Elaborazione di immagini
Image Processing

AA 2015-2016

Dipartimento di Informatica

Università degli Studi di Verona

General information

- Teacher: Gloria Menegaz
- Teaching assistant: Mauro Zucchelli
- Scheduling
  - Theory (4 CFU)
    - Mon. 14.30 to 16.30, room C
    - Wed. 11.30 to 13.30, room C
    - Laboratory (2 CFU)
    - Thu. 14.30 to 17.30, lab. Alpha
  - Tutoring (*ricevimento*)
    - by appointment (email)
  - Start and end dates
    - March 2nd, 2015 – beginning
      May , 2015 - end
- Exam
  - Oral/Written
  - Theory+lab
- Support
  - Slides of the course
- Books:
  - Digital image processing, Gonzalez-Woods
  - Digital image processing, Pratt
  - Student's notes

Contents

- Introduction (GW, Chapter 1)

Contents

- Introduction (GW, Chapter 1)
- Fourier Transform
- Sampling in 2D
- Quantization
- Filtering (linear filters)
- Edge detection
- Segmentation techniques
- Basics of pattern recognition
  - Clustering, classification
- Color imaging
- Hints for Wavelets and multiresolution
- The JPEG coding standard

Why do we process images?

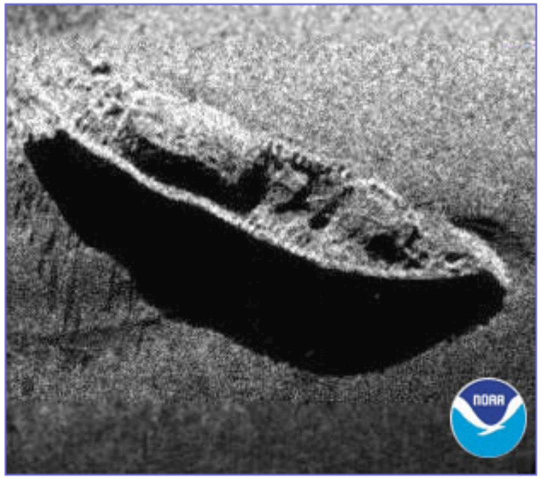
- To facilitate their storage and transmission
- To prepare them for display or printing
- To enhance or restore them
- To extract information from them
- To hide information in them

Image types

Optical (CCD)

radar (SAR)

underwater



infrared

medical (MRI)

Microarray images

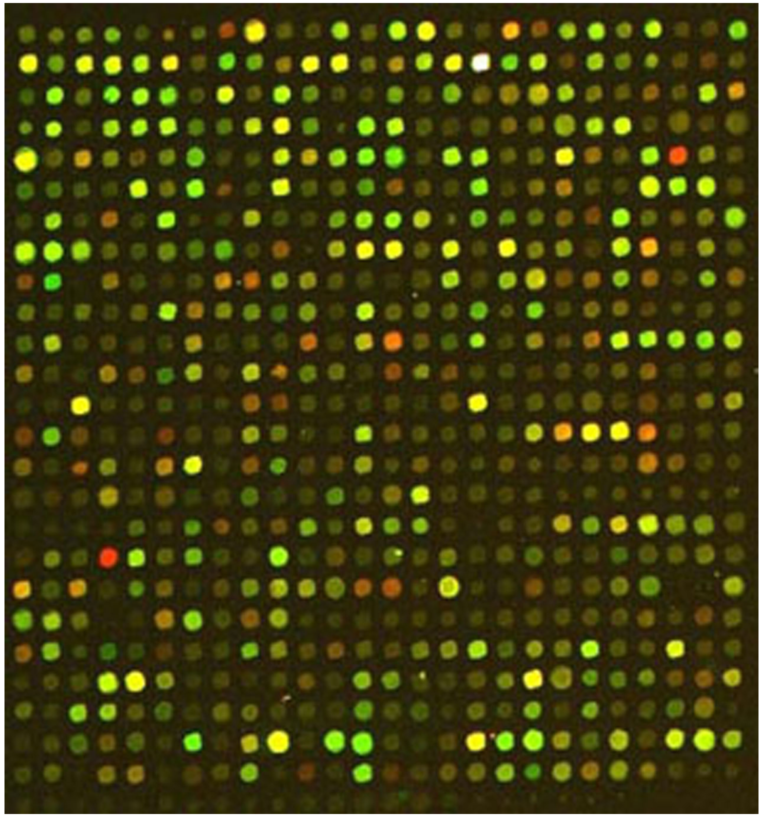


Image Processing Example

- Image Restoration

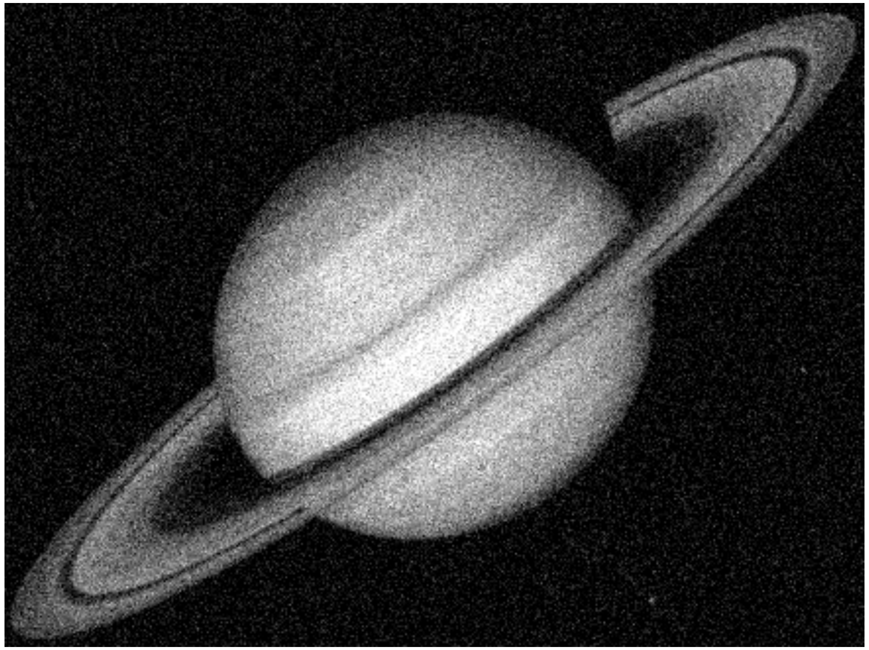
Original image

Blurred

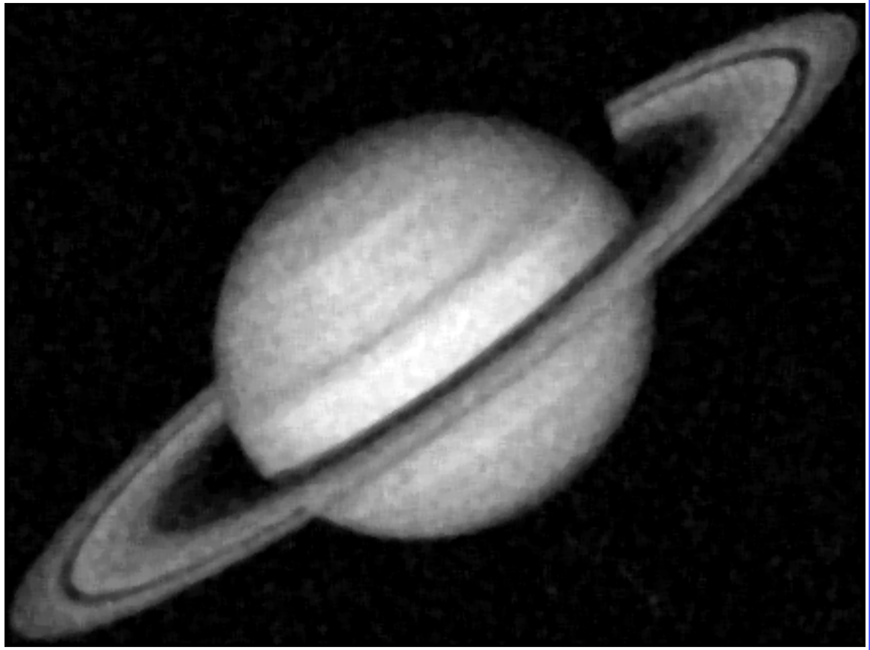
Restored by Wiener filter

Image Processing Example

- Noise Removal



Noisy image

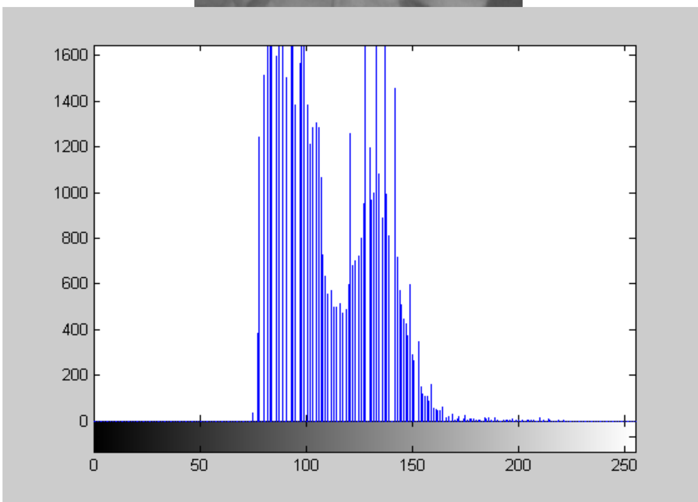


Denoised by Median filter

Image Processing Example

- Image Enhancement

Histogram equalization



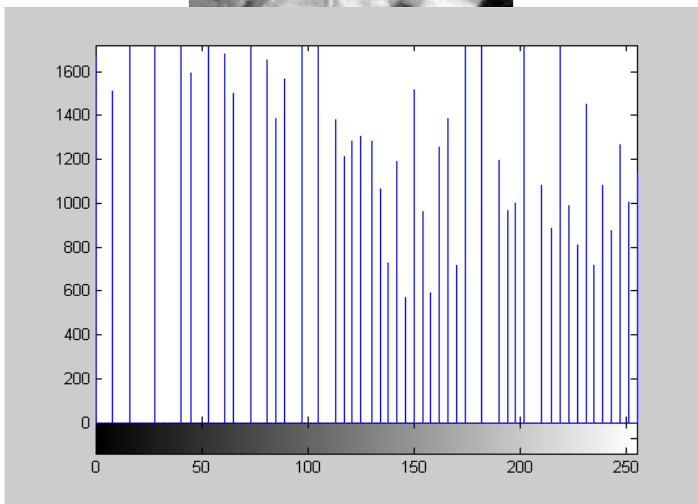


Image Processing Example

- Artifact Reduction in Digital Cameras

Original scene

Captured by a digital camera

Processed to reduce artifacts

Image Processing Example

- Image Compression

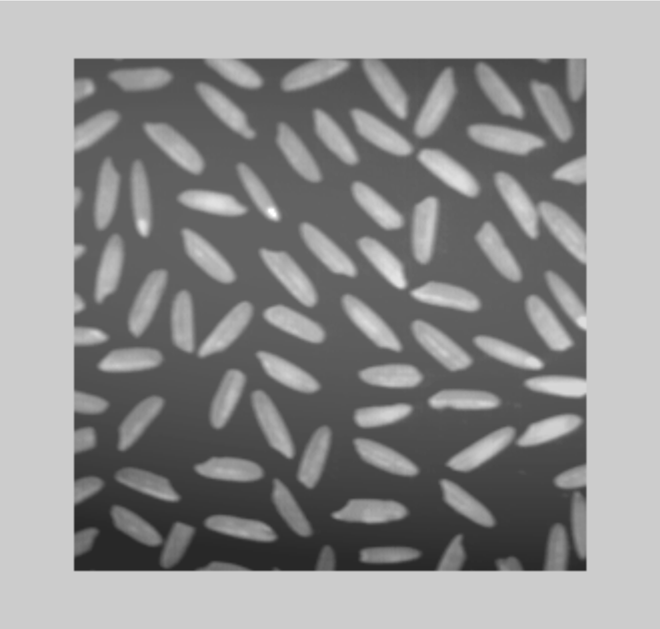
Original image 64 KB

JPEG compressed 15 KB

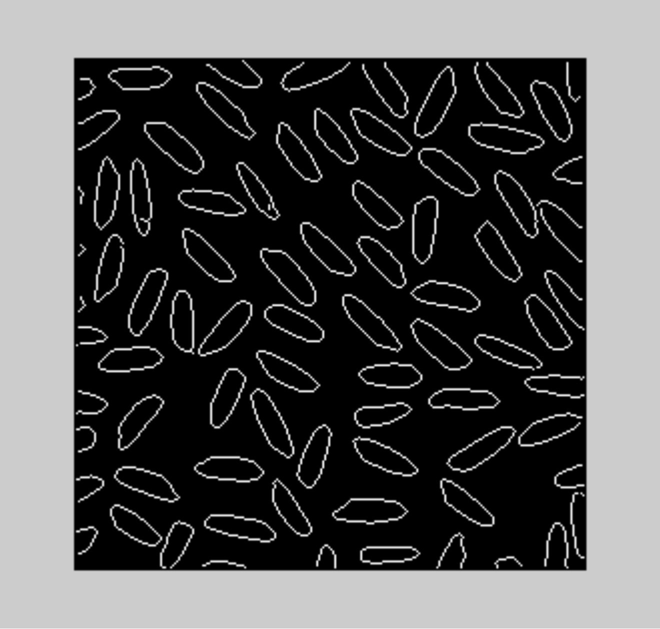
JPEG compressed 9 KB

Image Processing Example

- Object Segmentation



“Rice” image



Edges detected using Canny filter

Image Processing Example

- Resolution Enhancement

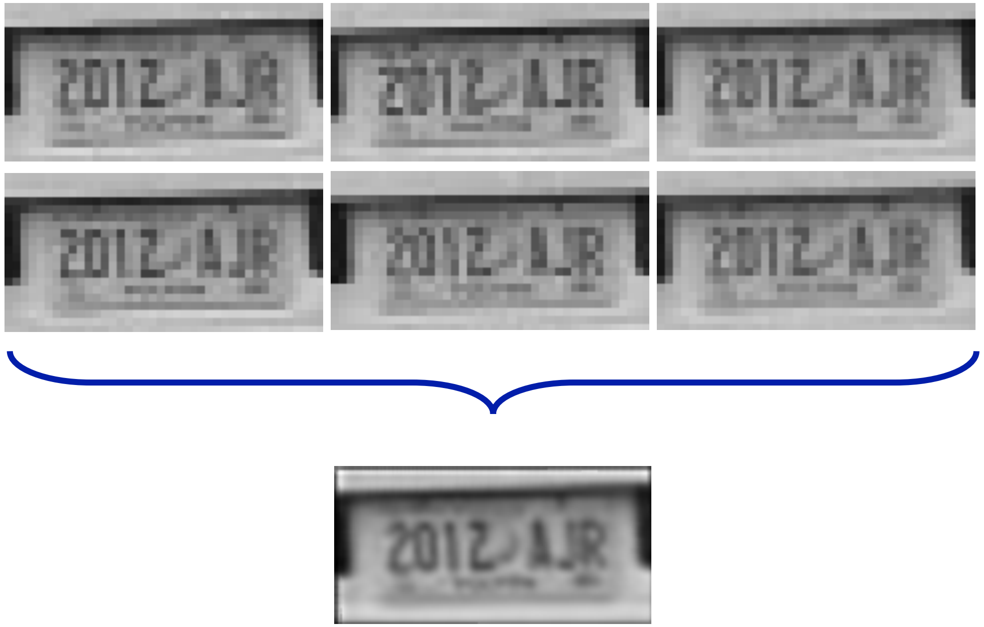


Image Processing Example

- Security and encryption
  - Watermarking

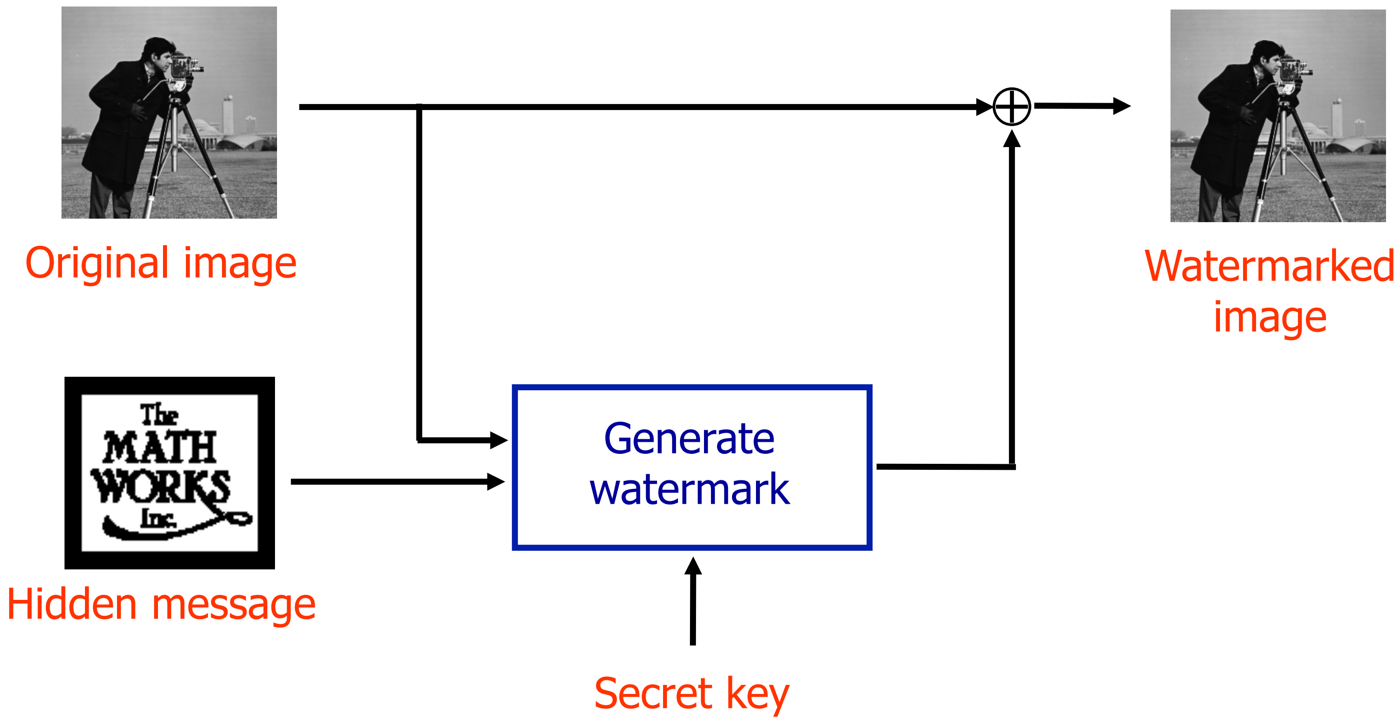


Image Processing Example

- Face Recognition

Image Processing Example

- Fingerprint Matching

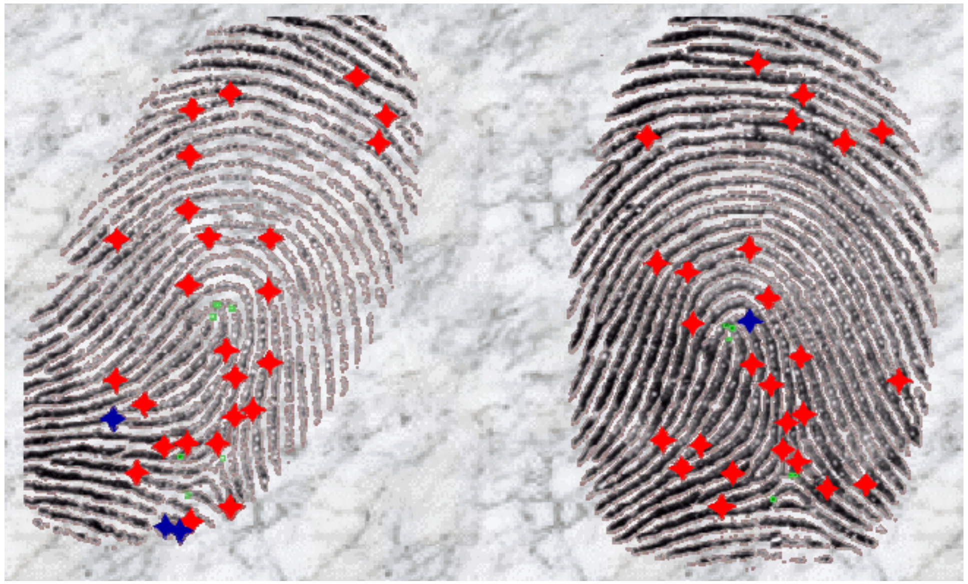


Image Processing Example

- Segmentation

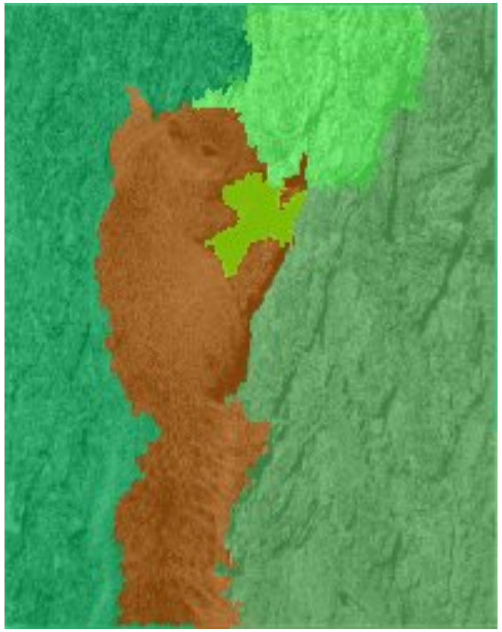
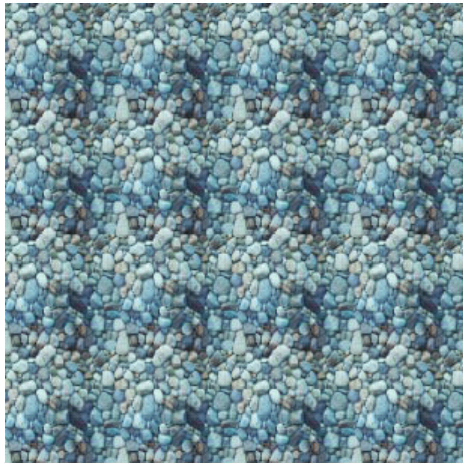


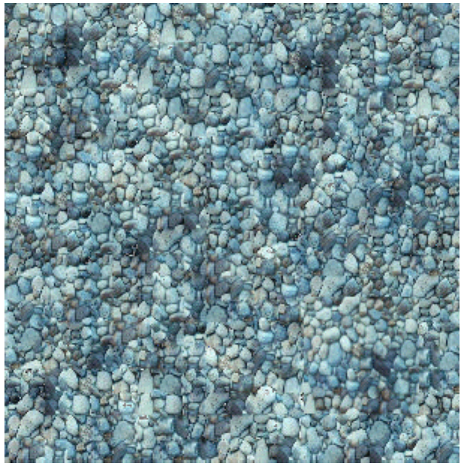
Image Processing Example

- Texture Analysis and Synthesis

Photo



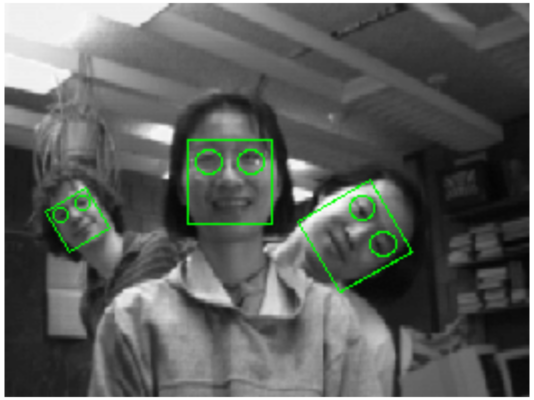
Pattern repeated



Computer generated

Image Processing Example

- Face detection and tracking



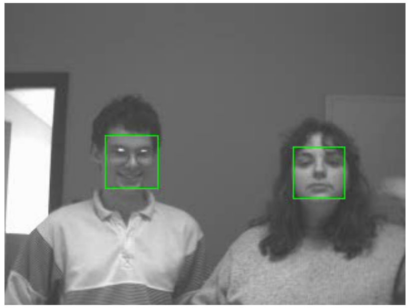


Image Processing Example

- Face Tracking

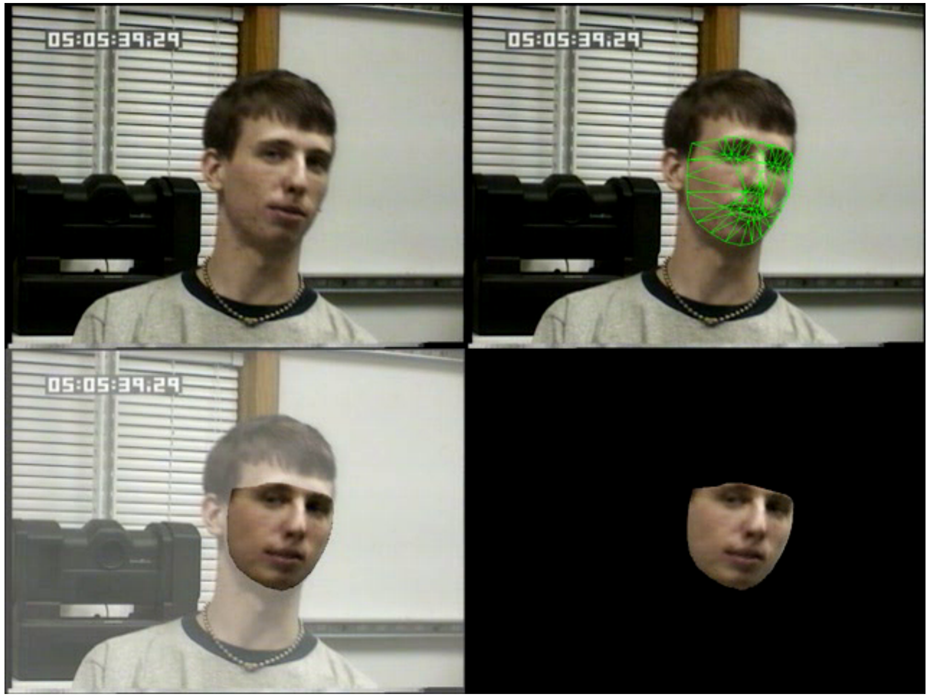
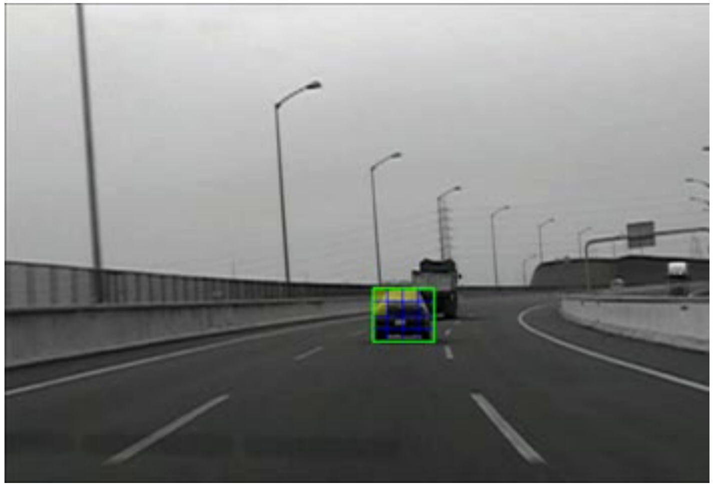


Image Processing Example

- Object Tracking



Computer vision

- Methods for estimating the *geometrical* and *dynamical* properties of the imaged scene based on the acquired images
  - Scene description based on image features
- Complementary to computer graphics
  - Get information about the 3D real world based on its 2D projections in order to automatically perform predefined tasks

Pattern Recognition

- Image interpretation
- Identification of basic and/or complex structures
  - implies pre-processing to reduce the intrinsic redundancy in the input data
  - knowledge-based
    - use of a-priori knowledge on the real world
    - stochastic inference to compensate for partial data
- Key to clustering and classification
- Applications
  - medical image analysis
  - microarray analysis
  - multimedia applications

Pattern Recognition

- Clustering
  - data analysis aiming at constructing and characterizing clusters (sets without prior knowledge)
- Feature extraction and selection
  - reduction of data dimensionality
- Classification
  - Structural (based on a predefined "syntax"):
    - each pattern is considered as a set of primitives
    - clustering in the form of parsing
  - Stochastic
    - Based on statistics (region-based descriptors)

Applications

- Efficiently manage different types of images
  - Satellite, radar, optical..
  - Medical (MRI, CT, US)
  - Image representation and modelling
- Quality enhancement
  - Image restoration
    - deblurring, denoising, hole filling
- Image analysis
  - Feature extraction and exploitation
- Image reconstruction from projections
  - scene reconstruction, CT, MRI
- Compression and coding

Typical issues

Context-independent

- Image resampling and interpolation
  - Sampling, quantization, filtering
- Visualization and rendering
- Multispectral imaging
  - Satellite, color
- Motion detection, tracking
- Automatic quality assessment
- Data mining
  - query by example

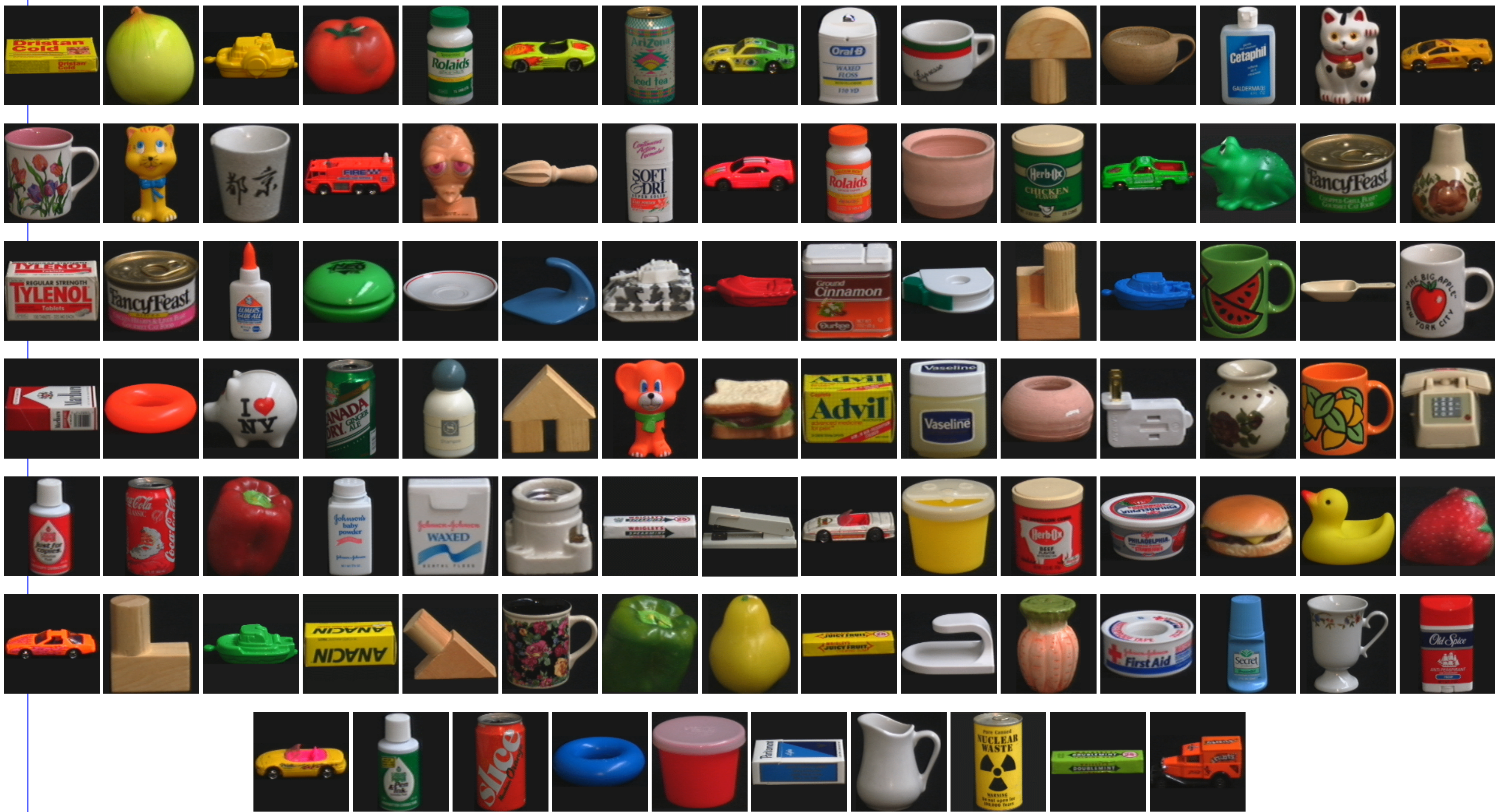
Medical imaging

- Image analysis
  - optical devices, MRI, CT, PET, US (2D to 4D)
- Image modeling
  - Analysis of hearth motion, models of tumor growth, computer assisted surgery
- Telemedicine
  - remote diagnosis, distributed systems, medical databases

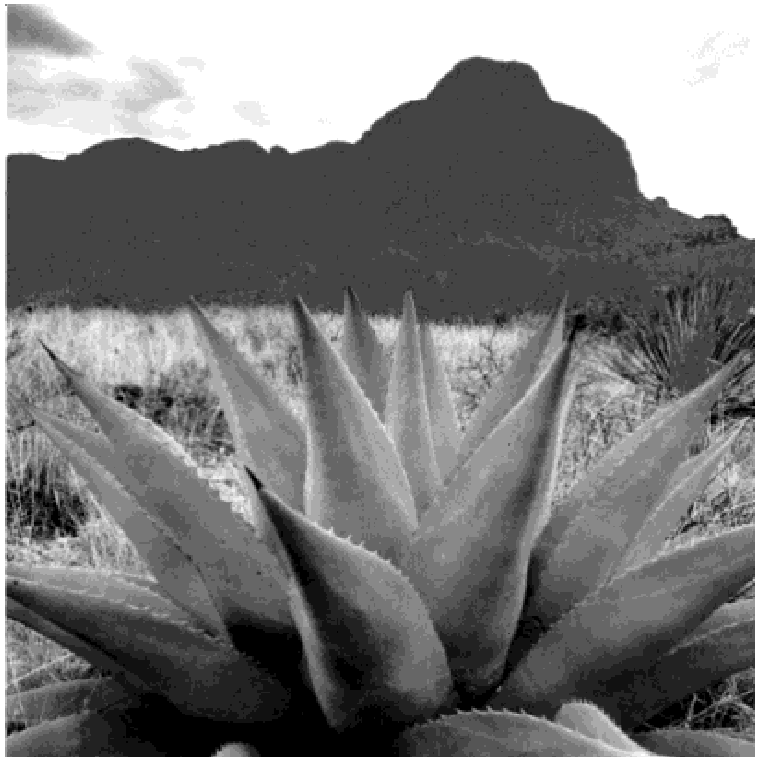
Other applications

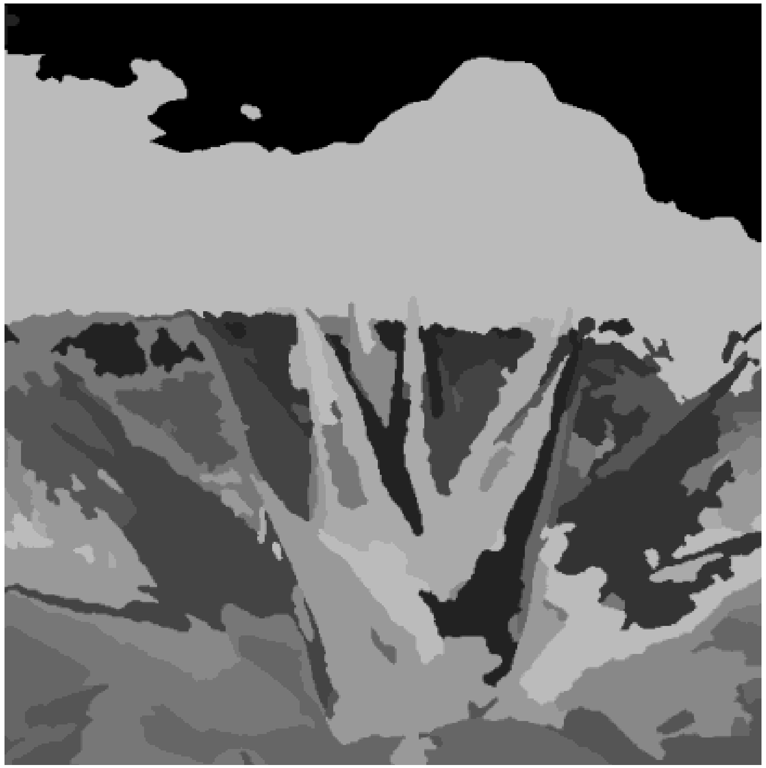
- Quality control
- Reverse engineering
- Surveillance (monitoring and detection of potentially dangerous situations)
- Social computing (face and gesture recognition for biometrics and behavioural analysis)
- Robotics (machine vision)
- Virtual reality
- Telepresence

Query by example



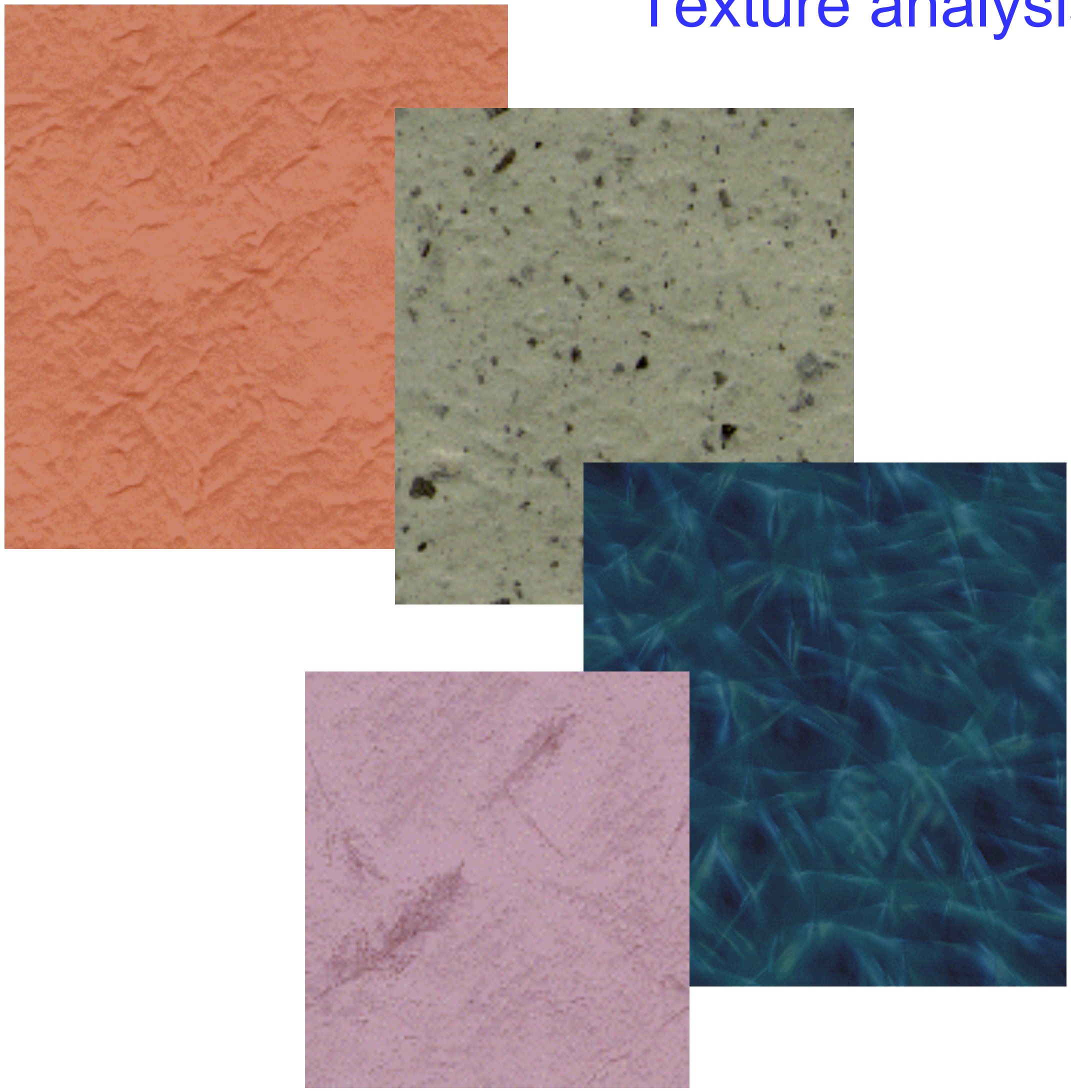
Segmentation



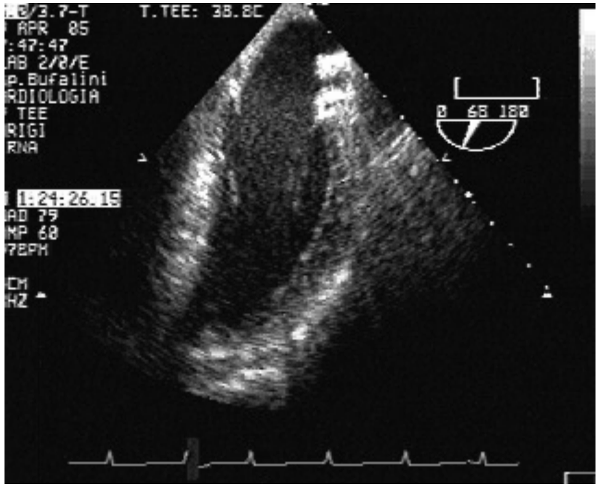


Face recognition

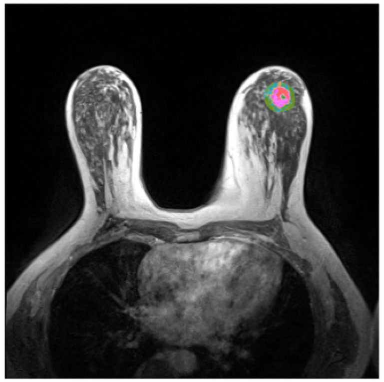
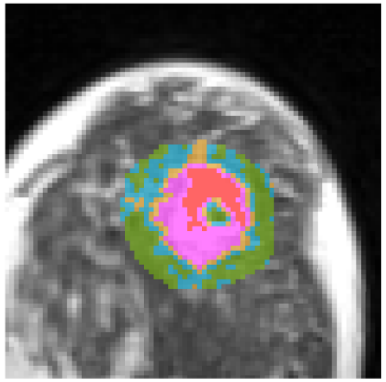
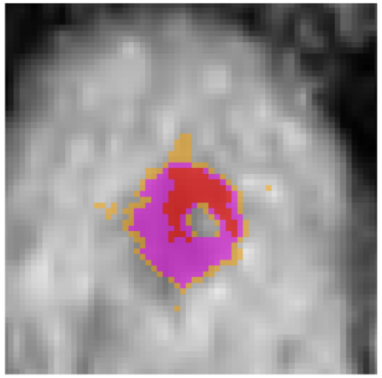
Texture analysis



Medical textures

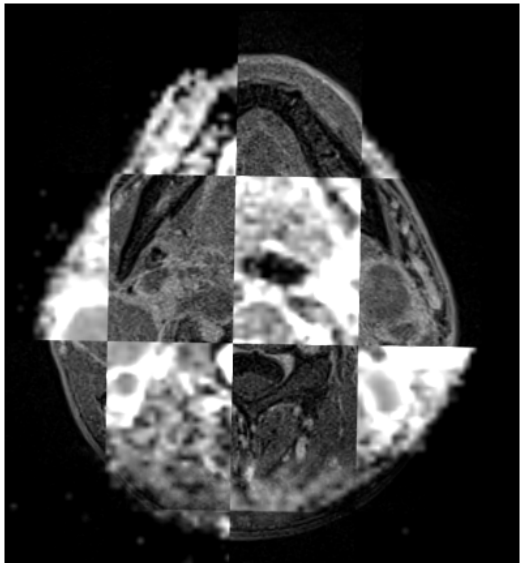


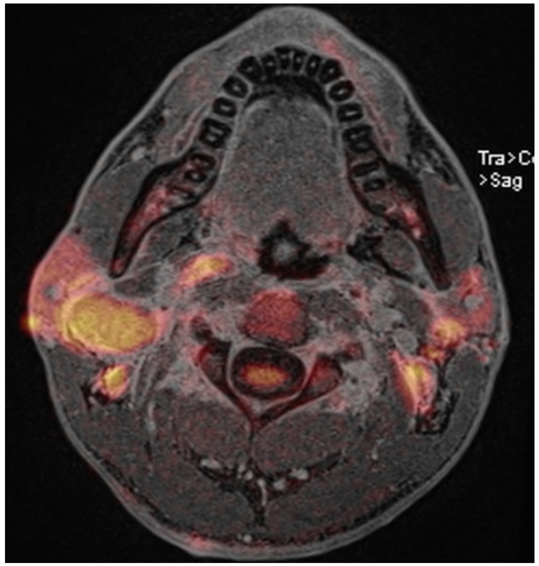
Medical Image Analysis

MI applications

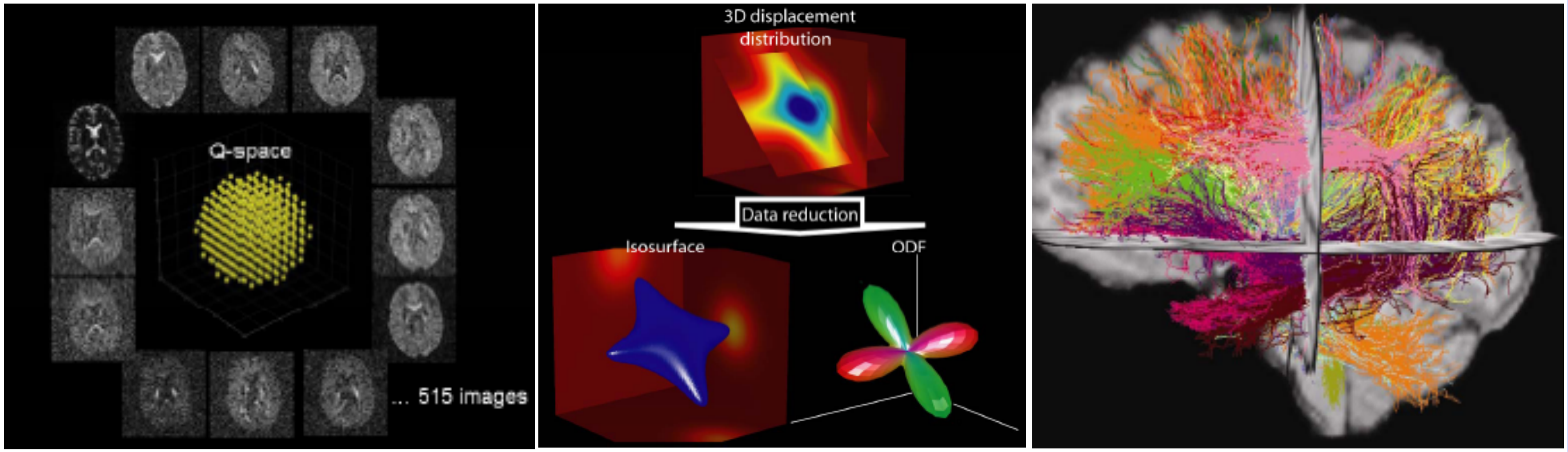
- Tumor identitication and staging



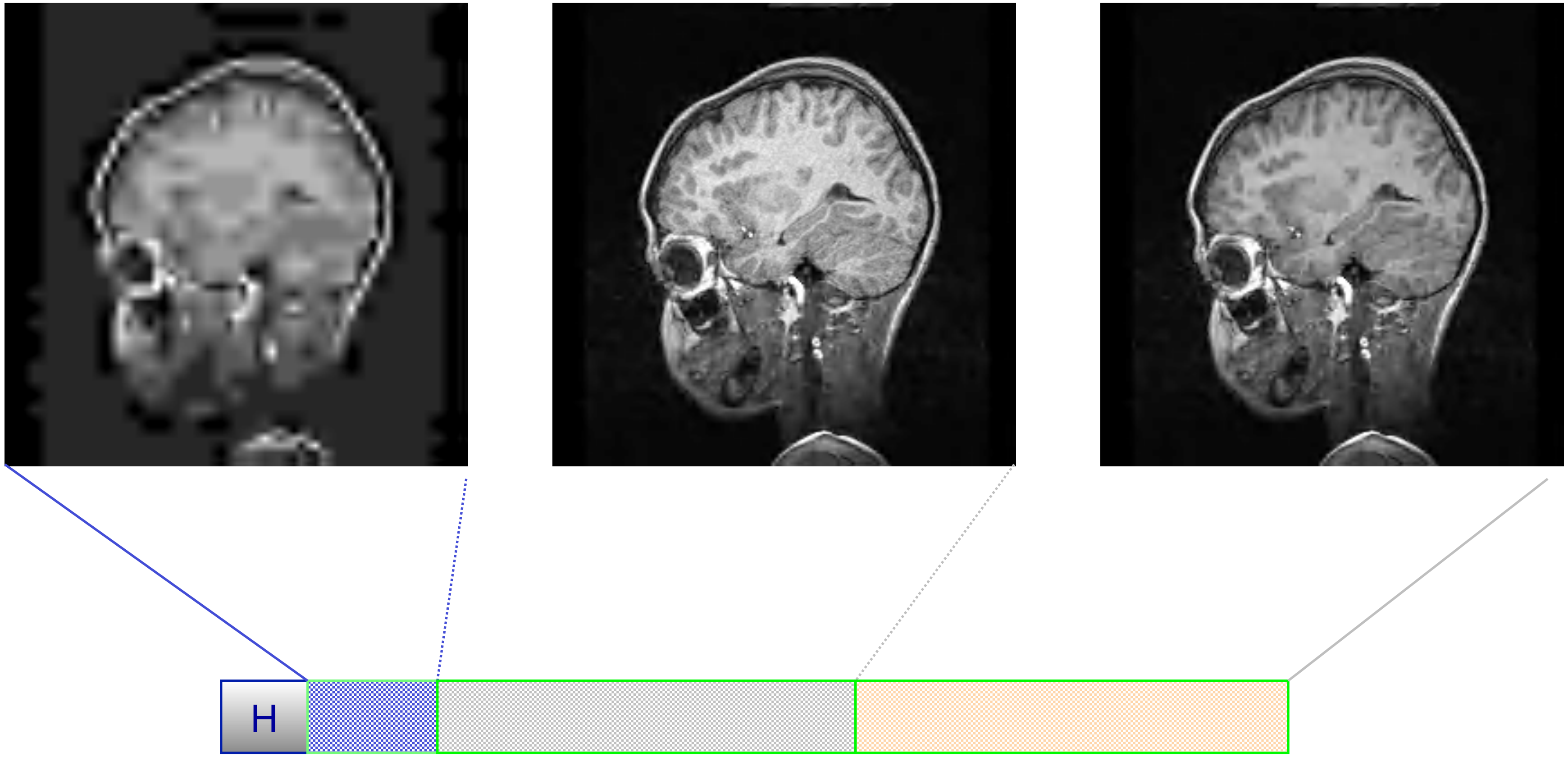


MI applications

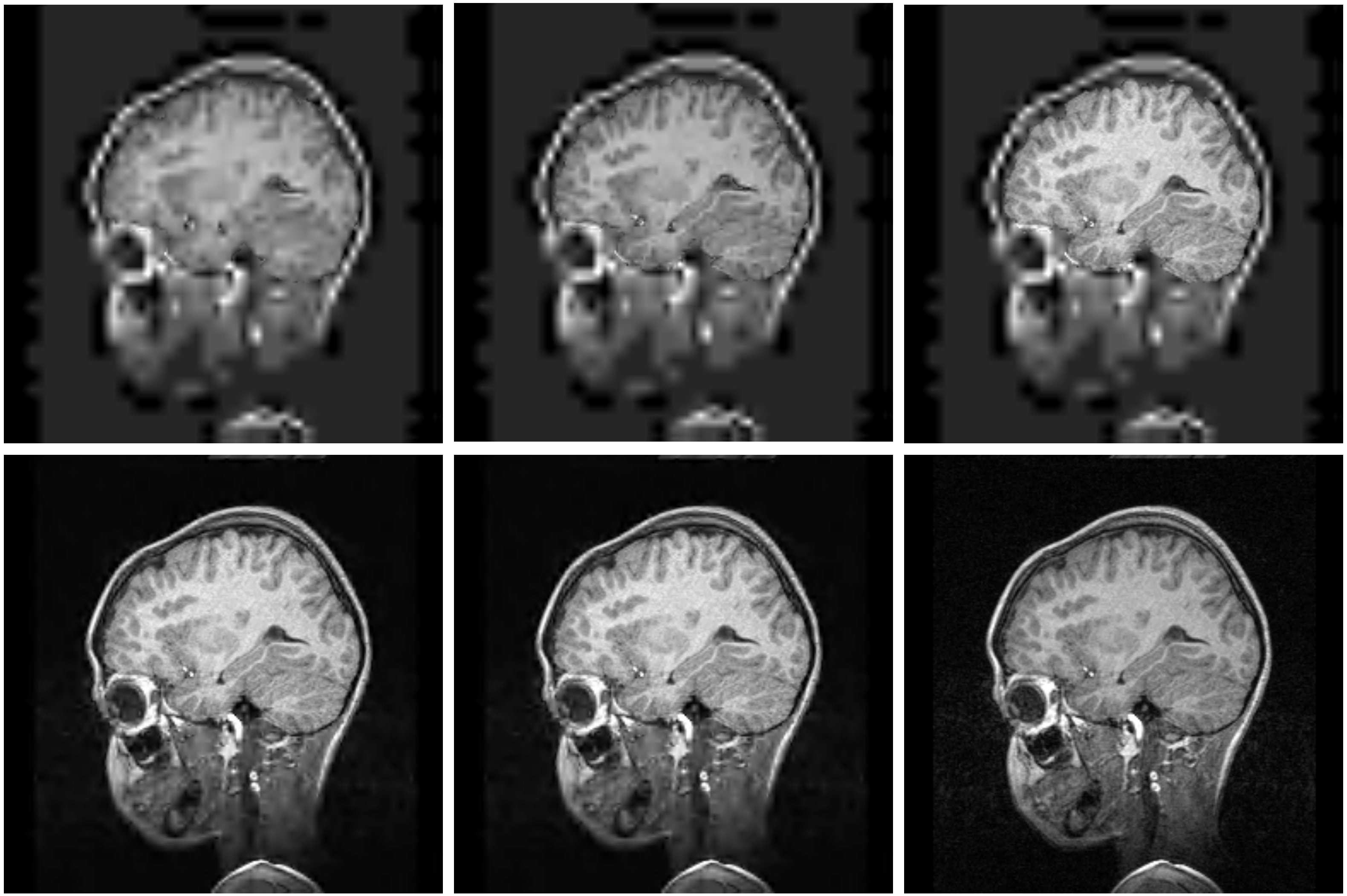
- Exploring brain anatomy by diffusion weighted MRI



Compression and coding

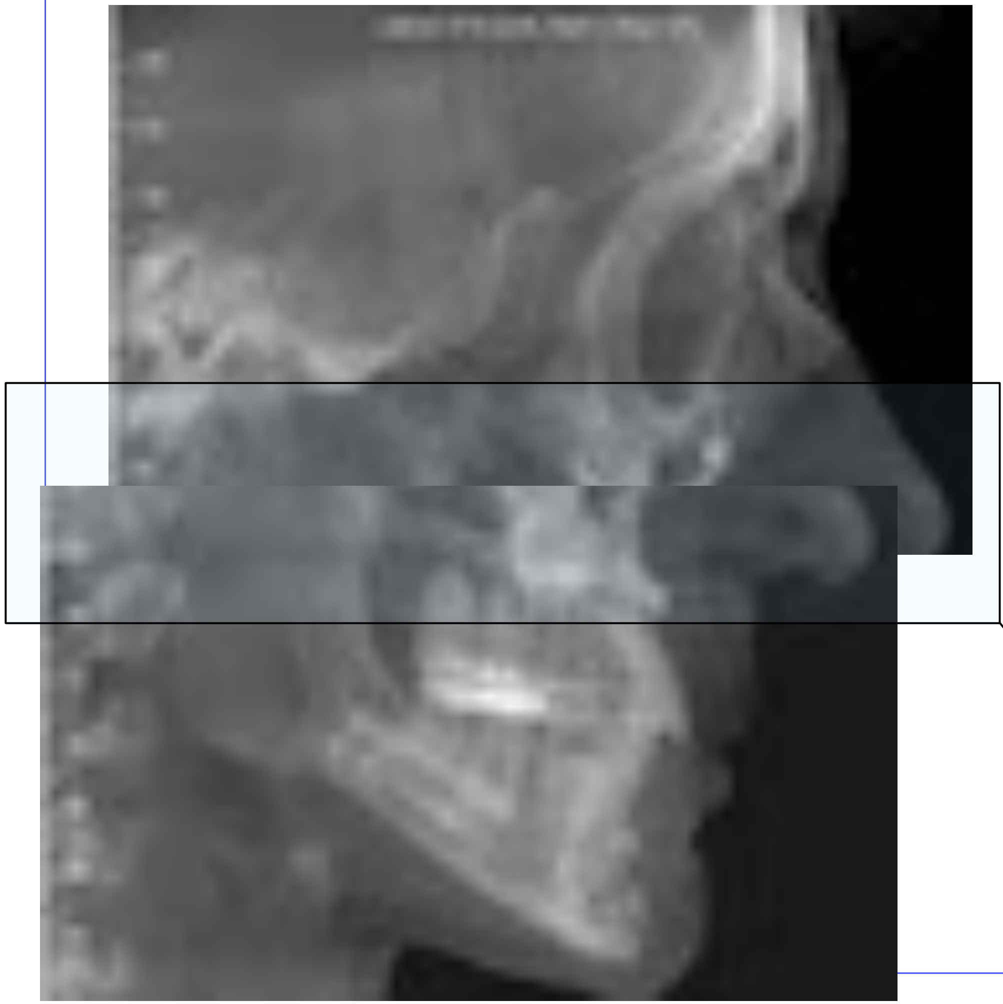


Object-based processing



Mosaicing

Volumetric stitching



Align the two images acquired at different times and blend them in the area of overlap obtaining a new high-resolution image

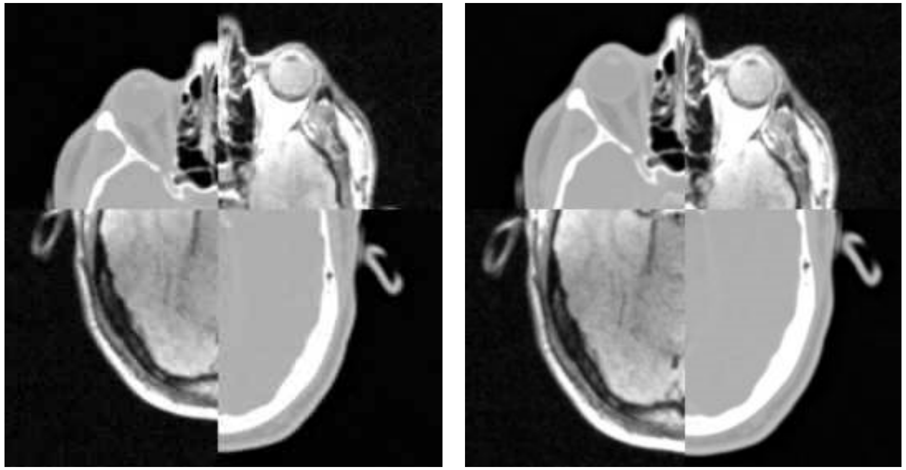
Overlap Area

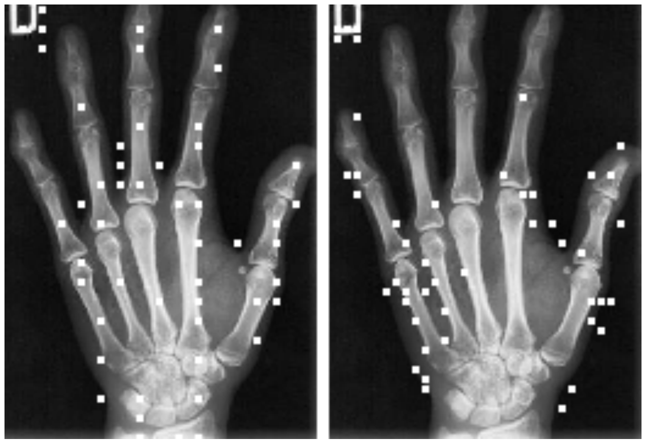
Volume stitching

Applications:

- maxillofacial
- chest x-ray
- where is required a large area of observation.

Image registration





One image is the reference (fixed image) and the other is the one to be matched (moving image). The registration process can be performed either on the entire set of voxels (pixels) or by choosing some target points according to some predefined criteria (feature points). These are usually pixels where there is high contrast like edges, junctions or user-defined landmarks (manually defined through a user interface).

Hierarchical Image Pyramid

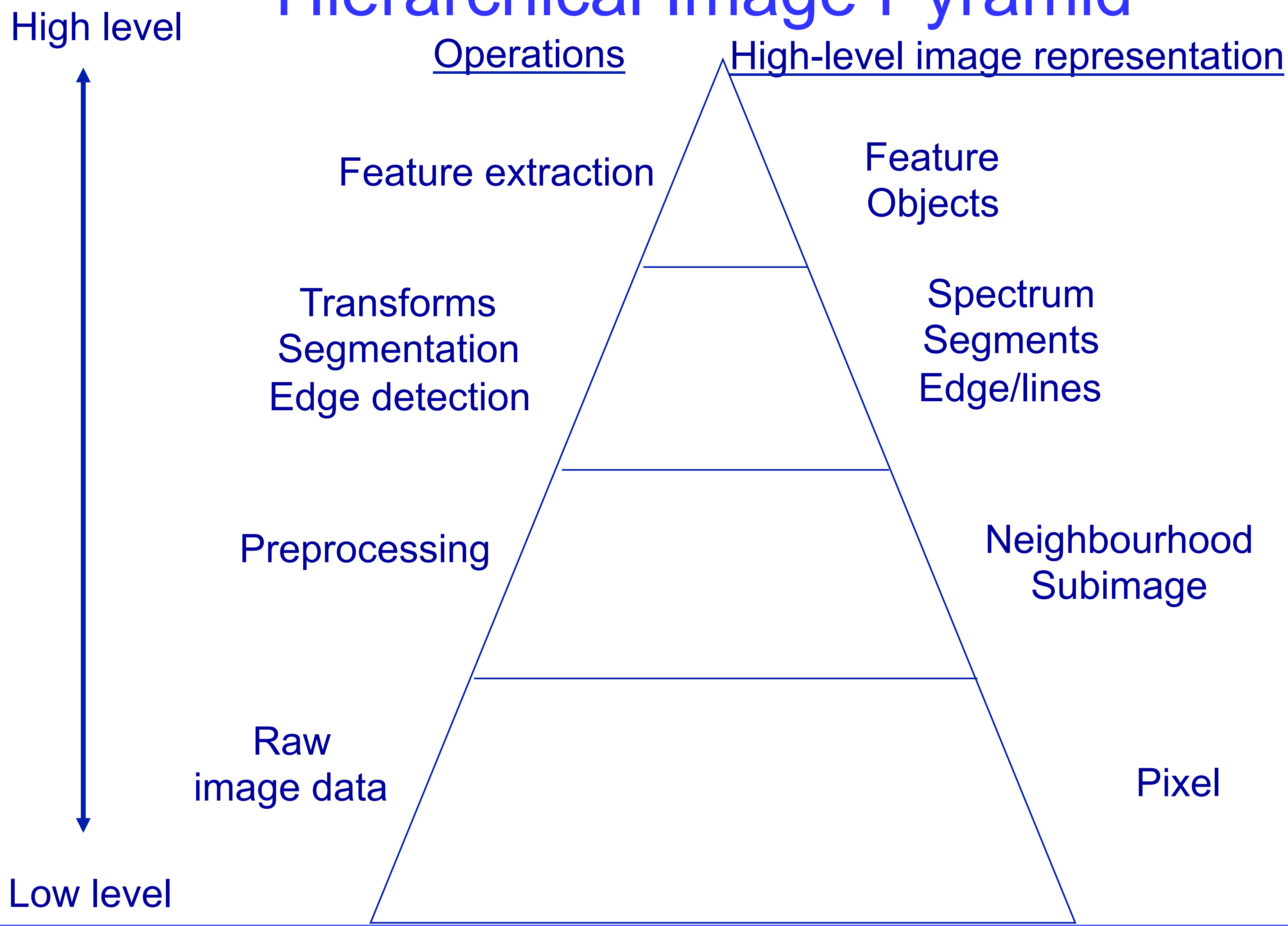
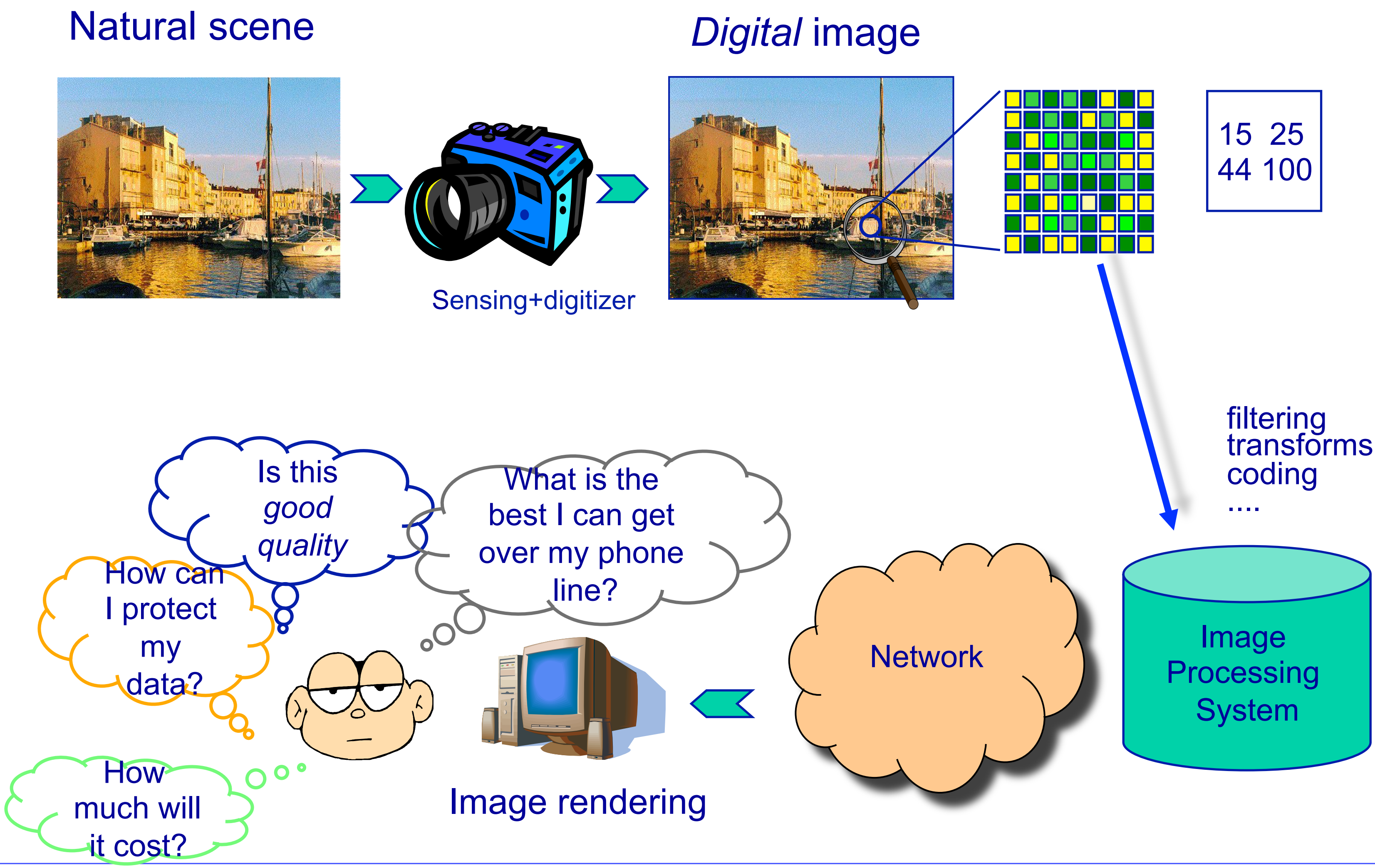


Image formation and fundamentals

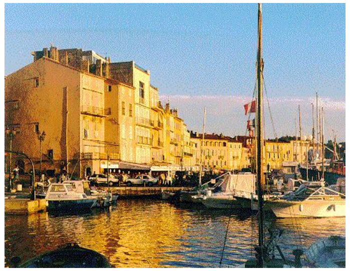
Gonzalez-Woods Chapter 1

Digital image acquisition



Sensing and digitization

Analogic image



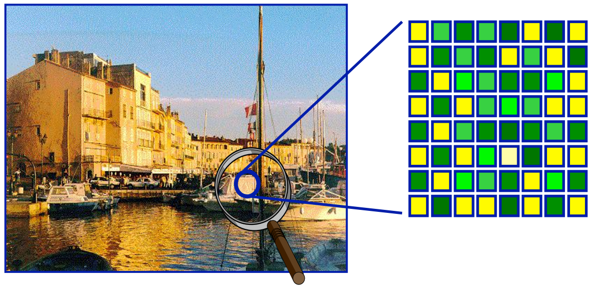
(capturing device)

A/D conversion

Sampling (2D)

Quantization

Digital image



{15,1,2}
{25,44,1}
….

Sensing and digitization

- Sensing device: device that is sensitive to the light emitted by the object we intend to image
- Digitizer: is a device for converting the output of the physical sensing device into digital form
  - For instance, in a digital video camera, the sensors produce an electrical output proportional to light intensity.The digitizer converts these outputs to digital data
- *Specialized image processing hardware* usually consists of the digitizer just mentioned, plus hardware that performs other primitive operations, such as an arithmetic logic unit (ALU), which performs arithmetic and logical operations in parallel on entire images.

Components of an image processing (IP) system

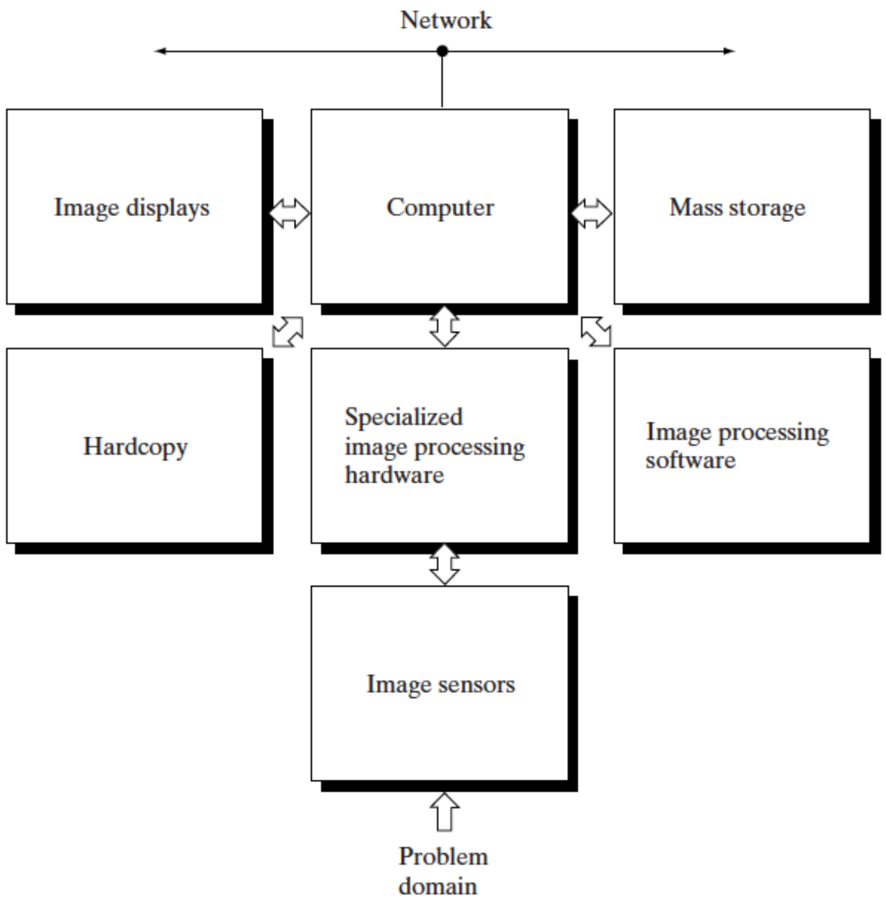


FIGURE 1.24 Components of a general-purpose image processing system.

Mass storage

- Mass storage capabilities is a must in image processing applications.
  - An image of size 1024*1024 pixels, in which the intensity of each pixel is an 8-bit quantity, requires one megabyte of storage space if the image is not compressed.
- When dealing with thousands, or even millions, of images, providing adequate storage in an image processing system can be a challenge
- **Units**
  - 1 byte = 8 bits = 2^3 bit
  - 1 KB (kilobyte) = 1024 bytes = 2^{10} byte = 2^{13} bit
  - 1 MB (mega byte) = 1024*1024 bytes = 2^{20} byte = 2^{23} bit

Gamma-ray imaging

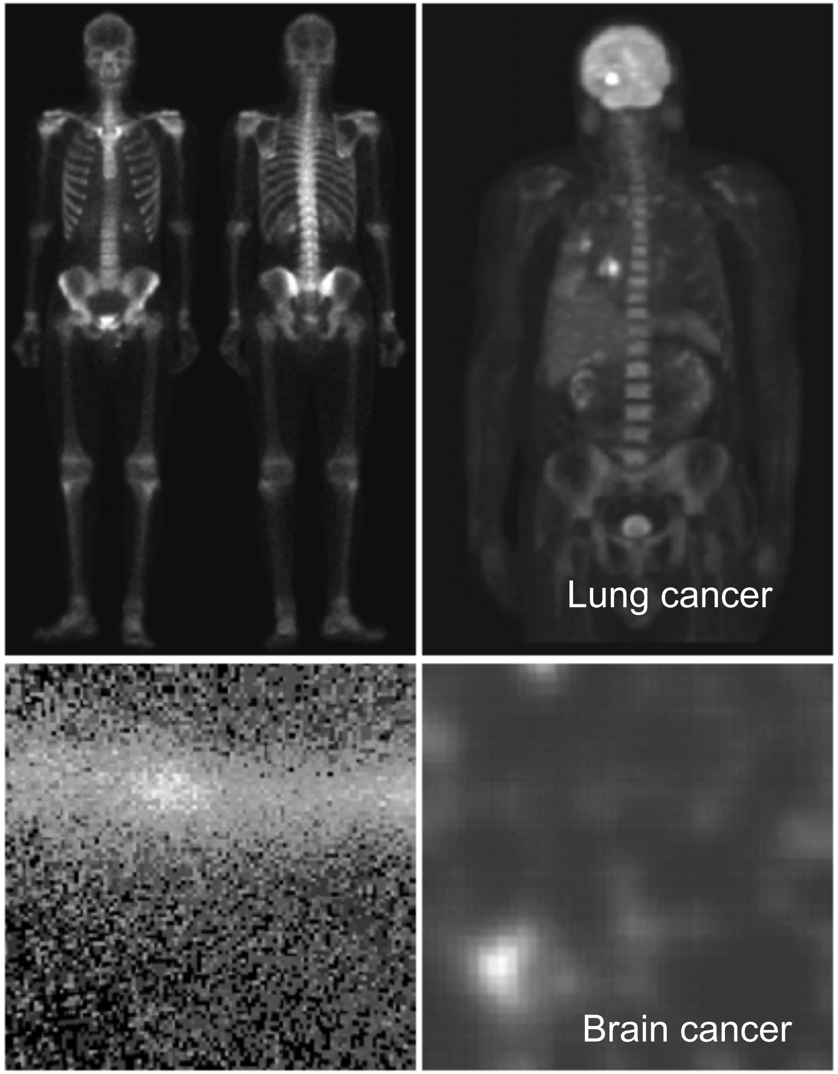
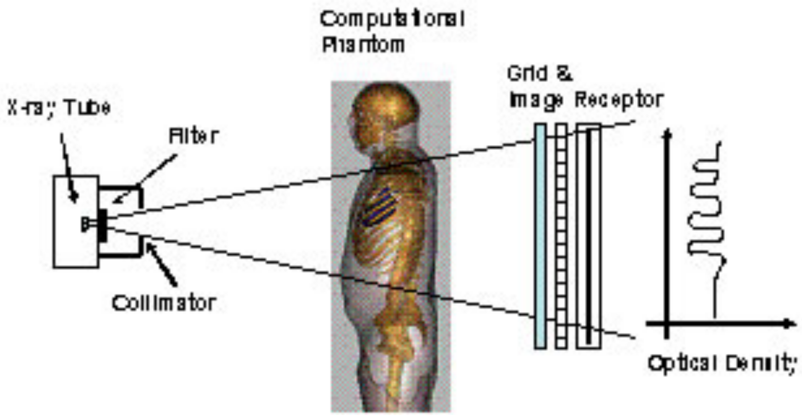


FIGURE 1.6 Examples of gamma-ray imaging. (a) Bone scan. (b) PET image. (c) Cygnus Loop. (d) Gamma radiation (bright spot) from a reactor valve. (Images courtesy of (a) G.E. Medical Systems, (b) Dr. Michael E. Casey, CTI PET Systems, (c) NASA, (d) Professors Zhong He and David K. Wehe, University of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

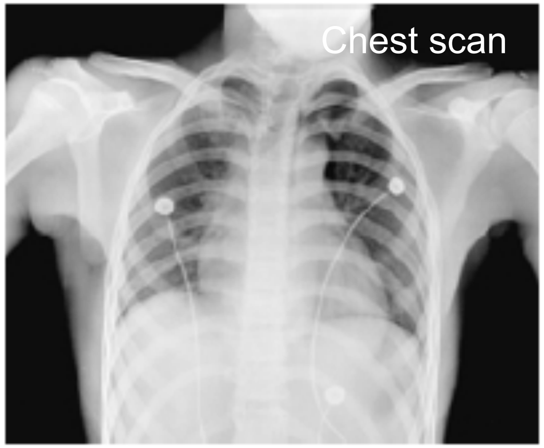
Bone scan and Positron emission tomography (PET). The patient is given a radioactive isotope that emits positrons as it decays. When a positron meets an electron, both are annihilated and two gamma rays are given off. These are detected and a tomographic image is created using the basic principles of tomography. The detector captures the gamma rays that are emitted.

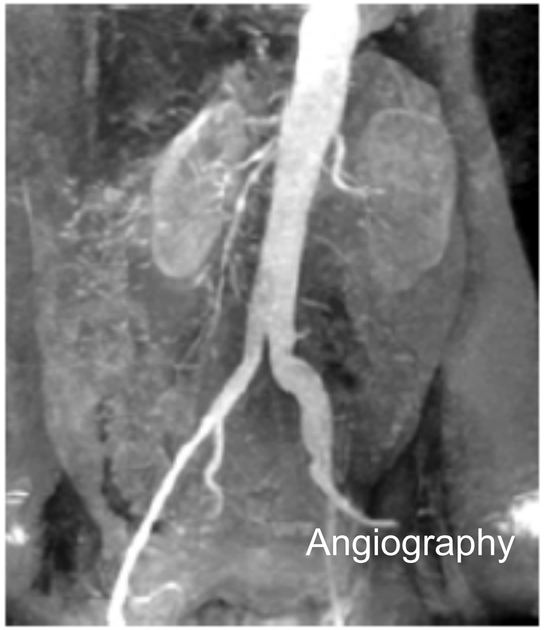
X-ray imaging

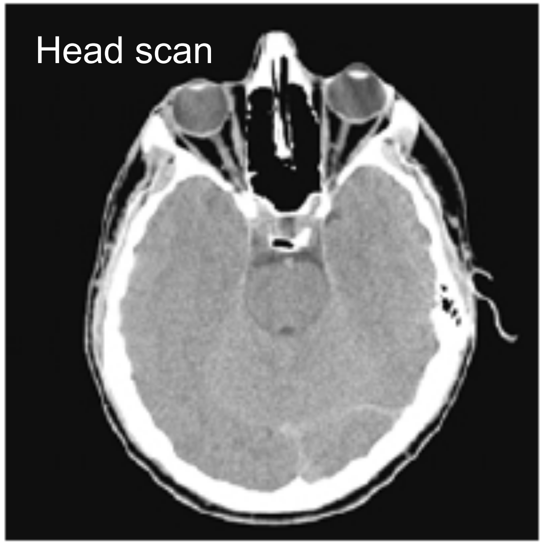
- An X-ray source is used. Electrons fly from the catode to the anode. When they hit a nucleus in the patient energy is released in the form of X-ray radiation.
- In digital radiography, digital images are obtained by one of two methods:
  - (1) by digitizing X-ray films;
  - or (2) by having the X-rays that pass through the patient fall directly onto devices (such as a phosphor screen) that convert X-rays to light.
  - The light signal in turn is captured by a light-sensitive digitizing system.

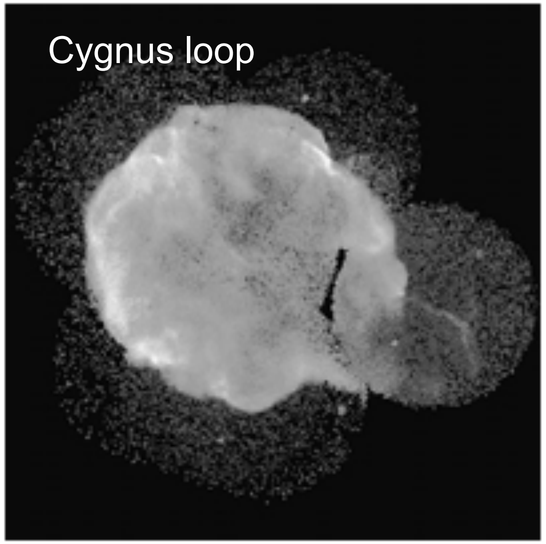


X-ray imaging









Multispectral imaging

- It was originally developed for space-based imaging.
- Acquired by remote sensing radiometers (RS)
- A multispectral image is one that captures image data at specific frequencies across the electromagnetic spectrum.
- The wavelengths may be separated by filters or by the use of instruments that are sensitive to particular wavelengths, including light from frequencies beyond the visible light range, such as infrared.
- Spectral imaging can allow extraction of additional information the human eye fails to capture with its receptors for red, green and blue.

Multispectral imaging

Includes several bands in the visual and infrared regions of the 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

Multispectral imaging

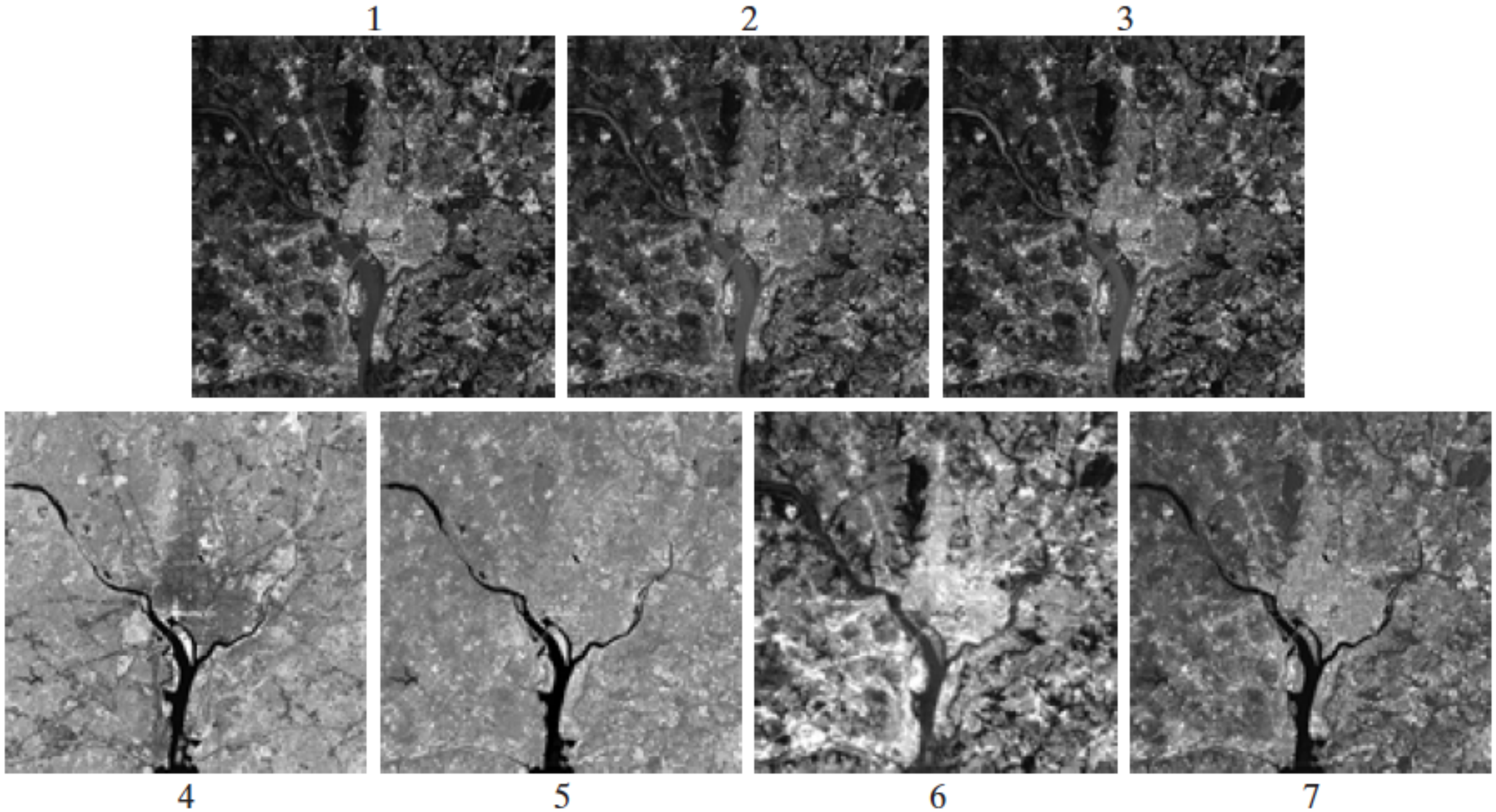


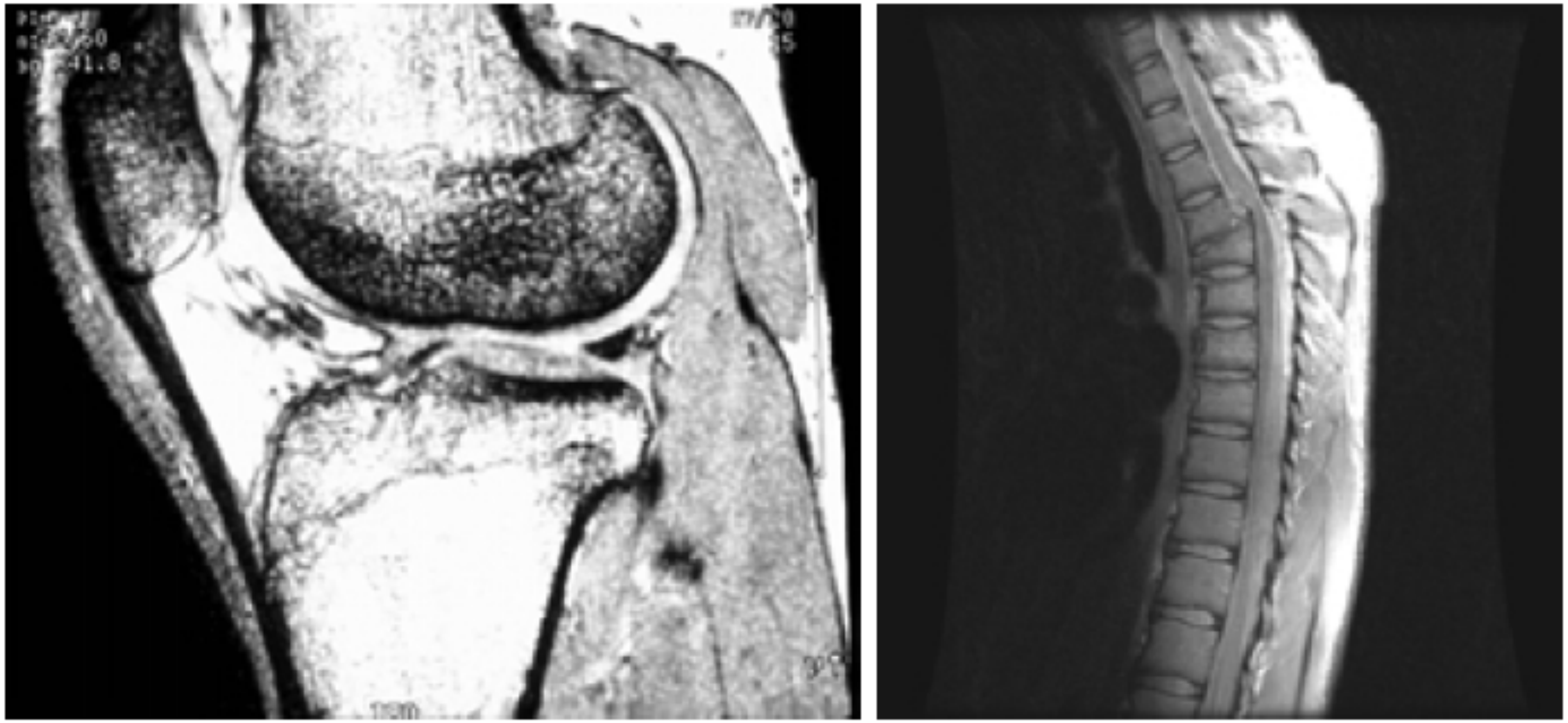
FIGURE 1.10 LANDSAT satellite images of the Washington, D.C. area. The numbers refer to the thematic bands in Table 1.1. (Images courtesy of NASA.)

Imaging in the radio band

- In medicine radio waves are used in magnetic resonance imaging (MRI). This technique places a patient in a powerful magnet and passes radio waves through his or her body in short pulses. Each pulse causes a responding pulse of radio waves to be emitted by the patient’ s tissues.
- The location from which these signals originate and their strength are determined by a computer, which produces a two-dimensional picture of a section of the patient.
- MRI can produce pictures in any plane.

Imaging in the radio band: MRI in humans

Imaging in the radio band: MRI in humans



a b

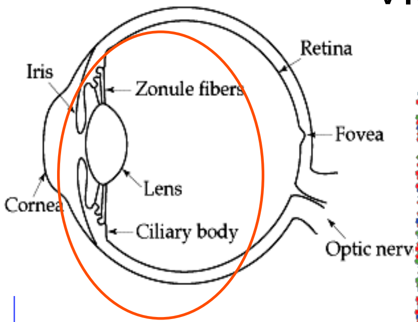
FIGURE 1.17 MRI images of a human (a) knee, and (b) spine. (Image (a) courtesy of Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School, and (b) Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

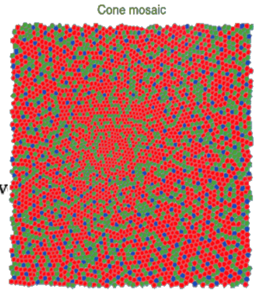
Digital image fundamentals

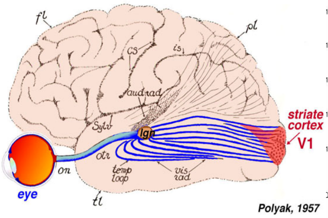
Gonzalez Woods Chapter 2

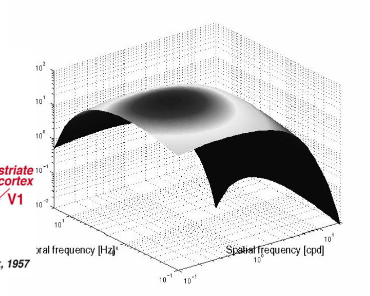
Visual and Imaging chain

VISUAL CHAIN

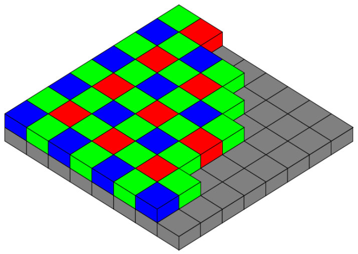


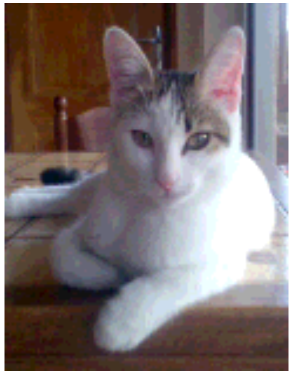






Optics	Sampling (A/D)	Multiscale representation	Filtering	Processing





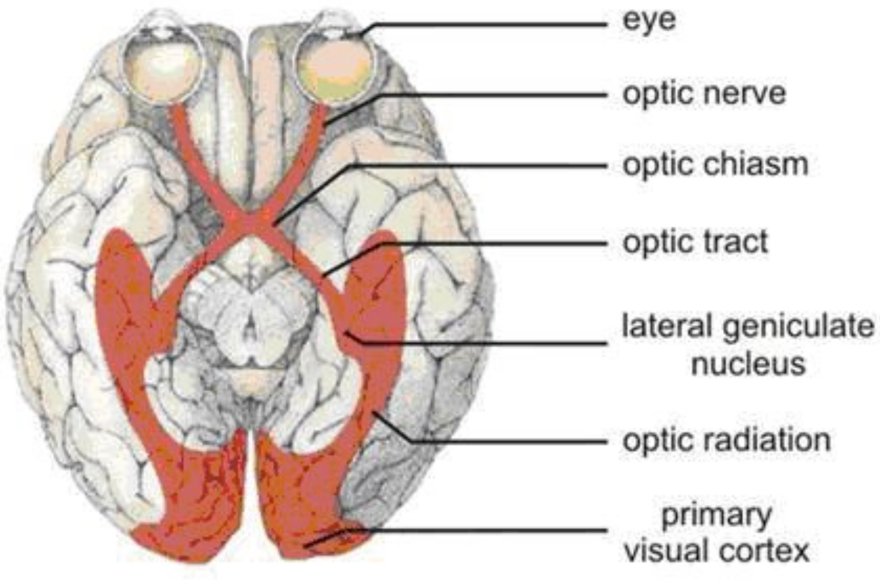
IMAGING CHAIN

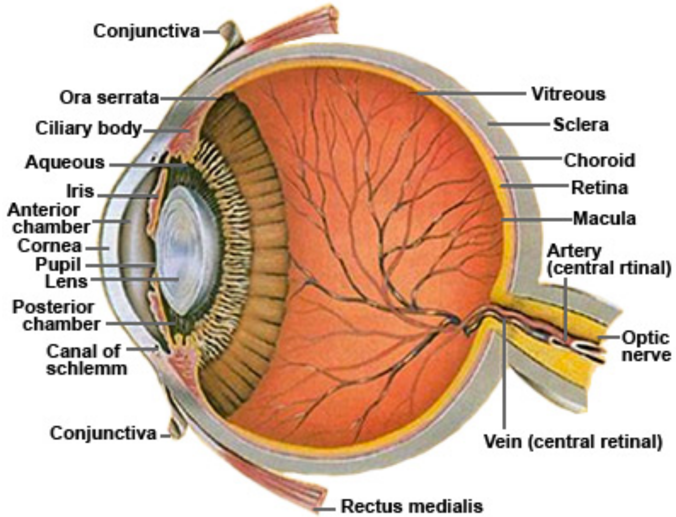
Overview of the chapter

- Section 2.1 briefly summarizes the mechanics of the human visual system, including image formation in the eye and its capabilities for brightness adaptation and discrimination.
- Section 2.2 discusses light, other components of the electromagnetic spectrum, and their imaging characteristics.
- Section 2.3 discusses imaging sensors and how they are used to generate digital images.
- Section 2.4 introduces the concepts of uniform image sampling and gray-level quantization.
- Section 2.5 deals with some basic relationships between pixels that are used throughout the book.
- Section 2.6 defines the conditions for linear operations. As noted in that section, linear operators play a central role in the development of image processing techniques.

The human visual system

- How human and electronic imaging compare in terms of resolution and ability to adapt to changes in illumination are not only interesting, they also are important from a practical point of view
- Vision involves the whole brain!





The eye

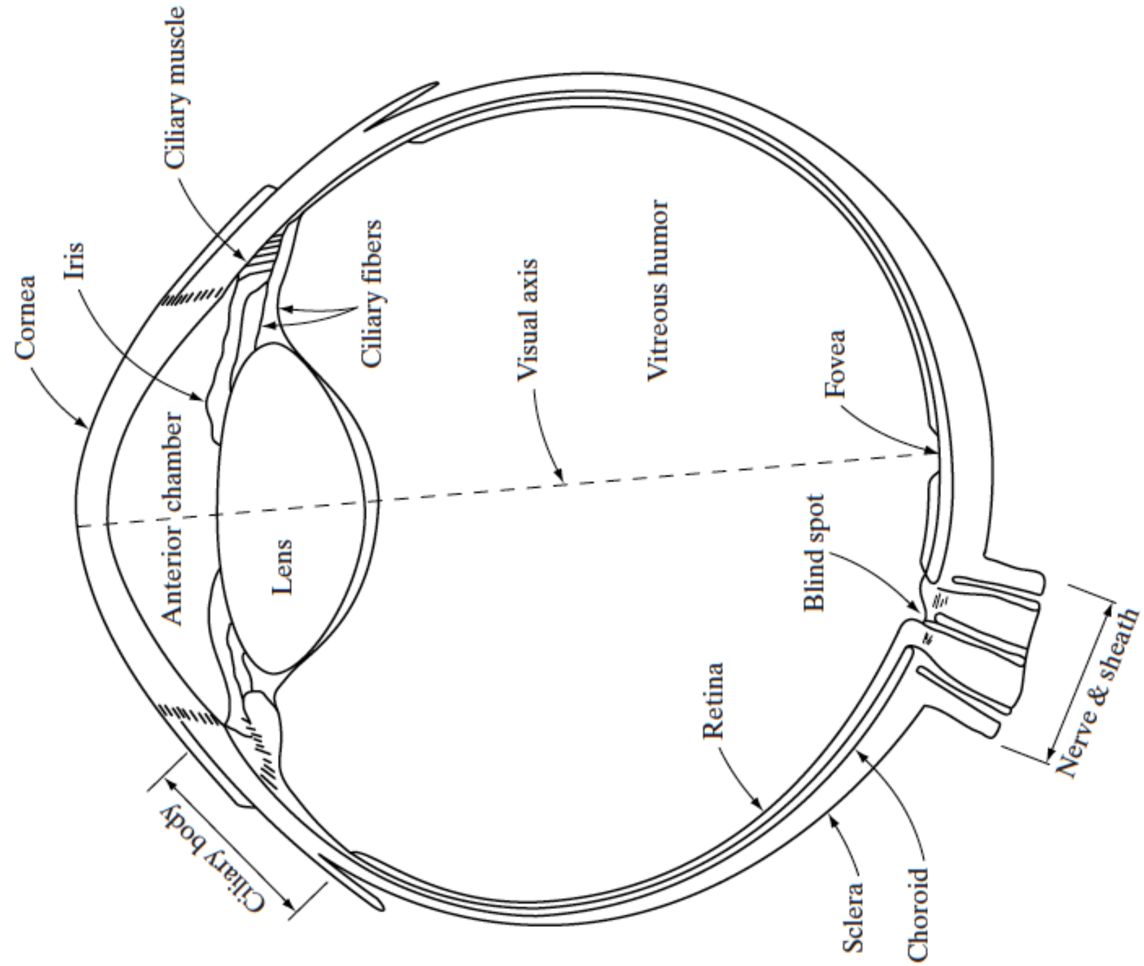
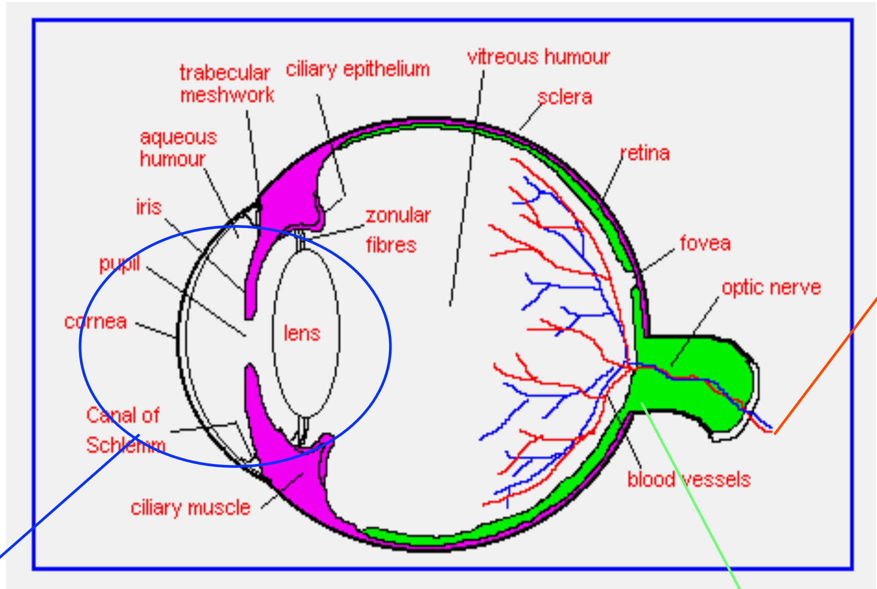

Ciliary muscle
Iris
Cornea
Ciliary fibers
Visual axis
Vitreous humor
Fovea
Anterior chamber
Lens
Blind spot
Retina
Nerve & sheath
Ciliary body
Sclera
Choroid

Image formation

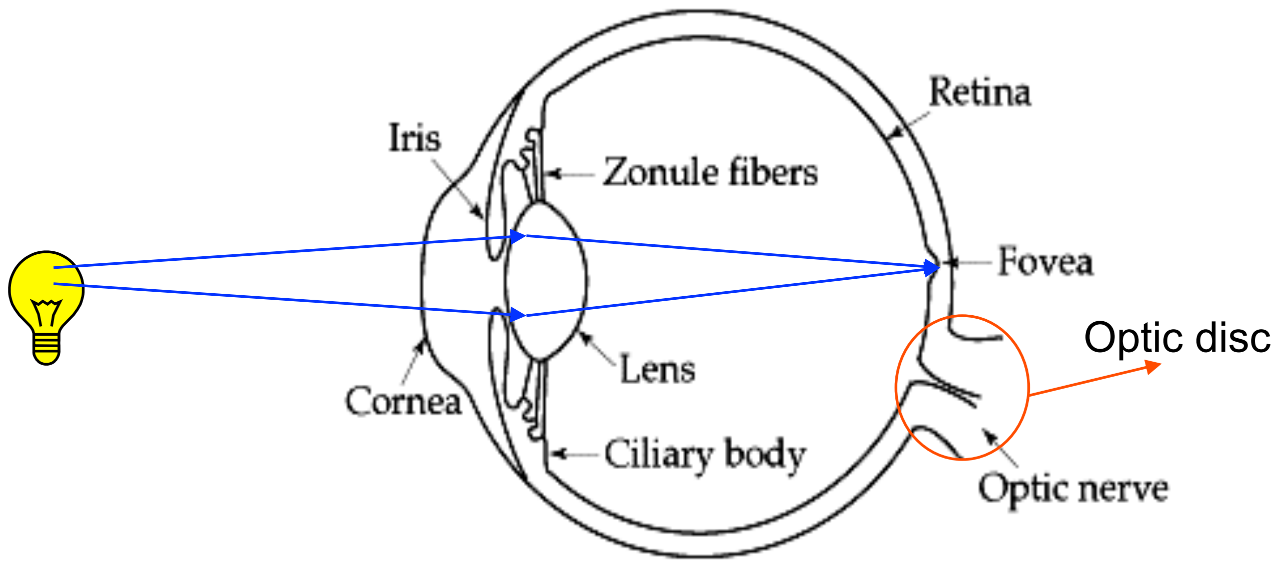


The neural responses are transformed into neural representations within the optic nerve which brings them to the brain to form other **cortical representations**

Cornea and lens focus the impinging light to the retina

The **photoreceptors** on the retina transpose the quanta into neural responses

Image formation



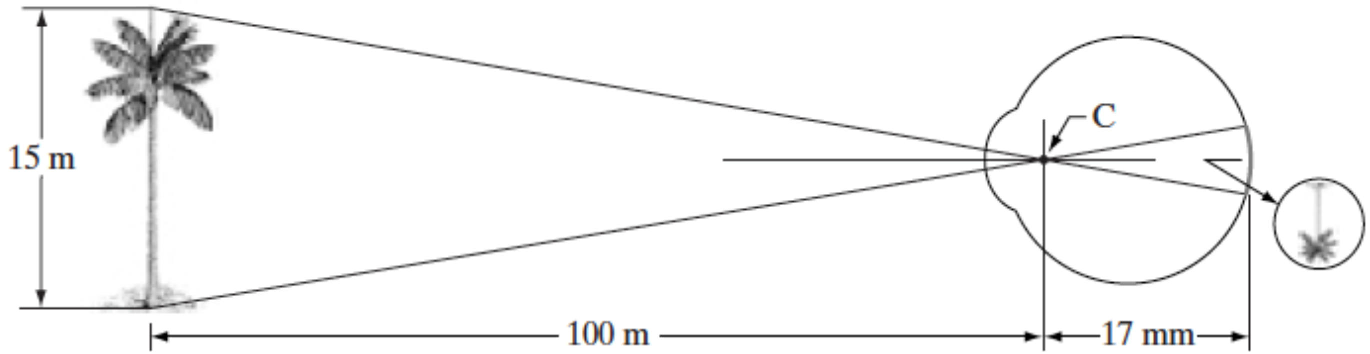
- The light entering the eye is brought to focus on the retina by the eye optics (cornea +lens)
  - The focus of the optical system must be kept on the retina in any condition
- The retina is a thin layer of neural tissue. It consists of different types of neurons. The axons of some of these are collected into the optic nerve.
- The optic nerve exits the retina at the *optic disc* to bring the signal to the brain for further processing

Image formation

- The distance between the center of the lens and the retina (called the focal length) varies from approximately 17 mm to about 14 mm
- The retinal image is reflected primarily in the area of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

h

FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.



15/100=h/17 or h=2.55 mm

The retina and eye optics

- The innermost membrane of the eye is the retina, which lines the inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina.
- There are two classes of receptors: cones and rods . The cones in each eye number between 6 and 7 million. They are located primarily in the central portion of the retina, called the fovea , and are highly sensitive to color. Humans can resolve fine details with these cones largely because each one is connected to its own nerve end.
- Muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea.

The retina and eye optics

- Cone vision is called photopic or bright-light vision.
- The number of rods is much larger: Some 75 to 150 million are distributed over the retinal surface. The larger area of distribution and the fact that *several rods are connected to a single nerve* end reduce the amount of detail discernible by these receptors.
- Rods serve to give a general, overall picture of the field of view. They are not involved in color vision and are sensitive to low levels of illumination. This phenomenon is known as scotopic or dim-light vision.
  - For example, objects that appear brightly colored in daylight when seen by moonlight appear as colorless forms because only the rod are stimulated.

Rods and cones

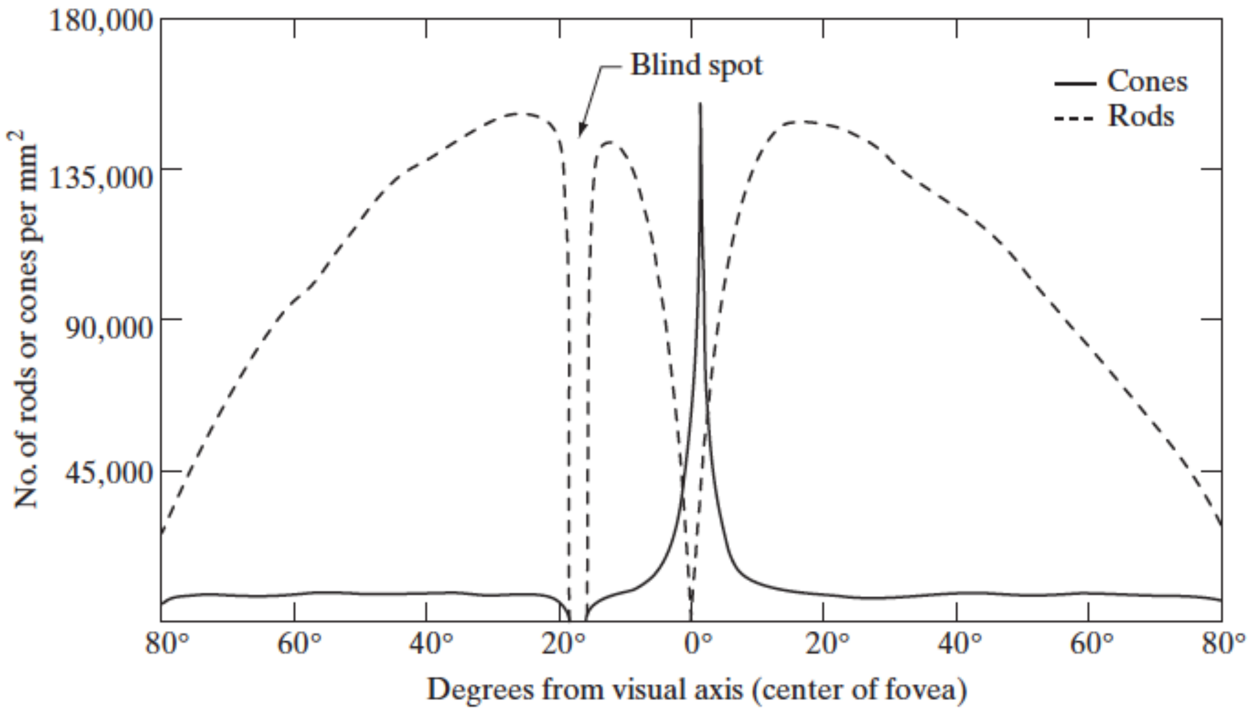


FIGURE 2.2
Distribution of rods and cones in the retina.

The fovea

- The fovea itself is a circular indentation in the retina of about 1.5 mm in diameter.
- By taking some liberty in interpretation, we can view the fovea as a square sensor array of size 1.5 mm*1.5 mm.
- The density of cones in that area of the retina is approximately 150,000 elements per mm2. Based on these approximations, the number of cones in the region of highest acuity in the eye is about 337,000 elements.
- Just in terms of raw resolving power, a charge-coupled device (CCD) imaging chip of medium resolution can have this number of elements in a receptor array no larger than 5 mm* 5 mm

Photoreceptor mosaic

- The retinal image is sampled by the photo-receptors of the retina
  – Discrete sampling grid

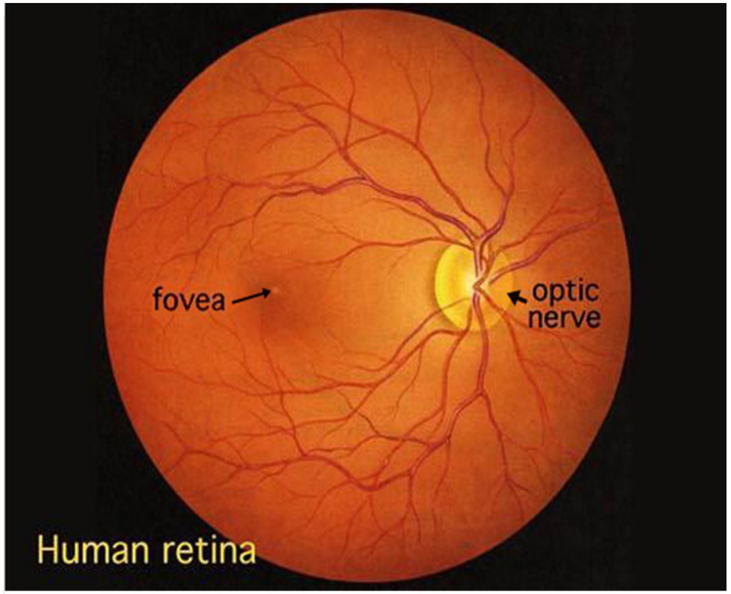


Fig. 1. Human retina as seen through an opthalmoscope.

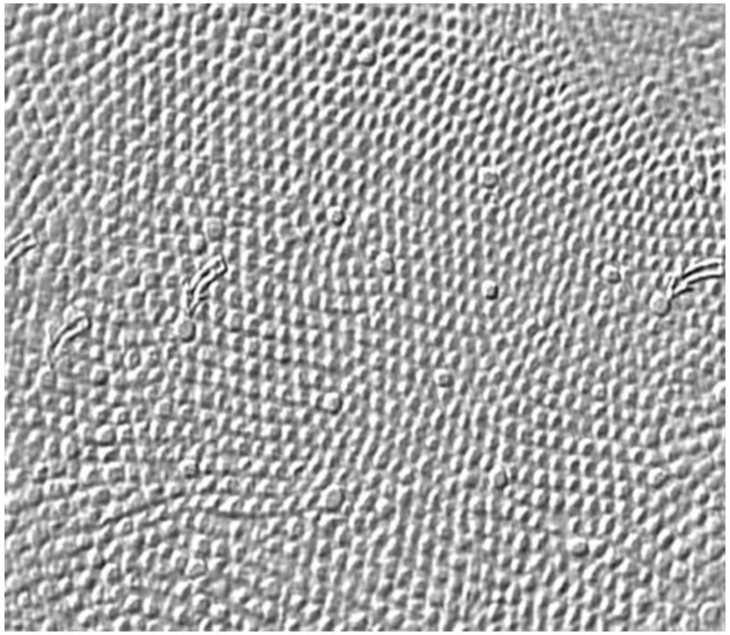
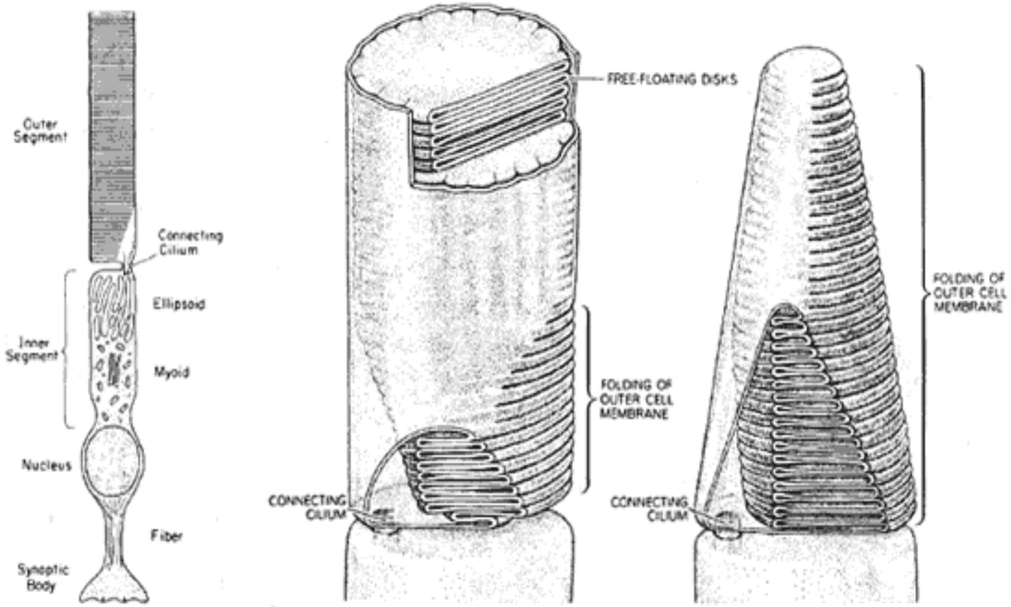


Fig. 13. Tangential section through the human fovea. Larger cones (arrows) are blue cones.

Cones and Rods shape



At the left is a generalized conception of the important structural features of a vertebrate photoreceptor cell. At the right are shown the differences between the structure of rod (left) and cone (right) outer segments. These diagrams are from Young (1970) and Young (1971).

Light and electromagnetic spectrum

- In 1666, Sir Isaac Newton discovered that when a beam of sunlight is passed through a glass prism, the emerging beam of light is not white but consists instead of a continuous spectrum of colors ranging from violet at one end to red at the other.
- The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy. Wavelength (λ) and frequency (v) are related by the expression $c=\lambda v$ where c is the speed of light (2.998*108 m s).
- The energy of the various components of the electromagnetic spectrum is given by the expressio **E=hv** where h is Planck' s constant.

The EM spectrum

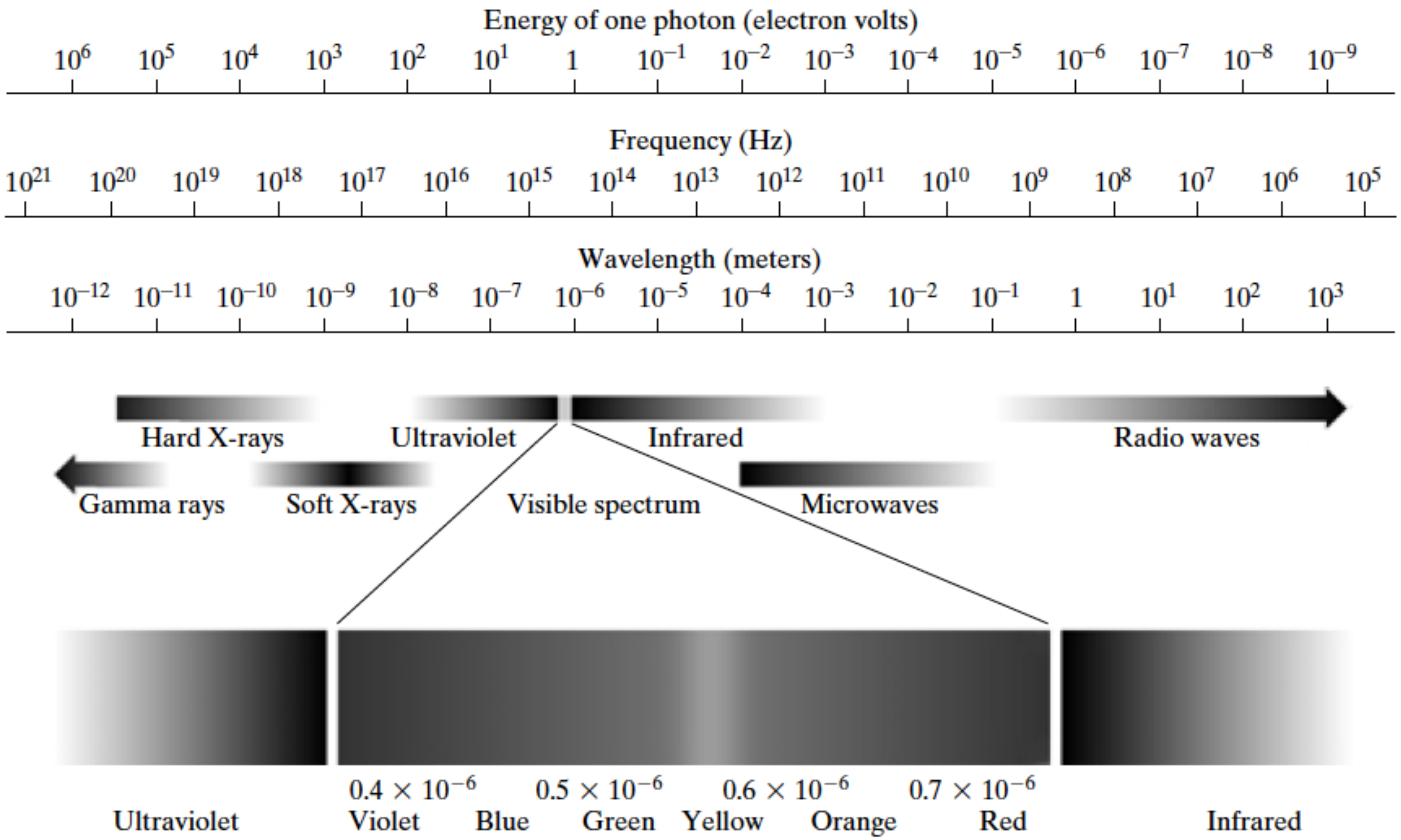
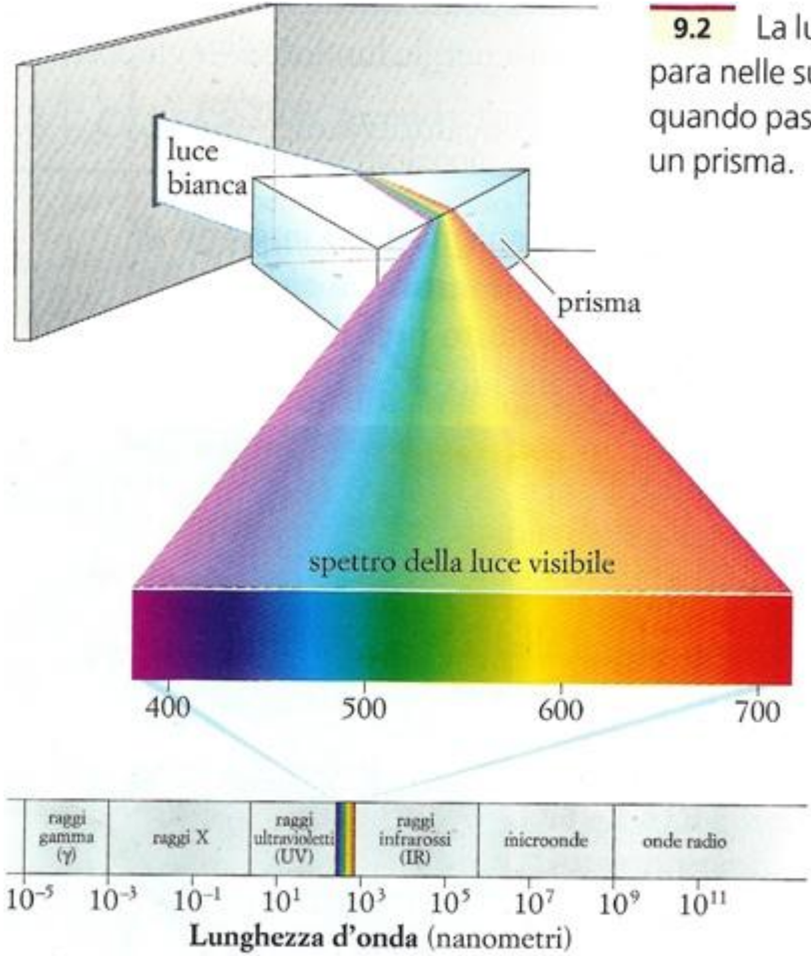


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Newton's prism

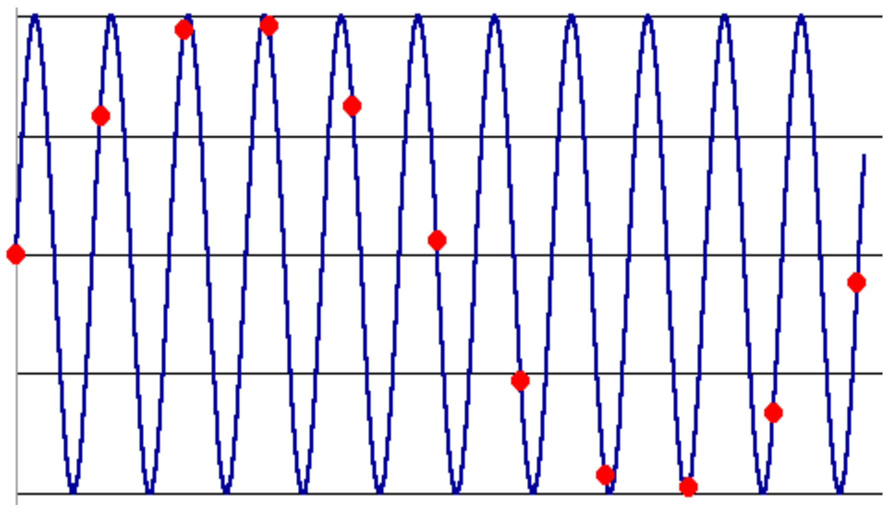


9.2 La luce bianca si separa nelle sue componenti quando passa attraverso un prisma.

EM units

- The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy. Wavelength **(λ)** and frequency **(v)** are related by the expression $c=\lambda v$ where **c** is the speed of light (**2.998*108** m s).
- The energy of the various components of the electromagnetic spectrum is given by the expression **E=hv** where h is Planck' s constant.
- The units of wavelength are meters, with the terms microns (denoted m and equal to $\mathbf{10^{-6}}$ m) and nanometers ($\mathbf{10^{-9}}$ m) being used just as frequently.
- Frequency is measured in Hertz (Hz), with one Hertz being equal to one cycle of a sinusoidal wave per second.
- A commonly used unit of energy is the electron-volt.

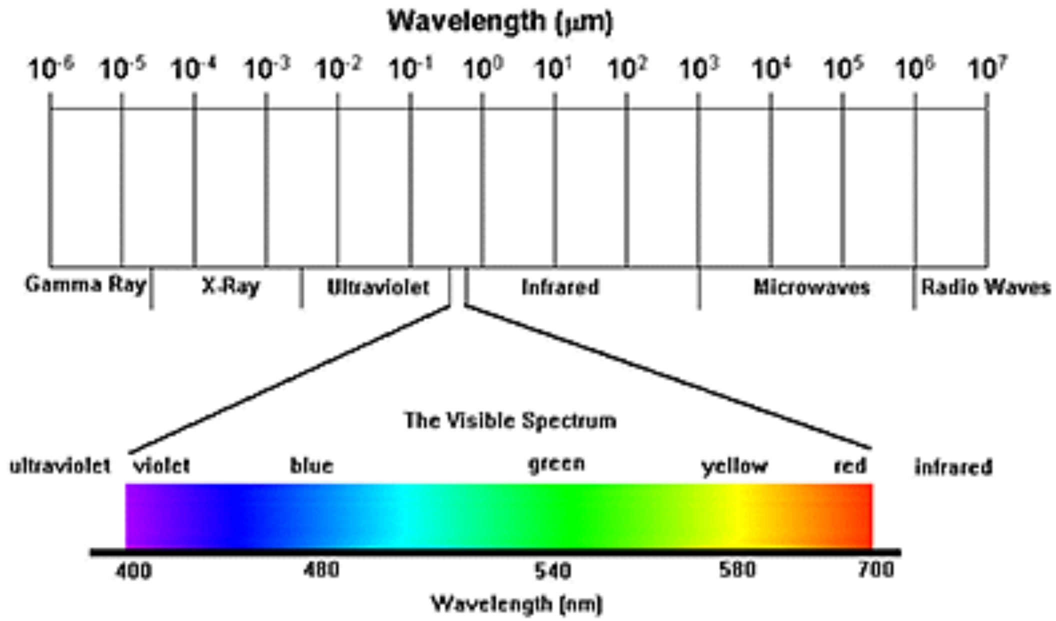
Sampling and Aliasing



Visible light

- Light is a particular type of electromagnetic radiation that can be seen and sensed by the human eye.
- The visible band of the electromagnetic spectrum spans the range from approximately 0.43 nm (violet) to about 0.79 nm (red).
- *For convenience*, the color spectrum is divided into six broad regions: violet, blue, green, yellow, orange, and red.
- Light that is void of color is called achromatic or monochromatic light. The only attribute of such light is its intensity , or amount. The term gray level generally is used to describe monochromatic intensity because it ranges from black, to grays, and finally to white.

Visible light



EM and perceptual units

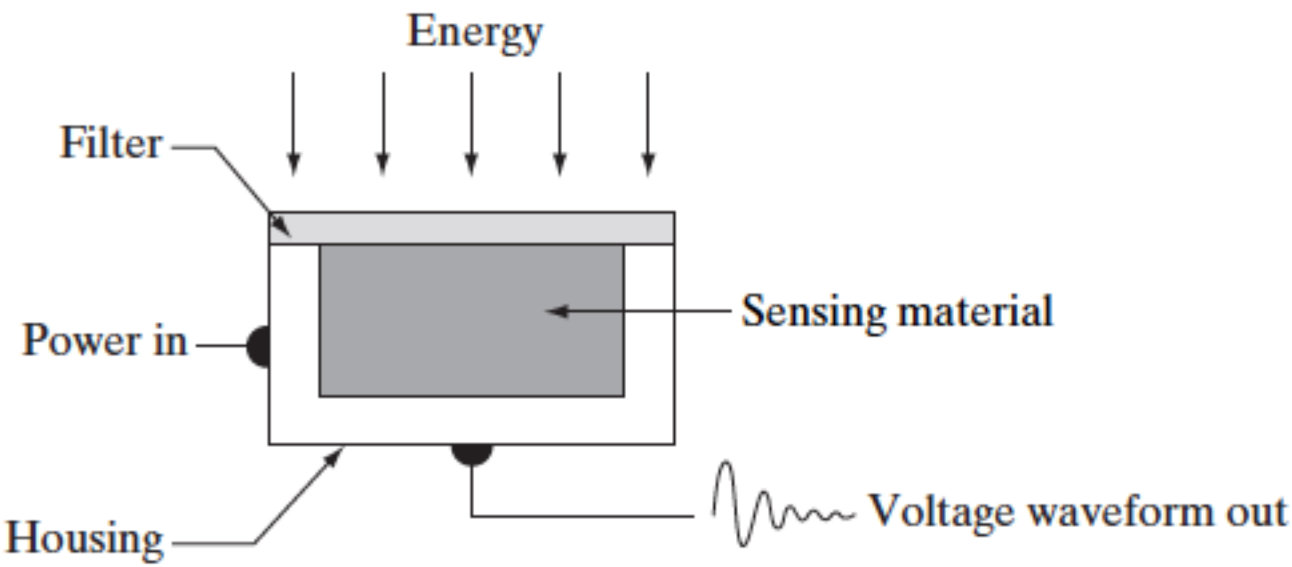
- **Radiance** is the total amount of energy that flows from the light source, and it is usually measured in watts (**W**).
- **Luminance**, measured in lumens (**lm**), gives a measure of the amount of energy **an observer perceives** from a light source.
  - For example, light emitted from a source operating in the far infrared region of the spectrum could have significant energy (radiance), but an observer would hardly perceive it; its luminance would be almost zero.
- **Brightness** is a subjective descriptor of light perception that is practically impossible to measure using a physical detector.

Image sensing and acquisition

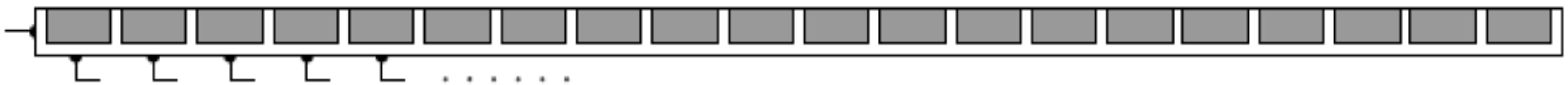
Image sensing and acquisition

- The types of images in which we are interested are generated by the combination of an "illumination" source and the reflection or absorption of energy from that source by the elements of the "scene" being imaged.
- Idea: the incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected. The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response.

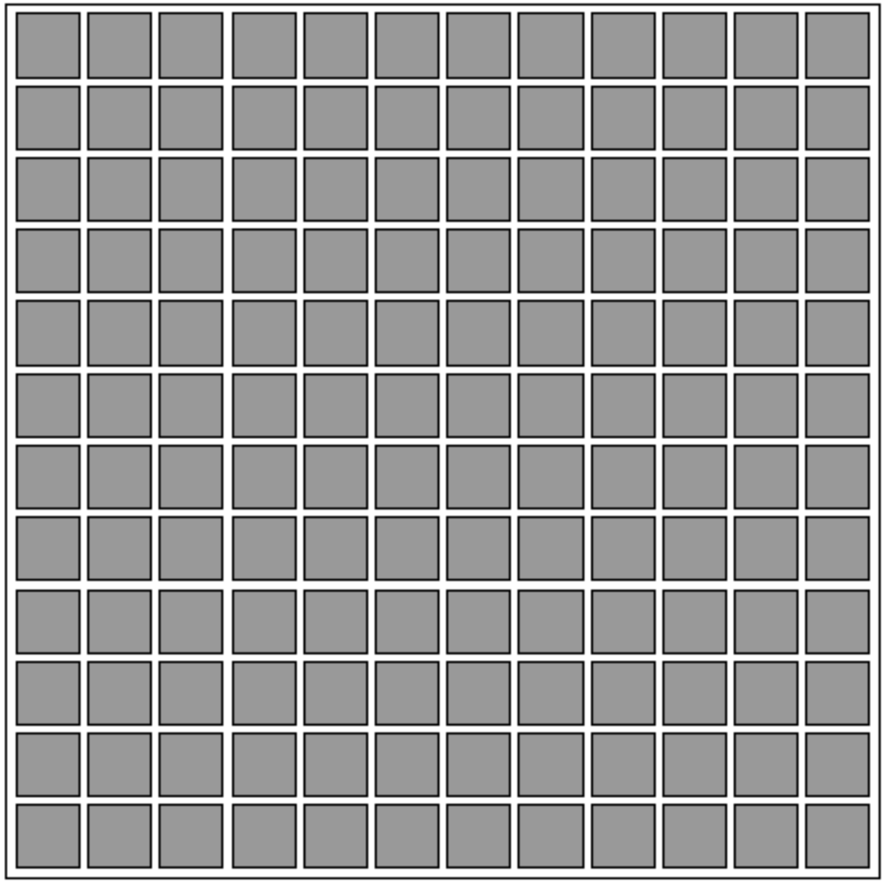
Imaging sensors



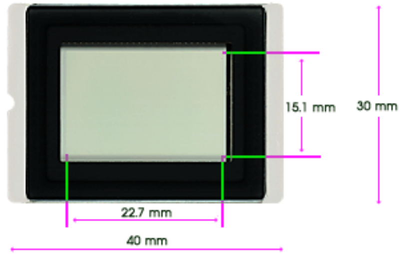
Line sensors

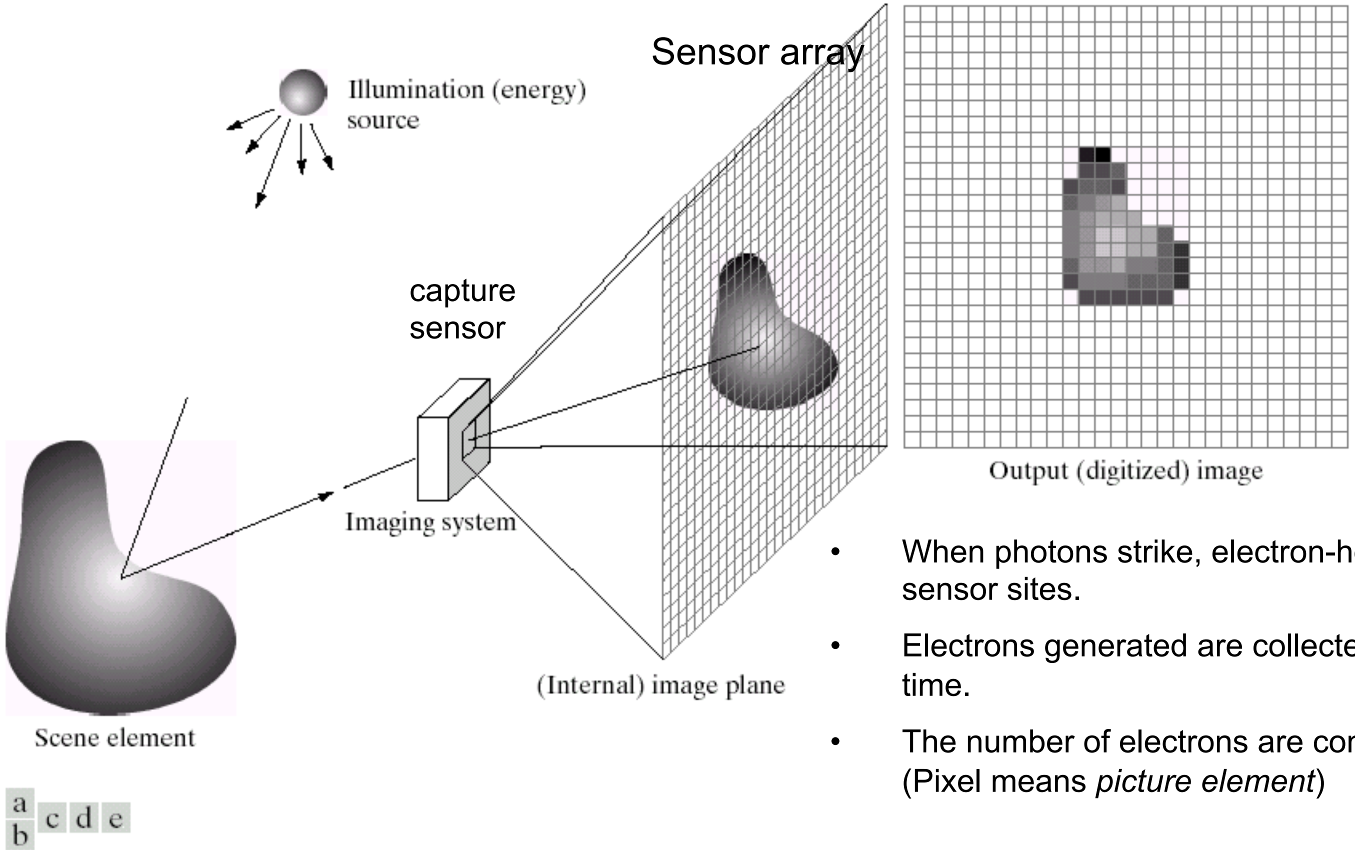


Array sensors



Digital Image Acquisition





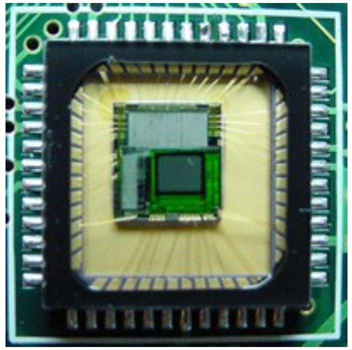


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

- When photons strike, electron-hole pairs are generated on sensor sites.
- Electrons generated are collected over a certain period of time.
- The number of electrons are converted to pixel values. (Pixel means *picture element*)

Image acquisition using sensor arrays

- The first function performed by the imaging system is to collect the incoming energy and focus it onto an image plane. If the illumination is light, the front end of the imaging system is a lens, which projects the viewed scene onto the lens focal plane.
- The sensor array, which is coincident with the focal plane, produces outputs proportional to the integral of the light received at each sensor.
- Digital and analog circuitry sweep these outputs and convert them to a (video) signal, which is then processed by another section of the imaging system.
- The output is a digital image.

Image formation model

- Image: two dimensional light intensity function denoted by $f(x,y)$ where the value of the amplitude f at the spatial coordinates (x,y) gives the intensity (luminance) of the graylevel image at point (x,y).
- As light is an electromagnetic wave, the luminance must be finite

$$0<f(x,y)<+\infty$$

- $f(x,y)$ results from the reflection of the light incident on a surface (illumination, $i(x,y)$) and can be written as

  $f(x,y)=i(x,y)*r(x,y)$

  Where $r(x,y)$ denotes the reflectance

  $0<i(x,y)<+\infty$

  $0<r(x,y)<1$: $r=0$ means total absorption and $r=1$ means total reflection

Image formation model

- We call the intensity of a monochrome image at any coordinates (x_0, y_0) the gray level (*l*) of the image at that point: $f(x_0, y_0)=l$
- l lies in the range $l_{min} \leq l \leq l_{max}$
- In theory, the only requirement on l_{min} is that it be positive, and on l_{max} that it be finite. In practice, $l_{min}=i_{min}r_{min}$ and $L_{max}=i_{max}r_{max}$
- The interval is called the gray scale. Common practice is to shift this interval numerically to the interval [0, L-1] , where l=0 is considered black and l=L-1 is considered white on the gray scale. All intermediate values are shades of gray varying from black to white.

Image sampling and quantization

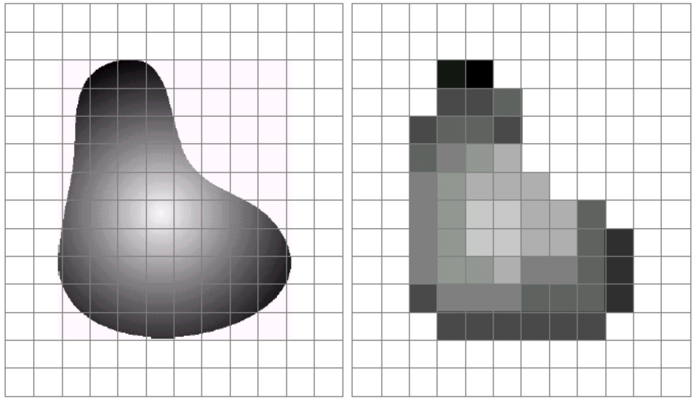


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Two types of discretization:

- There are finite number of pixels
  - Sampling → Spatial resolution
- The amplitude of pixel is represented by a finite number of bits
  - Quantization → Gray-scale resolution

Digital Image Acquisition

Take a look at this cross section

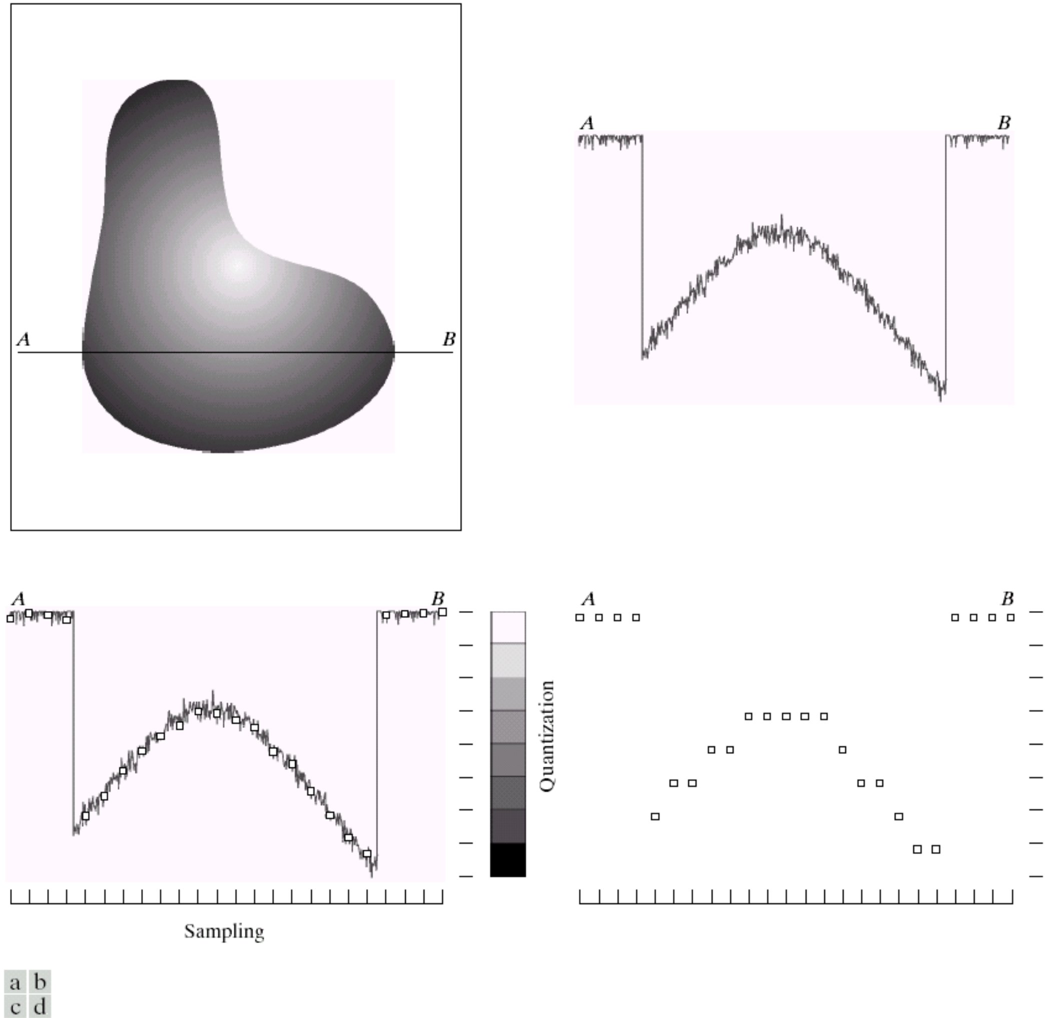


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Image sampling and quantization

- To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: **sampling and quantization**.
- Figure 2.16(a) shows a continuous image, f(x, y) , that we want to convert to digital form.
- An image may be continuous with respect to the x - and y – coordinates, and also in amplitude.
- To convert it to digital form, we have to sample the function in both coordinates and in amplitude.
- Digitizing the coordinate values is called sampling
- Digitizing the amplitude values is called quantization

Sampling and quantization

- In order to form a digital function, the gray-level values must be converted (quantized) into discrete quantities
- Sampling means that the values of the continuous function f(x,y) are retained only in specific positions (i,j) where $0 \leq i \leq N_x$ and $0 \leq j \leq N_y$, where N_x and N_y are integer values. The sampling topology depends on the spatial arrangement and size of the sensors that are used to acquire the image.
- Clearly, the quality of a digital image is determined to a large degree by the number of samples and discrete gray levels used in sampling and quantization.

Sampling exemple

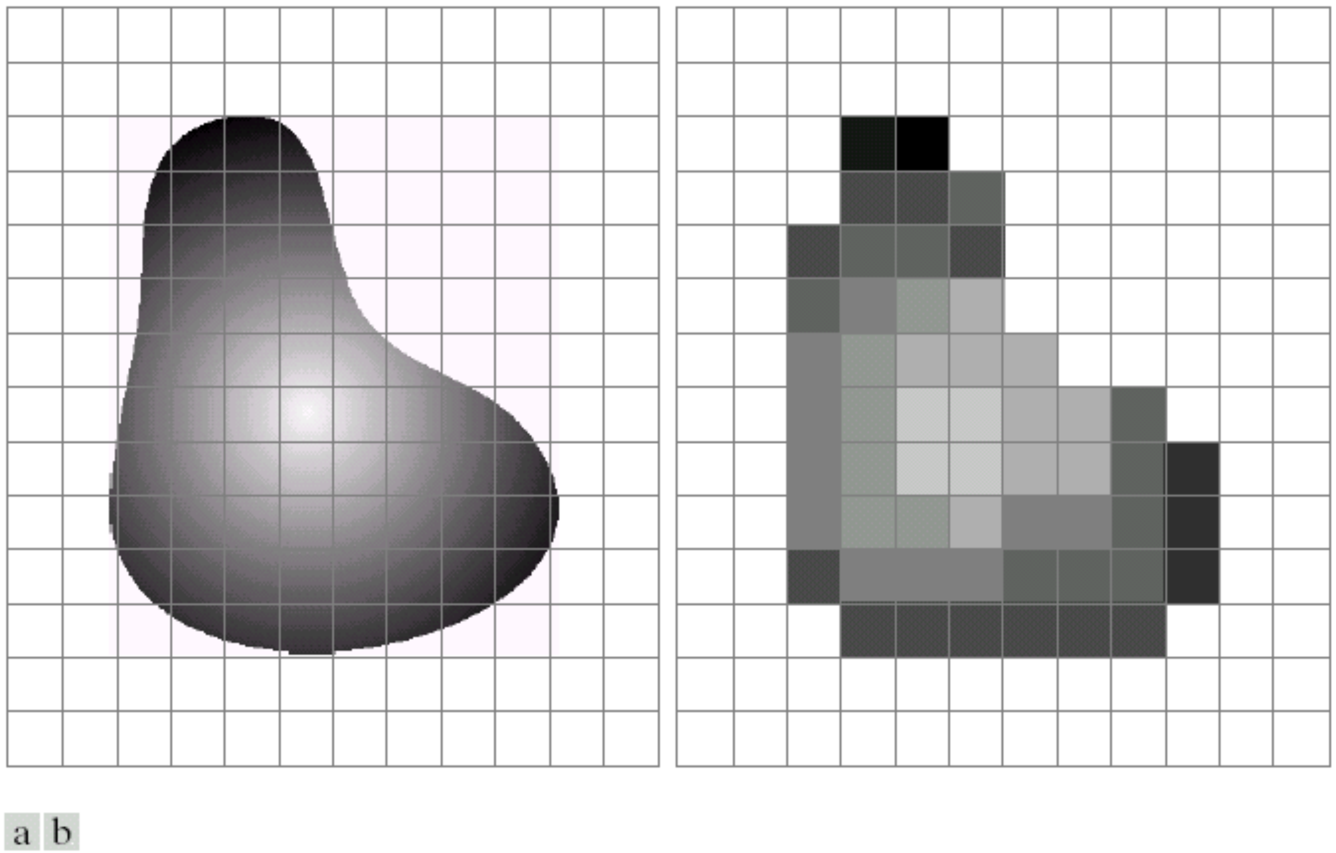
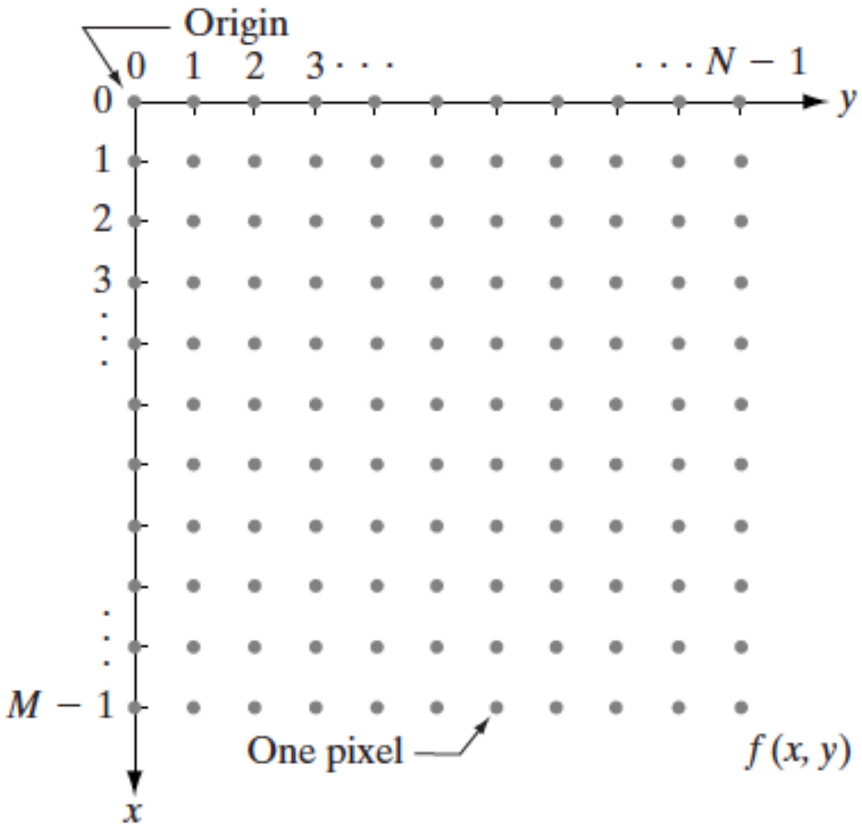


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Representing digital images



Digital images are matrices

- A digital image can be represented in matrix form.
  - Each element of this matrix array is called an image element, picture element, *pixel* , or *pel* .

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix}$$

Mathematical formulation

- Let Z and R denote the set of real integers and the set of real numbers, respectively. The sampling process may be viewed as partitioning the (x,y) plane into a grid, with the coordinates of the center of each grid being a pair of elements from the Cartesian product Z^2, which is the set of all ordered pairs of elements **(**z_i, z_j), with z_i and z_j being integers from Z.
- Hence, f(x, y) is a digital image if (x, y) are integers from Z^2 and f is a function that assigns a gray-level value (that is, a real number from the set of real numbers, R) to each distinct pair of coordinates (x, y).
  - If the gray levels also are integers (as usually is the case in this and subsequent chapters), Z replaces R, and a digital image then becomes a 2-D function whose coordinates and amplitude values are integers.

Choice of the values

- The number of gray levels is chosen to be a power of 2 for practical reasons: $L=2^n$, which generates grayvalues ranging from $I_{min}=0$ to $I_{max}=2^n-1$
  - We assume that the discrete levels are equally spaced and that they are integers in the interval **[0, L-1]** .
  - Sometimes the range of values spanned by the gray scale is called the *dynamic range* of an image
- The number of sampling points N, M is set by the sensor array.
- The number, **b**, of bits required to store a digitized image is $b=N*M*n$

A bit more on sampling and quantization

This is not covered by the GW chapter

Imparare i concetti

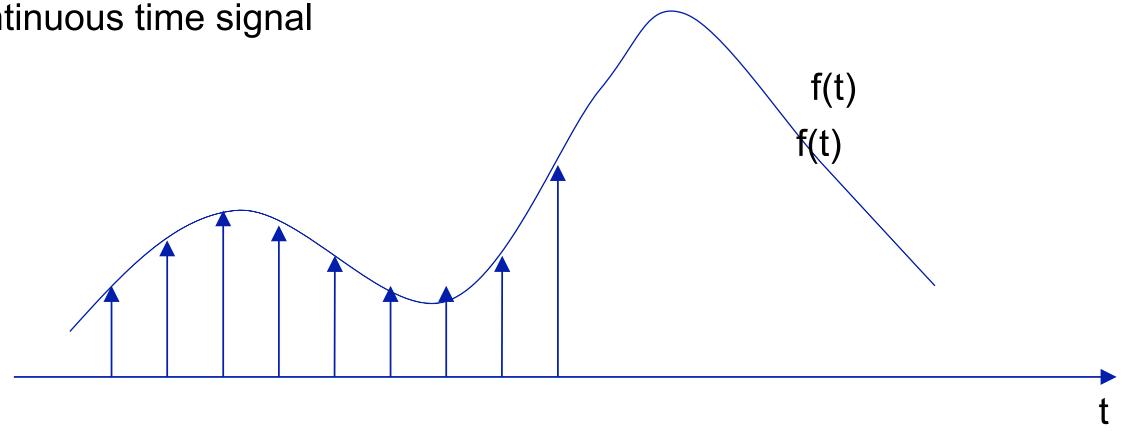
Le formule sono lasciate alla buona volontà…

Sampling in 1D

$$f[k] = f(kT_s) = f(t)\sum_k \delta(t - kT_s)$$

pettine

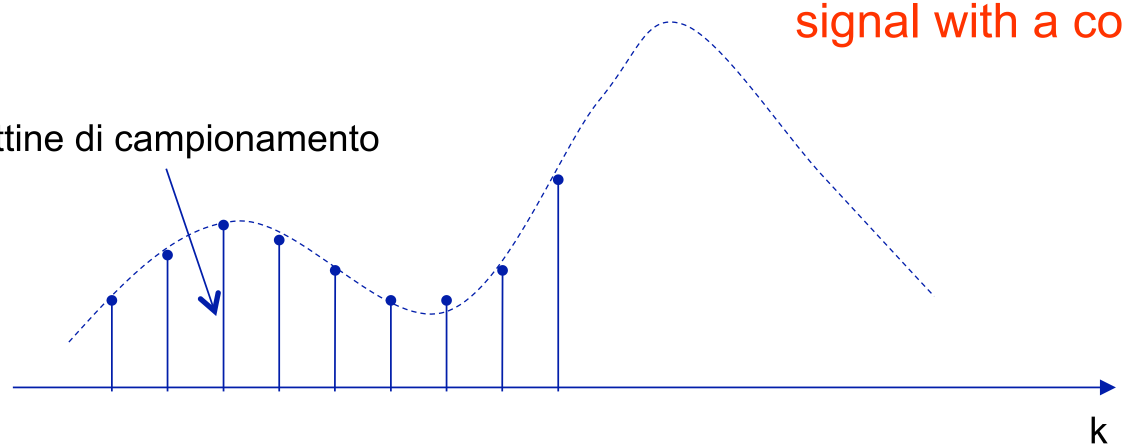
Continuous time signal



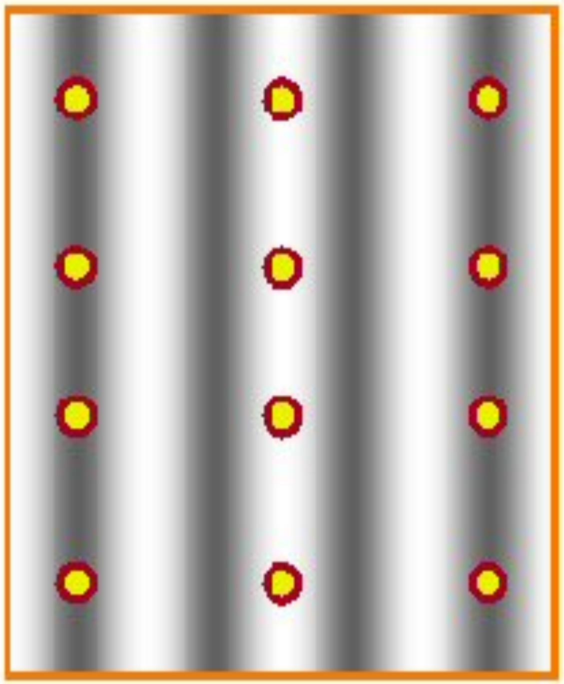
A sampled signal can be seen as the product of a continuos time signal with a comb.

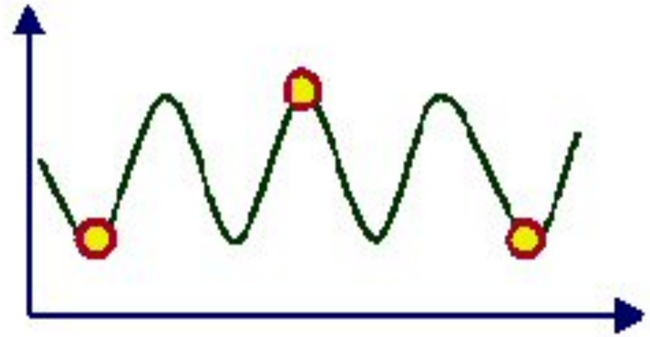
Discrete time signal

Pettine di campionamento



Nyquist theorem (1D)





At least 2 sample/period are needed to represent a sinusoid

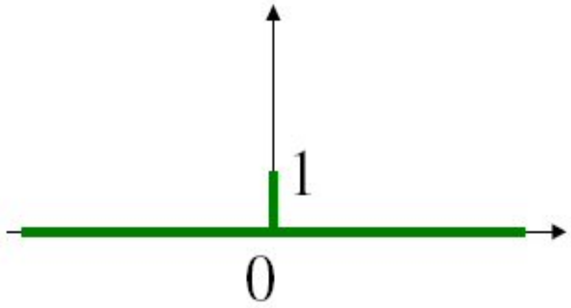
$$T_s \le \frac{1}{2}\frac{2\pi}{\omega_{\max}}$$

$$\omega_s = \frac{2\pi}{T_s} \ge 2\omega_{\max}$$

Delta pulse

- 1D Dirac pulse

$$\begin{cases} \delta(x) = 1 \text{ if } x=0 \\ \delta(x) = 0 \text{ else} \end{cases}$$

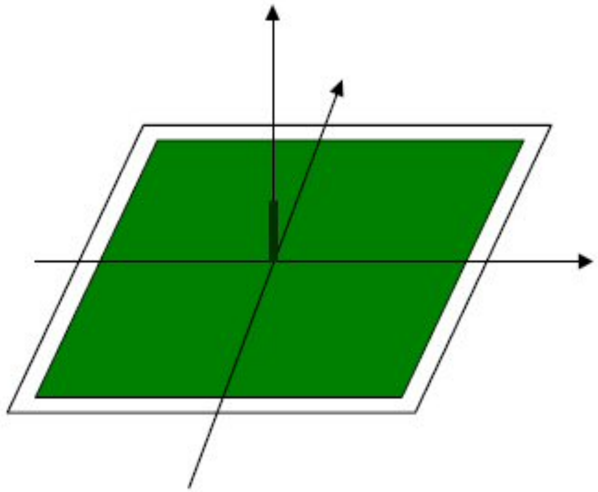


- 2D Dirac pulse

$$\begin{cases} \delta(x,y) = 1 \text{ if } x=0 \text{ and } y=0 \\ \delta(x,y) = 0 \text{ else} \end{cases}$$

which corresponds to :

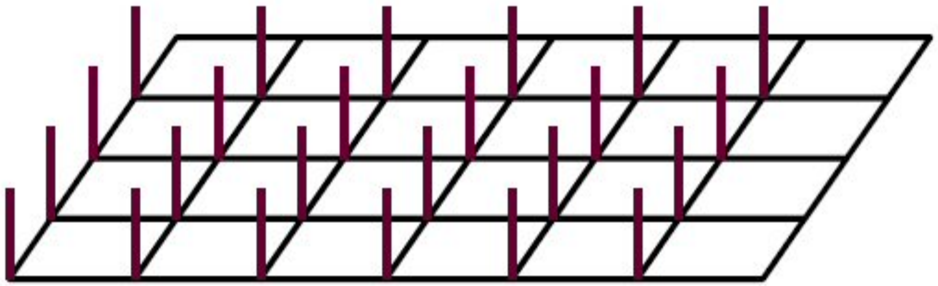
$$\delta(x,y) = \delta(x)\,\delta(y)$$



Dirac *brush*

- 1D sampling: Dirac comb (or Shah function)

- 2D sampling : Dirac « brush »

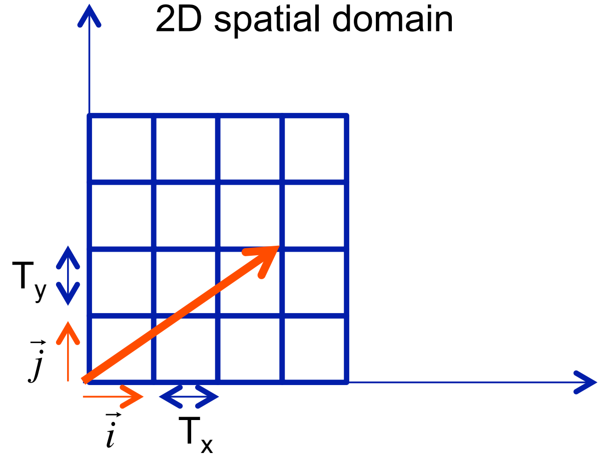


Nyquist theorem

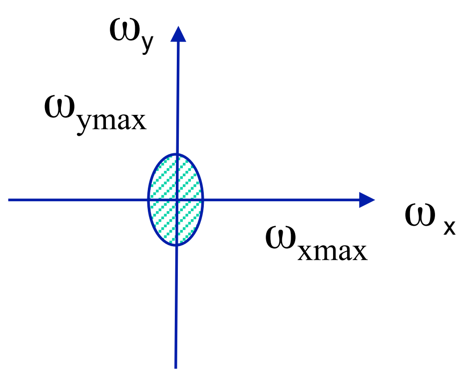
- Sampling in p-dimensions

$$s_T(\vec{x}) = \sum_{k \in Z^p} \delta(\vec{x} - (k_x T_x \vec{i} + k_y T_y \vec{j}))$$

$$f_T(\vec{x}) = f(\vec{x}) s_T(\vec{x})$$



2D Fourier domain

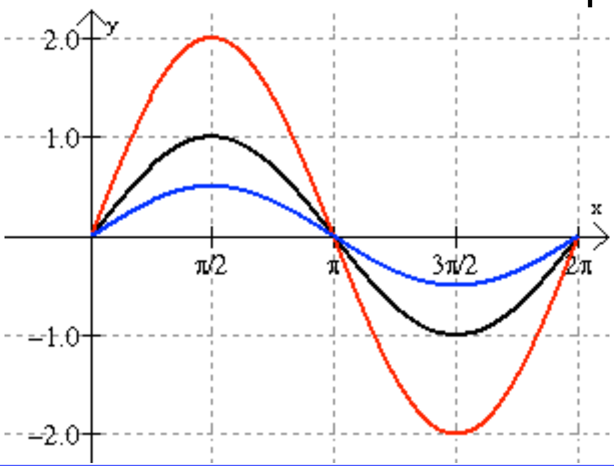


- Nyquist theorem

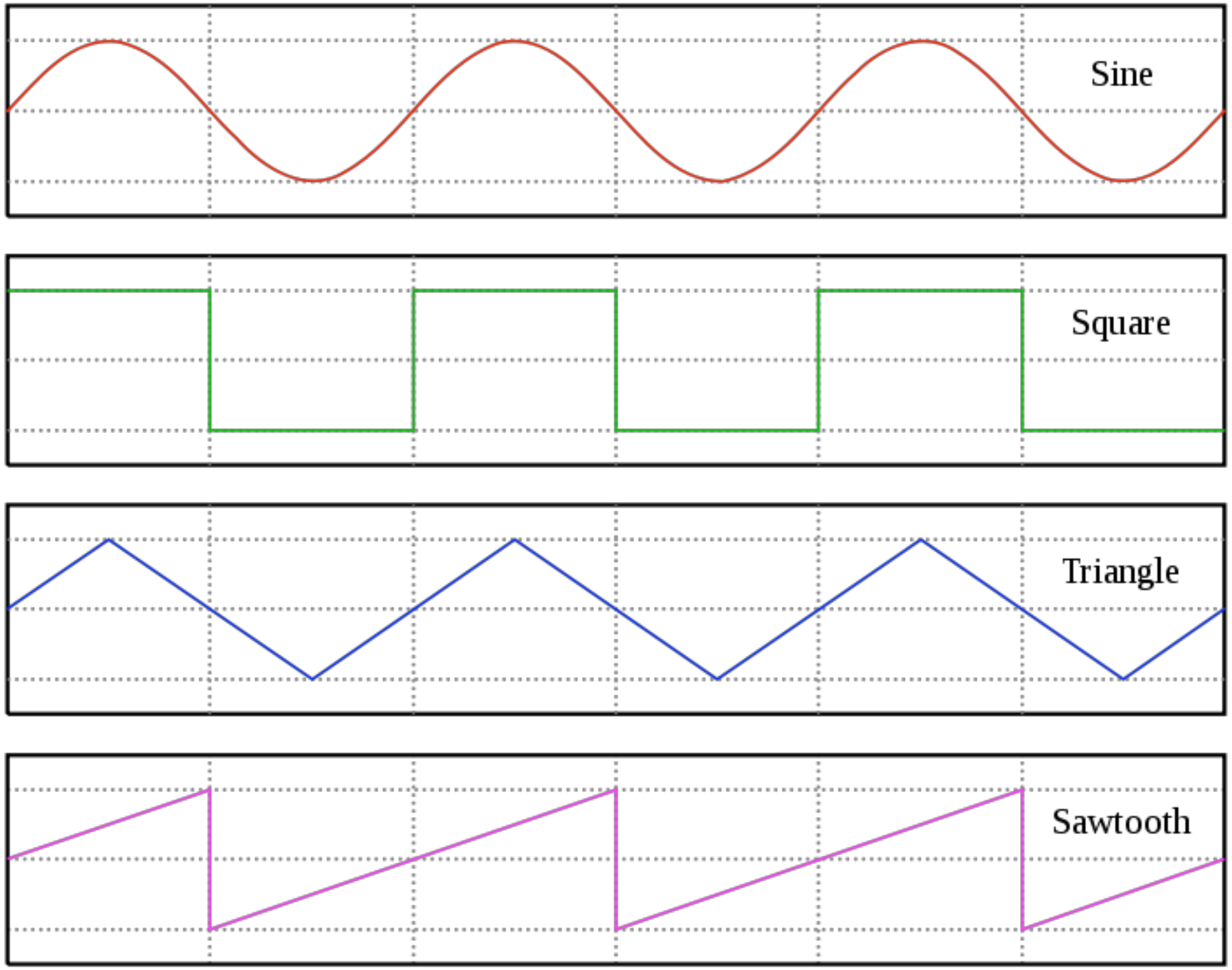
$$\begin{cases} \omega_x^s \geq 2\omega_{x\max} \\ \omega_y^s \geq 2\omega_{y\max} \end{cases} \Rightarrow \begin{cases} T_x^s \leq 2\pi \dfrac{1}{2\omega_{x\max}} \\ T_y^s \leq 2\pi \dfrac{1}{2\omega_{y\max}} \end{cases}$$

Nyquist (sampling) theorem

- At least two samples/period are needed to represent a sinusoid
- The number of samples per period that are needed to represent a generic periodic signal depend on how fast the signal changes within the period
  - Core concept: signal bandwidth - the larger the bandwidth, the faster the signal changes, the more samples are needed to represent it
  - Example: for a period=T, the sinusoid only needs 2 samples, while the box function would require infinite samples



Examples

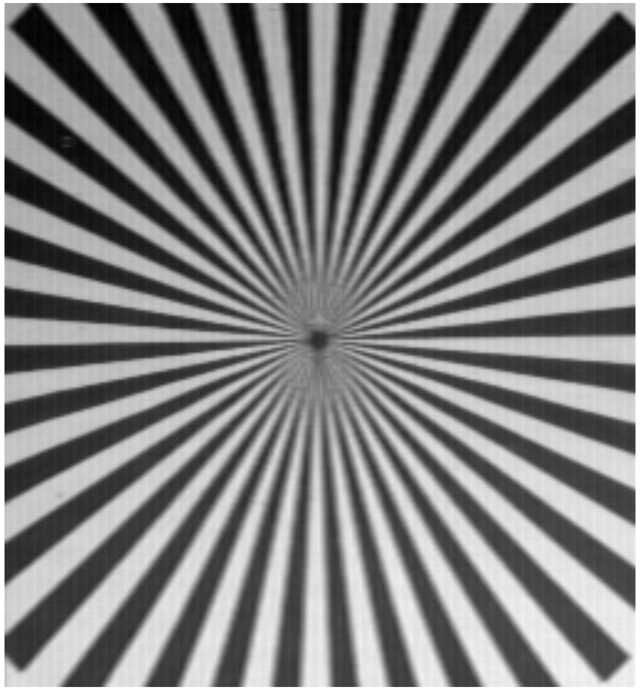


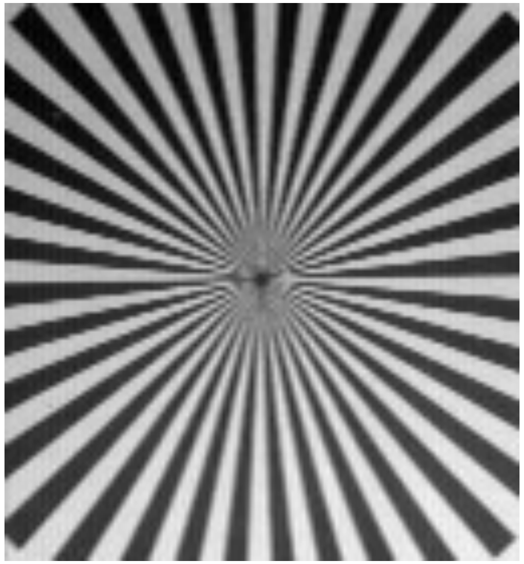
Bandwidth=1/period
2 samples/period
No error using 2 samples/period

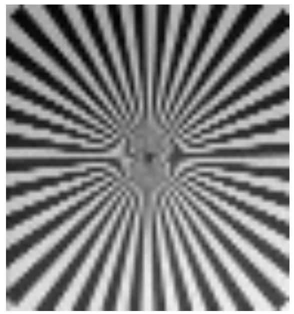
Bandwidth=infinity
infinite samples/period
If a finite number of samples per period is used, a very large error affects the signal representation

Bandwidth=infinity
infinite samples/period
BUT less error if the signal is represented by a finite number of samples per period

Spatial aliasing





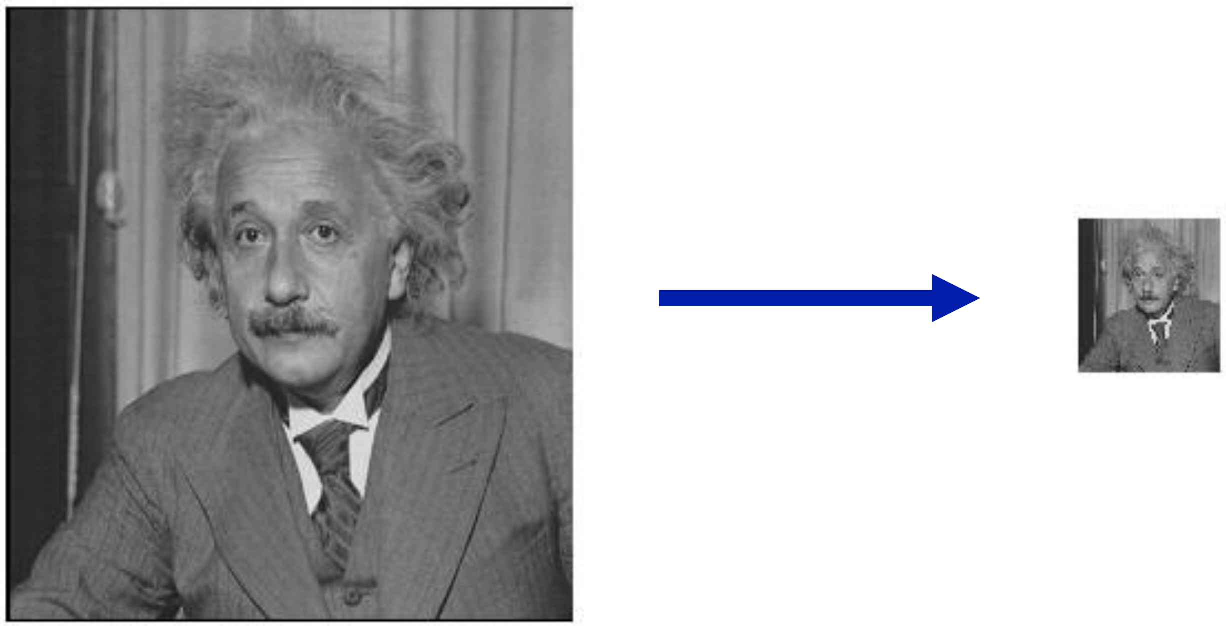


If the sampling criterion is not respected, some defects appear in the image. These are due to the fact that rapid signal changes cannot be correctly reproduced.

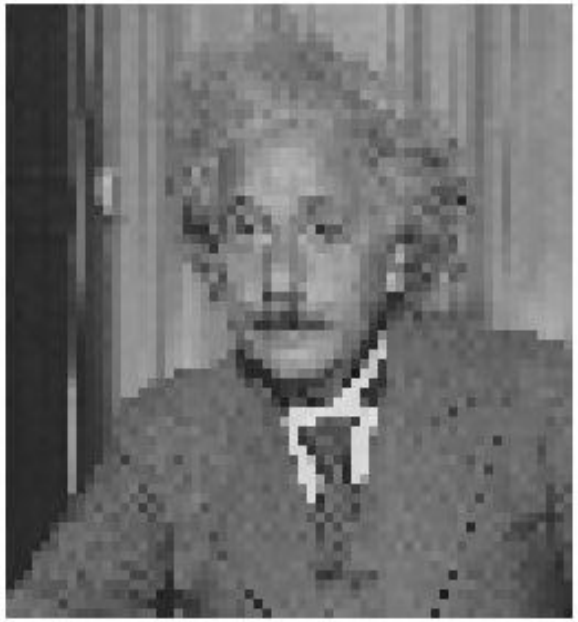
Resampling

- How to zoom and shrink a digital image
  - This topic is related to image sampling and quantization because **zooming** may be viewed as **oversampling**, while **shrinking** may be viewed as **undersampling**
- Change of the sampling rate
  - Increase of sampling rate: Interpolation or upsampling
    - **Blurring**, low visual resolution
  - Decrease of sampling rate: Rate reduction or downsampling
    - **Aliasing** and/or loss of spatial details

Downsampling

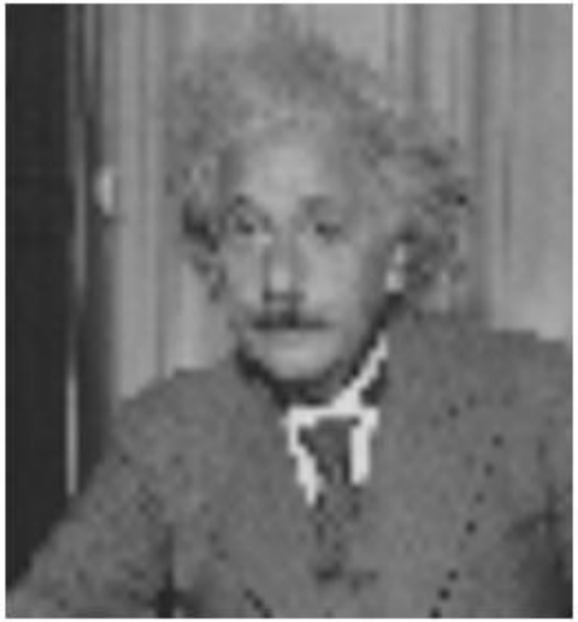


Upsampling or Interpolation



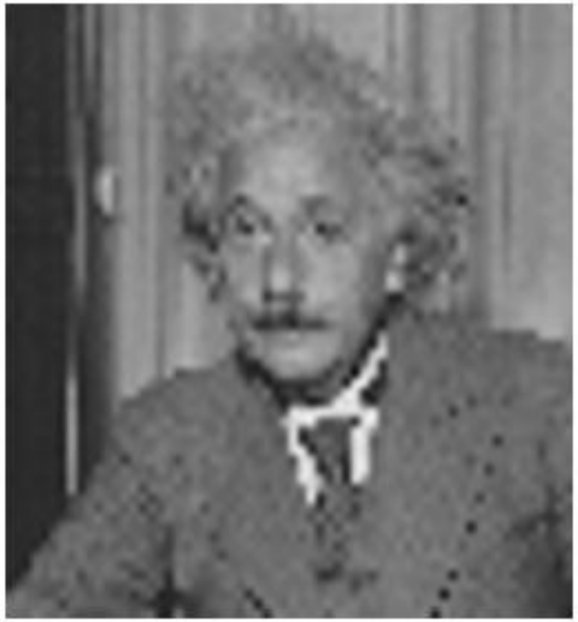
nearest neighbor (NN)

Upsampling or Interpolation



bilinear

Upsampling or Interpolation



bicubic

A hint to interpolation: zooming

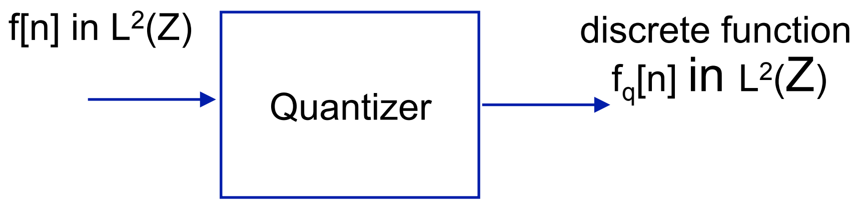
- Zooming requires two steps: the creation of new pixel locations, and the assignment of gray levels to those new locations.
  - Example: Suppose that we have an image of size 500*500 pixels and we want to enlarge it 1.5 times to 750*750 pixels. Conceptually, one of the easiest ways to visualize zooming is laying an imaginary 750*750 grid over the original image. Obviously, the spacing in the grid would be less than one pixel because we are fitting it over a smaller image. In order to perform gray-level assignment for any point in the overlay, we look for the closest pixel in the original image and assign its gray level to the new pixel in the grid. When we are done with all points in the overlay grid, we simply expand it to the original specified size to obtain the zoomed image.
  - This method of gray-level assignment is called **nearest neighbor interpolation**

Quantization

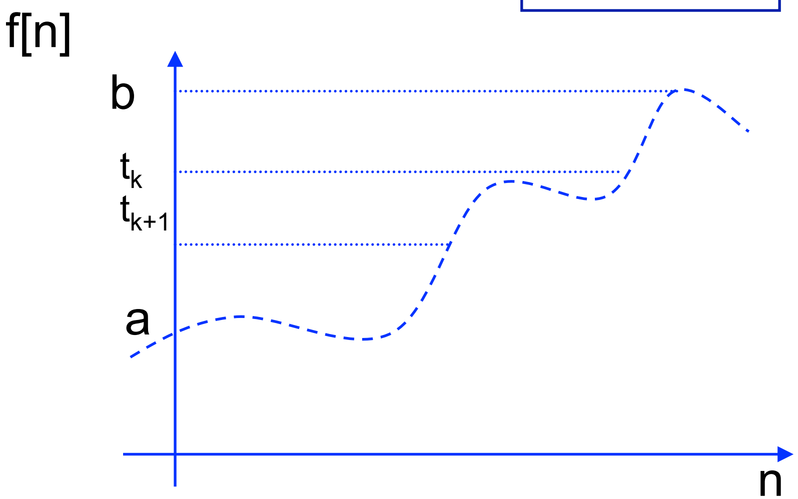
Quantization

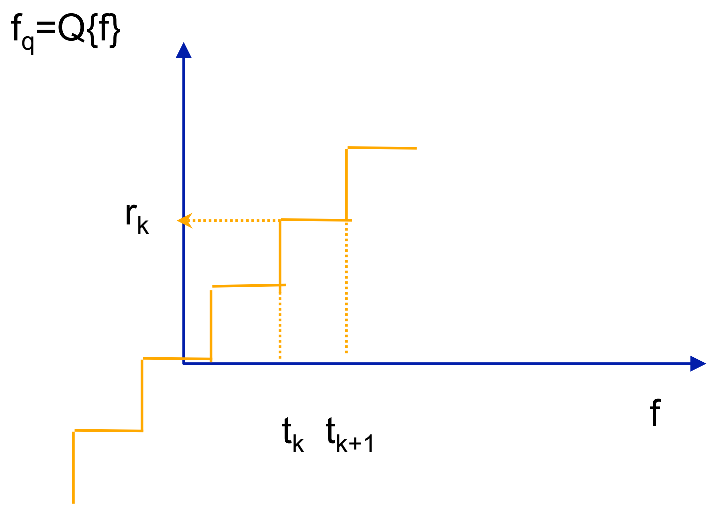
- A/D conversion ⇨ quantization

f: input function (sampled)
$F_{q=}Q\{f\}$: output function (sampled and quantized)



Uniform: all the quantization bins are equal



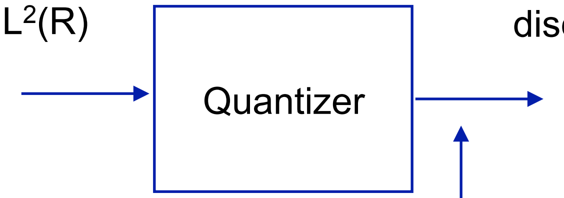


In the plot a continuous (non sampled) function is used for easyness of representation

Quantization

- A/D conversion ⇨ quantization

f in $L^2(R)$

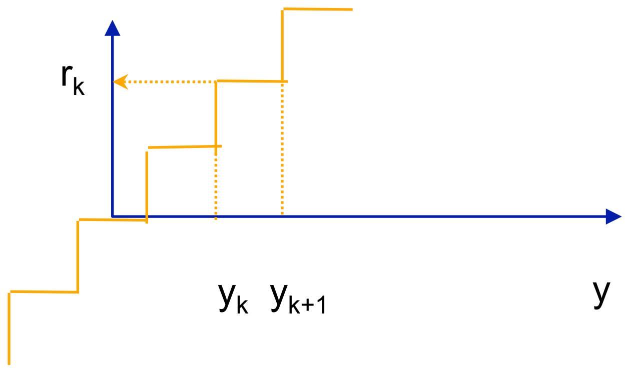


discrete function
f in $L^2(Z)$

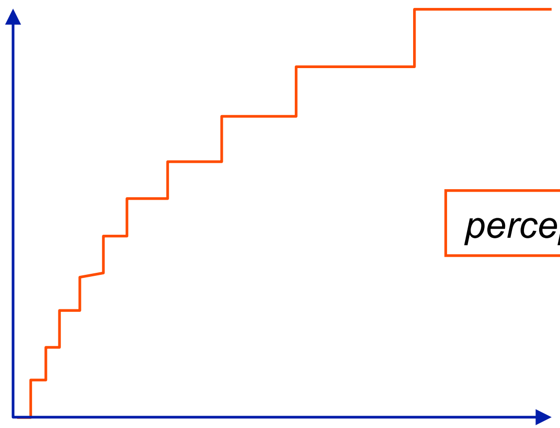
TRANSFER FUNCTION

uniform

X=Q{y}



f_q=Q{f}



perceptual

The quantization bin is not the same over the horizontal axis

Scalar quantization

- A scalar quantizer *Q* approximates *f* by $f_q = Q(f)$, which takes its values over a finite set.
- The quantization operation can be characterized by the mean square error (MSE) between the original and the quantized signals

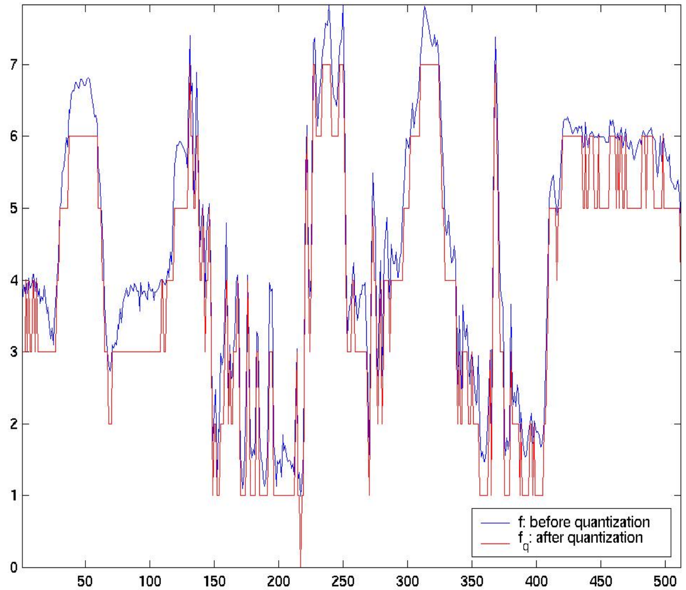
$$E = \sqrt{\frac{1}{NM}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left(f[i,j]-fq[i,j]\right)^2}$$

- Suppose that *f* takes its values in [*a, b*], which may correspond to the whole real axis. We decompose [*a, b*] in *K* intervals $\{(y_{k-1}, y_k]\}_{1\leq k \leq}$ with $y_0 = a$ and $y_K = b$.
- A scalar quantizer approximates all $x \in (y_{k-1}, y_k]$ by x_k:

$$\forall x \in \left(y_{k-1}, y_k\right], \quad Q(x) = x_k$$

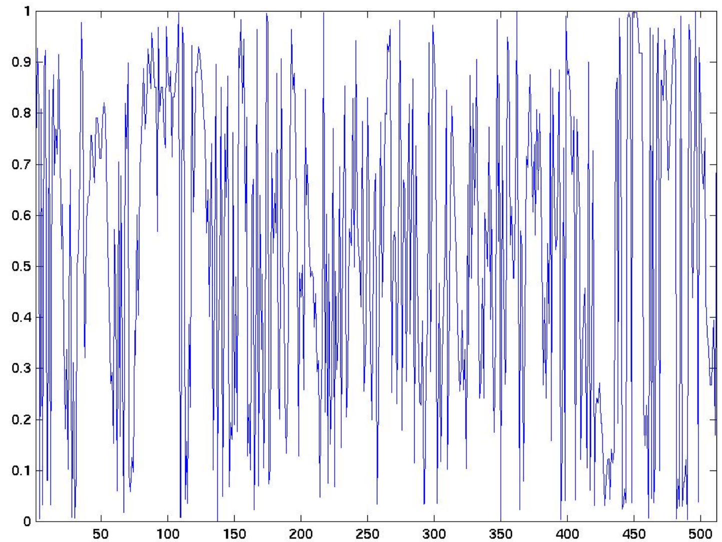
Quantization

Signal before (blue) and after quantization (red) Q



Equivalent noise: $n = f_q - f$

additive noise model: $f_q = f + n$



Quantization

original

5 levels

10 levels

50 levels

A different error measure

- Distortion measure
  - The distortion is measured as the expectation of the mean square error (MSE) difference between the original and quantized signals. The log of the inverse of the MSE (multiplied by 255) gives the **Peak Signal to Noise Ratio (PSNR)**. This is the most widespread measure of image quality

$$PSNR = 20\log_{10}\frac{255}{MSE} = 20\log_{10}\frac{255}{\frac{1}{N\times M}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(I_1[i,j]-I_2[i,j]\right)^2}}$$

- Bottleneck: lack of correlation with perceived image quality
  - Even though this is a very natural way for the quantification of the quantization artifacts, it is not representative of the *visual annoyance* due to the majority of common artifacts.
- Visual models are used to define perception-based image quality assessment metrics

Example

- The PSNR does not allow to distinguish among different types of distortions leading to the same RMS error between images with very different features
- The MSE between images (b) and (c) is the same, so it is the PSNR. However, the visual annoyance of the artifacts is different

Back to GW

Digital Image Acquisition

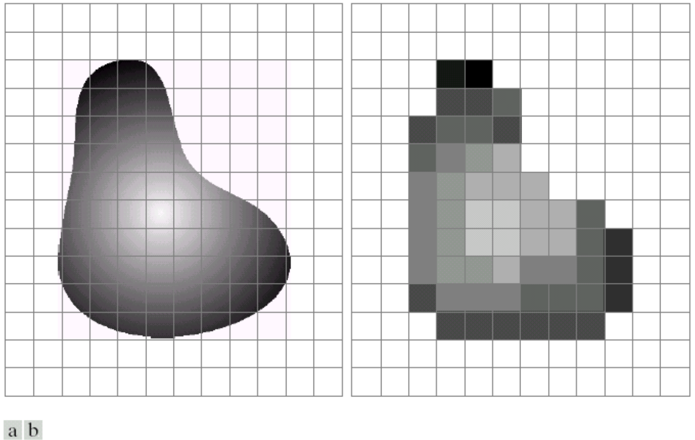
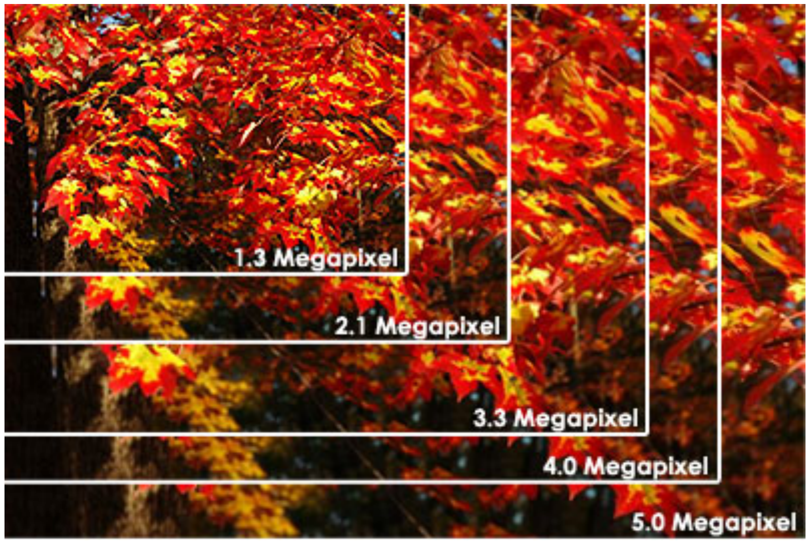
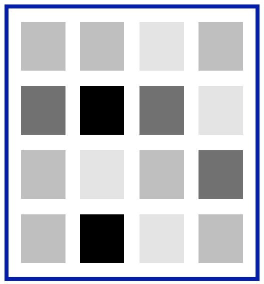


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



- **256x256** - Found on very cheap cameras, this resolution is so low that the picture quality is almost always unacceptable. This is 65,000 total pixels.
- **640x480** - This is the low end cameras. This resolution is ideal for e-mailing pictures or posting pictures on a Web site.
- **1216x912** - This is a "megapixel" image size -- 1,109,000 total pixels -- good for printing pictures.
- **1600x1200** - With almost 2 million total pixels, this is "high resolution." You can print a 4x5 inch print taken at this resolution with the same quality that you would get from a photo lab.
- **2240x1680** - Found on 4 megapixel cameras -- the current standard -- this allows even larger printed photos, with good quality for prints up to 16x20 inches.
- **4064x2704** - A top-of-the-line digital camera with 11.1 megapixels takes pictures at this resolution. At this setting, you can create 13.5x9 inch prints with no loss of picture quality.

Example: greylevel images



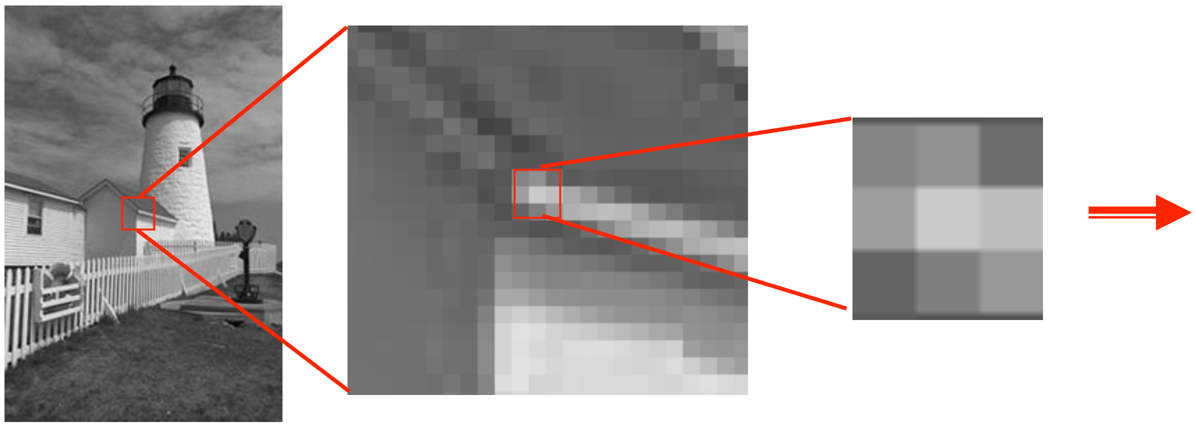
100	100	200	90
50	0	50	200
100	200	100	50
100	0	200	100

Images : Matrices of numbers
Image processing : Operations among numbers
bit depth : number of bits/pixel
N bit/pixel : 2^{N-1} shades of gray (typically N=8)

Matrix Representation of Images

- A digital image can be written as a matrix

$$x[n_1, n_2] = \begin{bmatrix} x[0,0] & x[0,1] & \mathrm{L} & x[0,N-1] \\ x[1,0] & x[1,1] & \mathrm{L} & x[1,N-1] \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ x[M-1,0] & \mathrm{L} & \mathrm{L} & x[M-1,N-1] \end{bmatrix}_{MxN}$$



$$\begin{bmatrix} 35 & 45 & 20 \\ 43 & 64 & 52 \\ 10 & 29 & 39 \end{bmatrix}$$

Summary

- Digital images
  - Sampling+quantization
- Sampling
  - Determines the graylevel value of each pixel
    - Pixel = picture element
- Quantization
  - Reduces the resolution in the graylevel value to that set by the machine precision
- Images are stored as matrices of unsigned chars

Spatial and gray level resolution

- When an actual measure of physical resolution relating pixels and the level of detail they resolve in the original scene are not necessary, it is not uncommon to refer to an **L** -level digital image of size **M*N** as having a spatial resolutio of **M*N** pixels and a gray-level resolution of **L** levels.
- Not precise definition

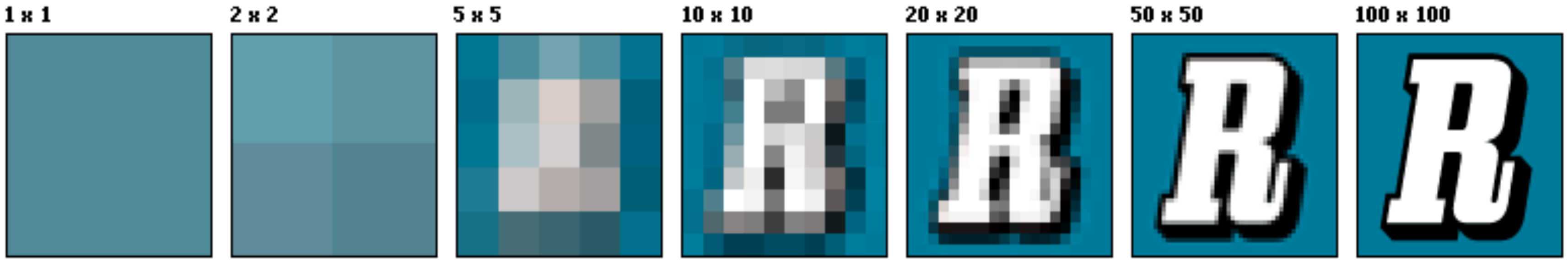
Spatial and Gray level resolution

- Sensor resolution (CCD): Dots Per Inch (DPI)
  - Number of individual dots that can be placed within the span of one linear inch (2.54 cm)
- Image resolution
  - Pixel resolution: NxM
  - Spatial resolution: Pixels Per Inch (PPI)
  - Spectral resolution: bandwidth of each spectral component of the image
    - Color images: 3 components (R,G,B channels)
    - Multispectral images: many components (ex. SAR images)
  - Radiometric resolution: Bits Per Pixel (bpp)
    - Greylevel images: 8, 12, 16 bpp
    - Color images: 24bpp (8 bpp/channel)
  - Temporal resolution: for movies, number of frames/sec
    - Typically 25 Hz (=25 frames/sec)

About resolution

- **Sampling** is the principal factor determining the **spatial resolution** of an image.
  - Basically, spatial resolution determines the smallest discernible detail in an image.
- **Gray-level resolution** similarly refers to the smallest discernible change in gray level, but, as noted in Section 2.1.3, measuring **discernible changes** in gray level is a highly subjective process

Example: Image - pixel resolution



Example: Image – Spatial resolution

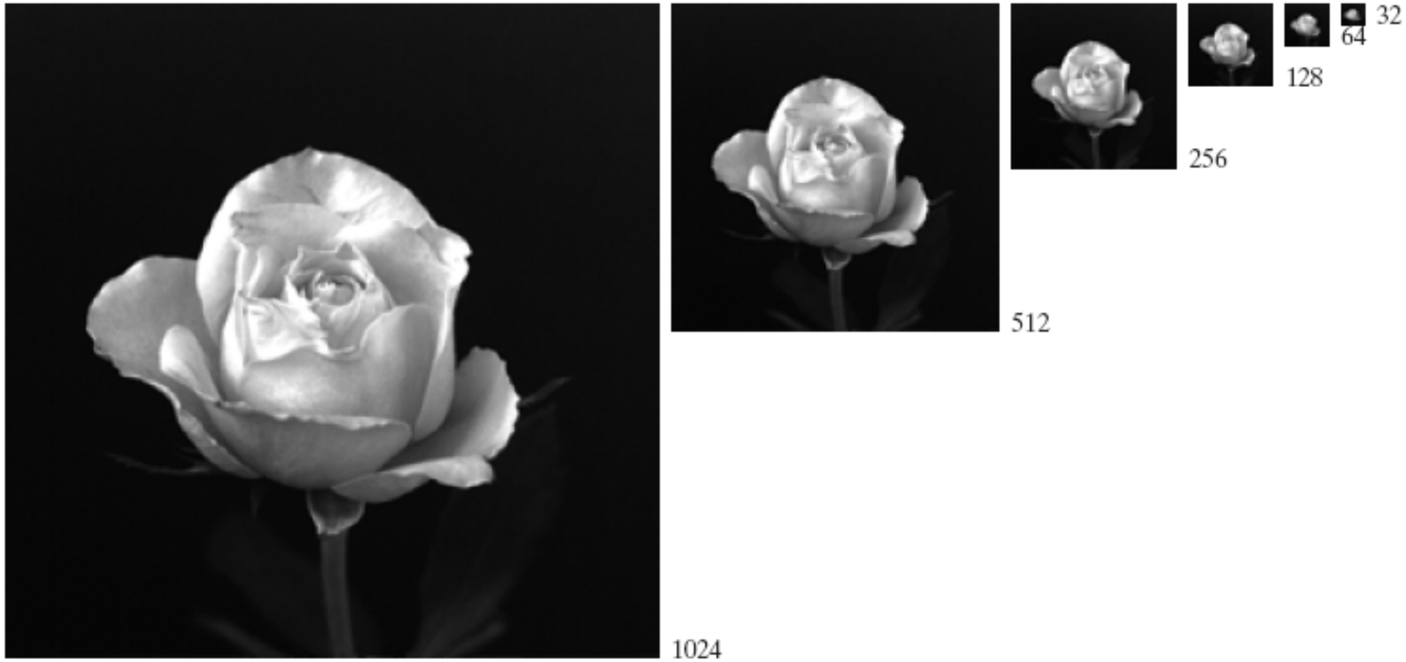


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Example: Image – spatial resolution

Don't confuse image size and resolution.

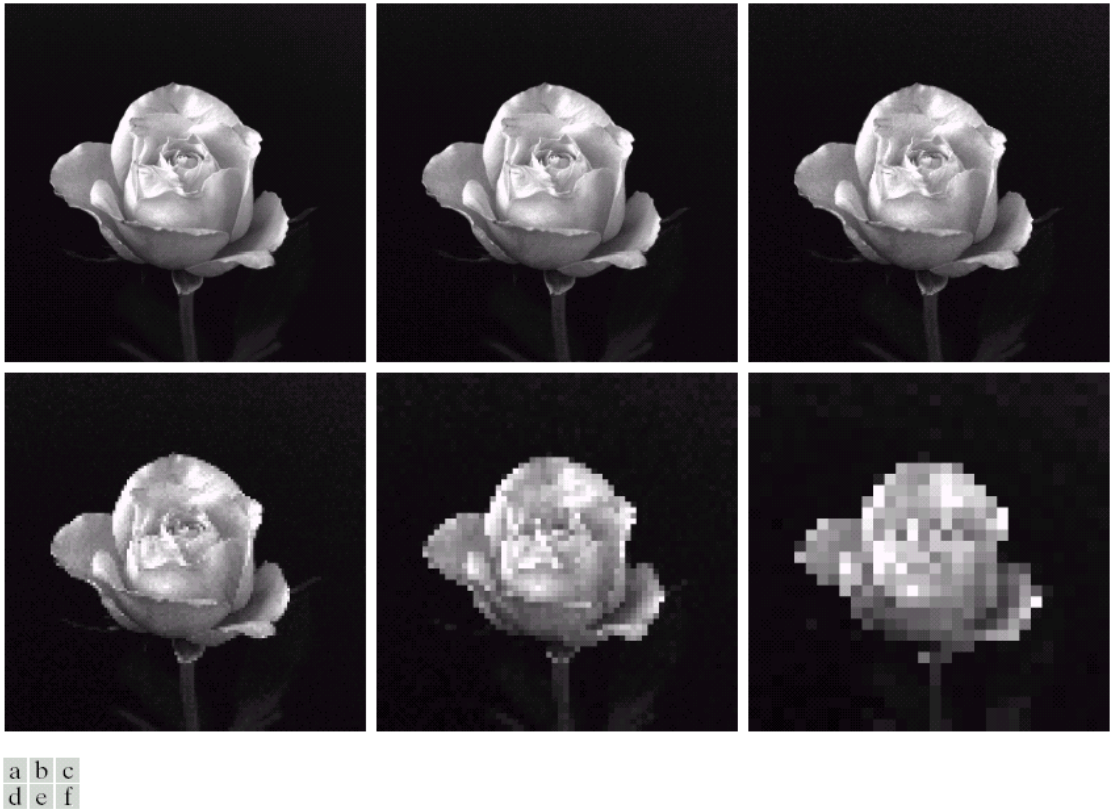


FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

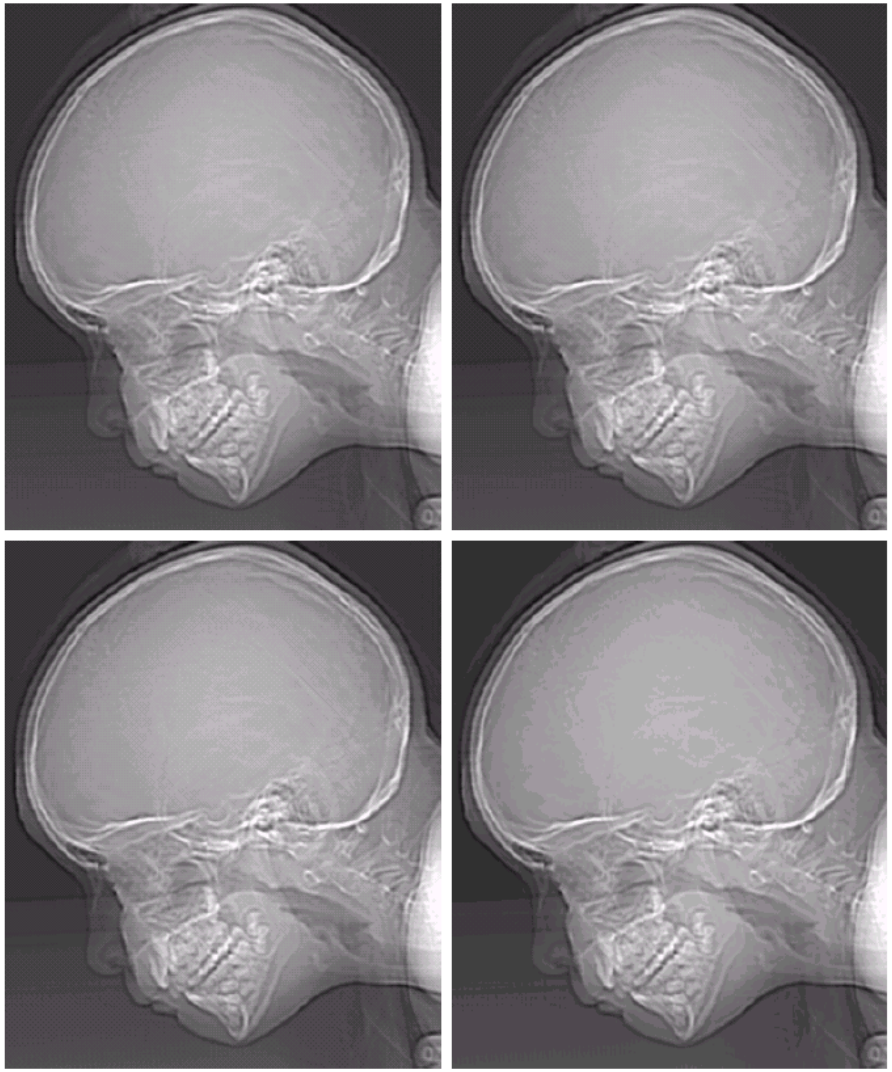
Bit Depth – Image – radiometric (grayscale) Resolution

8 bits

7 bits

6 bits

5 bits



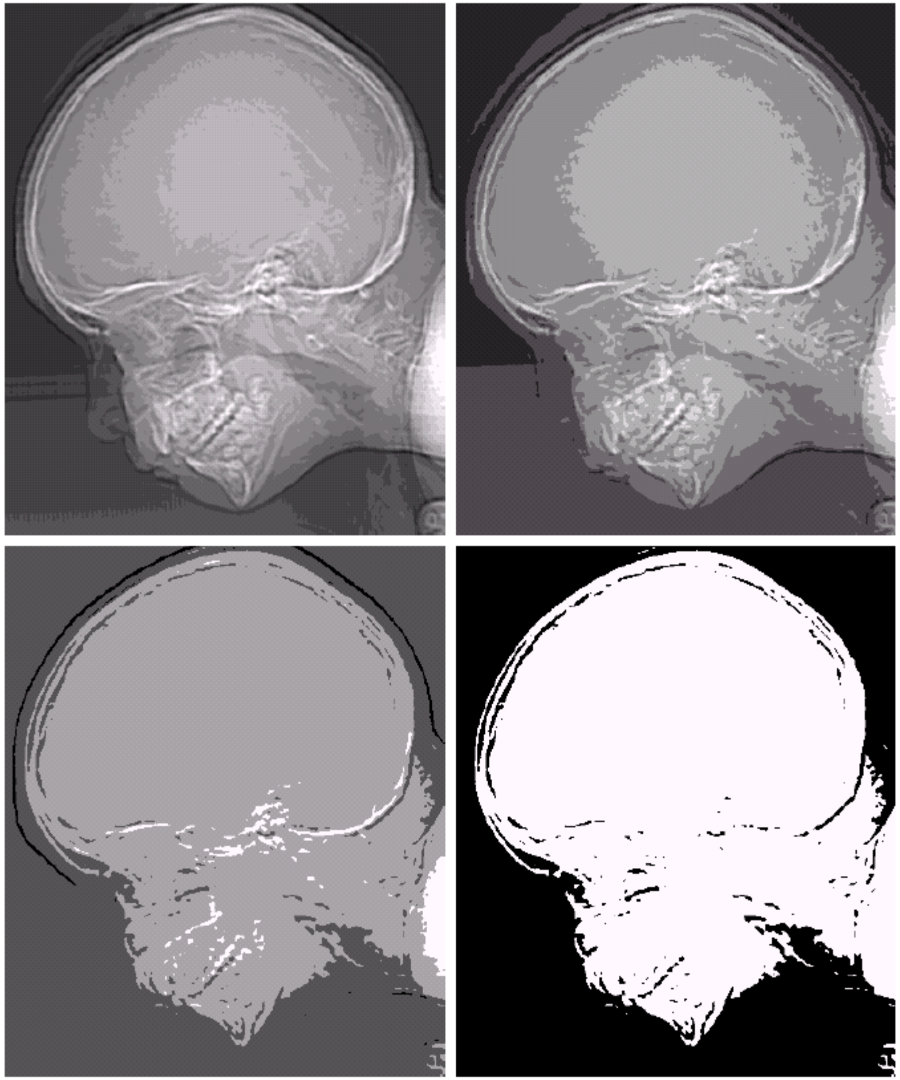
a b
c d

FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

Bit Depth – Grayscale Resolution

e f
g h

4 bits



3 bits

2 bits

1 bit

FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

Basic relationships between pixels

- An image is denoted by f(x, y). When referring in this section to a particular pixel, we use lowercase letters, such as p and q.
- Neighbors of a pixel
  - A pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by

    $$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$$

  - This set of pixels, called the **4-neighbors of p** , is denoted by $N_4(p)$. Each pixel is at unit distance from (x, y), and some of the neighbors of p lie outside the digital image if (x, y) is on the border of the image.
  - The four **diagonal neighbors** of p have coordinates

    $$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

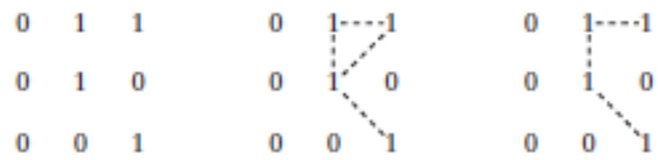
  - and are denoted by $N_D(p)$. These points, together with the 4-neighbors, are called the 8-neighbors of p, denoted by $N_8(p)$.

Adjacency, Connectivity, Regions, and Boundaries

- Used to evaluate the connectivity among pixels
- To establish if two pixels are **connected**, it must be determined if they are **neighbors** and if their gray levels satisfy a specified criterion of **similarity** (say, if their gray levels are equal)
- Let V be the set of gray-level values used to define adjacency.
  - Concept: in a binary image, 4-connected pixels are adjacent if they have the same value
  - In a grayscale image, the idea is the same but V can contain more elements: in the adjacency of pixels with a range of possible gray-level values 0 to 255, set V could be any subset of these 256 values

Types of adjacency

- Types of adjacency
  - 4-adjacency. Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$
  - 8-adjacency. Two pixels p and q with values from V are 8-adjacent if q is in the set N_8(p
  - m-adjacency **(mixed adjacency)**: two pixels p and q with values in V are mixed adjacent if
  - (i) q is in $N_4(p)$, *or*
  - (ii) q is is $N_D(p)$ *and* the set overlap($N_4(p)$, $N_4(q)$) has no pixels whos values are from V



a b c

FIGURE 2.26 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

Digital path or curve

- Two image subsets S1 and S2 are adjacent if some pixel in S1 is adjacent to some pixel in S2 .
- Let S represent a subset of pixels in an image.Two pixels p and q are said to be connected in S if there exists a path between them consisting entirely of pixels in S. For any pixel p in S , the set of pixels that are connected to it in S is called a *connected component* of S.
  - If it only has one connected component, then set S is called a connected set .

- A *digital path* or *curve* from pixel p(x,y) to pixel q(s,t) is a sequence of distinct pixels of coordinates

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

where $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, and pixels (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$. In this case, n is the *length* of the path. If $(x_0, y_0) = (x_n, y_n)$, the path is a *closed* path. We can define 4-, 8-, or m-paths de-

Regions

- Let R be a subset of pixels in an image. We call R a region of the image if R is a *connected* set.
- The **boundary** (also called border or contour) of a region R is the set of pixels in the region that have one or more neighbors that are not in R.
  - If R happens to be an entire image (which we recall is a rectangular set of pixels), then its boundary is defined as the set of pixels in the first and last rows and columns of the image.This extra definition is required because an image has no neighbors beyond its border. Normally, when we refer to a region, we are referring to a subset of an image, and any pixels in the boundary of the region that happen to coincide with the border of the image are included implicitly as part of the region boundary.

Edges

- Edges and region boundaries are different conceps
- Region boundaries are a *global* concept, as they form a closed path around regions
- Edges consist of pixels with sharp graylevel gradients or, more in general, gradients in other features (region-based processing), so they are a *local* concept
- It is possible to link edge points into edge segments, and sometimes these segments are linked in such a way that correspond to boundaries, but this is not always the case.The one exception in which edges and boundaries correspond is in binary images where they coincide

Distances

- For pixels p, q , and z , with coordinates (x, y), (s, t) , and (v, w) , respectively, D is a distance function or metric if

(a) D(p, q)≥0, D(p, q)=0 iff p=q,

(b) D(p, q)=D(q, p), and

(c) D(p,z)≤D(p, q)+D(q, z).

- Euclidean distance between p and q is defined as

$$D(p,q)=\sqrt{(x-s)^2+(y-t)^2}$$

- D_4 distance (also called **city-block distance**) between p and q

$$D(p,q)=|x-s|+|y-v|$$

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

$D_4 \le 2$

Distances

- D_8 or chessboard distance

$$D(p,q) = \max\{|x-s|,|y-v|\}$$

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

$$D_8 \le 2$$

Image file formats

- Many image formats (about 44)
- BMP, lossless
- TIFF, lossless/lossy
- GIF (Graphics Interchange Format)
  - Lossless, 256 colors, copyright protected
- JPEG (Joint Photographic Expert Group)
  - Lossless and lossy compression
  - 8 bits per color (red, green, blue) for a 24-bit total
- PNG (Portable Network Graphics)
  - Freewere
  - supports truecolor (16 million colours)
- ... more to come ..