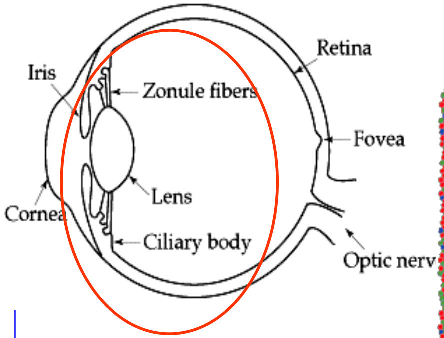
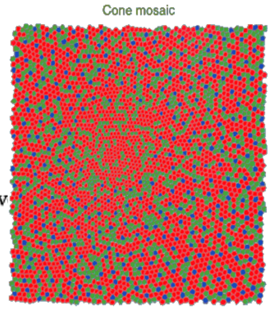
Digital image fundamentals

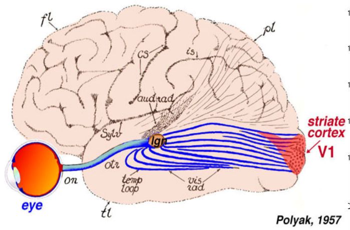
Gonzalez Woods Chapter 2

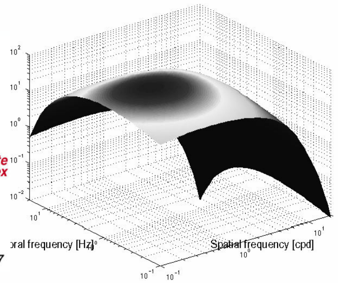
Visual and Imaging chain

VISUAL CHAIN

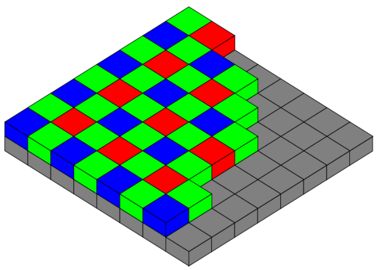




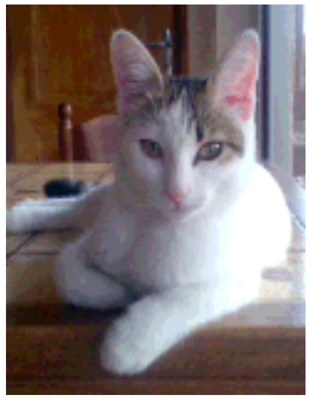




Optics	Sampling (A/D)	Multiscale representation	Filtering	Processing



IMAGING CHAIN



The EM spectrum

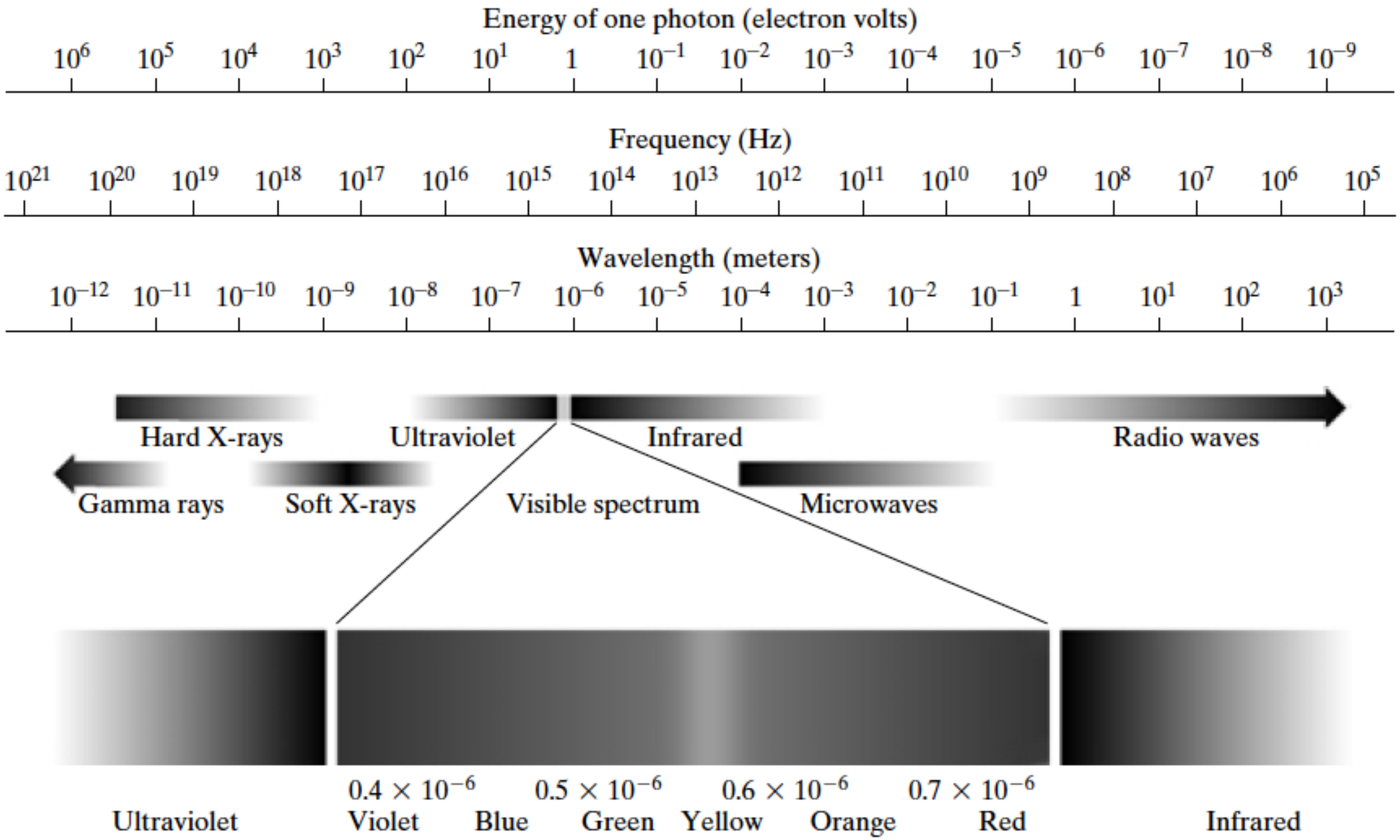
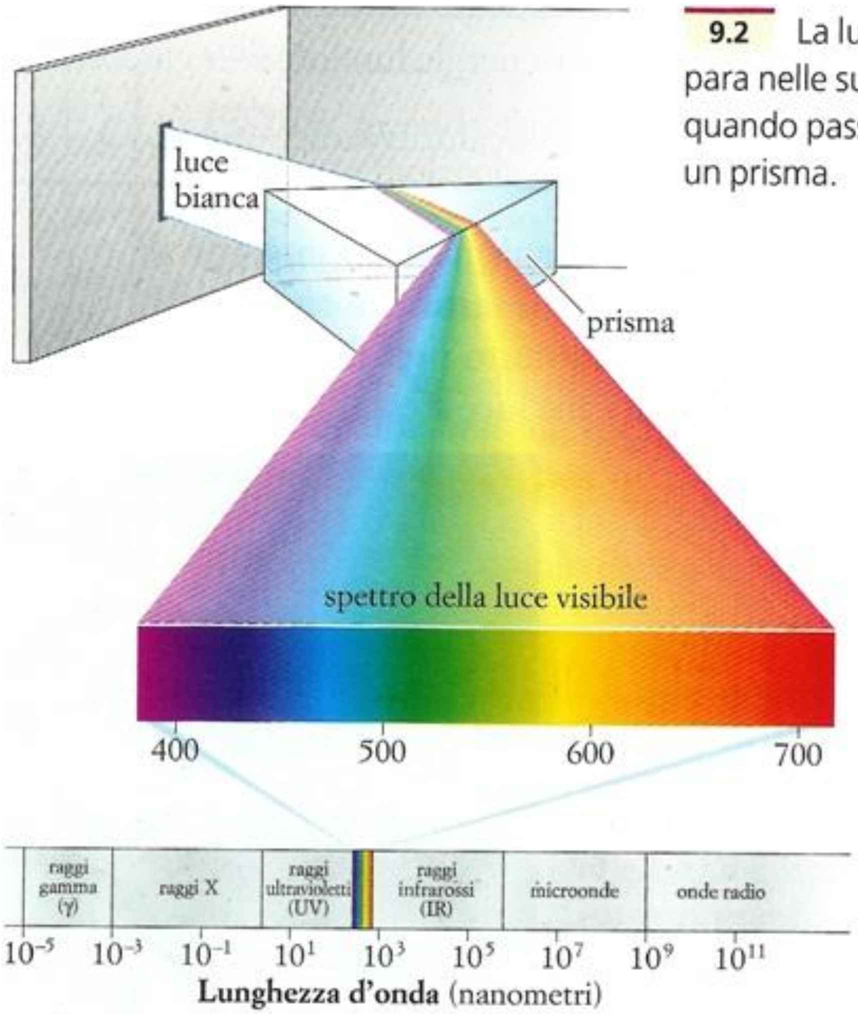


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Newton's prism



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

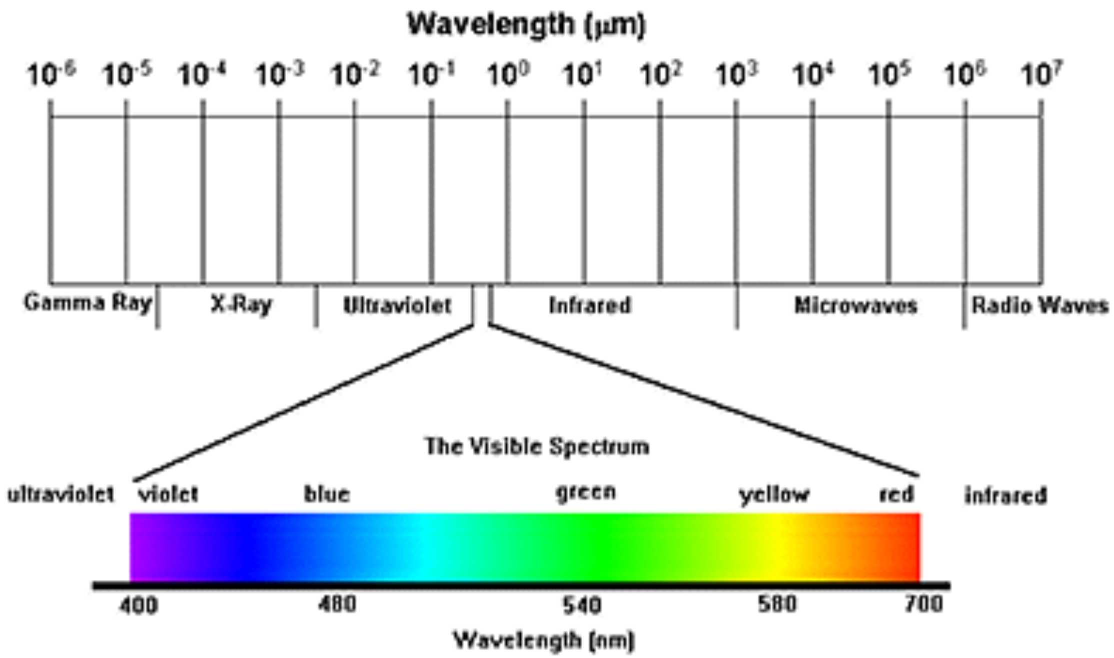
EM units

- 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 **10^{-6}** m) and nanometers (**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.

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

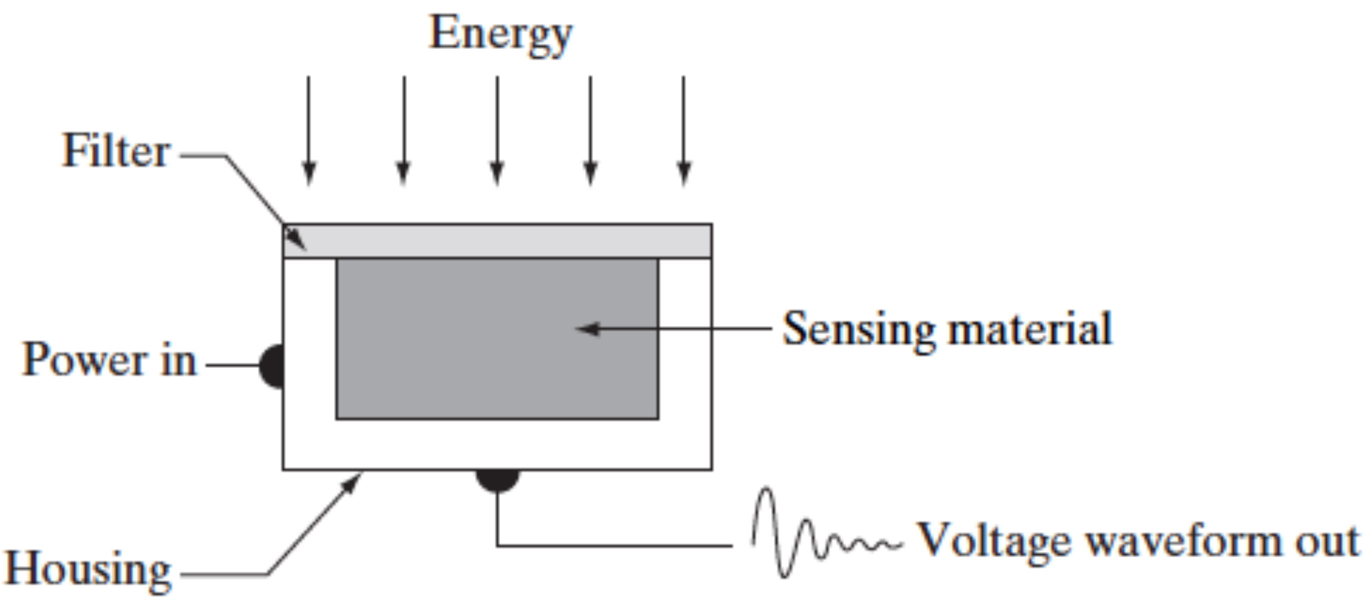
- **Radiance** is the total amount of energy that flows from the light source, and it is usually measured in watts (**W**).
- **Luminance**, measured in lumens (**lm**), gives a measure of the amount of energy **an observer perceives** from a light source.
  - For example, light emitted from a source 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

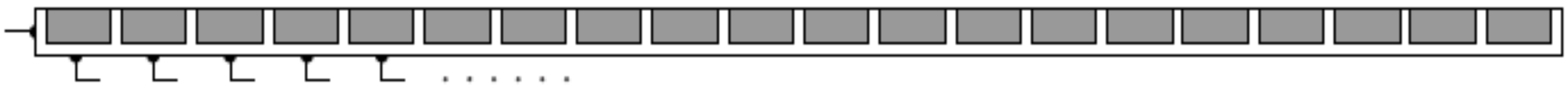
Image sensing and acquisition

- 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.

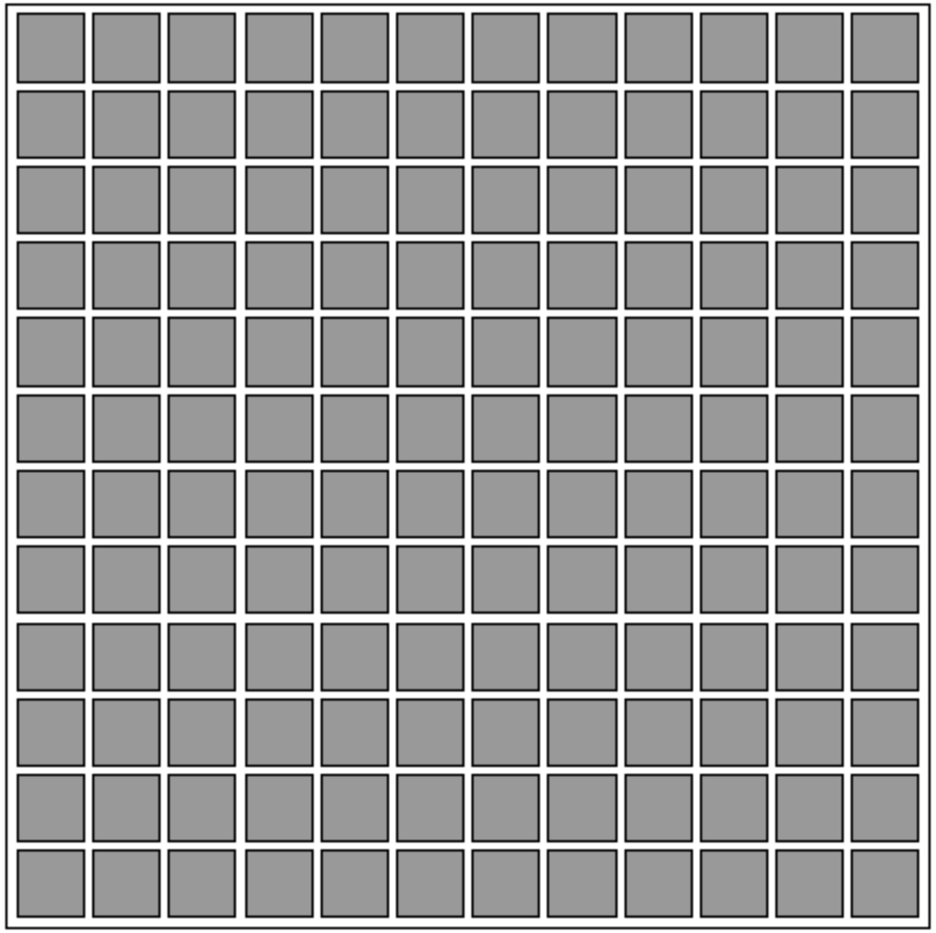
Imaging sensors



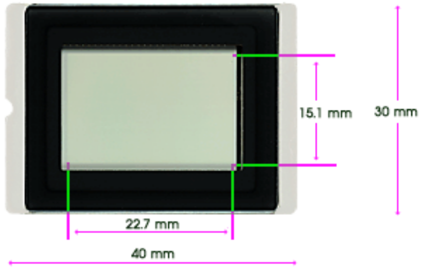
Line sensors

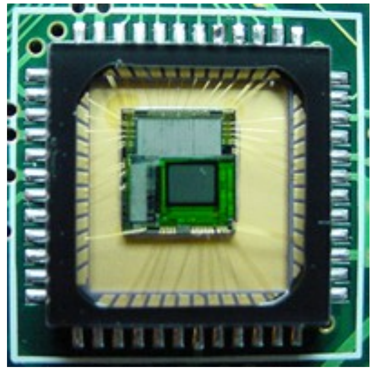


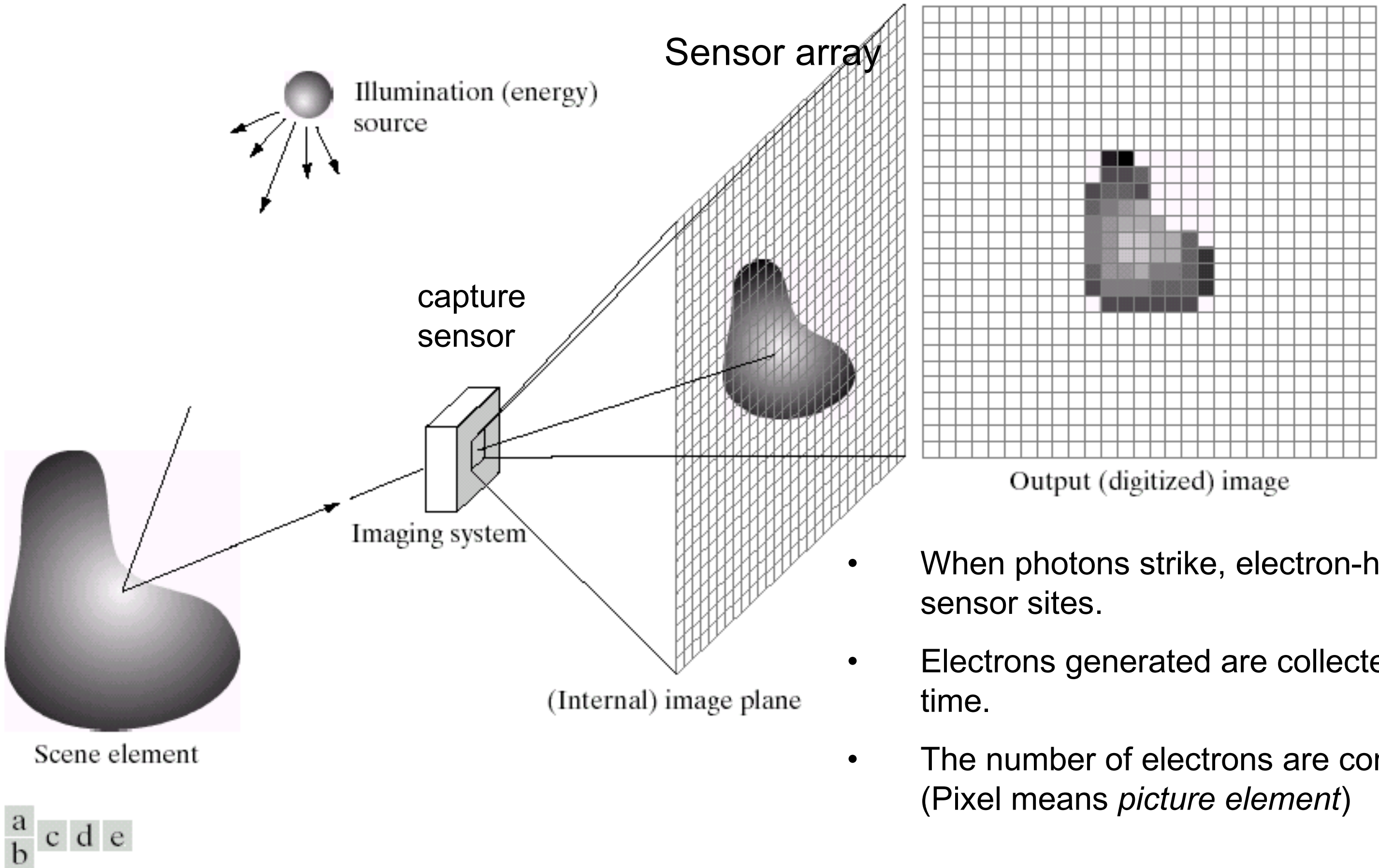
Array sensors



Digital Image Acquisition







- 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*)

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.

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

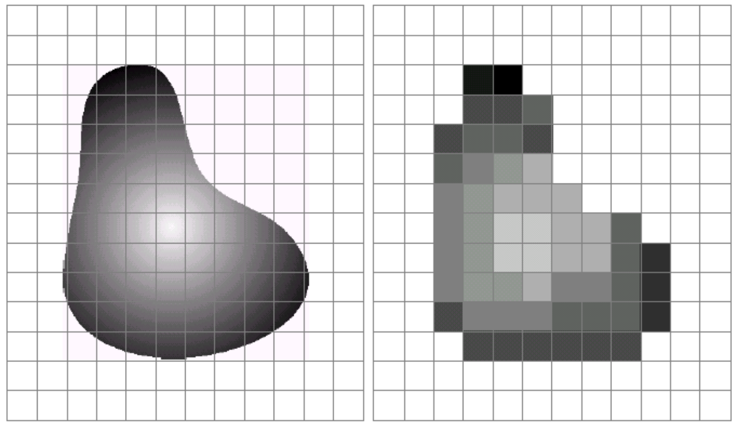
  $0<i(x,y)<+\infty$

  $0<r(x,y)<1$: 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



a b

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

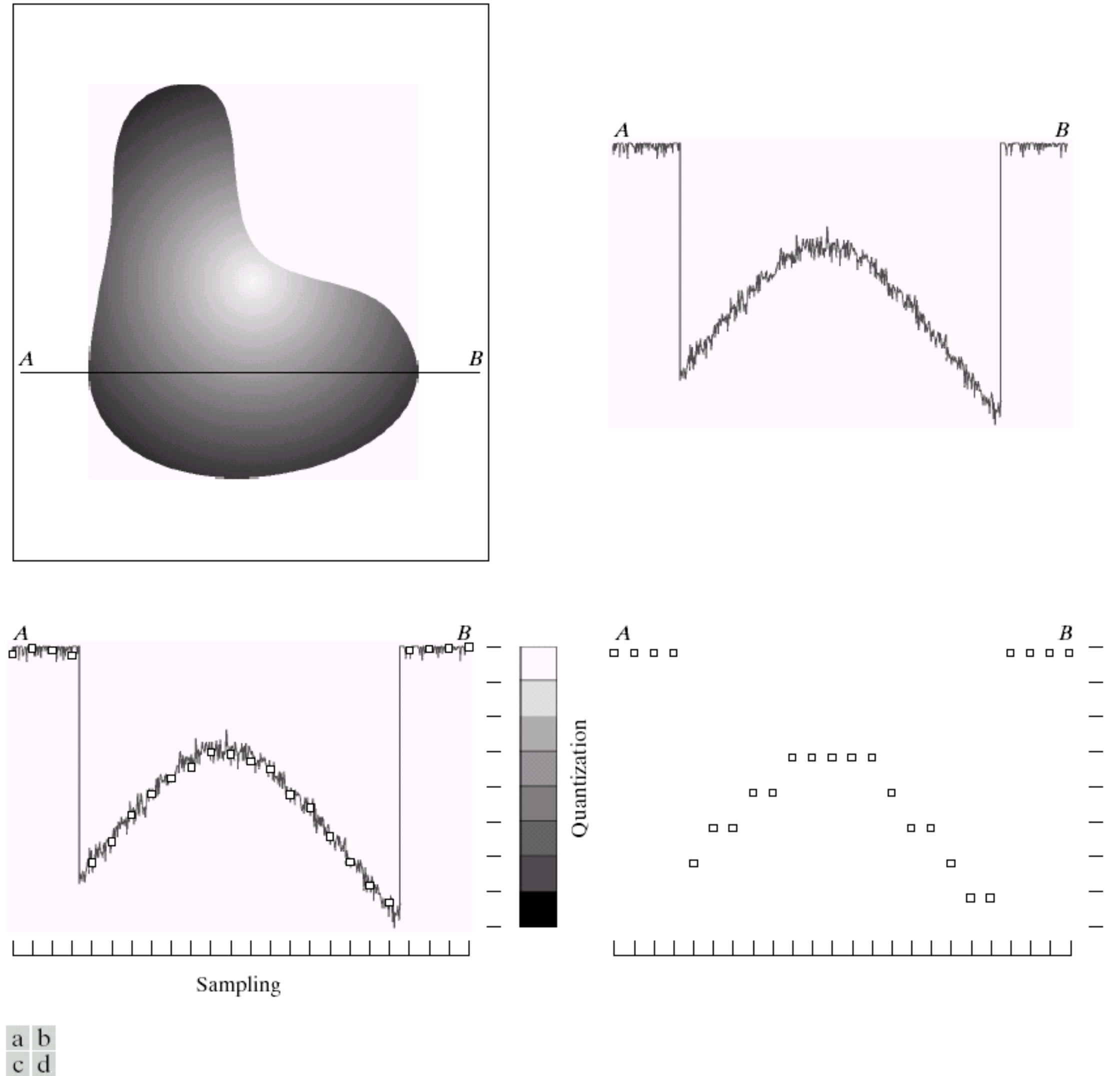


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

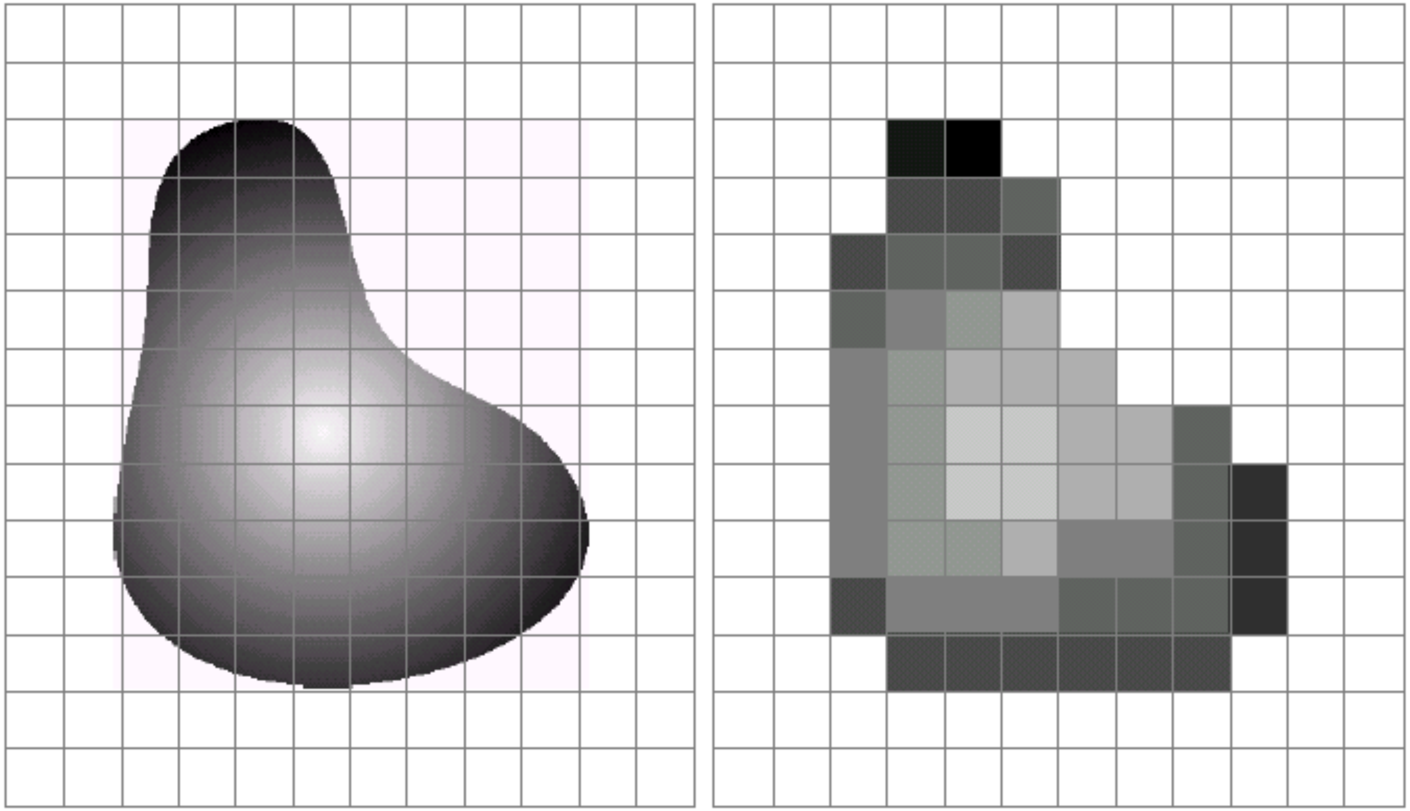
Image sampling and quantization

- 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 $0 \le i \le N_x$ and $0 \le j \le N_y$, 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.

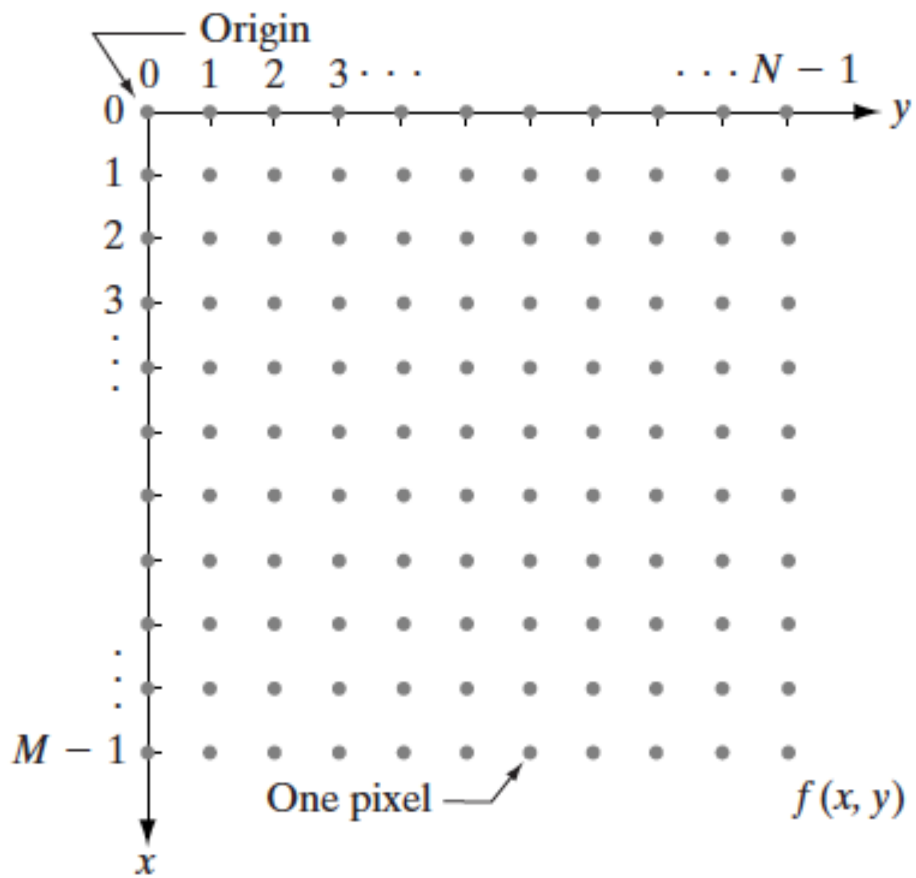
Sampling example



a b

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 xy 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.

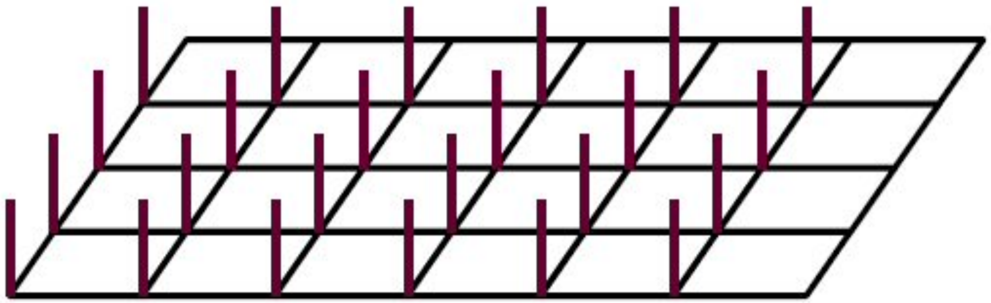
Choice of the values

- 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 $b=N*M*n$

Dirac *brush*

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

- 2D sampling : Dirac « brush »



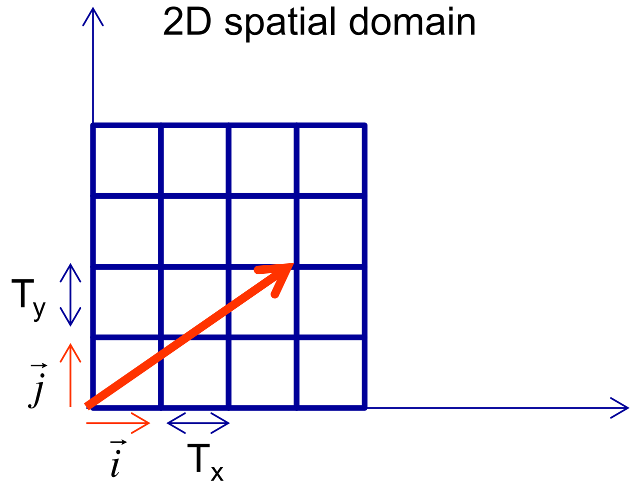
Nyquist theorem

- Sampling in p-dimensions

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

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

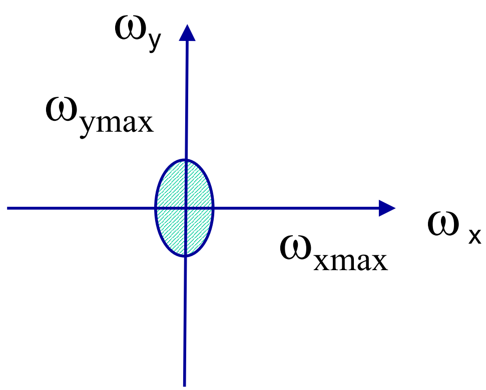
2D spatial domain



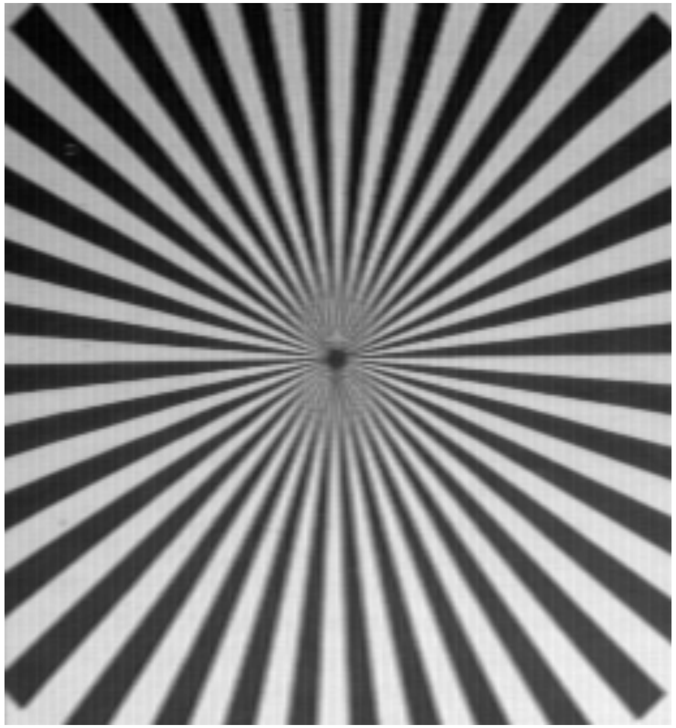
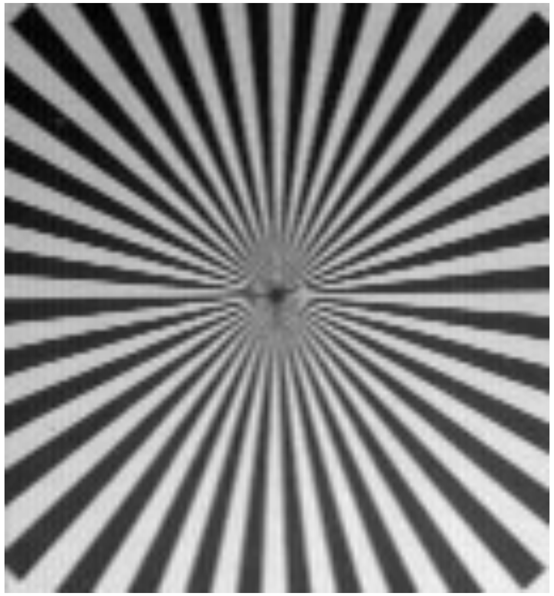
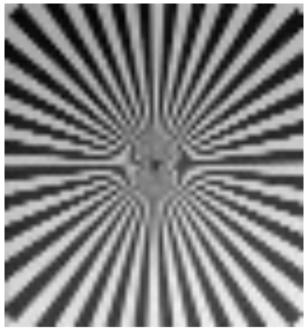
- Nyquist theorem

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

2D Fourier domain



Spatial aliasing

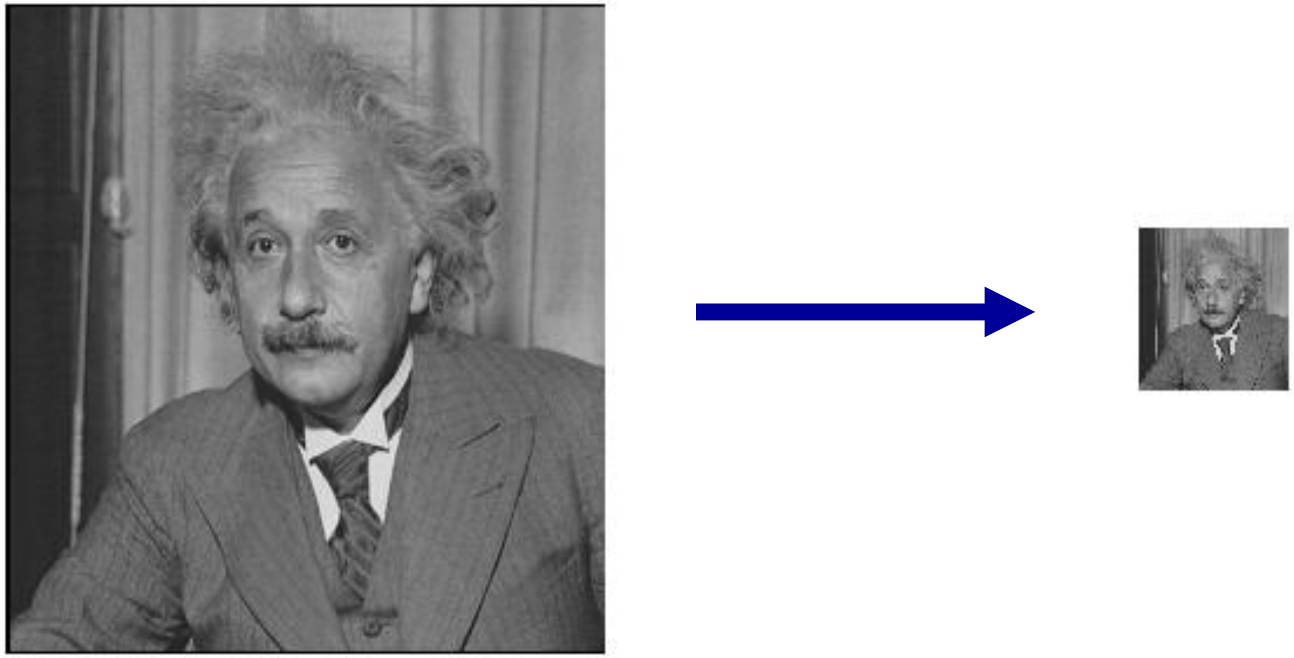




If the sampling criterion is not respected, some defects appear in the image. These are due to the fact that rapid signal changes cannot be correctly reproduced.

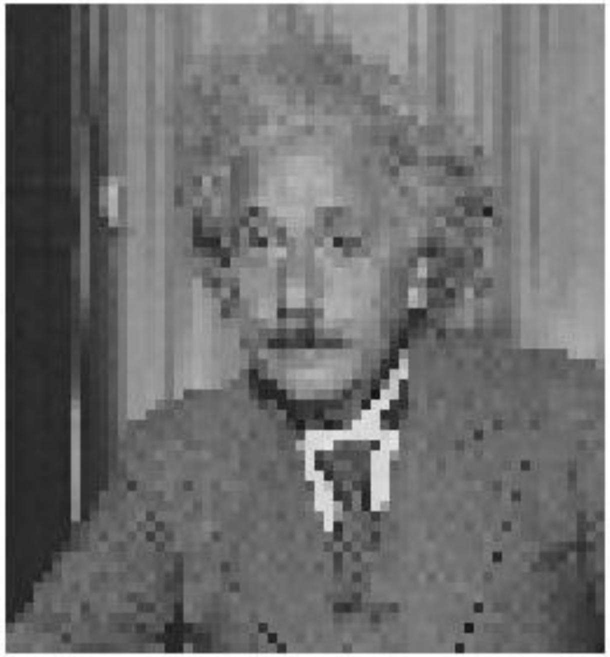
Resampling

- 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

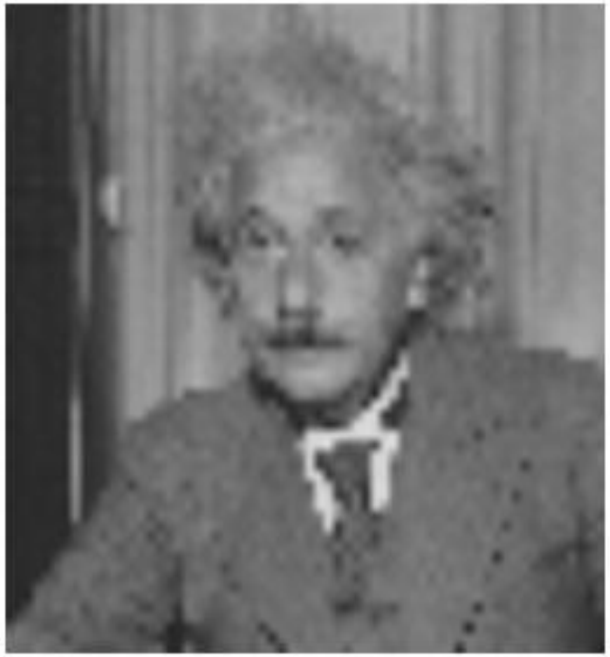


Upsampling or Interpolation



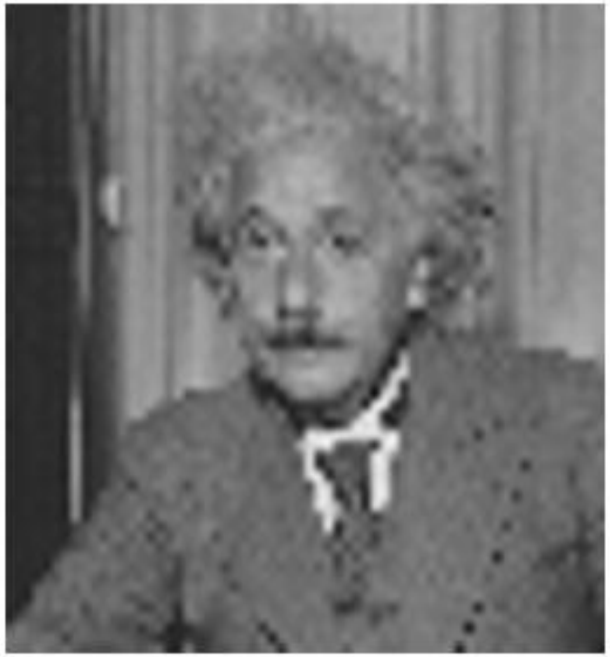
nearest neighbor (NN)

Upsampling or Interpolation



bilinear

Upsampling or Interpolation



bicubic

A hint to interpolation: zooming

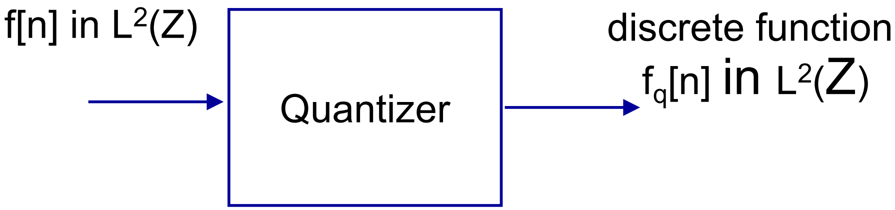
- Zooming requires two steps: the creation of new pixel locations, and the assignment of gray levels to those new locations.
    - Example: Suppose that we have an image of size 500*500 pixels and we want to enlarge it 1.5 times to 750*750 pixels. Conceptually, one of the easiest ways to visualize zooming 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

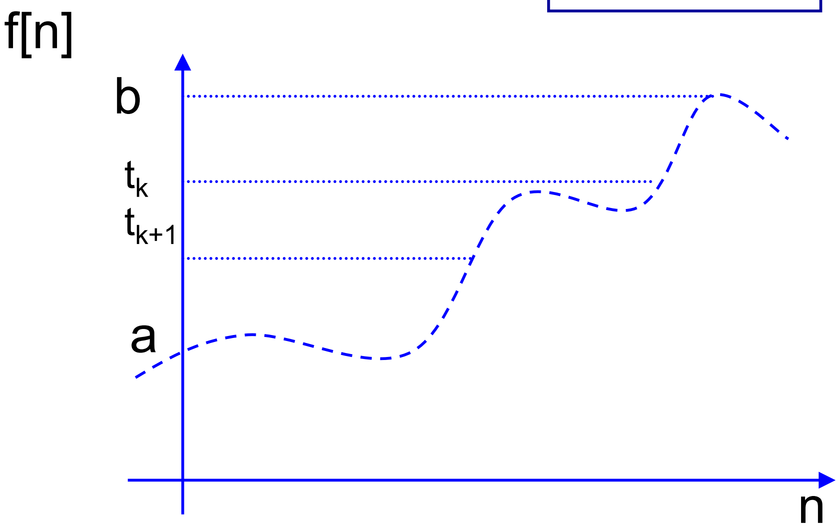
Quantization

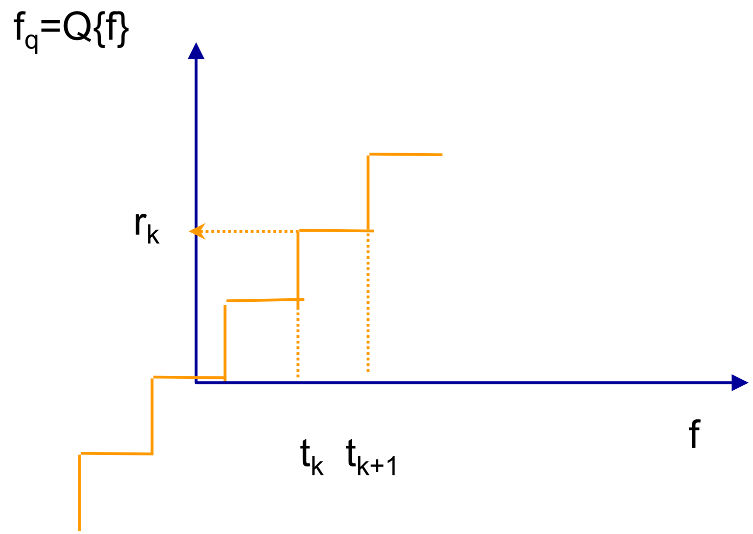
- A/D conversion ⇨ quantization

  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

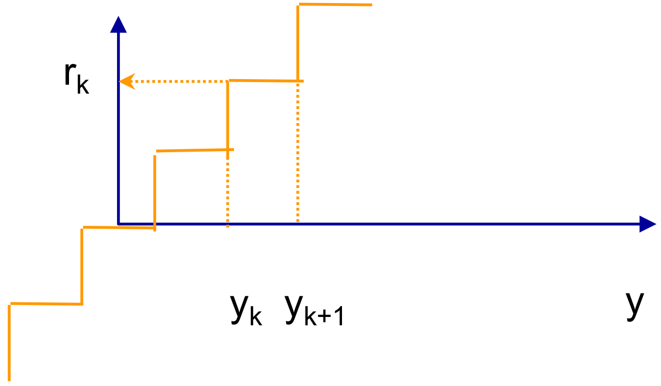
- A/D conversion ⇨ quantization

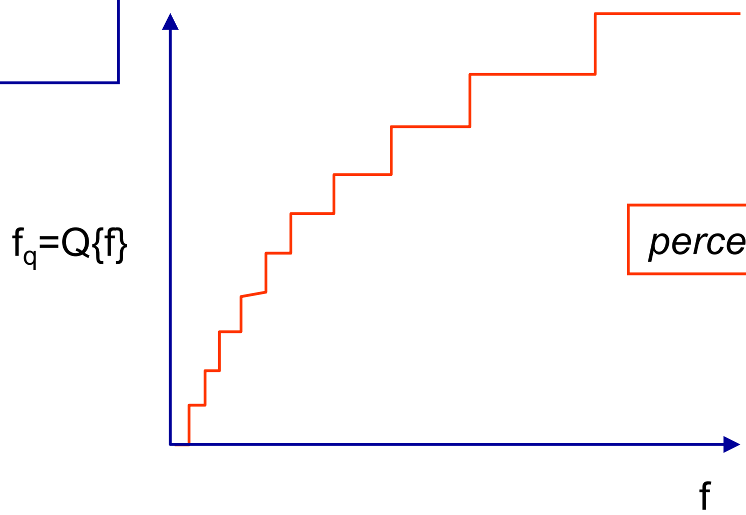
f in $L^2(R)$ → Quantizer → discrete function f in $L^2(Z)$

TRANSFER FUNCTION

uniform

$X=Q\{y\}$





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

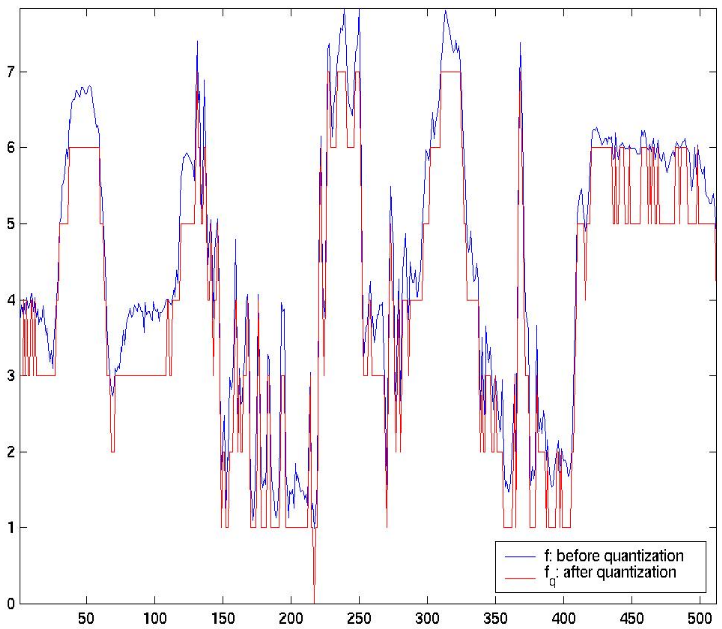
$$E = \sqrt{\frac{1}{NM}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left(f[i,j]-fq[i,j]\right)^2}$$

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

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

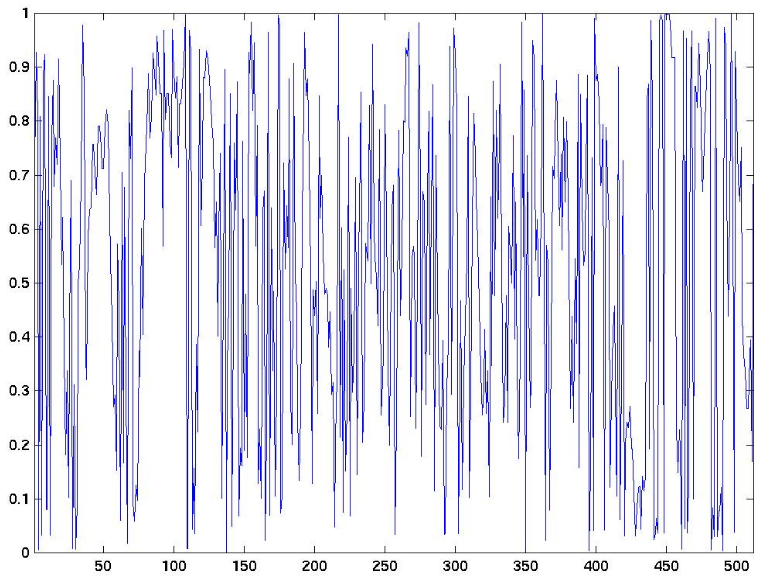
Quantization

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

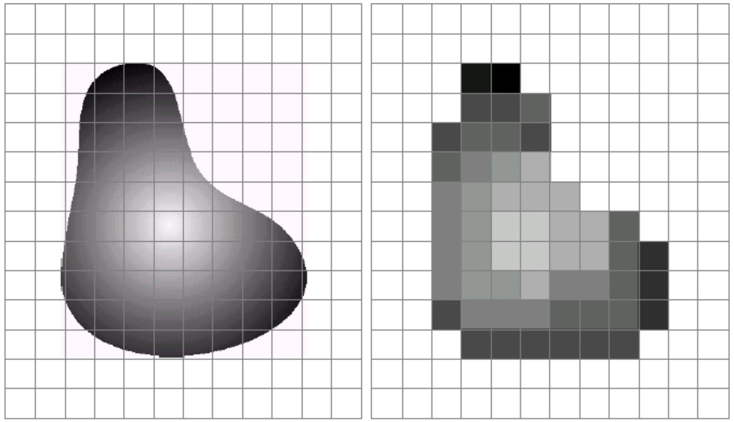
$$PSNR = 20\log_{10}\frac{255}{MSE} = 20\log_{10}\frac{255}{\frac{1}{N\times M}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(I_1[i,j]-I_2[i,j]\right)^2}}$$

- 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

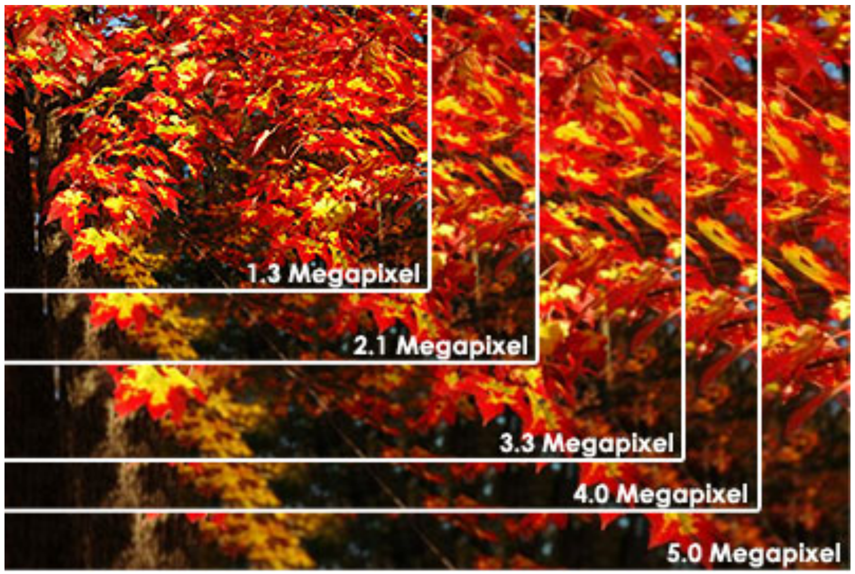
- The PSNR does not allow to distinguish among different types of distortions leading to the same RMS error between images with very different features
- The MSE between images (b) and (c) is the same, so it is the PSNR. However, the visual annoyance of the artifacts is different

Digital Image Acquisition

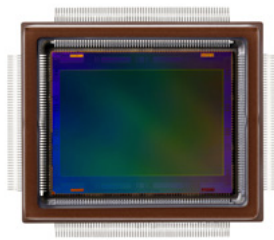


a b

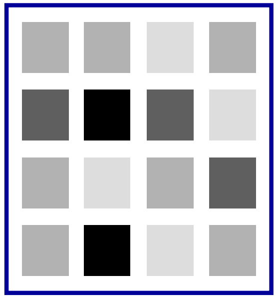
FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



- **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.
- **19580 x 12600**: nuovo CMOS APS-H (29.2 x 20.2mm) fino a 250 Megapixel



Example: greylevel images



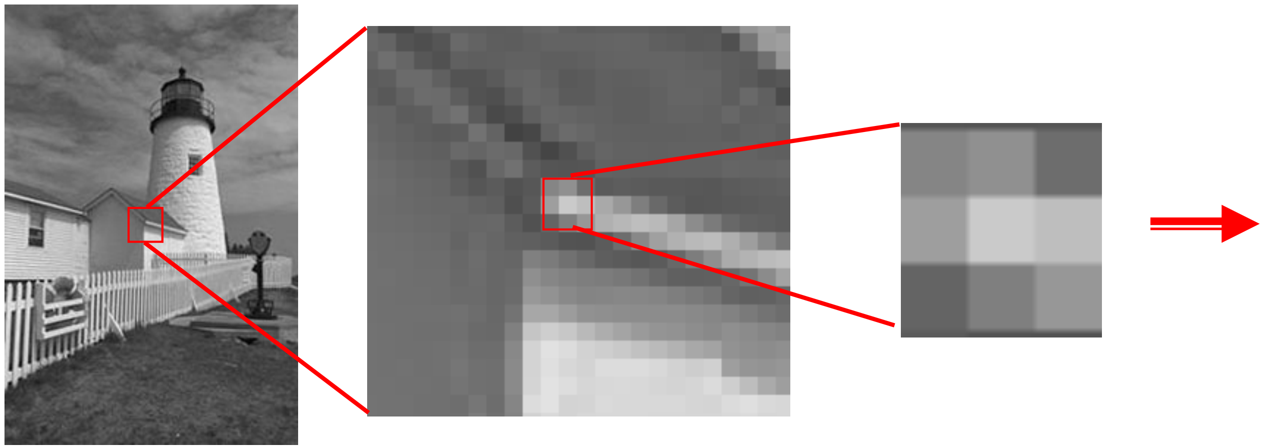
100	100	200	90
50	0	50	200
100	200	100	50
100	0	200	100

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

$$x[n_1, n_2] = \begin{bmatrix} x[0,0] & x[0,1] & \cdots & x[0, N-1] \\ x[1,0] & x[1,1] & \cdots & x[1, N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x[M-1,0] & \cdots & \cdots & x[M-1, N-1] \end{bmatrix}_{MxN}$$



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

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

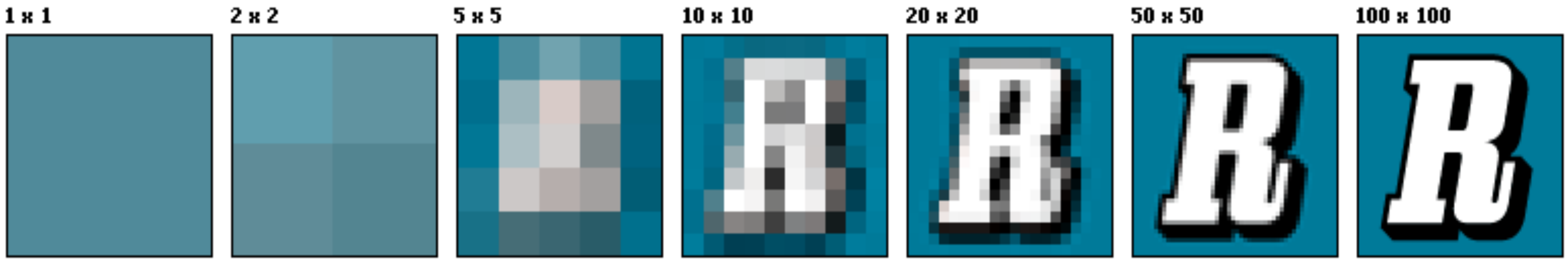
Spatial and Gray level resolution

- 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 (pixels/inch)

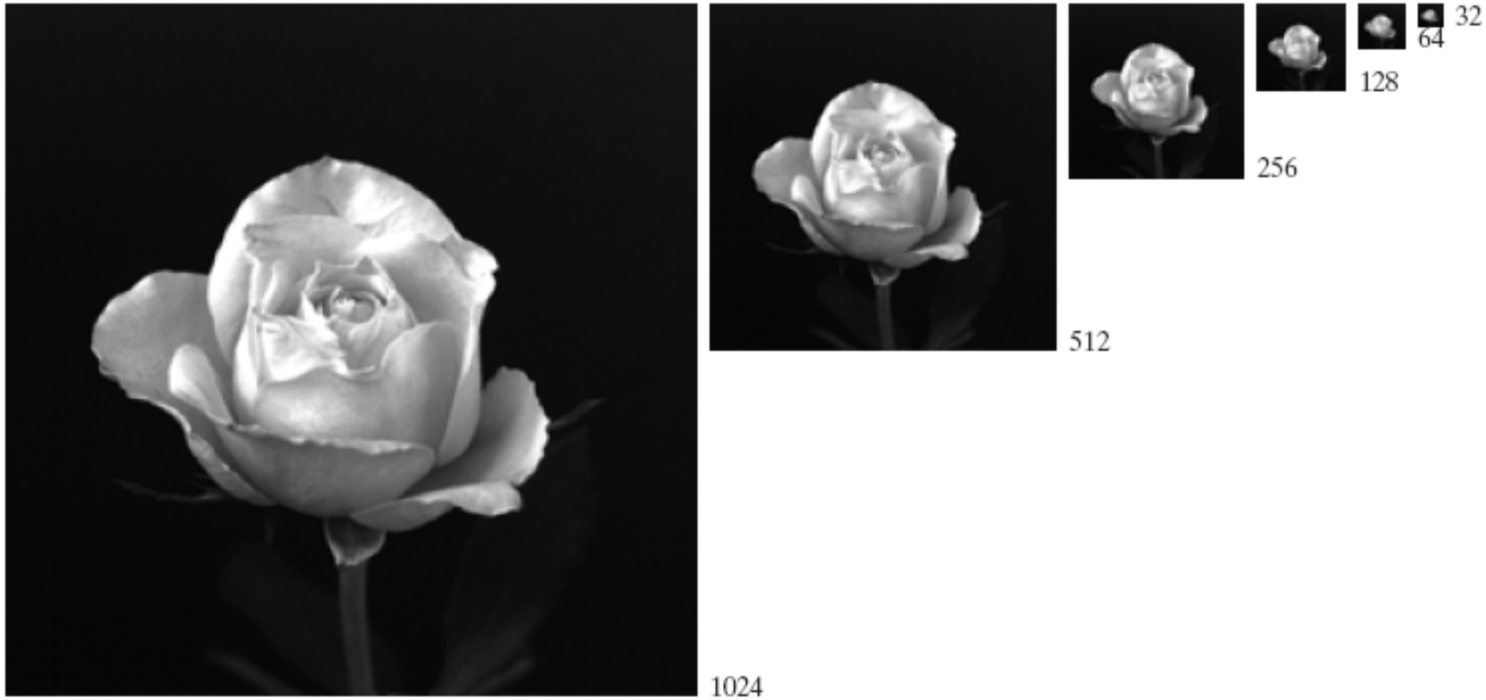
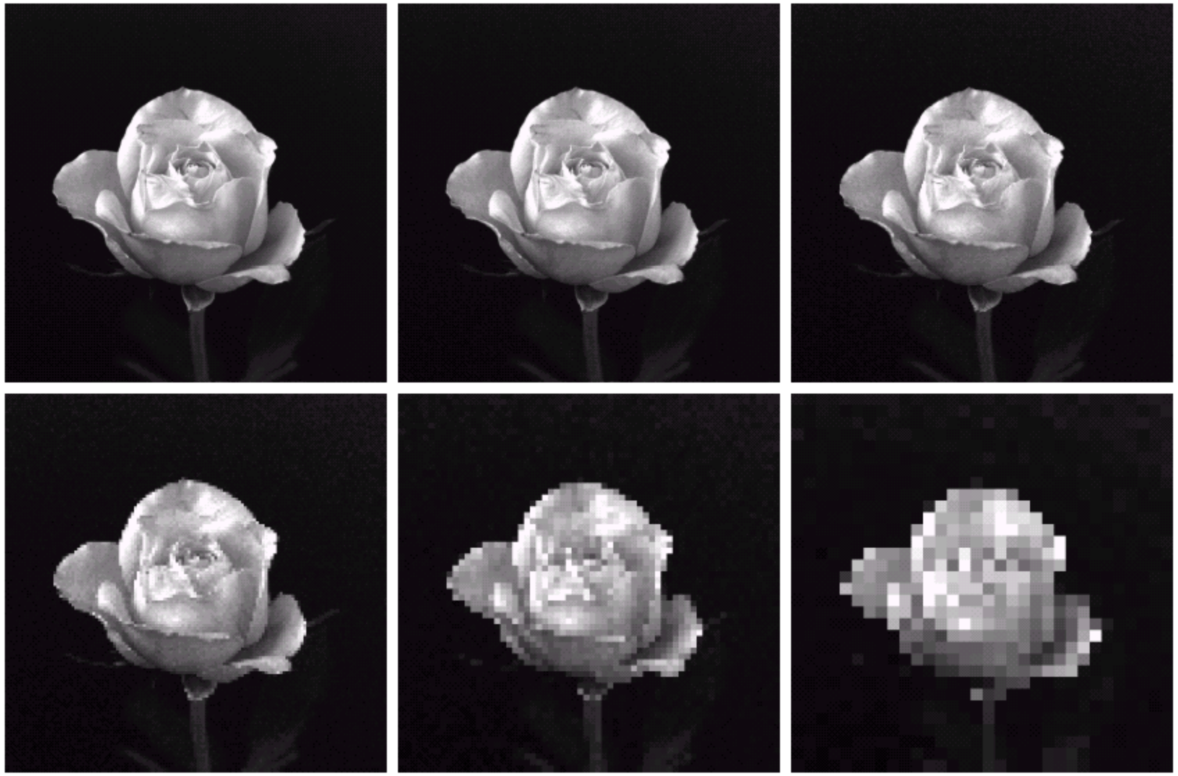


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Example: Image – spatial resolution

Don’t confuse image size and resolution.



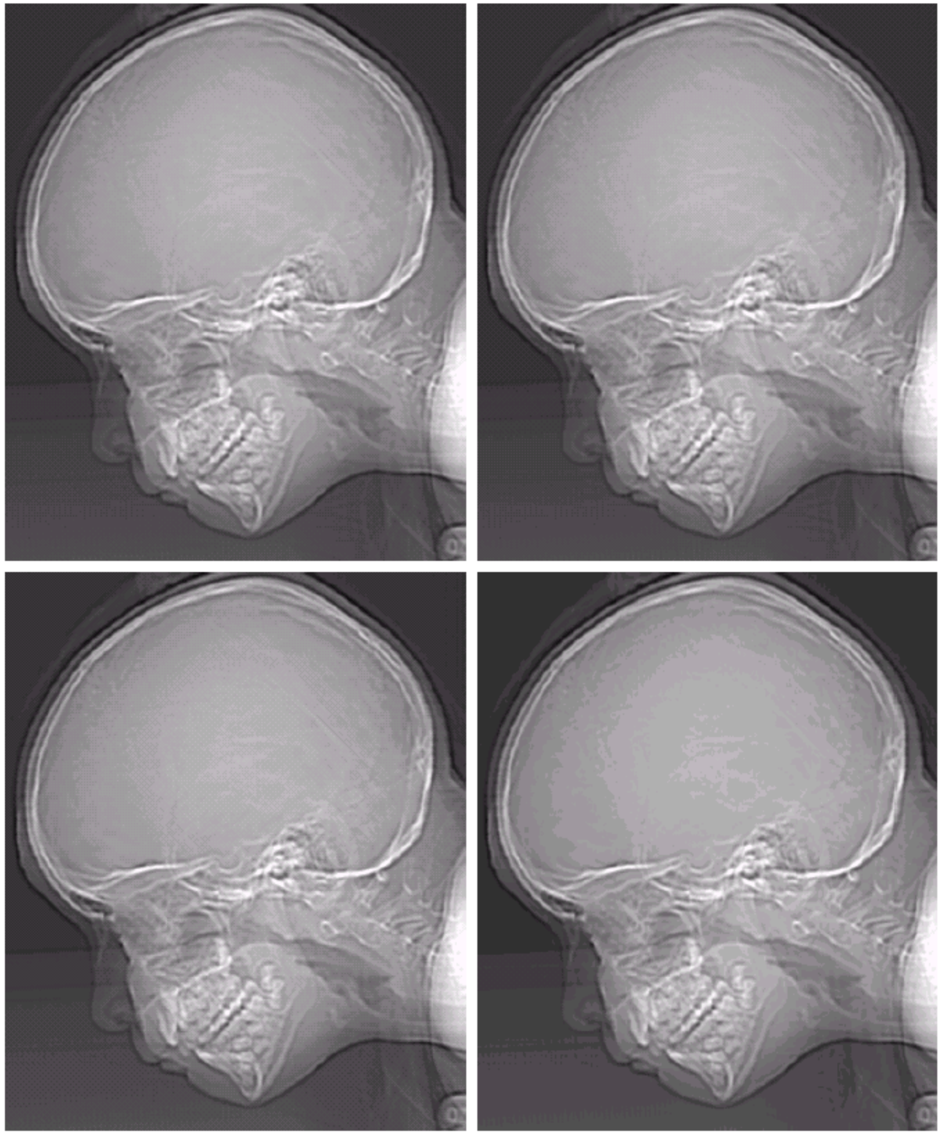
a	b	c
d	e	f

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

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.)

4 bits

3 bits

2 bits

1 bit