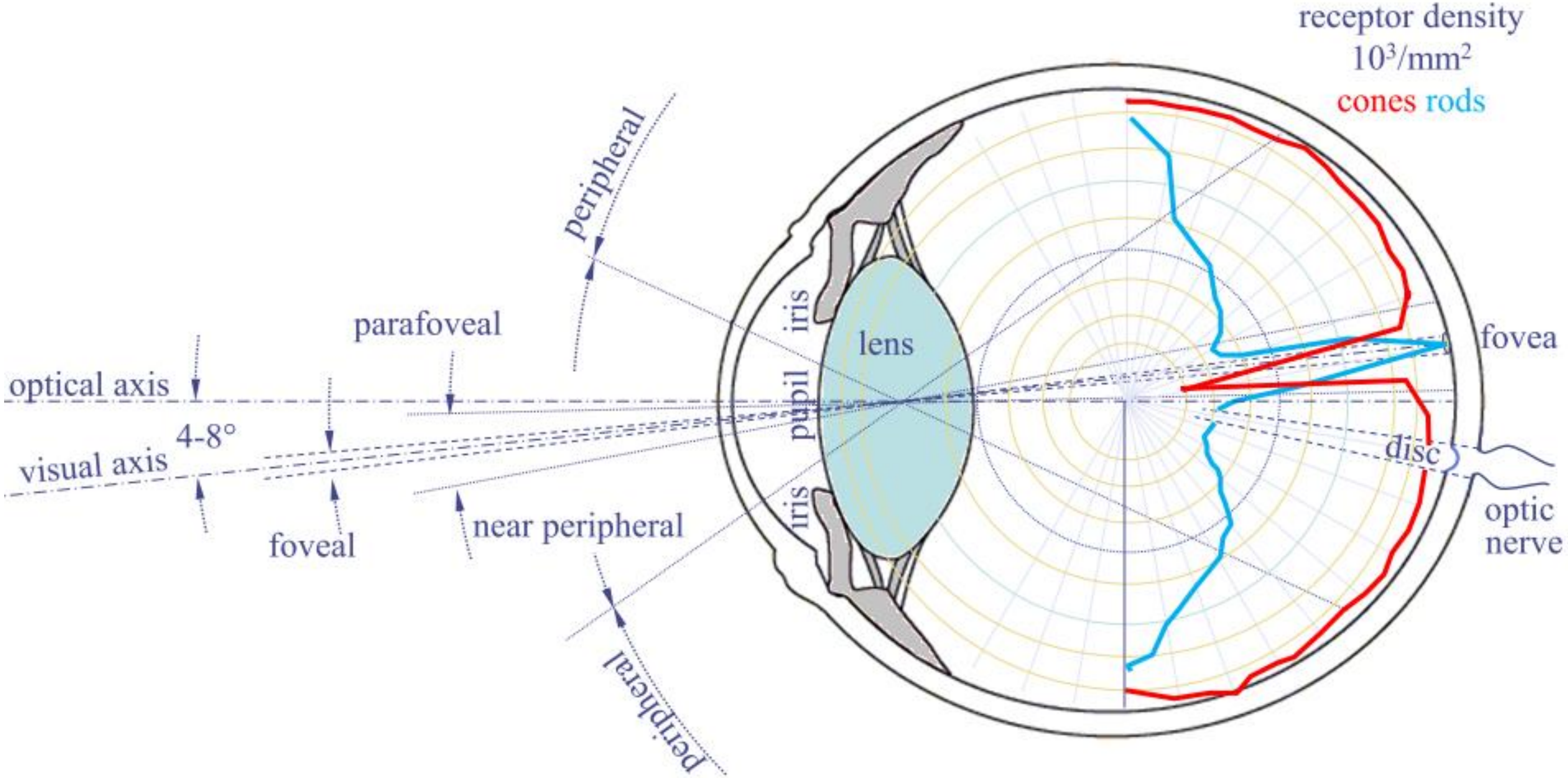

Niepce, "La Table Servie," 1822



Machine vision phylogenesis

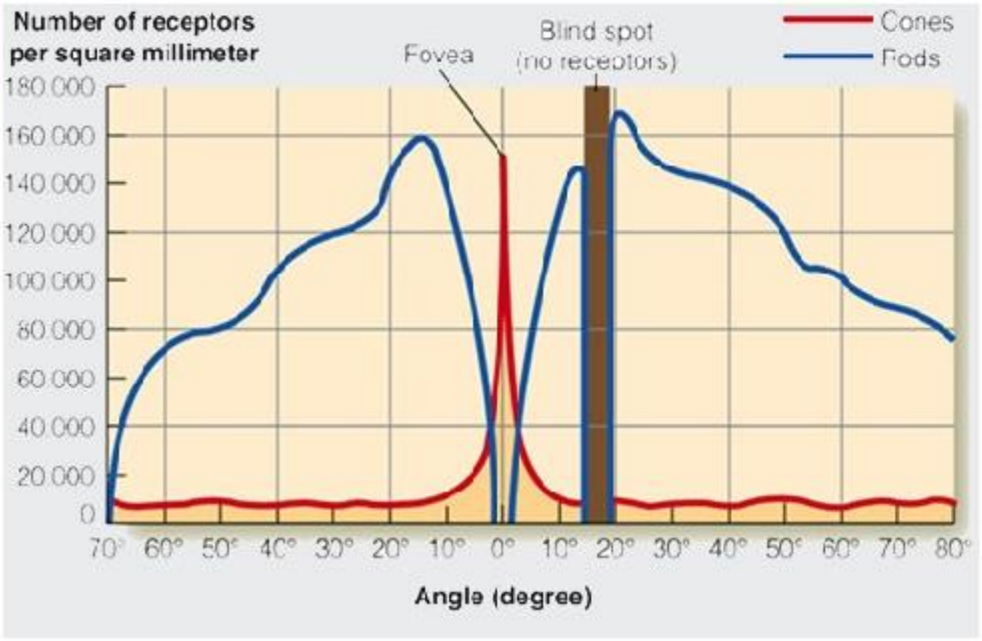


The human eye anatomy and visual field



The retina structure

Density distribution of rod and cone receptors

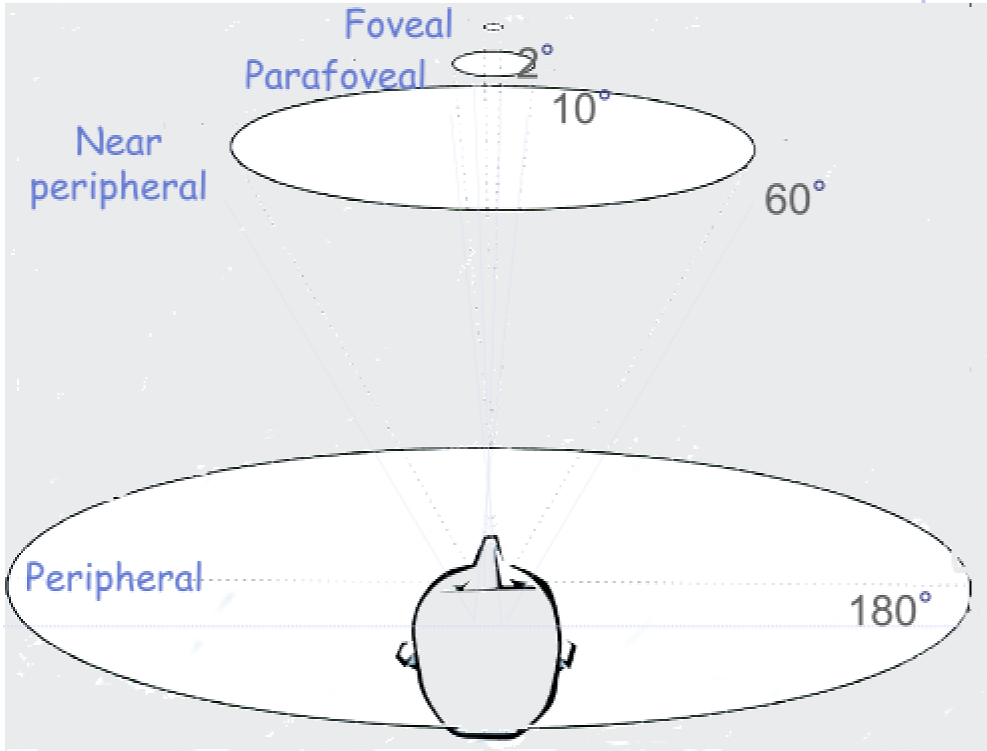


Loss of resolution with eccentricity in retina

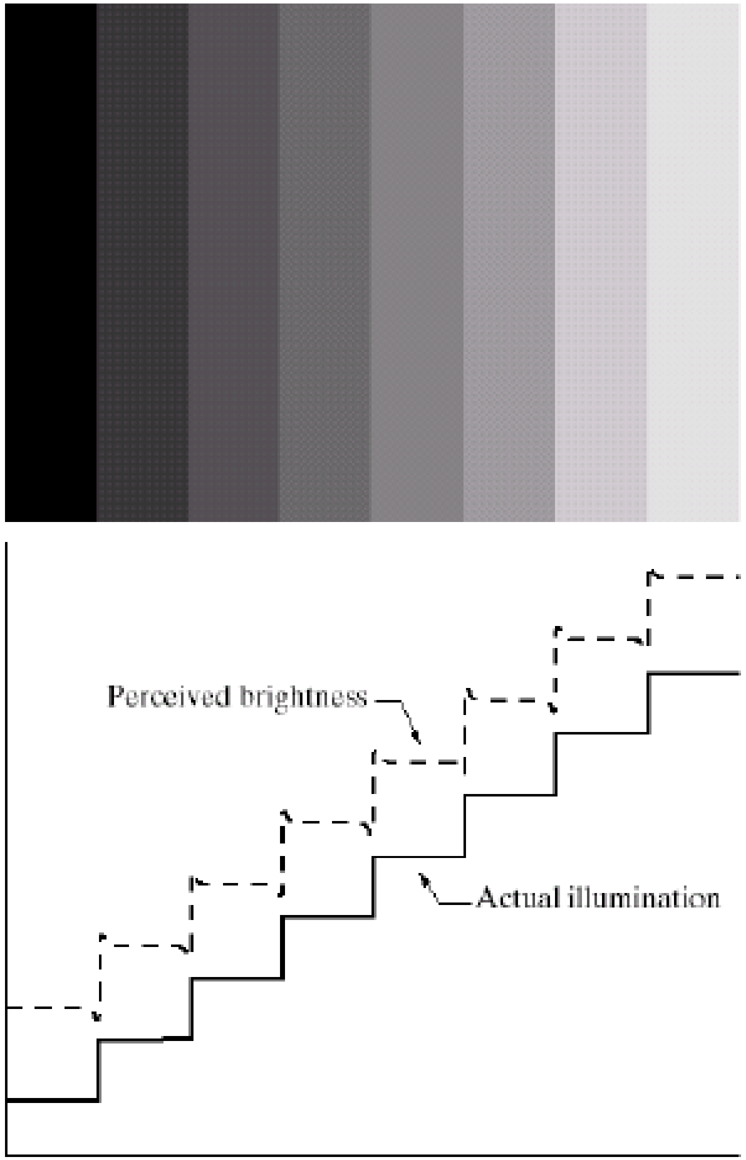
Derived from Duchowski 2002, p. 34

The functional field of view

- *Fovea*: central area of retina – provides the sharpest vision; when we are looking at something, we are directing our eyes so image is projected onto fovea
- *Parafoveal*: region surrounding fovea corresponding to retinal area from 2° to 10° off-centre; previews foveal information
- *Peripheral*: region of retina outside the 10° area – increased sensitivity to motion detection; reacts to flashing objects and sudden movements
- The regions are asymmetric, e.g. in reading the so-called perceptual span (size of the effective vision), is 3-4 letter spaces to the left of fixation and 14-15 letter spaces to the right; 10° of visual angle is roughly equivalent to 3-4 letter spaces



Human perception



a
b

FIGURE 2.7 (a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

Human perception

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Optical illusions

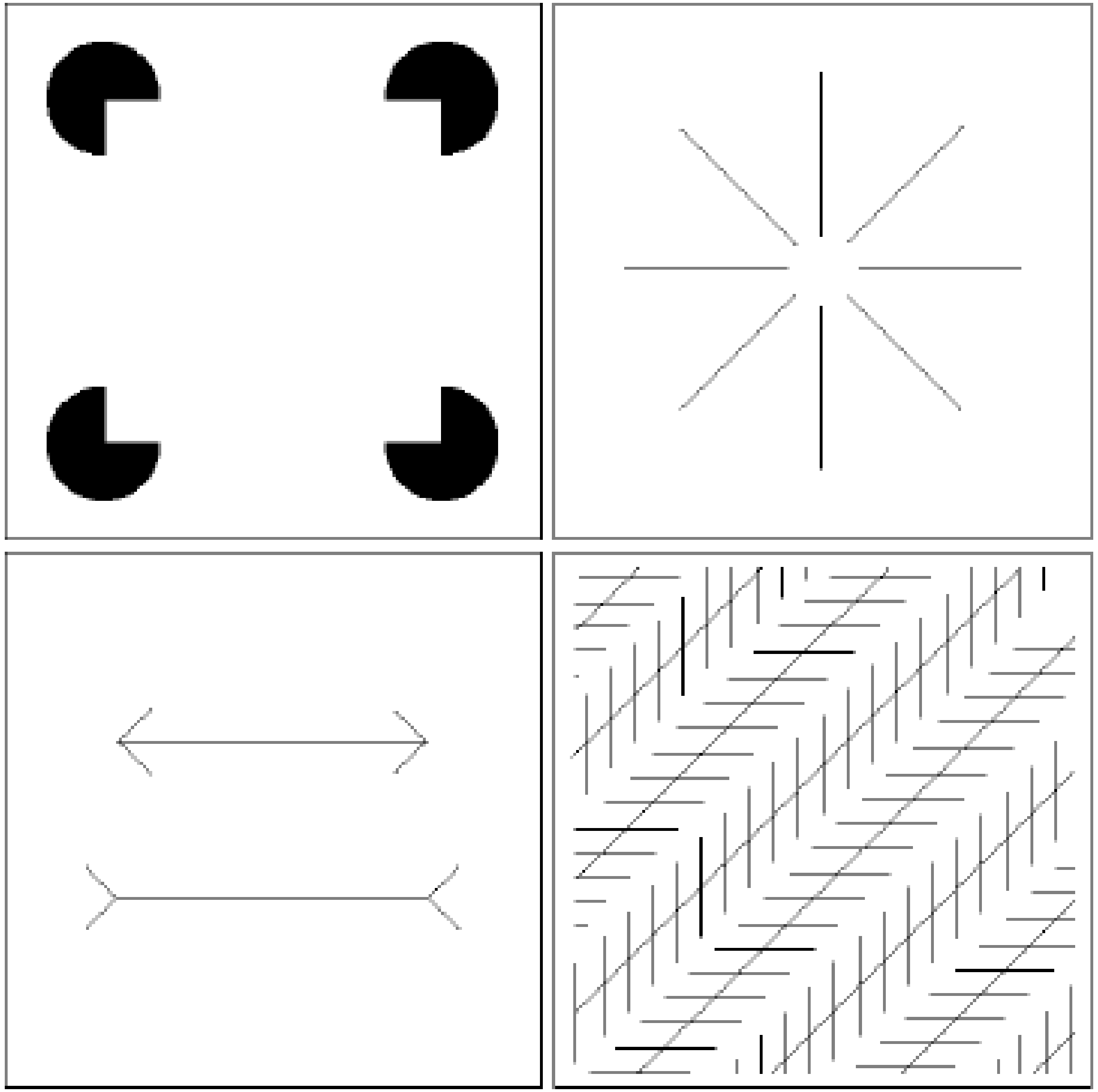


FIGURE 2.9 Some well-known optical illusions.

Electromagnetic spectrum

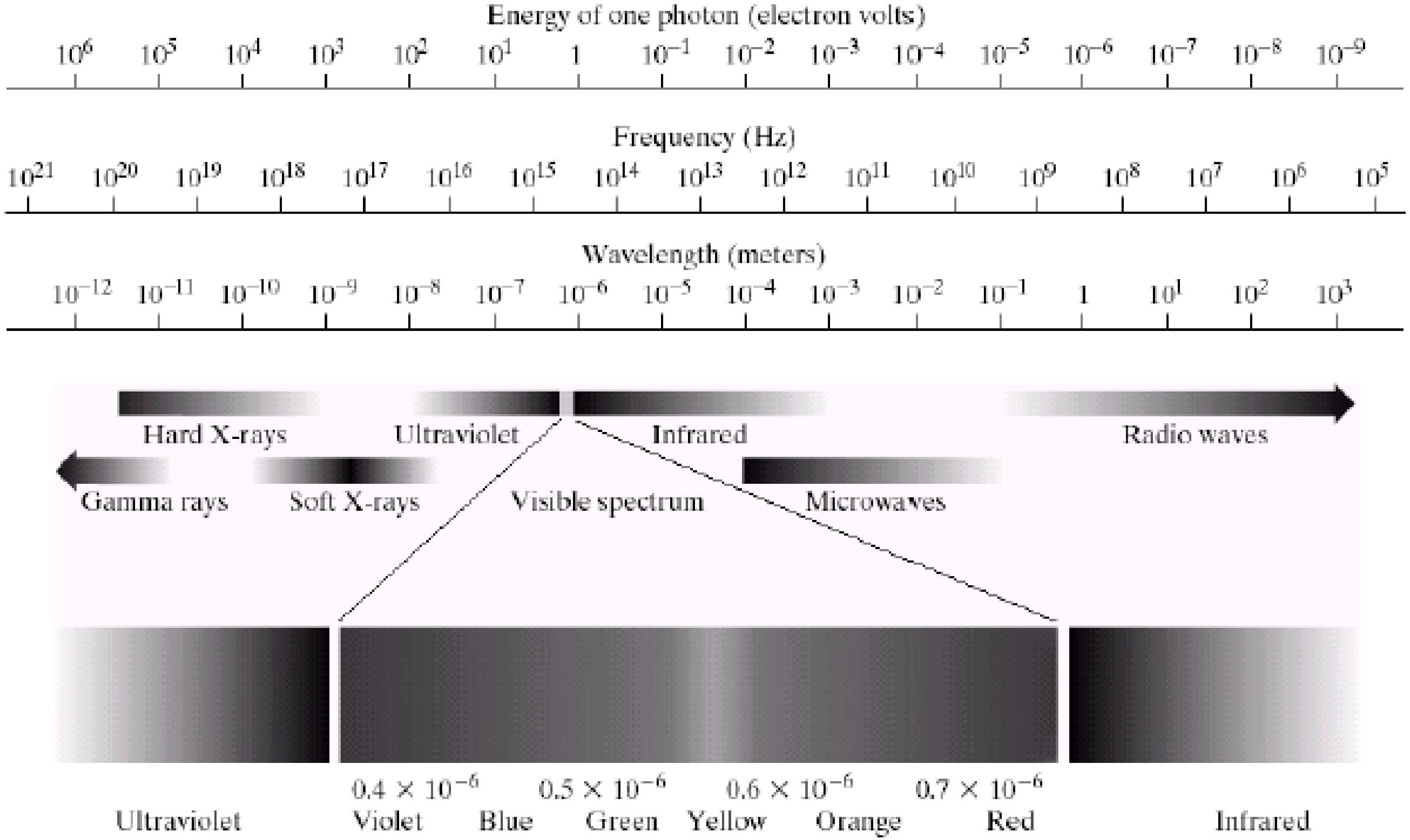
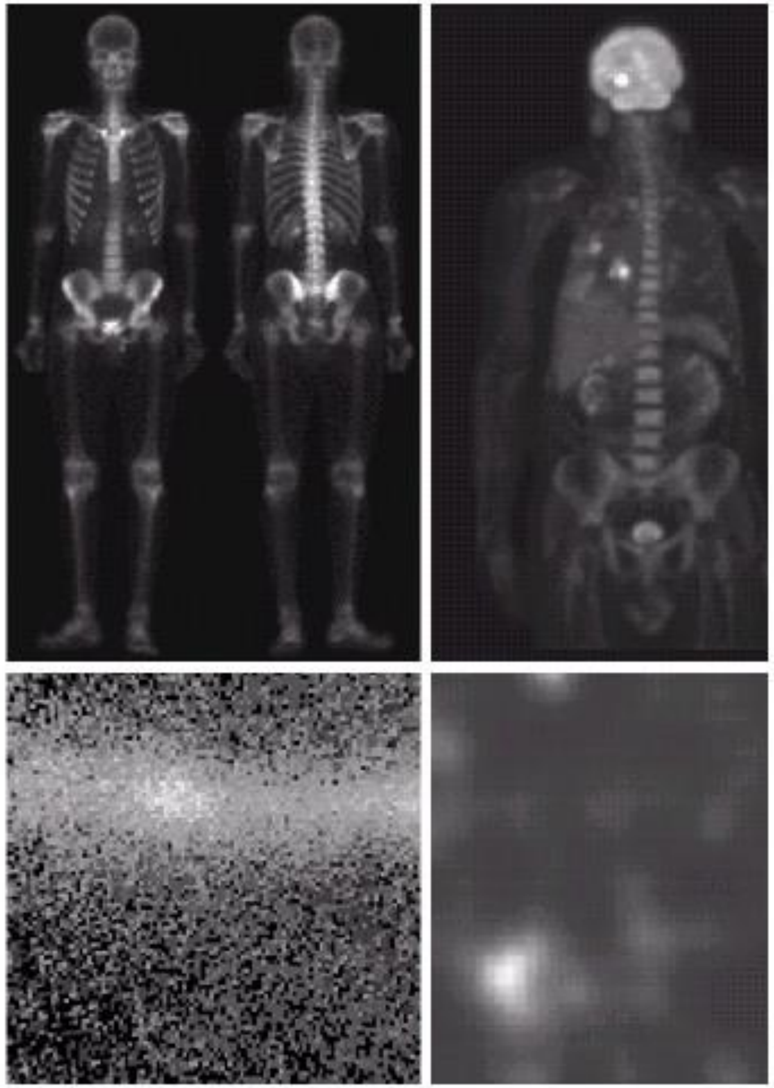


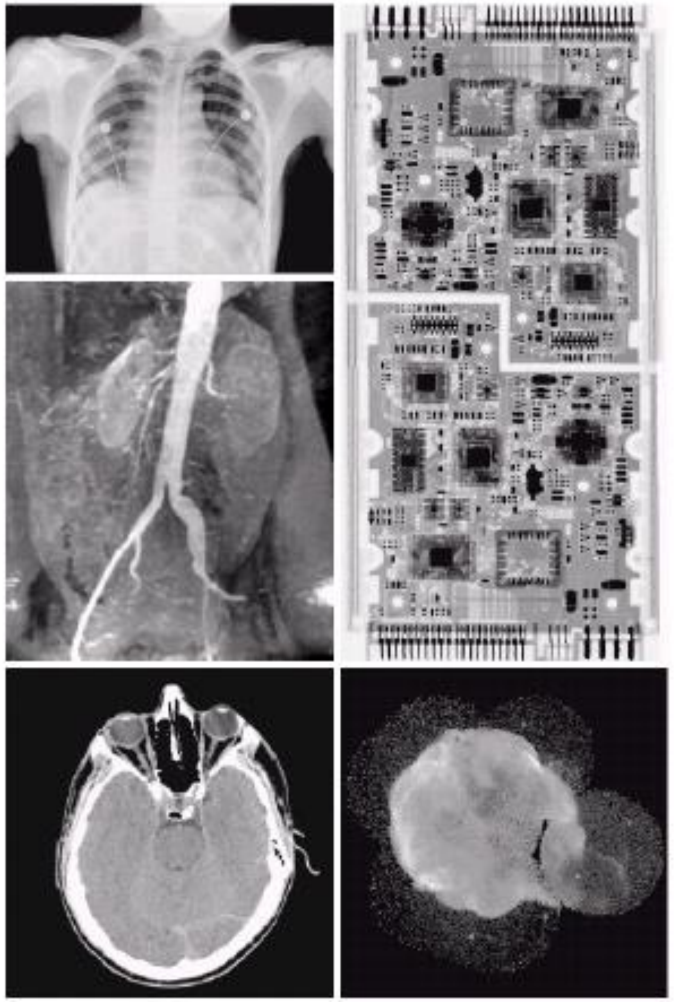
FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Electromagnetic spectrum

a b
c d

FIGURE 1.6 Examples of gamma-ray imaging. (a) Bone scan. (b) PET image. (c) Cygnus Loop. (d) Gamma radiation (bright spot) from a reactor valve. (Images courtesy of (a) G.E. Medical Systems, (b) Dr. Michael E. Casey, CTI PET Systems, (c) NASA, (d) Professors Zhong He and David K. Wehe, University of Michigan.)

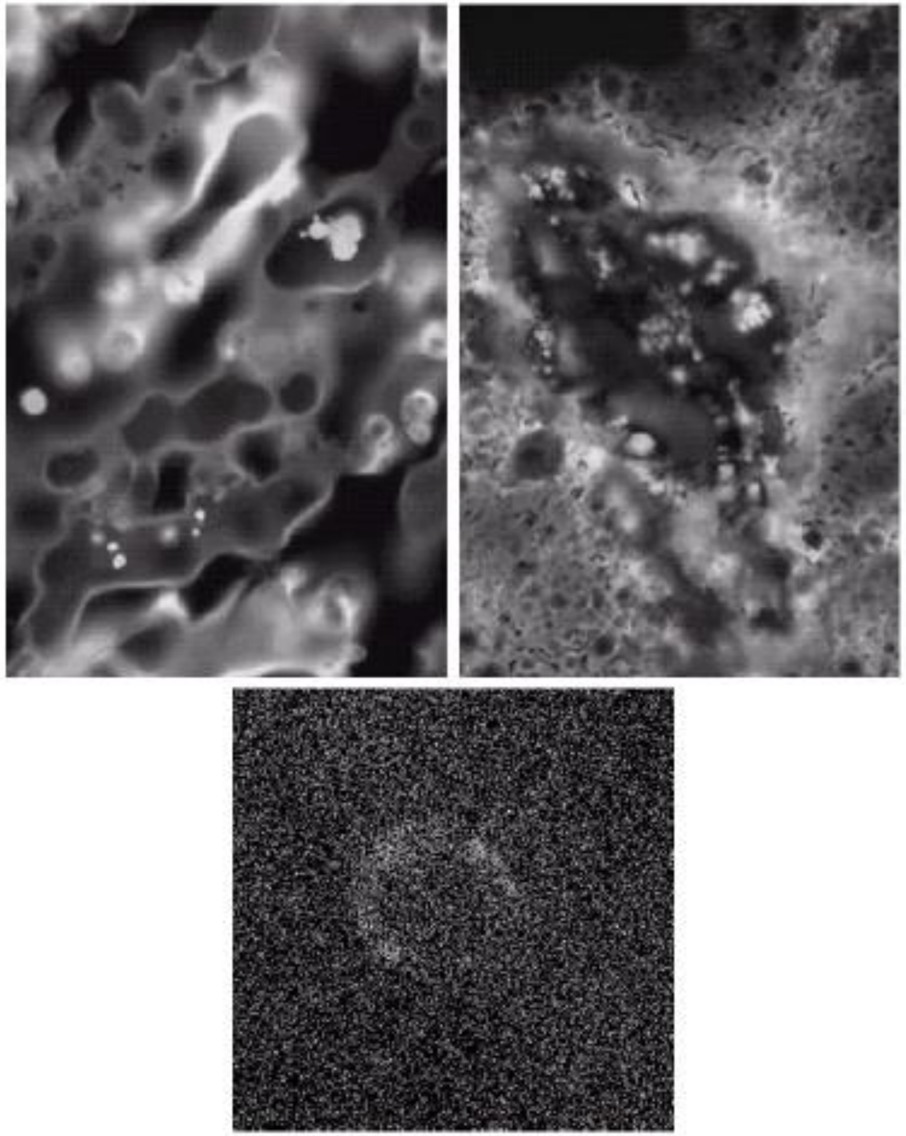




a d
b
c e

FIGURE 1.7 Examples of X-ray imaging. (a) Chest X-ray. (b) Aortic angiogram. (c) Head CT. (d) Circuit boards. (e) Cygnus Loop. (Images courtesy of (a) and (c) Dr. David R. Pickens, Dept. of Radiology & Radiological Sciences, Vanderbilt University Medical Center, (b) Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School, (d) Mr. Joseph E. Pascente, Lixi, Inc., and (e) NASA.)

Electromagnetic spectrum



a b
c

FIGURE 1.8 Examples of ultraviolet imaging. (a) Normal corn. (b) Smut corn. (c) Cygnus Loop. (Images courtesy of (a) and (b) Dr. Michael W. Davidson, Florida State University, (c) NASA.)

Electromagnetic spectrum

TABLE 1.1 Thematic bands in NASA's LANDSAT satellite.

Band No.	Name	Wavelength (μm)	Characteristics and Uses
1	Visible blue	0.45–0.52	Maximum water penetration
2	Visible green	0.52–0.60	Good for measuring plant vigor
3	Visible red	0.63–0.69	Vegetation discrimination
4	Near infrared	0.76–0.90	Biomass and shoreline mapping
5	Middle infrared	1.55–1.75	Moisture content of soil and vegetation
6	Thermal infrared	10.4–12.5	Soil moisture; thermal mapping
7	Middle infrared	2.08–2.35	Mineral mapping

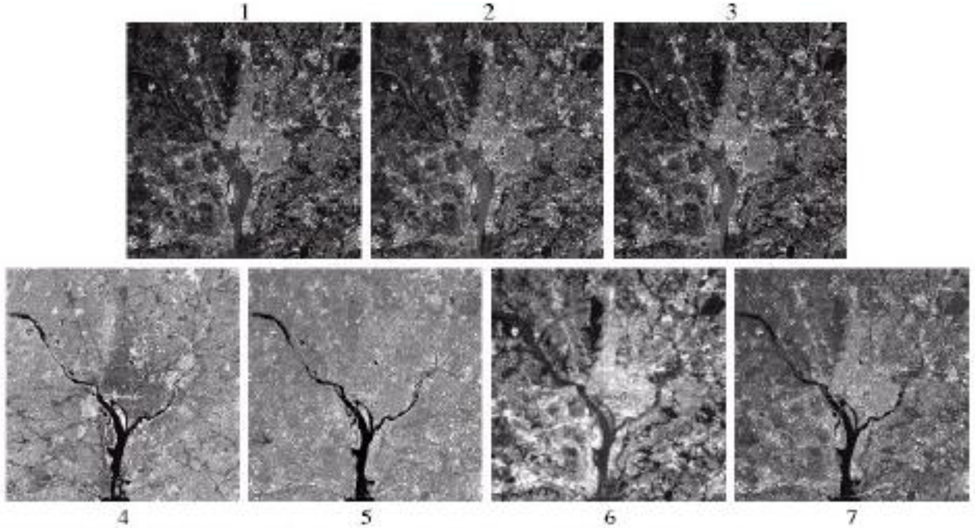


FIGURE 1.10 LANDSAT satellite images of the Washington, D.C. area. The numbers refer to the thematic bands in Table 1.1. (Images courtesy of NASA.)

Electromagnetic spectrum

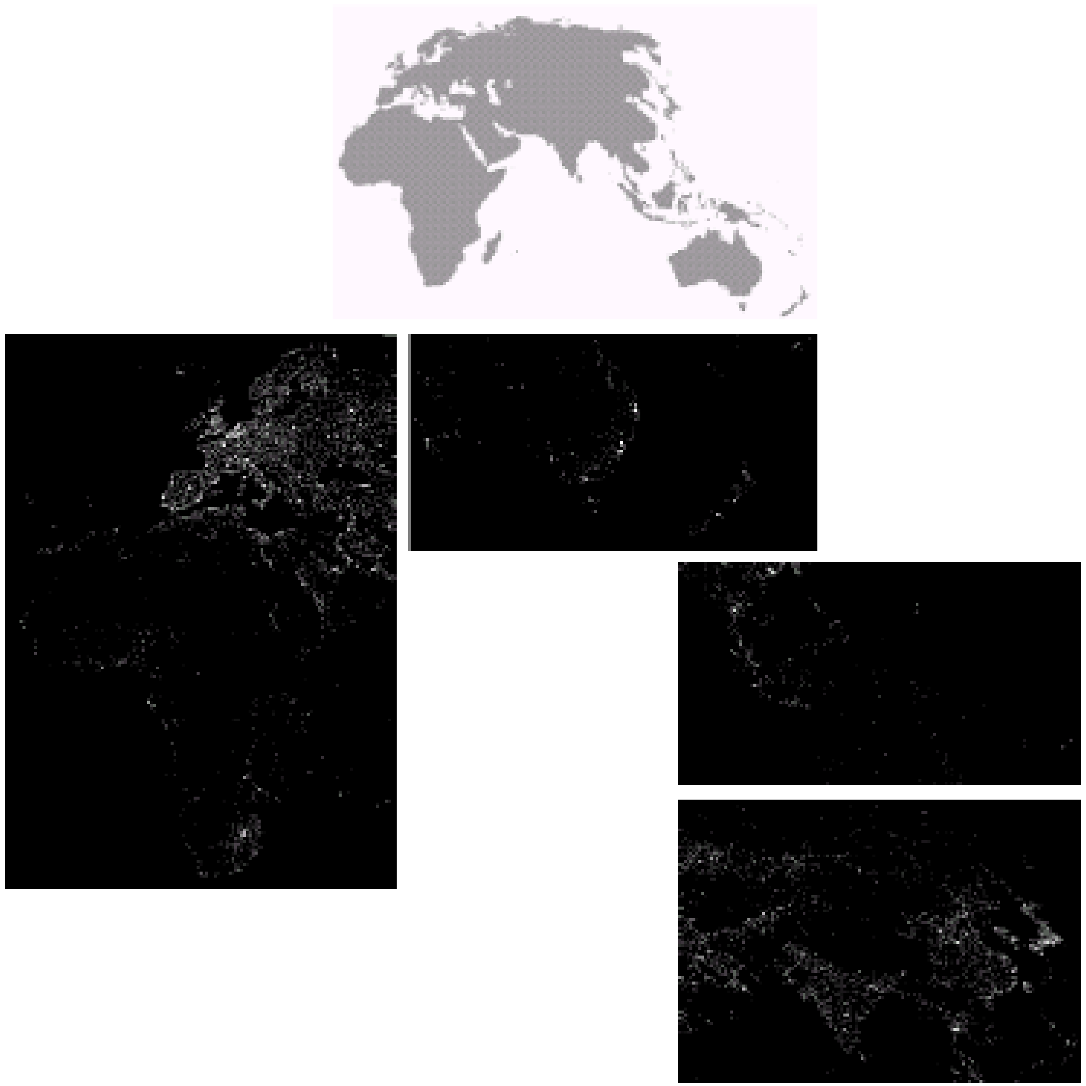


FIGURE 1.13 Infrared satellite images of the remaining populated part of the world. The small gray map is provided for reference. (Courtesy of NOAA.)

Electromagnetic spectrum

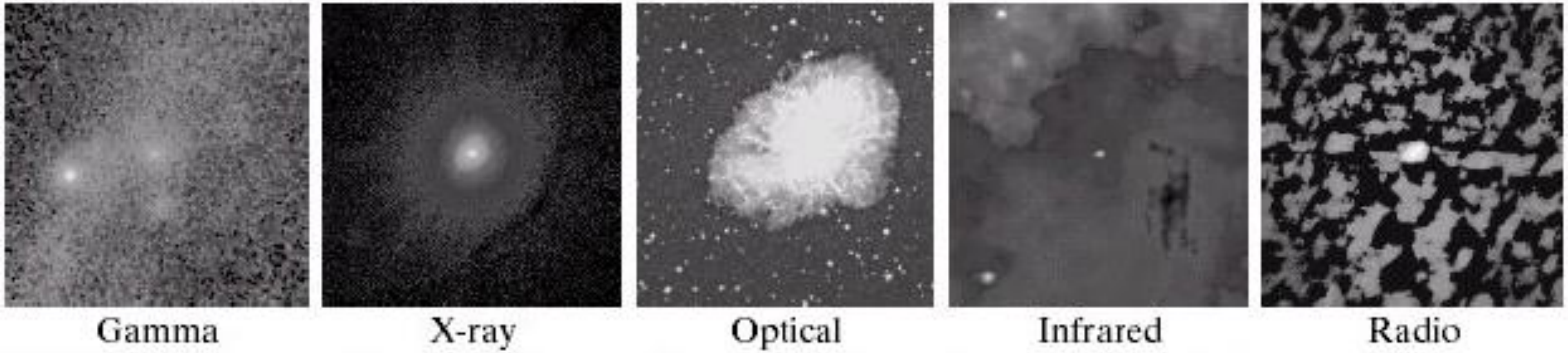


FIGURE 1.18 Images of the Crab Pulsar (in the center of images) covering the electromagnetic spectrum. (Courtesy of NASA.)

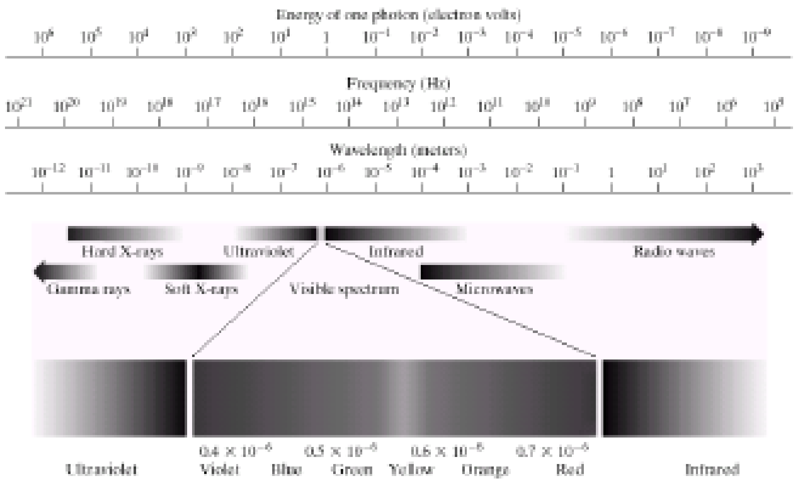


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Ultrasound imaging

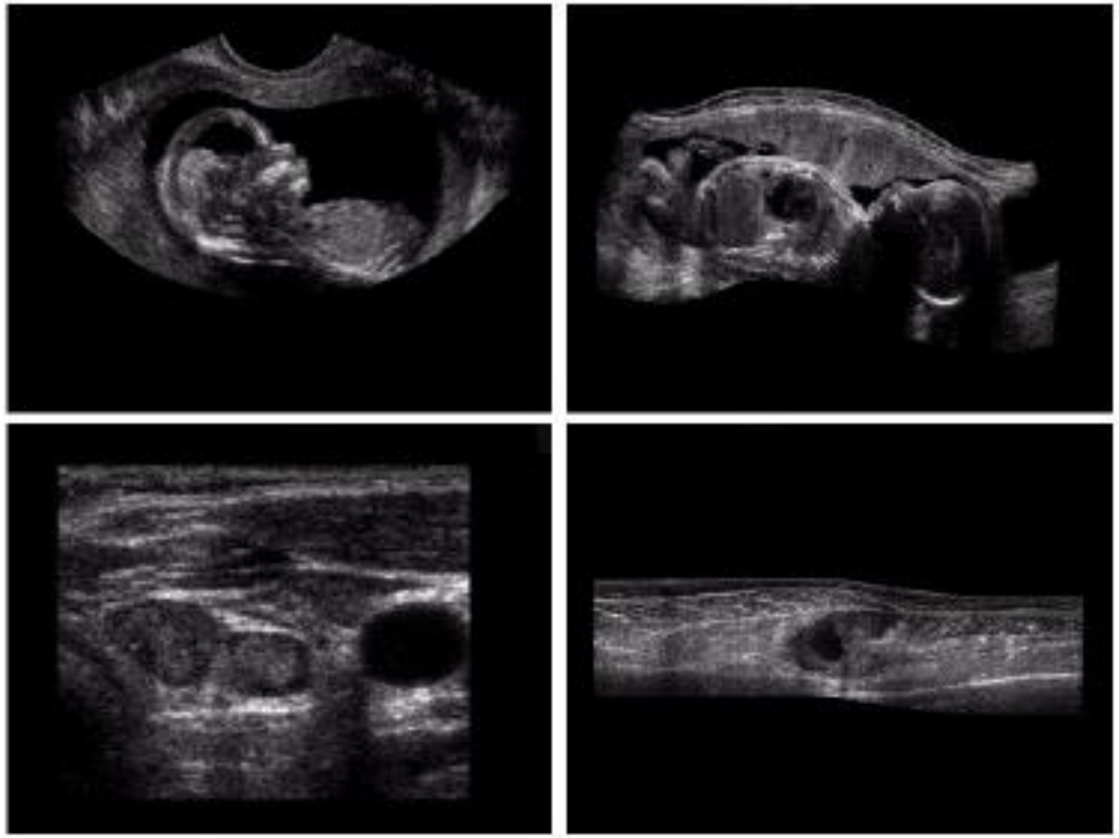
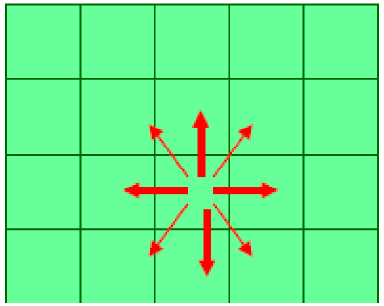
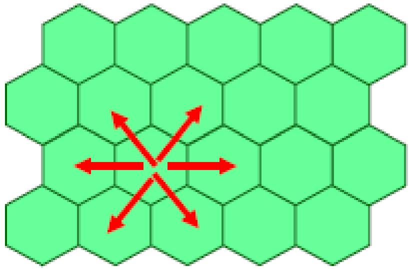


FIGURE 1.20 Examples of ultrasound imaging. (a) Baby. (2) Another view of baby. (c) Thyroids. (d) Muscle layers showing lesion. (Courtesy of Siemens Medical Systems, Inc., Ultrasound Group.)

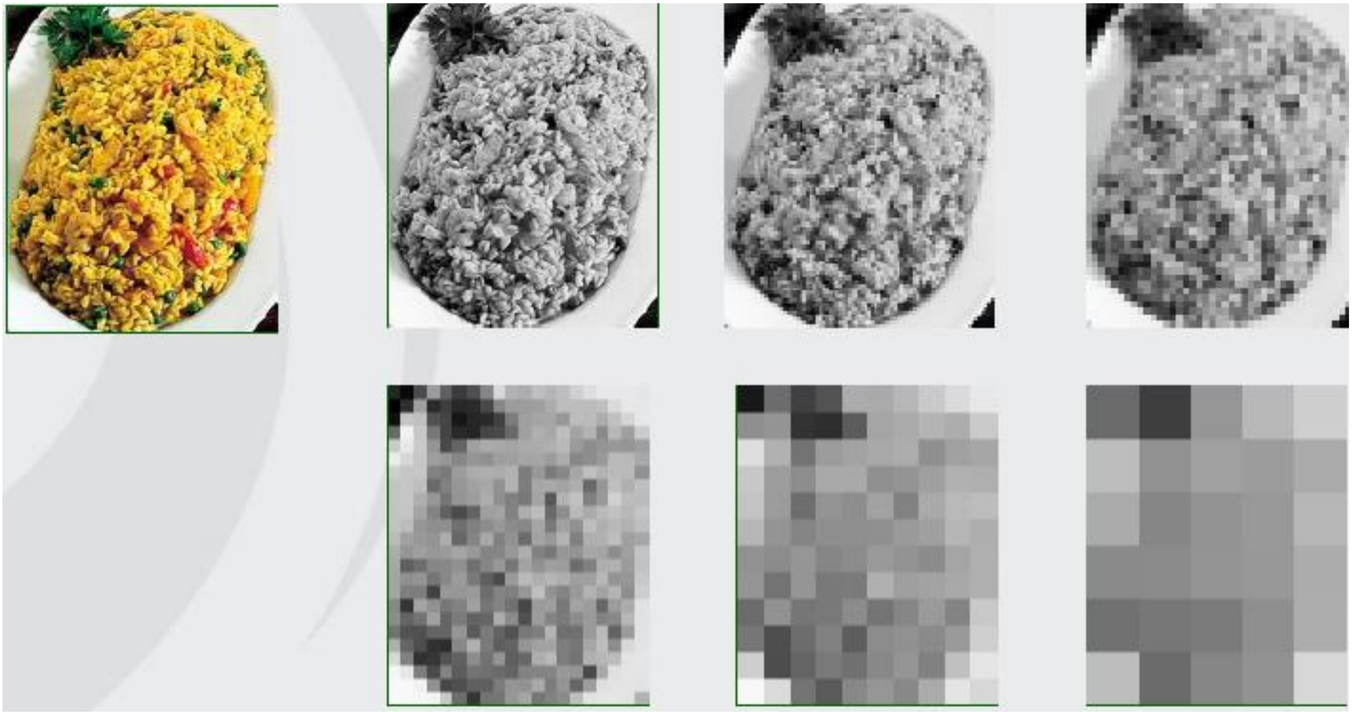
Spatial resolution

- The spatial resolution is related to the dimention of the details that can be detected
- The resolution cell is the smallest area with an associated value in a digital image
- The cell is usually a square (but sametimes other shapes are used)
- The pixel corresponds to the elementary cell





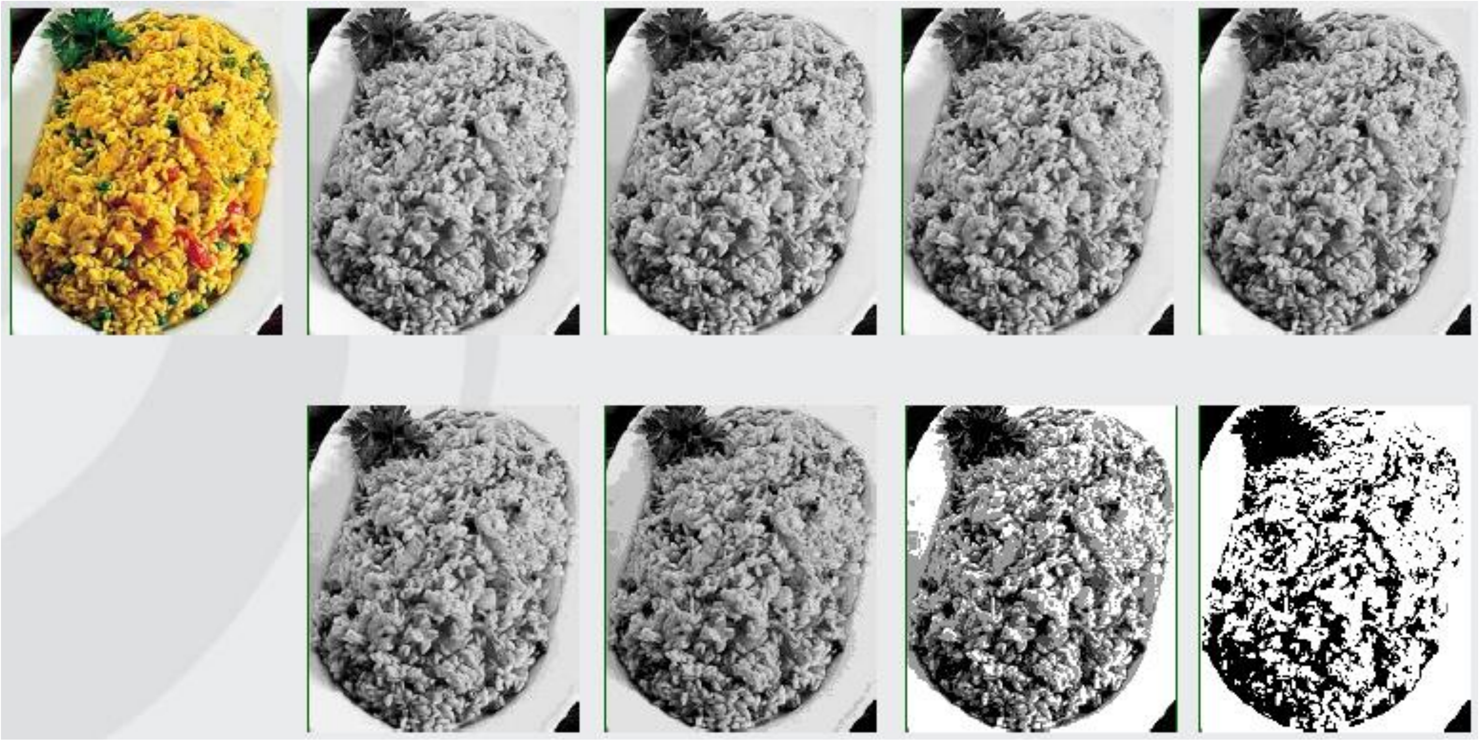
Spatial resolution



Color depth

- The color depth is the number of bits of each pixel
- A binary image is an image where each pixel can have only two values: (0, 1), (false, true), (object, background)
  - A binary image uses only a bit for each pixel
- A gray image is an image that uses larger ranges
  - Some common values: [0,63], [0,255], [0,1023] (6, 8, 10 bit)
- A human being can deal with 8 bits

Gray scale resolution



Color images

- The color images usually memorize 3 values for each pixel (red channel, green channel, blue channel)
- Each pixel usually use 1 byte (8 bits) so we can have 256x256x256 different colors (~16 millions)
- A human being is not able to discriminate so much colors

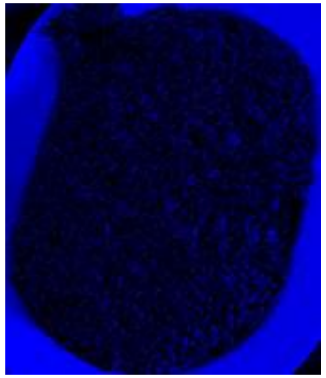
Color images

- Color image

- Red channel

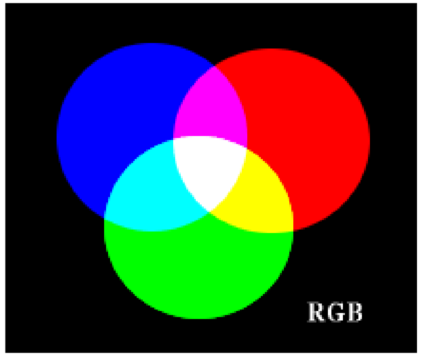
- Green channel

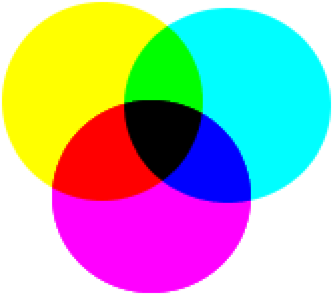
- Blue channel



Color models

- There are many modes to deal with colors
- They are related to the final task
  - RGB - monitors
  - CMYK – cyan, magenta, yellow, black - printers





Set of usable colors

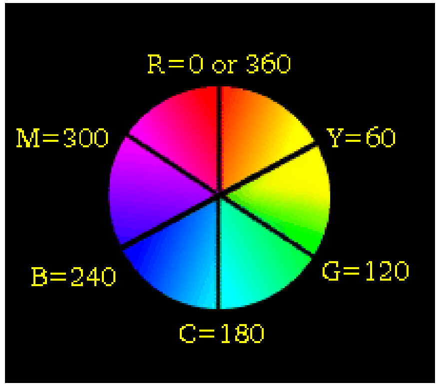
- The colors of a monitor are not the same of printable colors

Color models

- YIQ – luminance, inphase, quadrature – tv color
- HIS – hue, saturation, intensity
- HSV – hue, saturation, value
- HSB – hue, saturation, brightness

HSV

- **0°: 255, 0, 0**
- **60°: 255, 255, 0**
- **120°: 0, 255, 0**
- **180°: 0, 255, 255**
- **240°: 0, 0, 255**
- **300°: 255, 0, 255**



Color images

- A possible choice to limit the memory use a reduced number of colors is used (8, 4, 1 bits each pixel)
- So also a color LUT (look up table) is memorized

Color images

- Original image

- 256 colors

- 16 colors

- 8 colors

Color lut

Pixel value

red	green	blue
R1	G1	B1
R2	G2	B2
R3	G3	B3
R4	G4	B4
R5	G5	B5
R6	G6	B6

Visualize value (R4, G4, B4)

BMP images

```
typedef struct {
        short magic; /* "BM" */
        long file_dim; /* file dimension  */
        long l0; /* 0 */
        long header_dim; /* header dimension  */
        long l40; /* 40 */
        long xsize; /* image width */
        long ysize; /* image height */
        short nchan; /* 1 */
        short zsize; /* 1-4-8-24-32 */
        long compression; /* 0 -> no compression */
        long data_dim; /* data dimension */
        long xppi;
        long yppi;
        long colors; /* lut dimension */
        long colors1;
} bmp_header;
```

File structure

Header

LUT

Pixel values

PGM (portable gray map) images

File structure:

An ASCII Header (humen readable):

«P5» (magic number)

width height

Maximum pixel value (usually 255)

An arbitrary number of comments lines may be present (beginning with ‘#’)

Image data: 1 byte each pixel

PGM (ascii) images

File structure:

An ASCII Header (humen readable):

«P2» (magic number)

width height

Maximum pixel value (usually 255)

An arbitrary number of comments lines may be present (beginning with ‘#’)

Image data: 1 human readable number each pixel

PPM (portable pixel map) images

File structure:

An ASCII Header (humen readable):

«P6» (magic number)

width height

Maximum pixel value (usually 255)

An arbitrary number of comments lines may be present (beginning with ‘#’)

Image data: 3 bytes for each pixel (RGB)

PPM (ascii) images

File structure:

An ASCII Header (humen readable):

«P3» (magic number)

width height

Maximum pixel value (usually 255)

An arbitrary number of comments lines may be present (beginning with ‘#’)

Image data: 3 human readable numbers each pixel

GIF images

File structure:

«GIF89a» (magic number)

A Header (width, height, number of colors):

Color lut

Compressed image data

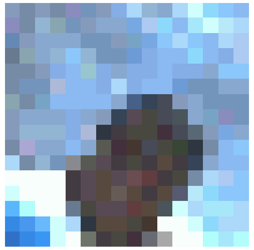
GIF images

- PPM image 290 Kb
- GIF image 53 Kb

- Lossless compression: it is possible to reconstruct the original image data (if the number of colors is at most 256)

JPG images

- The image is subdivided in blocks of 16x16 pixels
- An analysis in the frequence domain is done and high frequence componentsare eliminated (humans do not well recognize)
- For visualization the result is good



JPG images

- PPM image 290 Kb
- JPG image 25 Kb

- Lossy compression: it is not possible to reconstruct the original image data
- The compression level is a parameter of the transformation process

```
magick rose: -quality 80% rose.jpg
```

ARGB images

- Sometimes pixel values are memorize as integer values of 32 bits
- In this case it is used a fourth channel (alpha channel). It is used to memorize the degree of visibility of the pixel: 0 value corresponds to a transparent pixel, 255 to a opaque pixel
- Alpha channel can be used in Java images, in PNG images and in BMP imgages (obviously they are only examples).