Ψηφιακή Επεξεργασία Εικόνας (ΨΕΕ) – MYE037

Εαρινό εξάμηνο 2023-2024

Digital Image Fundamentals – Οι Θεμελιώδεις αρχές των ψηφιακών εικόνων

Άγγελος Γιώτης

a.giotis@uoi.gr

Images taken from:
R. Gonzalez and R. Woods. Digital Image Processing, Prentice Hall, 2008.
Digital Image Processing course by Brian Mac Namee, Dublin Institute of Technology.

Digital Image Fundamentals

"Those who wish to succeed must ask the right preliminary questions"

Aristotle

Contents

This lecture will cover:

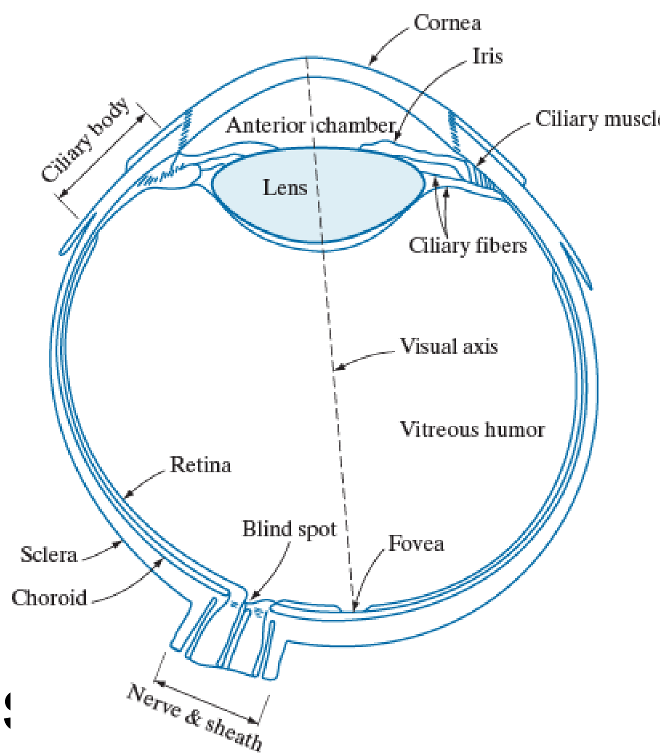
- The human visual system
- Light and the electromagnetic spectrum
- Image representation
- Image sensing and acquisition
- Sampling, quantisation and resolution

Human Visual System

- The best vision model we have!
- Knowledge of how images form in the eye can help us with processing digital images
- We will take just a whirlwind tour of the human visual system

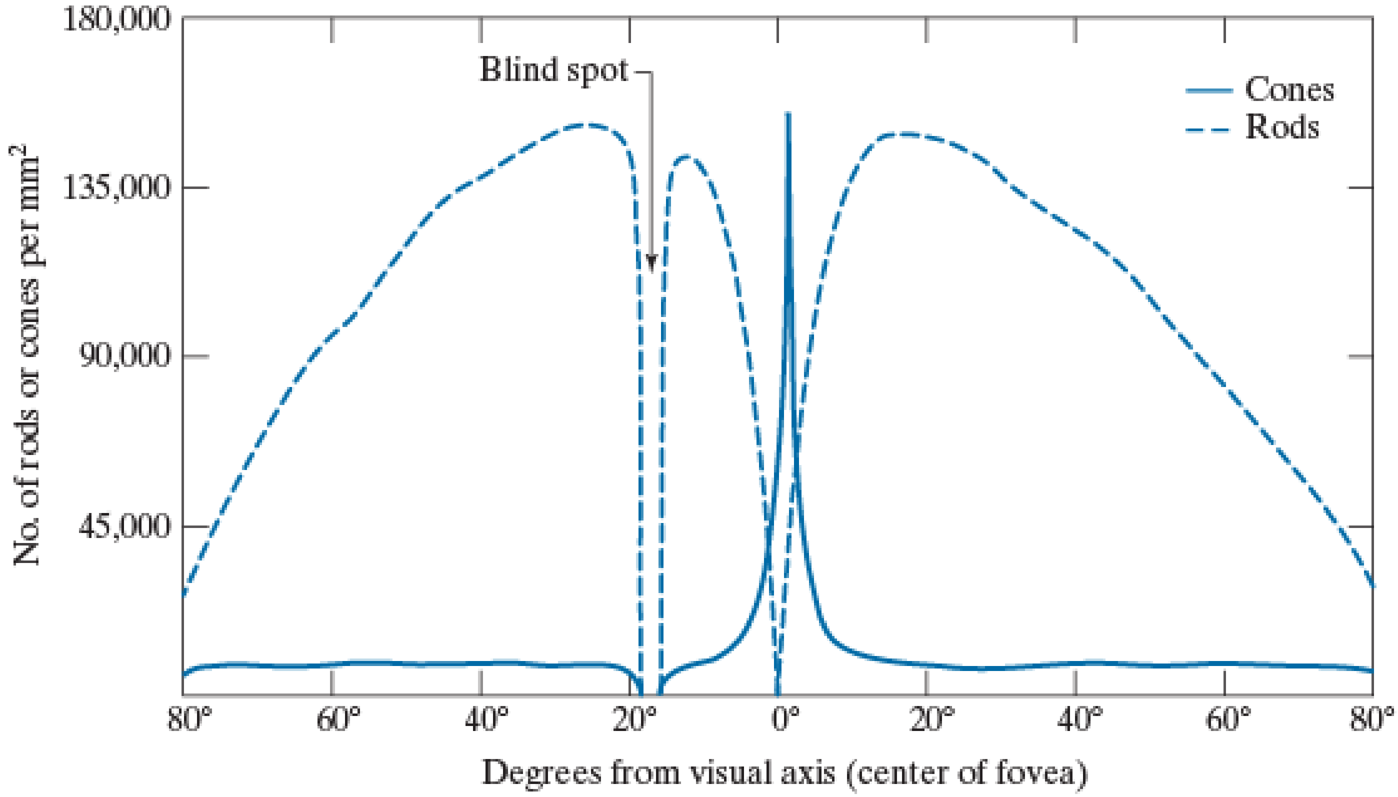
Structure Of The Human Eye

- The lens focuses light from objects onto the retina
- The retina is covered with light receptors called cones (6-7 million) and rods (75-150 million)
- Cones are concentrated around the fovea and are very sensitive to colour
- Rods are more spread out and are sensitive to low levels of illumination



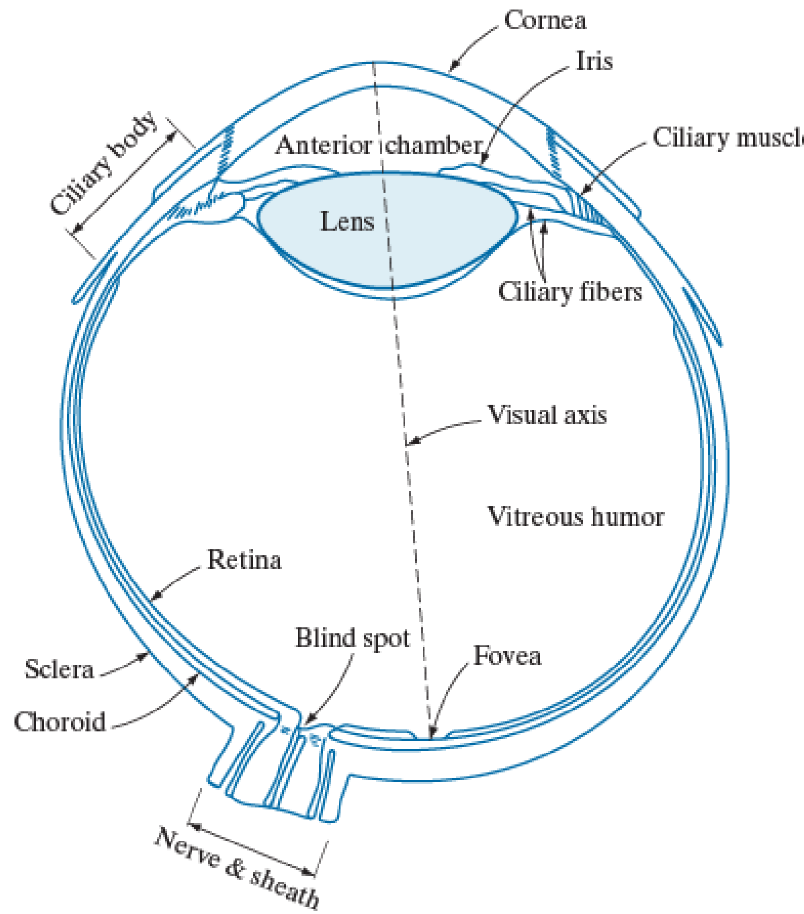
Structure Of The Human Eye (cont.)

Density of cones and rods across a section of the right eye



Structure Of The Human Eye (cont.)

- Each cone is connected to each own nerve end.
  - They can resolve fine details.
  - Sensitive to color (*photopic* vision)
- Many rods are connected to a single nerve end
  - Limited resolution with respect to cones
  - Not sensitive to color
  - Sensitive to low level illumination (*scotopic* vision)



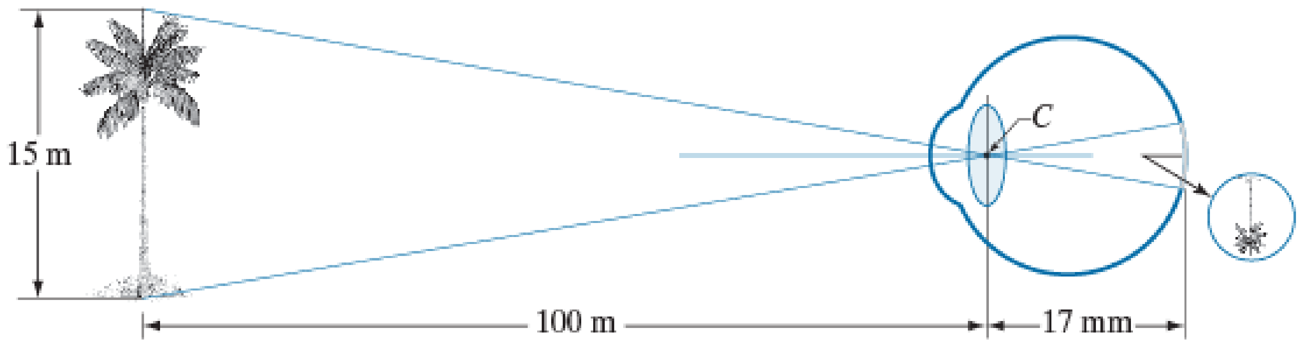
Blind-Spot Experiment

- Draw an image similar to that below on a piece of paper (the dot and cross are about 6 inches apart)

- Close your right eye and focus on the cross with your left eye
- Hold the image about 20 inches away from your face and move it slowly towards you
- The dot should disappear!

Image Formation In The Eye

- Muscles within the eye can be used to change the shape of the lens allowing us focus on objects that are near or far away (in contrast with a camera where the distance between the lens and the focal plane varies)
- An image is focused onto the retina causing rods and cones to become excited which ultimately send signals to the brain



Brightness Adaptation & Discrimination

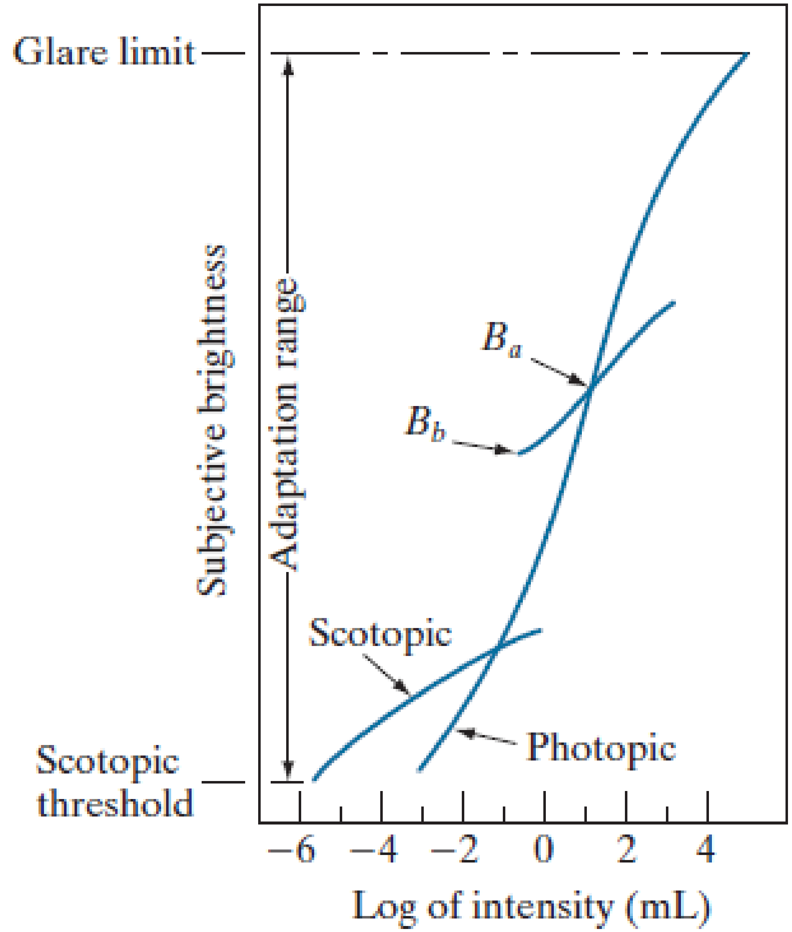
- The human visual system can perceive approximately 10^{10} different light intensity levels.
- At any time instance, we can only discriminate between a much smaller number – *brightness adaptation*.
- Similarly, the perceived intensity of a region is related to the light intensities of the regions surrounding it.

Subjective Brightness Perception

- Subjective brightness perceived by the human visual system, follows a logarithmic function relative to light intensity.
- The human visual system can adapt to a wide range of intensities, approximately 10^6 times, from scotopic to photopic vision.

Subjective Brightness Perception

- The transition from low-light (scotopic) to well-lit (photopic) vision occurs gradually over a range from 0.001 to 0.1 millilambert.
- This transition is depicted by the double branches of the adaptation curve within the specified range.

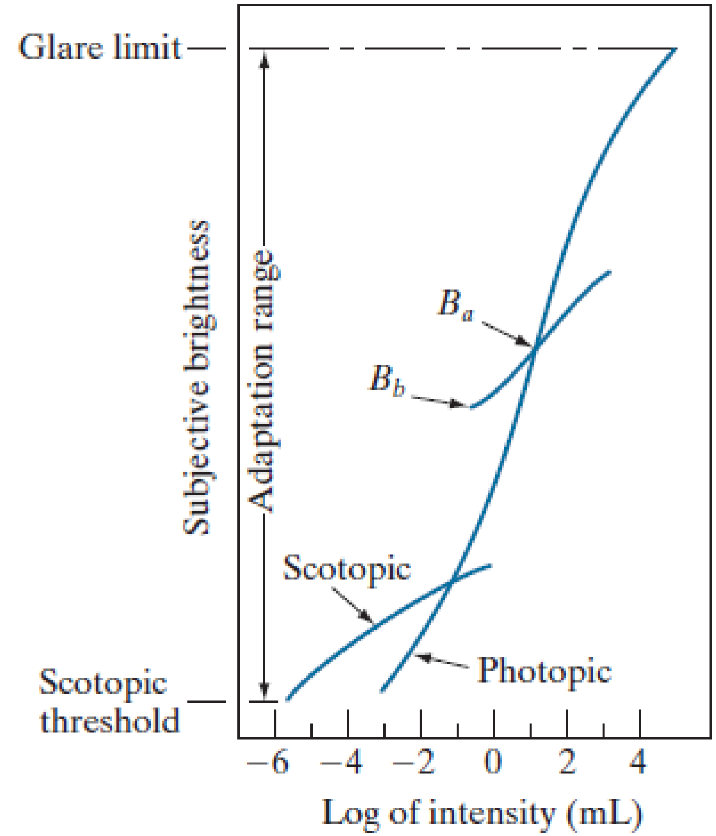


Brightness Adaptation

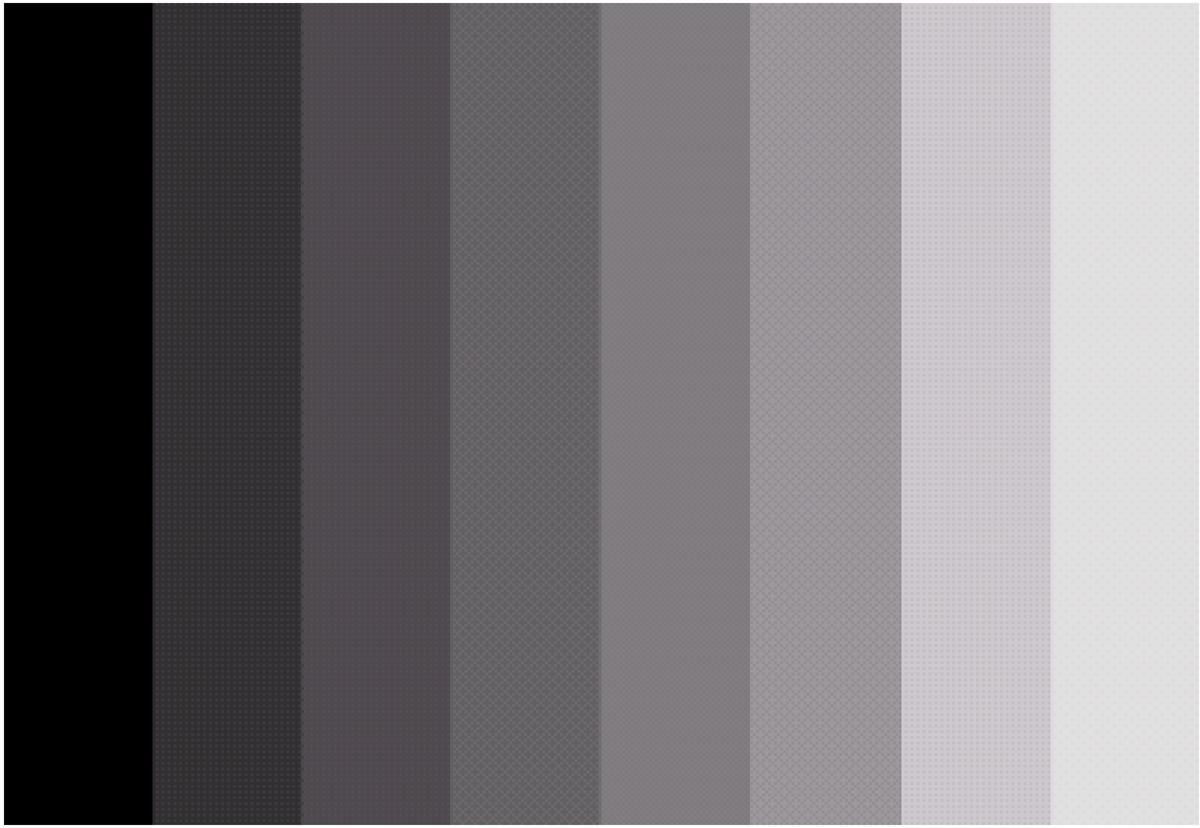
- The key point in interpreting the impressive dynamic range is that the visual system cannot operate over such a range simultaneously.
- Rather, it accomplishes this large variation by changing its overall sensitivity, a phenomenon known as *brightness adaptation*.
- The total range of distinct intensity levels the eye can discriminate simultaneously is rather small when compared with the total adaptation range.

Brightness Adaptation Level

- The current sensitivity level of the visual system under specific conditions.
- This adaptation level (brightness B_a) corresponds to a specific range of subjective brightness perceived by the eye.
- The range of perceived brightness, represented by a short intersecting curve, is limited.
- At the lower end of this range is a level B_b, below which all stimuli are perceived as indistinguishable blacks.
- The upper portion of the curve is not constrained, but extending it too far loses significance as higher intensities would elevate the adaptation level

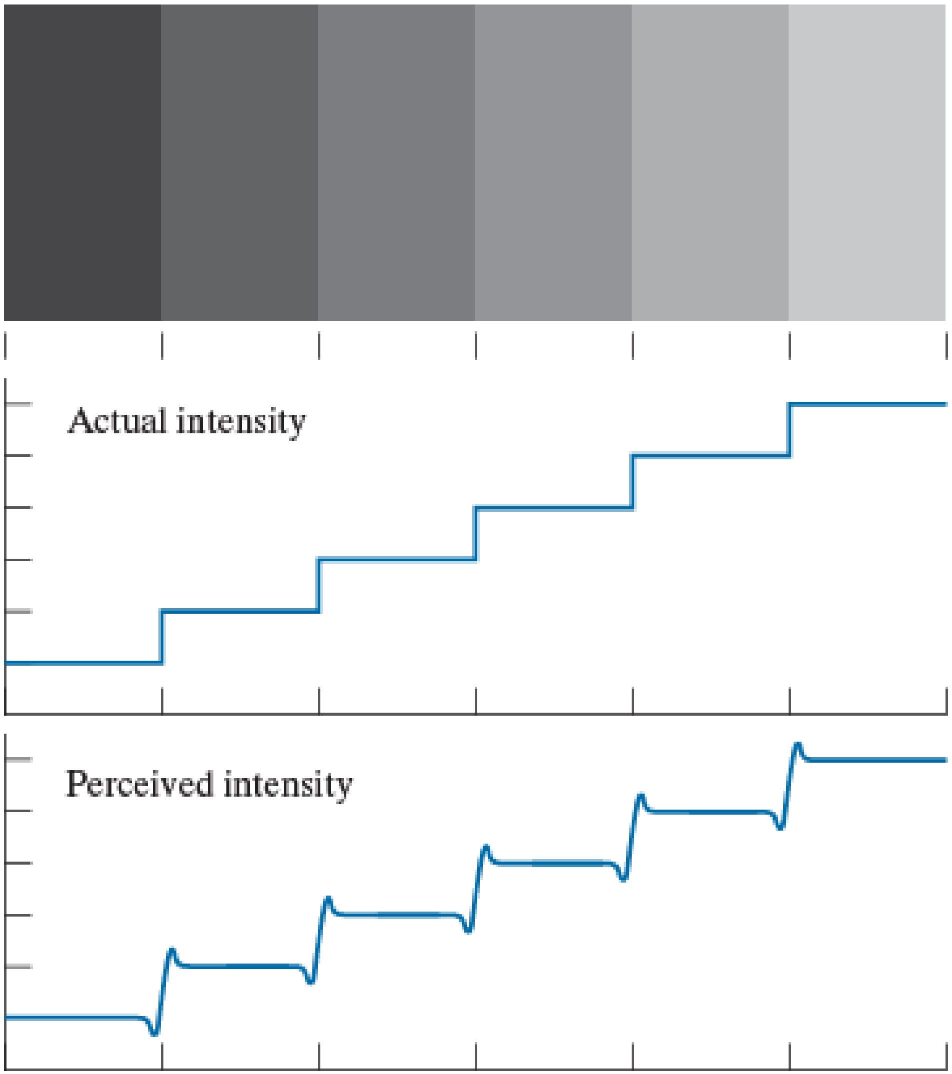


Brightness Adaptation & Discrimination (cont…)

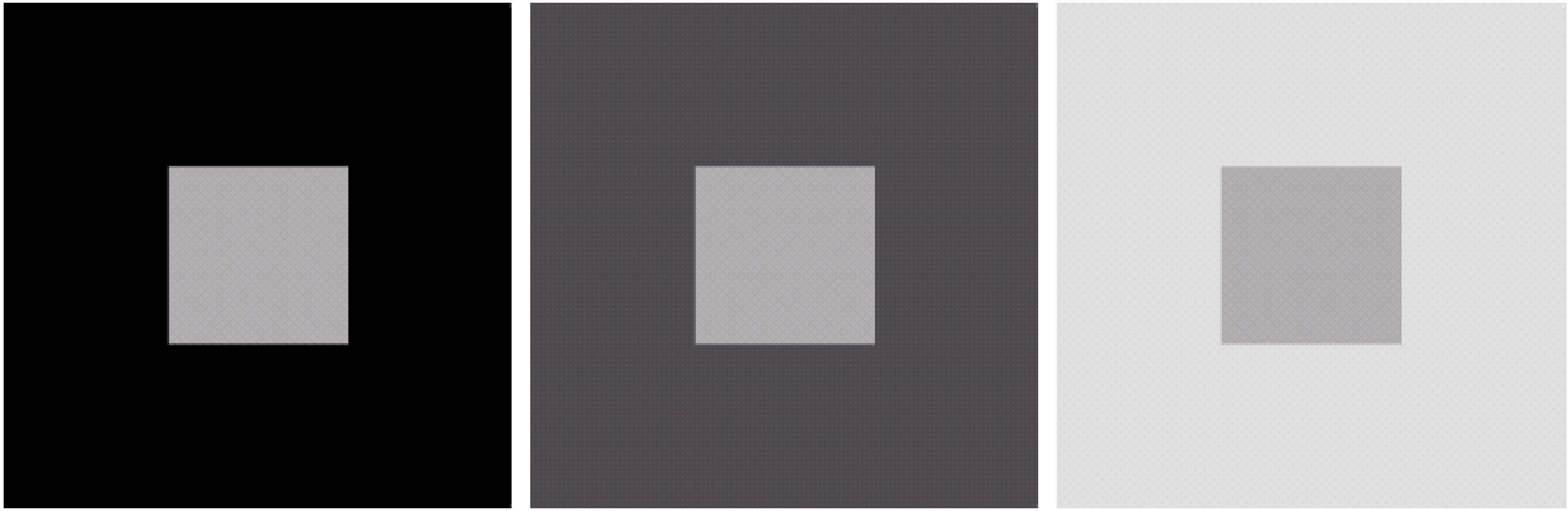


An example of Mach bands

Brightness Adaptation & Discrimination (cont…)



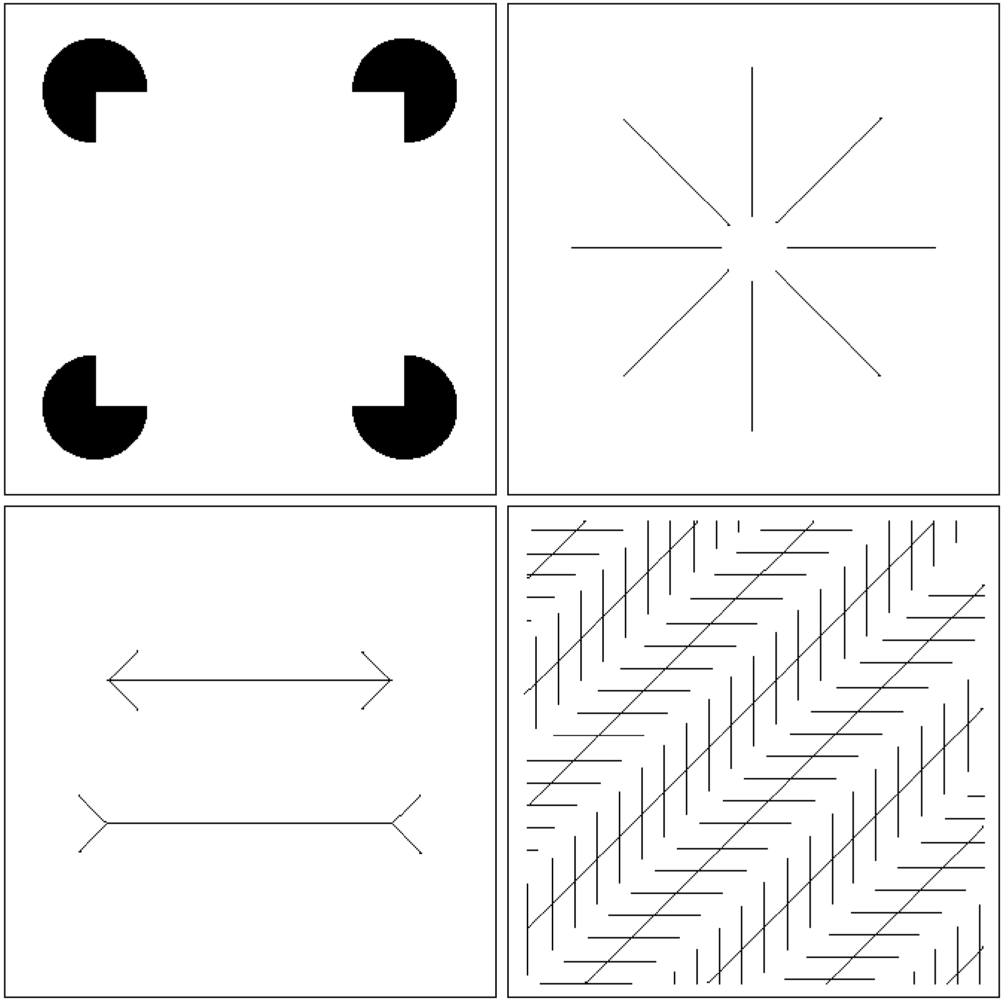
Brightness Adaptation & Discrimination (cont…)



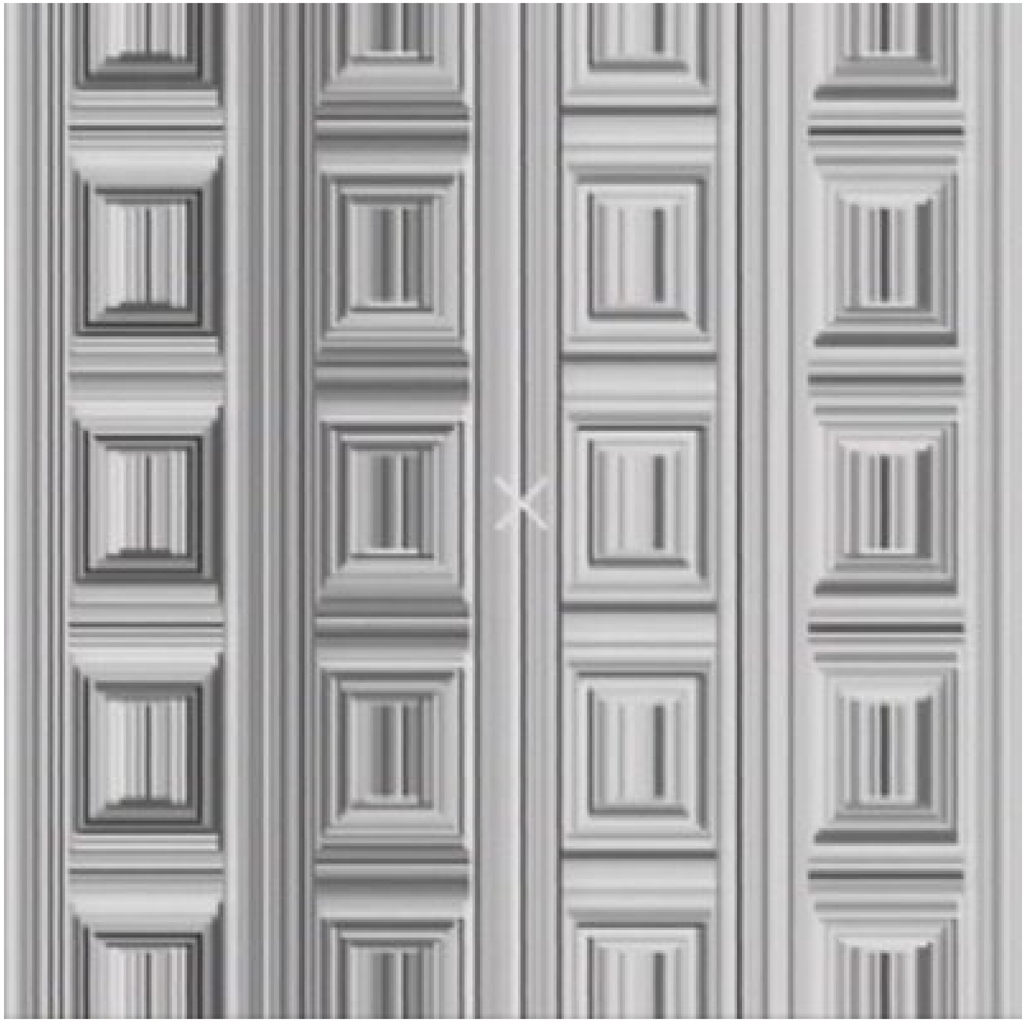
An example of *simultaneous contrast*
(square has constant brightness across shapes)

Optical Illusions

- Our visual system plays many interesting tricks on us

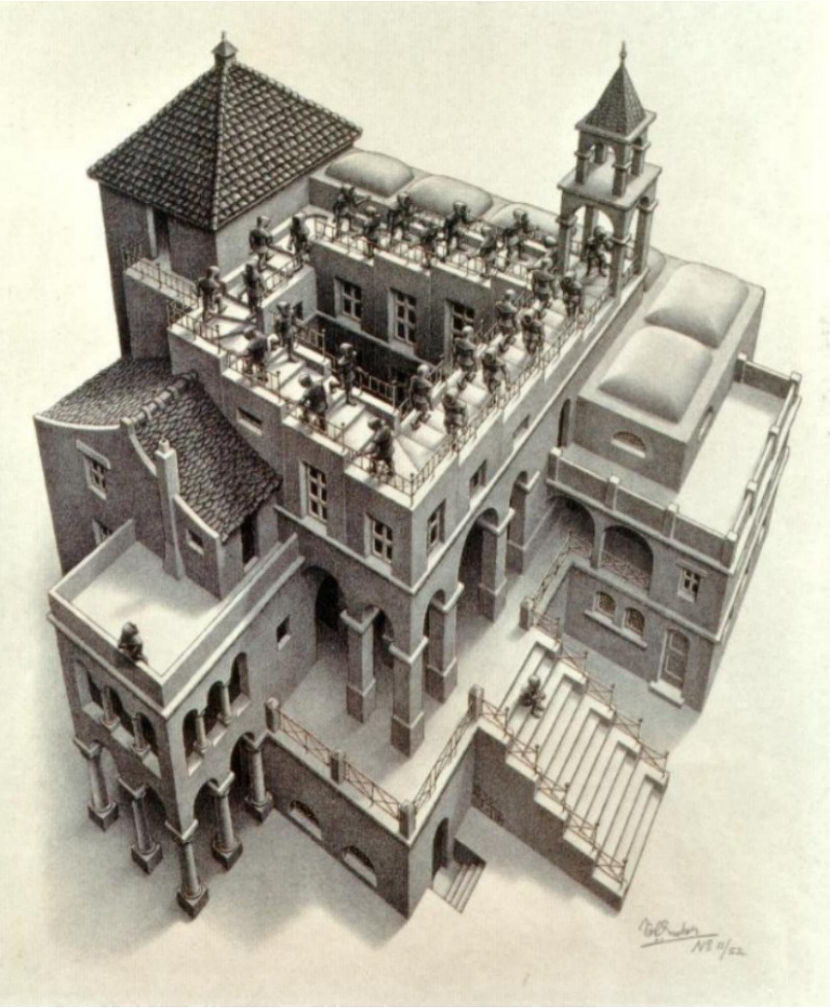


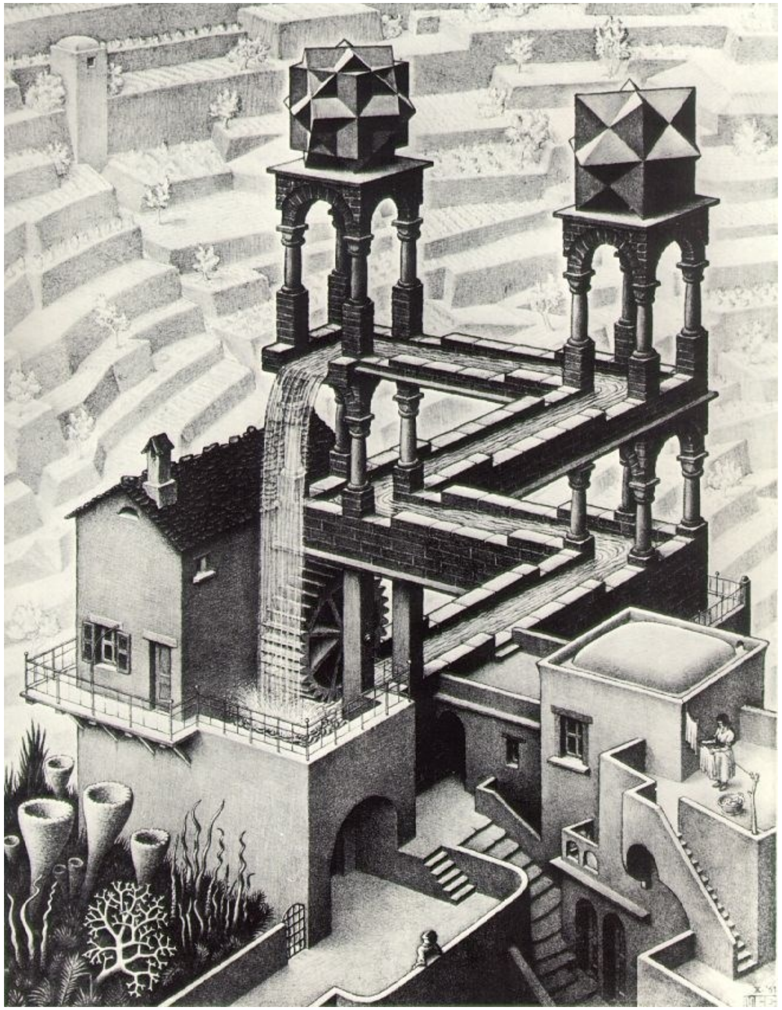
Optical Illusions (cont…)



Stare at the cross in the middle of the image and think circles

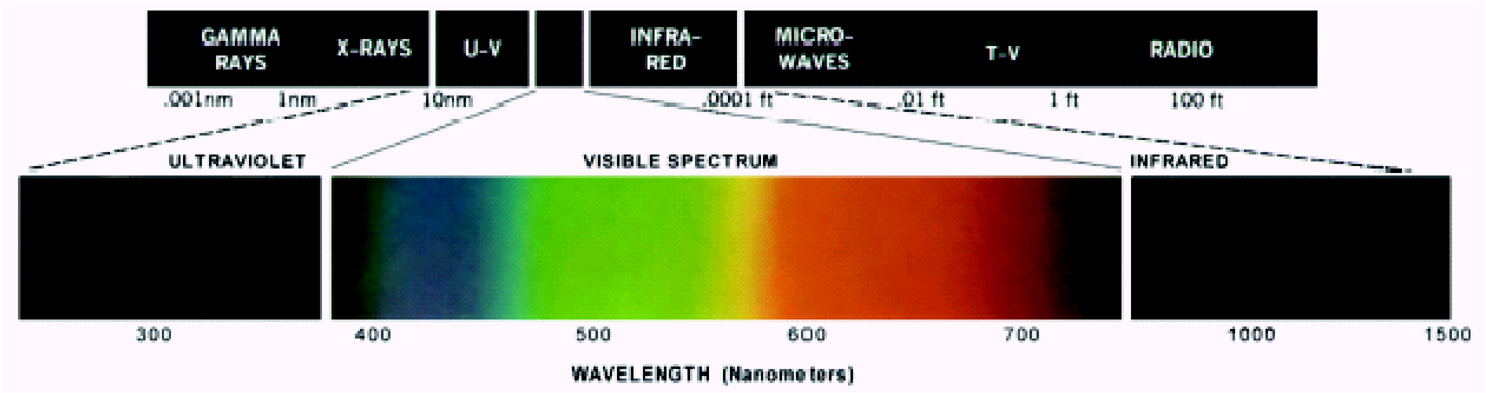
Optical Illusions (cont…)





Light And The Electromagnetic Spectrum

- Light is just a particular part of the electromagnetic spectrum that can be sensed by the human eye
- The electromagnetic spectrum is split up according to the wavelengths of different forms of energy



Reflected Light

- The colours that we perceive are determined by the nature of the light reflected from an object
- For example, if white light is shone onto a green object most wavelengths are absorbed, while green light is reflected from the object

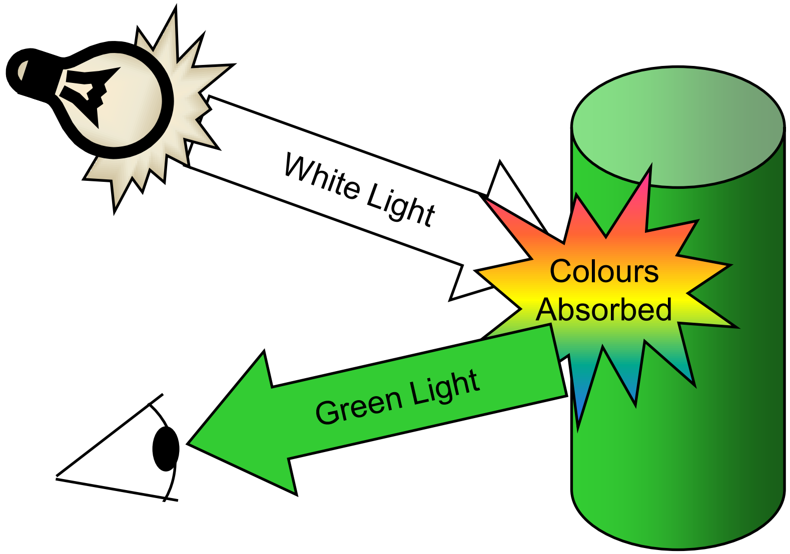


Image Acquisition

Images are typically generated by *illuminating a scene* and absorbing the energy reflected by the objects in that scene

Typical notions of illumination and scene can be way off:

- X-rays of a skeleton
- Ultrasound of an unborn baby
- Electro-microscopic images of molecules

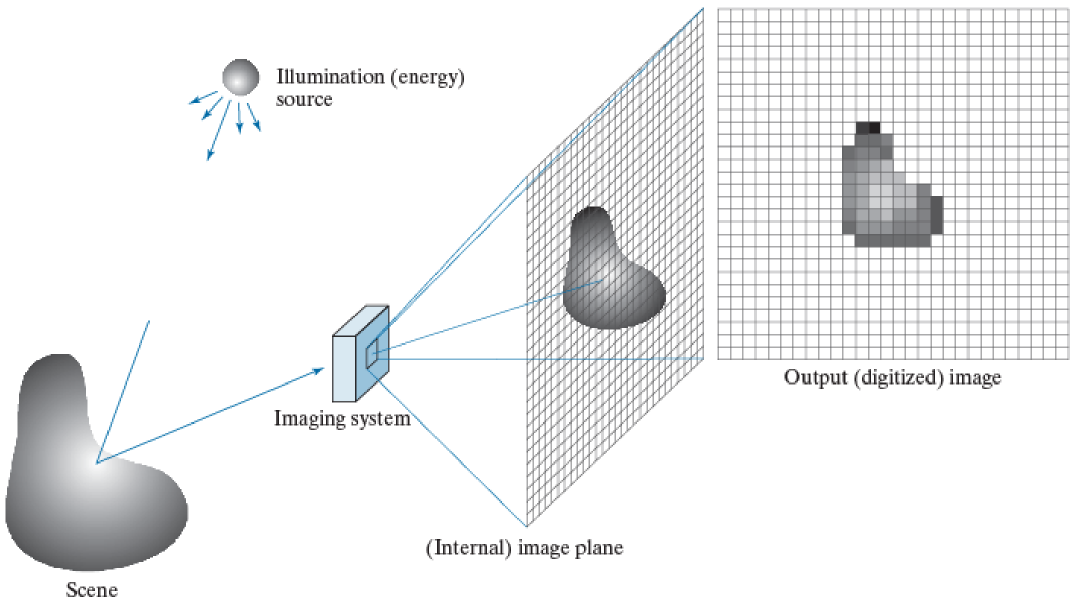
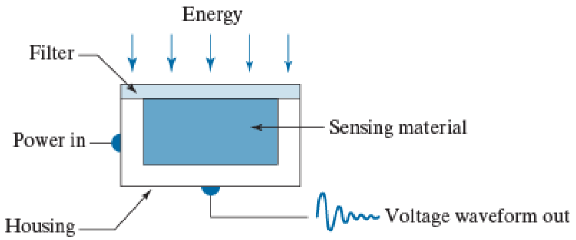
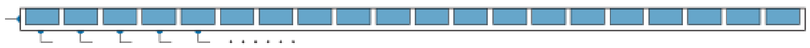


Image Sensing and Acquisition

- Sensors transform the incoming energy into voltage and the output of the sensor is digitized.
- Top: single sensing element
- Middle: Line (of image) sensors
- Bottom: 2D array (of image) sensors





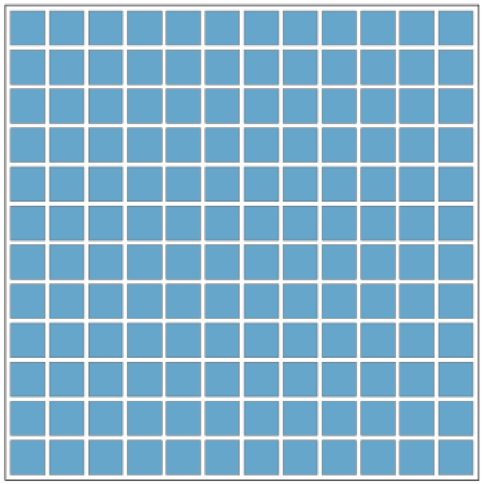


Image Sensing and Acquisition

- Using a filter in front of a sensor enhances its selectivity by favoring specific wavelengths of light, i.e., an optical green-transmission filter emphasizes light within the green band of the color spectrum.
- 2-D image generated by relative displacements in both the x- and y axis between the sensor and the area being imaged.

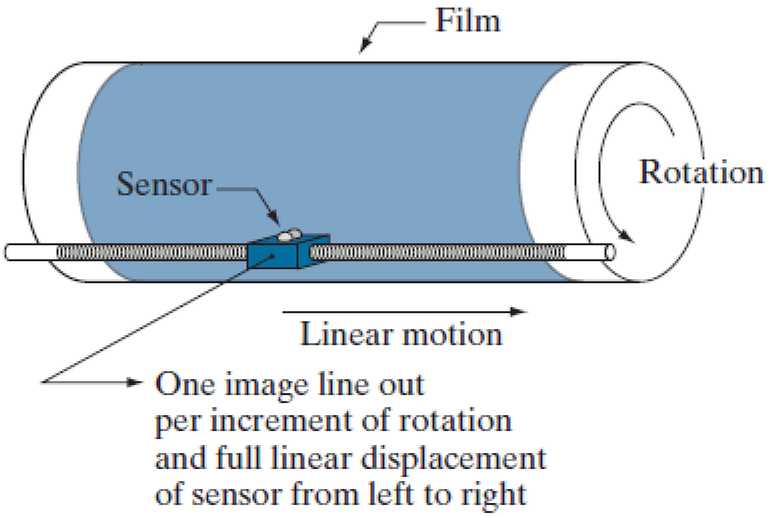


Image Sensing

Using Sensor Strips and Rings

Image acquisition using a linear sensor strip.

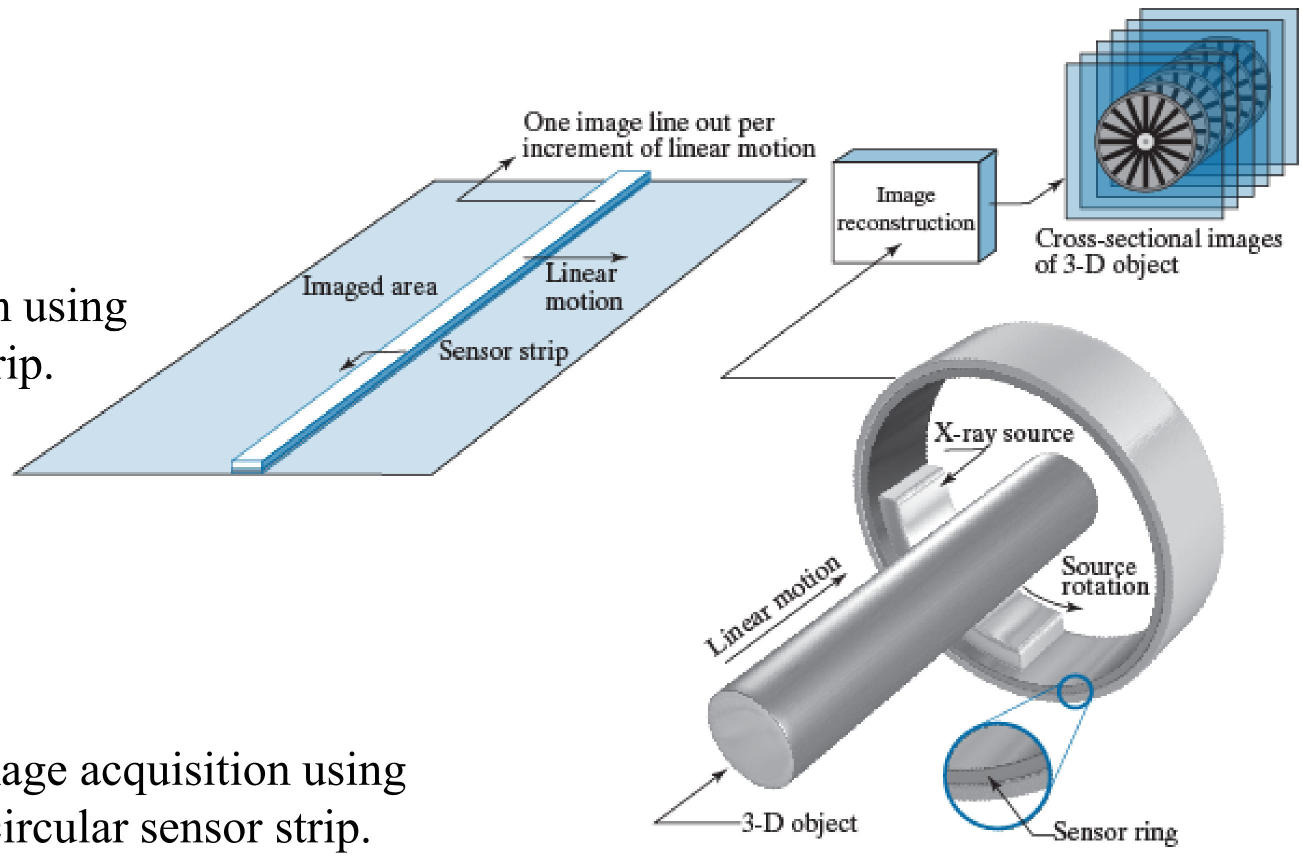
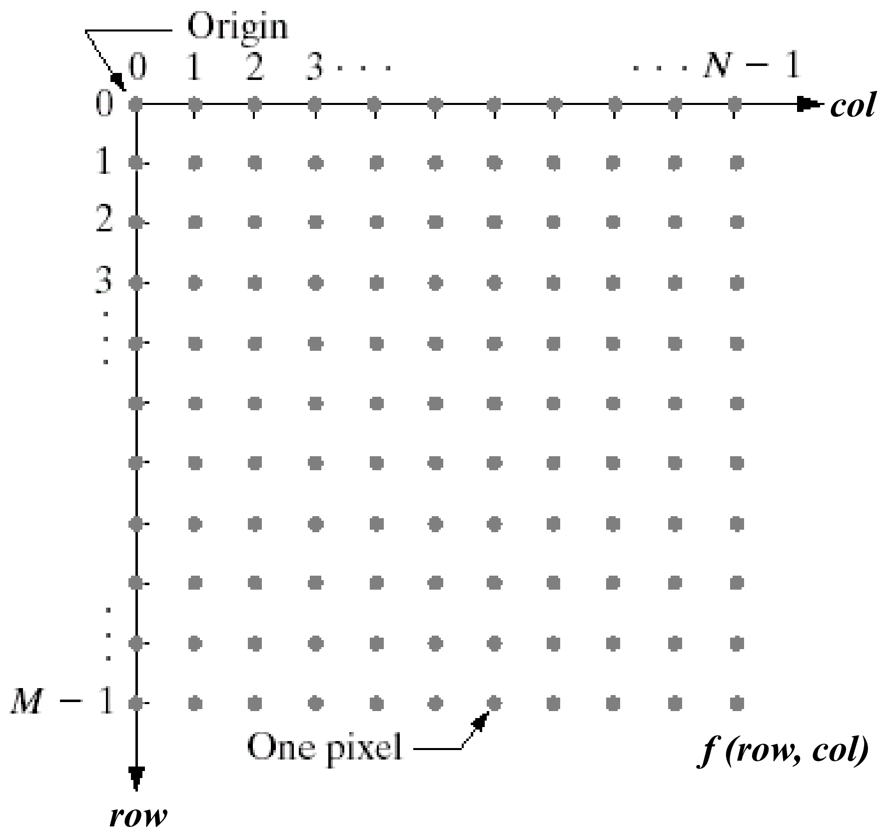


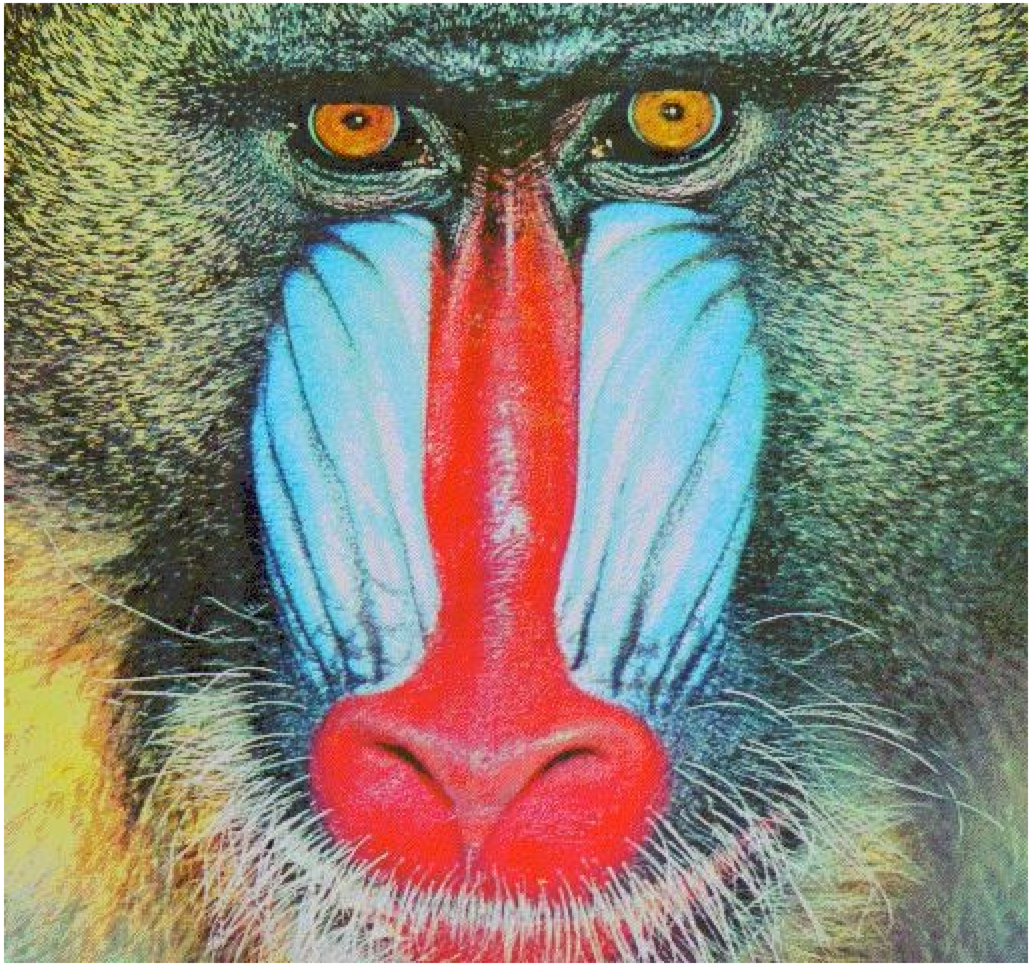
Image acquisition using a circular sensor strip.

Image Representation

- A digital image is composed of M rows and N columns of pixels each storing a value
- Pixel values are in the range 0-255 (black-white)
- Images can easily be represented as matrices



Colour images



Colour images

Image Sampling And Quantisation

- A digital sensor can only measure a limited number of **samples** at a **discrete** set of energy levels
- Quantisation is the process of converting a continuous **analogue** signal into a digital representation of this signal

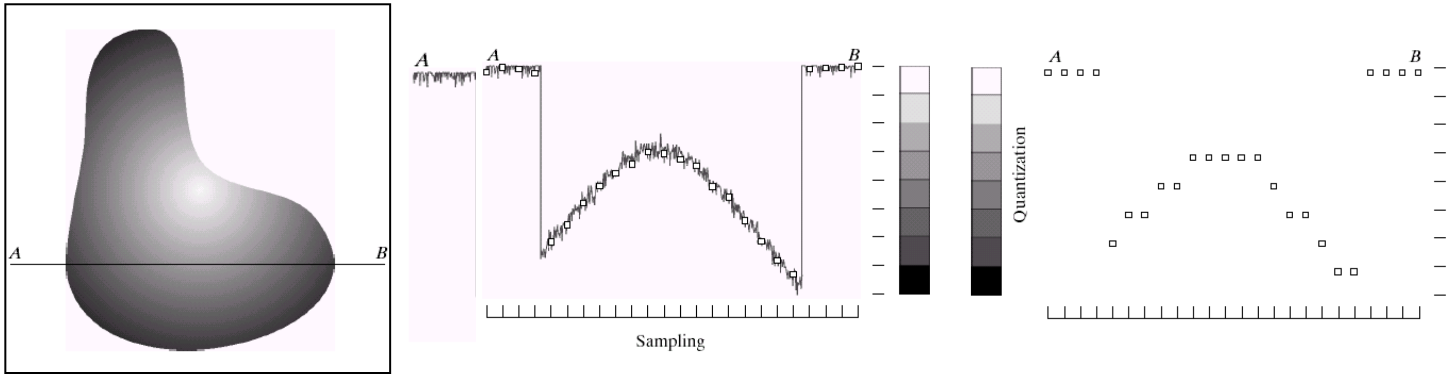


Image Sampling And Quantisation (cont…)

- Remember that a digital image is always only an **approximation** of a real world scene

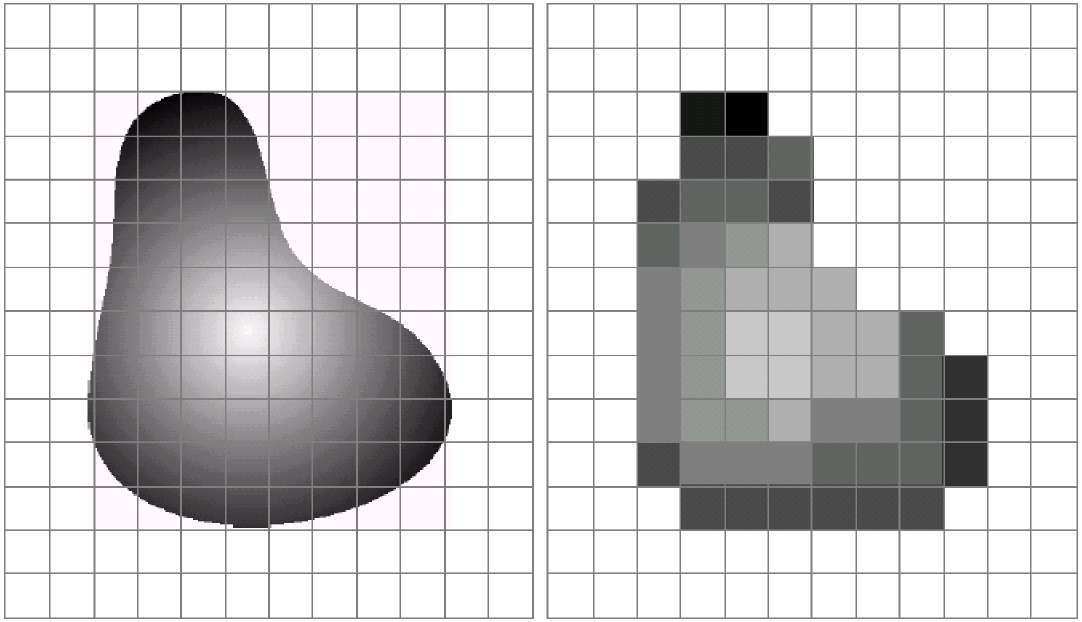
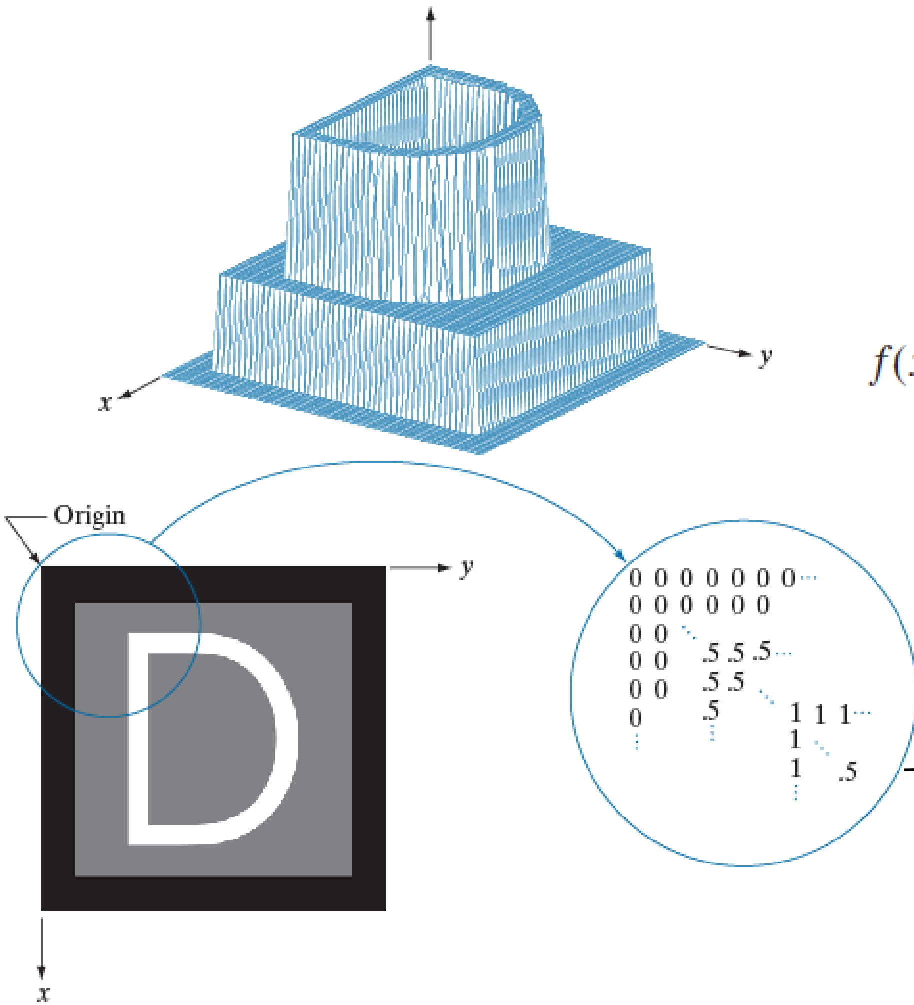


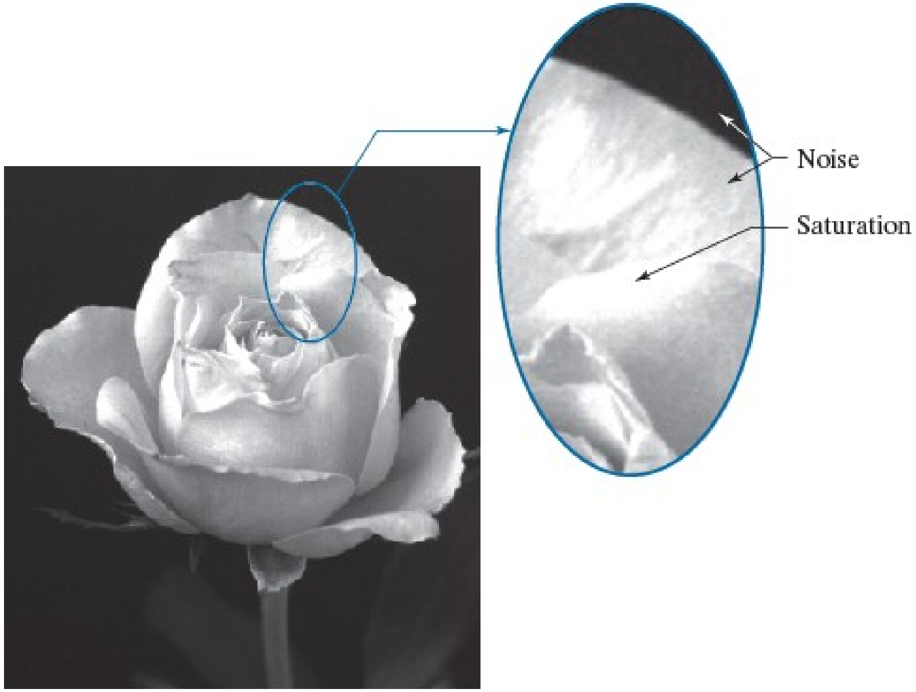
Image Representation



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

Saturation & Noise

- **Dynamic range**: The ratio of the maximum (*saturation*) to the minimum (*noise*) detectable intensity of the imaging system.
- Noise generally appear as a grainy texture pattern in the darker regions and masks the lowest detectable true intensity level



Spatial Resolution

- The spatial resolution of an image is determined by how sampling was carried out
- Spatial resolution simply refers to the smallest discernable detail in an image
  - Vision specialists will often talk about pixel size
  - Graphic designers will talk about dots per inch (DPI)

Spatial Resolution (cont…)

Spatial Resolution (cont…)

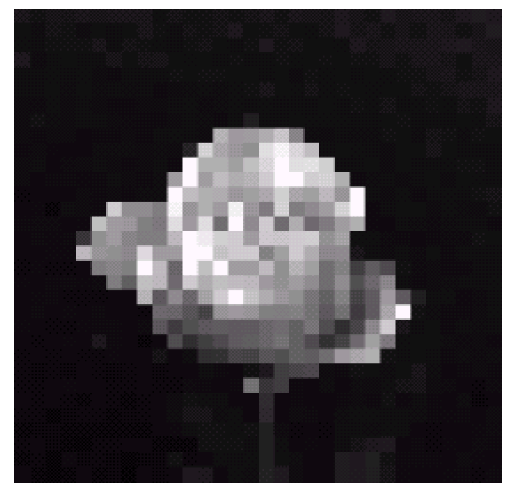
1024 * 1024

512 * 512

256 * 256

128 * 128

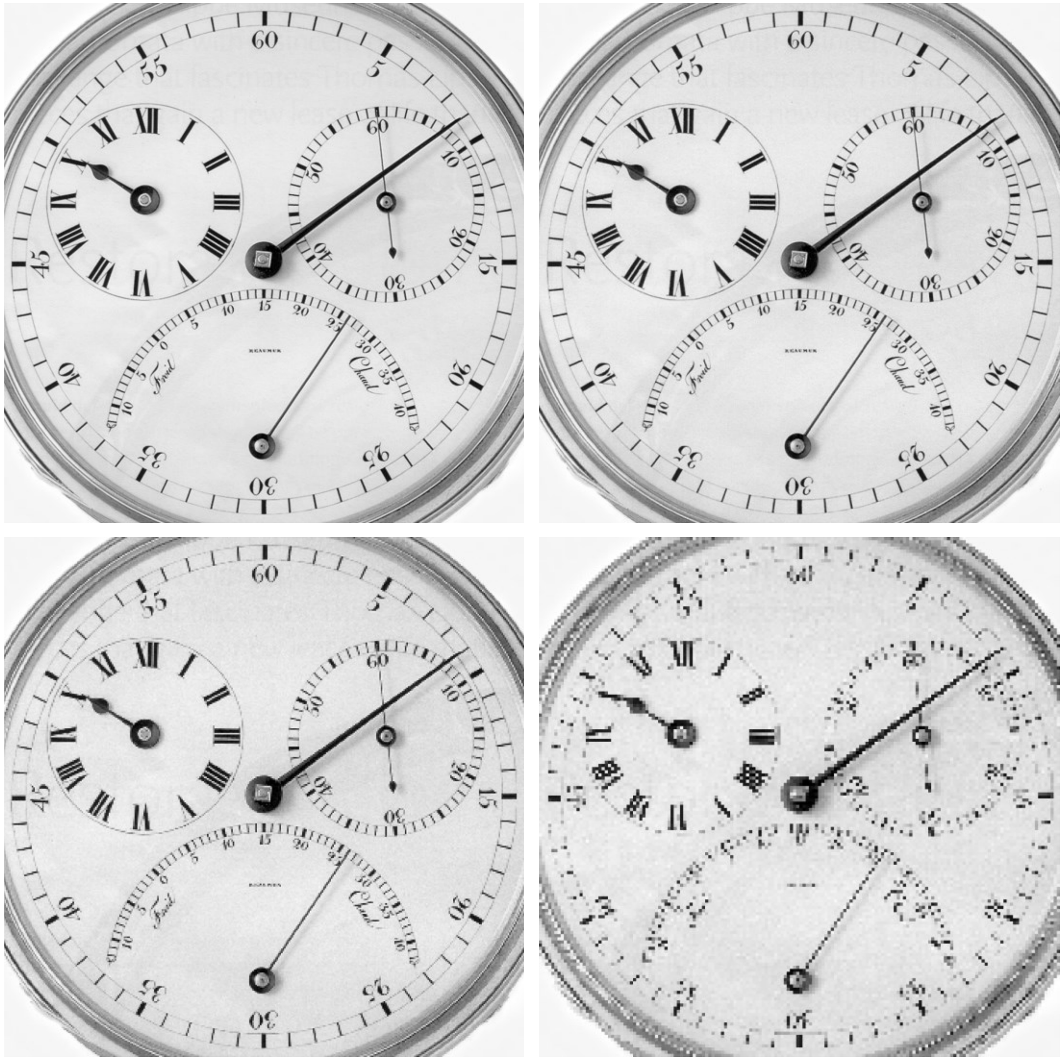
64 * 64



32 * 32

Spatial Resolution (cont…)

- Effects of reducing spatial resolution. The images shown are at:
  (a) 930 dpi,
  (b) 300 dpi,
  (c) 150 dpi, and
  (d) 72 dpi.



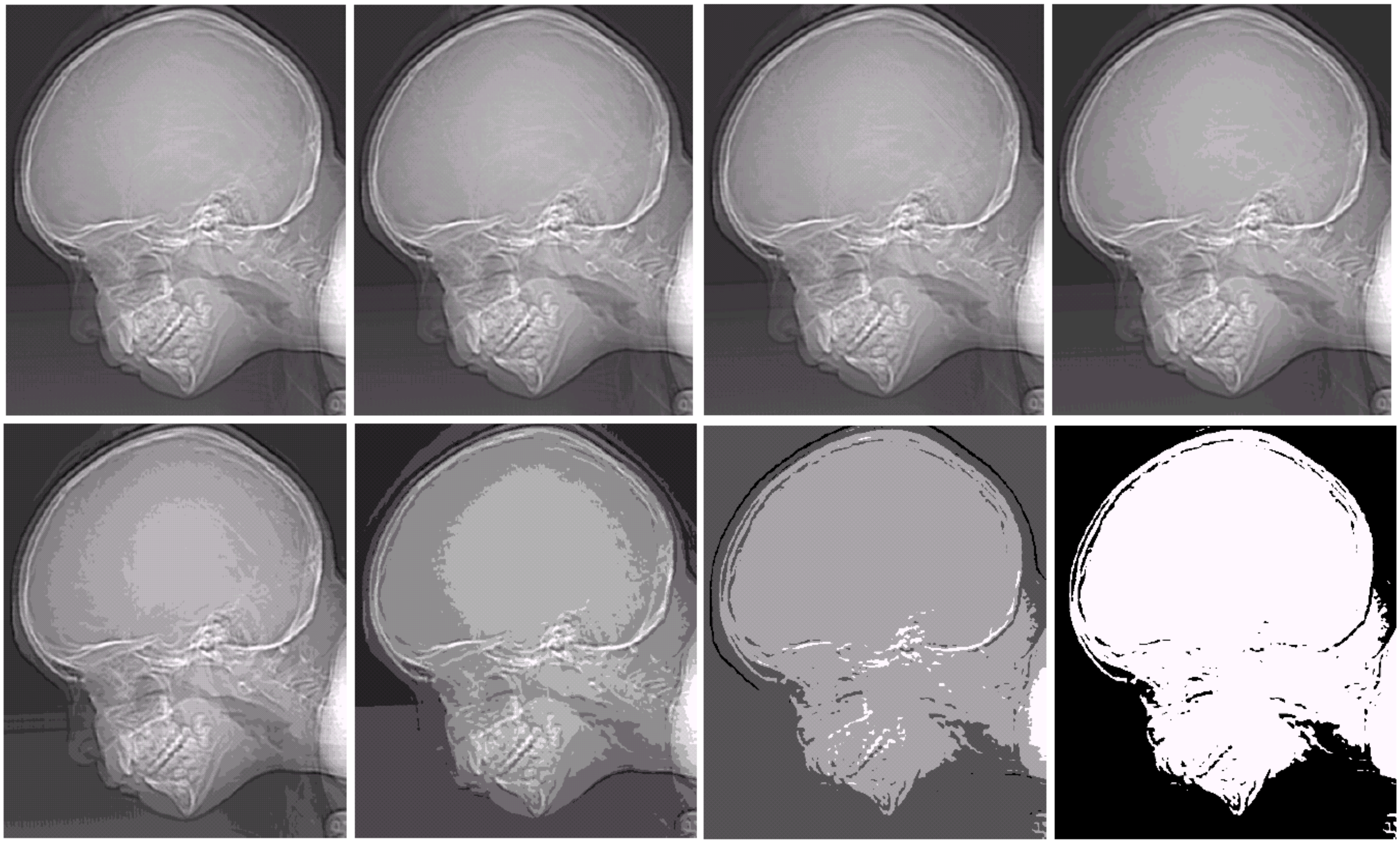
Intensity Level Resolution

- Intensity level resolution refers to the number of intensity levels used to represent the image
  - The more intensity levels used, the finer the level of detail discernable in an image
  - Intensity level resolution is usually given in terms of the number of bits used to store each intensity level

Number of Bits	Number of Intensity Levels	Examples
1	2	0, 1
2	4	00, 01, 10, 11
4	16	0000, 0101, 1111
8	256	00110011, 01010101
16	65,536	1010101010101010

Intensity Level Resolution (cont…)

256 grey levels (8 bits per pixel) 128 grey levels (7 bpp) 64 grey levels (6 bpp) 32 grey levels (5 bpp)



16 grey levels (4 bpp) 8 grey levels (3 bpp) 4 grey levels (2 bpp) 2 grey levels (1 bpp)

Intensity Level Resolution (cont…)

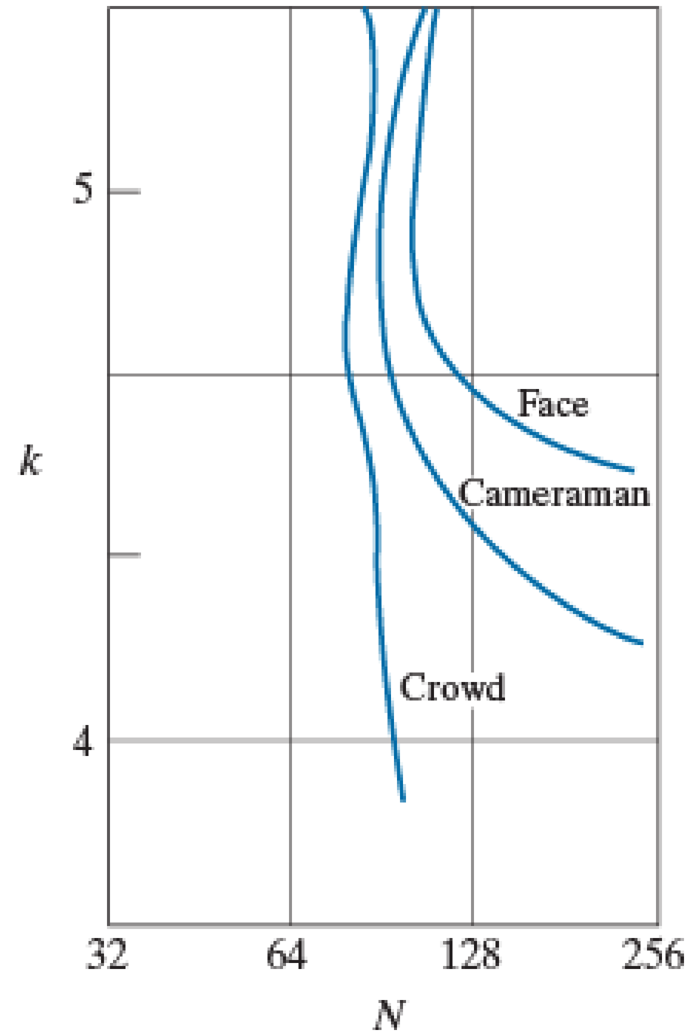
Low Detail

Medium Detail

High Detail

Intensity Level Resolution (cont…)

- $b = N^2 k$
- *Isopreference* curves represent the dependence between intensity and spatial resolutions.
  - Points lying on a curve represent images of "equal" quality as described by observers.
  - The curves become more vertical as the degree of detail increases (a lot of detail need less intensity levels).



Resolution: How Much Is Enough?

The big question with resolution is always *how much is enough*?

- This all depends on what is in the image and what you would like to do with it
- Key questions include
  - Does the image look aesthetically pleasing?
  - Can you see what you need to see within the image?

Resolution: How Much Is Enough? (cont…)

The picture on the right is fine for counting the number of cars, but not for reading the number plate

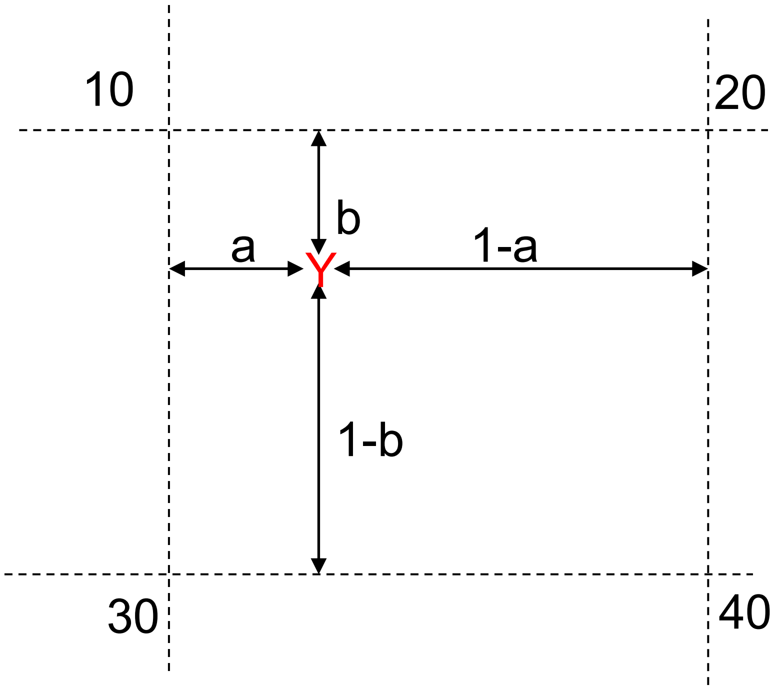
Interpolation

- The process of using known data to estimate values at unknown locations
- Basic operation for shrinking, zooming, rotation and translation
  - e.g. a 500x500 image has to be enlarged by 1.5 to 750x750 pixels
  - Create an imaginary 750x750 grid with the same pixel spacing as the original and then shrink it to 500x500
  - The 750x750 shrunk pixel spacing will be less than the spacing in the original image.
  - Pixel values have to be determined in between the original pixel locations

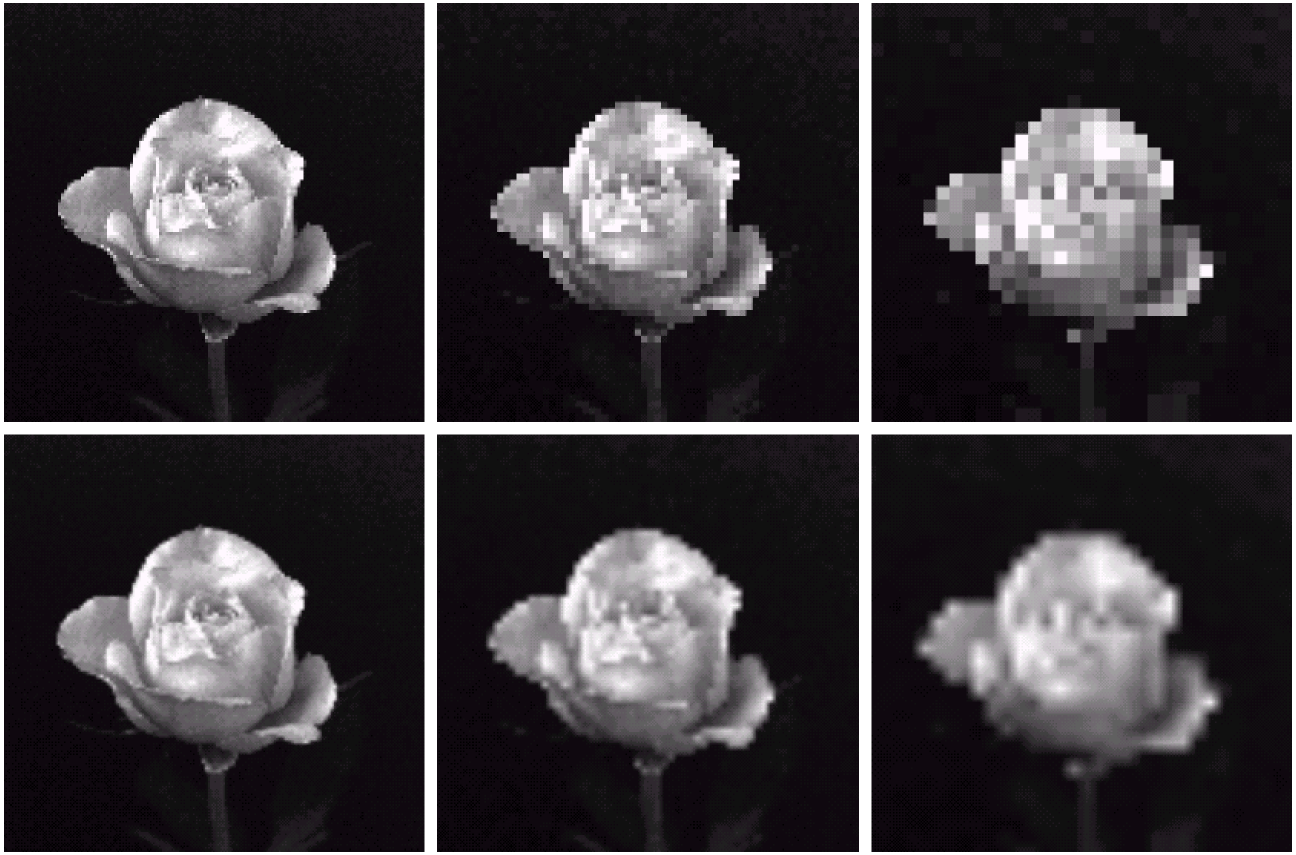
Interpolation (cont.)

- How to determine pixel values
  - Nearest neighbour
  - Bilinear
  - Bicubic
  - 2D sinc
- Nearest Neighbour
  - Whats is the value of Y;

10



Interpolation (cont...)



a b c
d e f

FIGURE 2.25 Top row: images zoomed from 128 × 128, 64 × 64, and 32 × 32 pixels to 1024 × 1024 pixels, using nearest neighbor gray-level interpolation. Bottom row: same sequence, but using bilinear interpolation.

Distances between pixels

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

$$a)\, D(p,q) \geq 0 \ (D(p,q) = 0 \ iff \ \ p = q),$$
$$b)\, D(p,q) = D(q,p),$$
$$c)\, D(p,z) \leq D(p,q) + D(q,z).$$

- The Euclidean distance between p and q is defined as:

$$D_e(p,q) = \left[(x-s)^2 + (y-t)^2 \right]^{\frac{1}{2}}$$

Distances between pixels (cont.)

- The city-block (Manhattan) or D_4 distance between p and q is defined as:

$$D_4(p,q) = |x-s| + |y-t|$$

- Pixels having the city-block distance from a pixel (x,y) less than or equal to some value T form a diamond centered at (x,y) . For example, for $T=2$:

		2		
	2	**1**	**2**	
2	**1**	**0**	**1**	**2**
	2	**1**	**2**	
		2		

Distances between pixels (cont.)

- The chessboard or D_8 distance between p and q is defined as:

$$D_8(p,q) = \max(|x-s|,|y-t|)$$

- Pixels having the D_8 distance from a pixel (x,y) less than or equal to some value T form a square centered at (x,y). For example, for $T=2$:

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**

Mathematical operations used in digital image processing

- Arithmetic operations (e.g image subtraction pixel by pixel)
- Matrix and vector operations
- Linear (e.g. sum) and nonlinear operations (e.g. min and max)
- Set and logical operations
- Spatial and neighbourhood operations (e.g. local average)
- Geometric spatial transformations (e.g. rotation)

Image subtraction

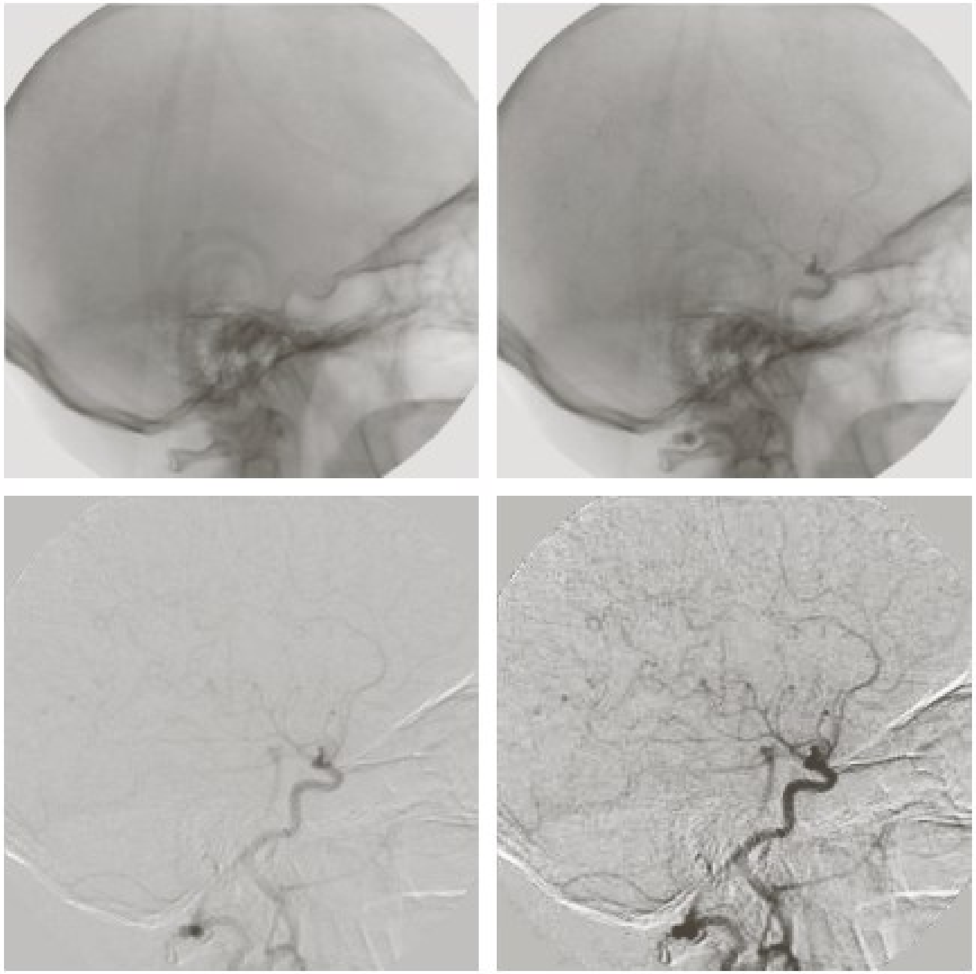


FIGURE 2.28 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

Image multiplication

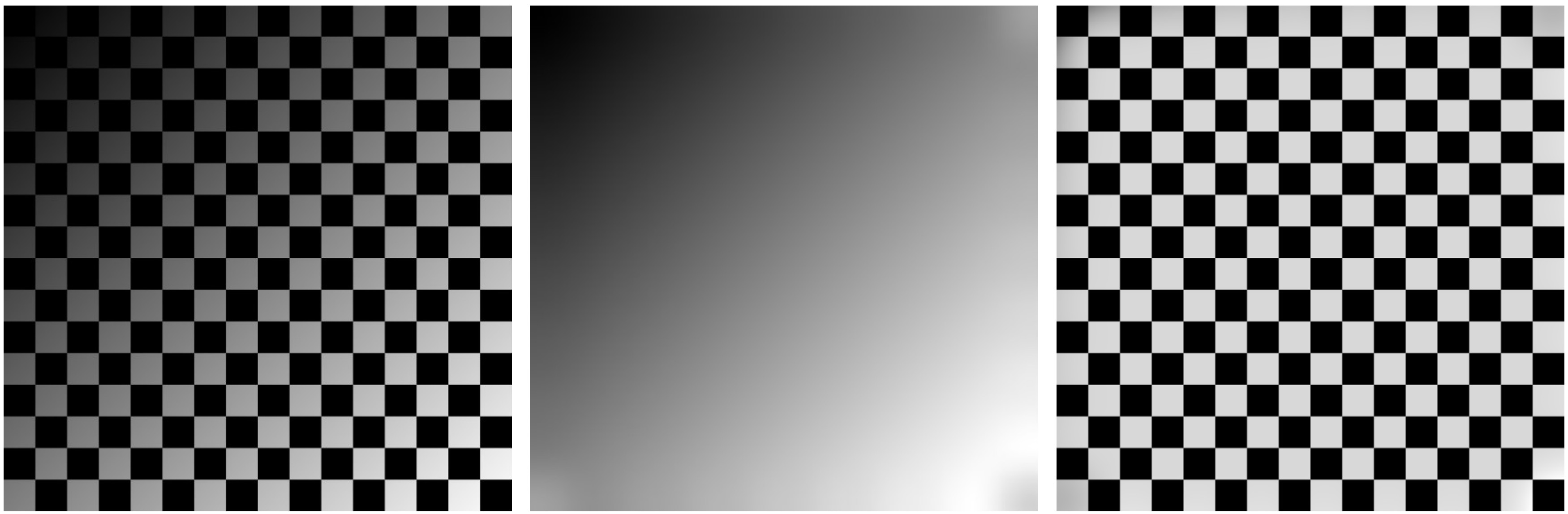


FIGURE Shading correction. (a) Shaded test pattern. (b) Estimated shading pattern. (c) Product of (a) by the reciprocal of (b). (See Section 3.5 for a discussion of how (b) was estimated.)

Image multiplication (cont.)

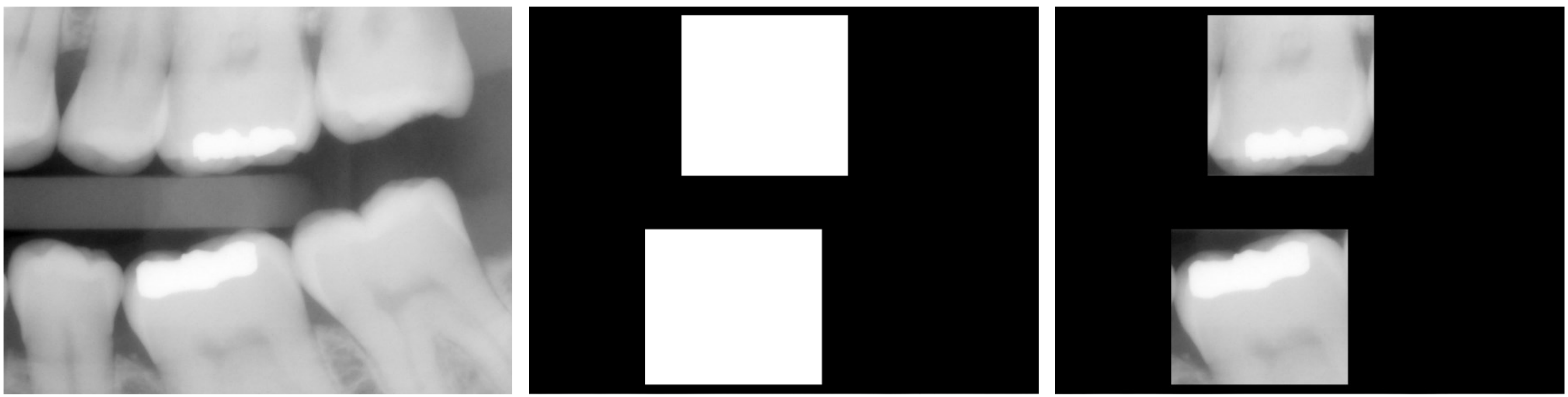
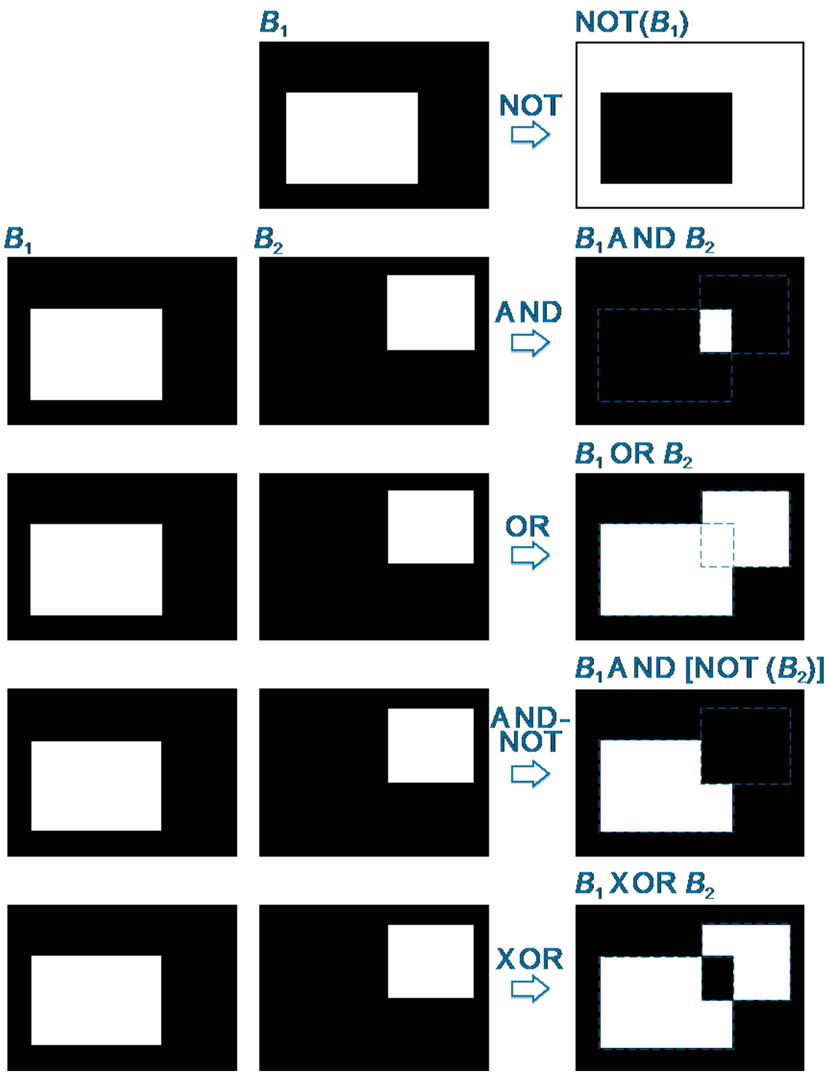
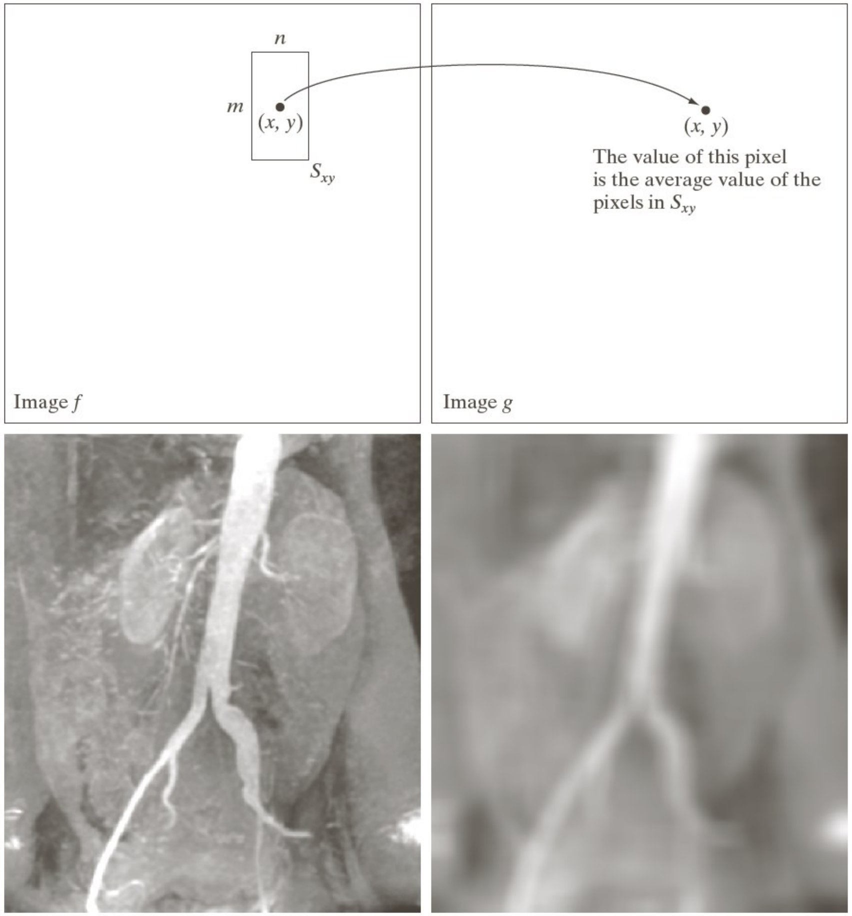


FIGURE (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

Logical operator



Neighbourhood operation



A note on arithmetic operations

- Most images are displayed at 8 bits (0-255).
- When images are saved in standard formats like TIFF or JPEG the conversion to this range is automatic.
- However, the approach used for the conversion depends on the software package.
  - The difference of two images is in the range [-255, 255] and the sum is in the range [0, 510].
  - Many packages simply set all negative values to 0 and all values exceeding 255 to 255 which is undesirable.

A note on arithmetic operations (normalization)

- An approach that guarantees that the full range is captured into a fixed number of bits is the following:
- At first, make the minimum value of the image equal to zero:

$$f_m = f - \min(f)$$

- Then perform intensity scaling to $[0, K]$

$$f_s = \frac{f_m}{\max(f_m)} K$$

Geometric spatial transformations

- A common geometric transformation is the *affine* transform

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

- It can be used to express transformations such as rotate, scale and sheer an image depending on the value of the elements of $\mathbf{T}$,
- except translation, which would require that a constant 2-D vector be added to the right side of the equation.

Geometric spatial transformations

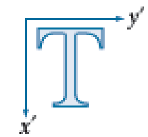
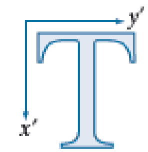
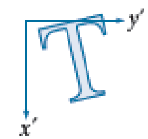
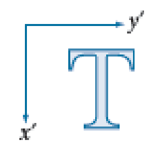
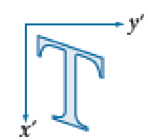
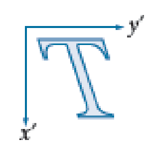
- However, it is possible to use homogeneous coordinates to express all four affine transformations using a single 3 × 3 matrix in the following general form:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2\text{-}45)$$

- To avoid empty pixels we implement the inverse mapping
- Interpolation is essential

Geometric spatial transformations (cont.)

TABLE 2.3 Affine transformations based on Eq. (2-45).

Transformation Name	Affine Matrix, A	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
Scaling/Reflection (For reflection, set one scaling factor to −1 and the other to 0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
Rotation (about the origin)	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x\cos\theta - y\sin\theta$ $y' = x\sin\theta + y\cos\theta$	
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

Geometric spatial transformations (cont.)

- The effects and importance of interpolation in image transformations

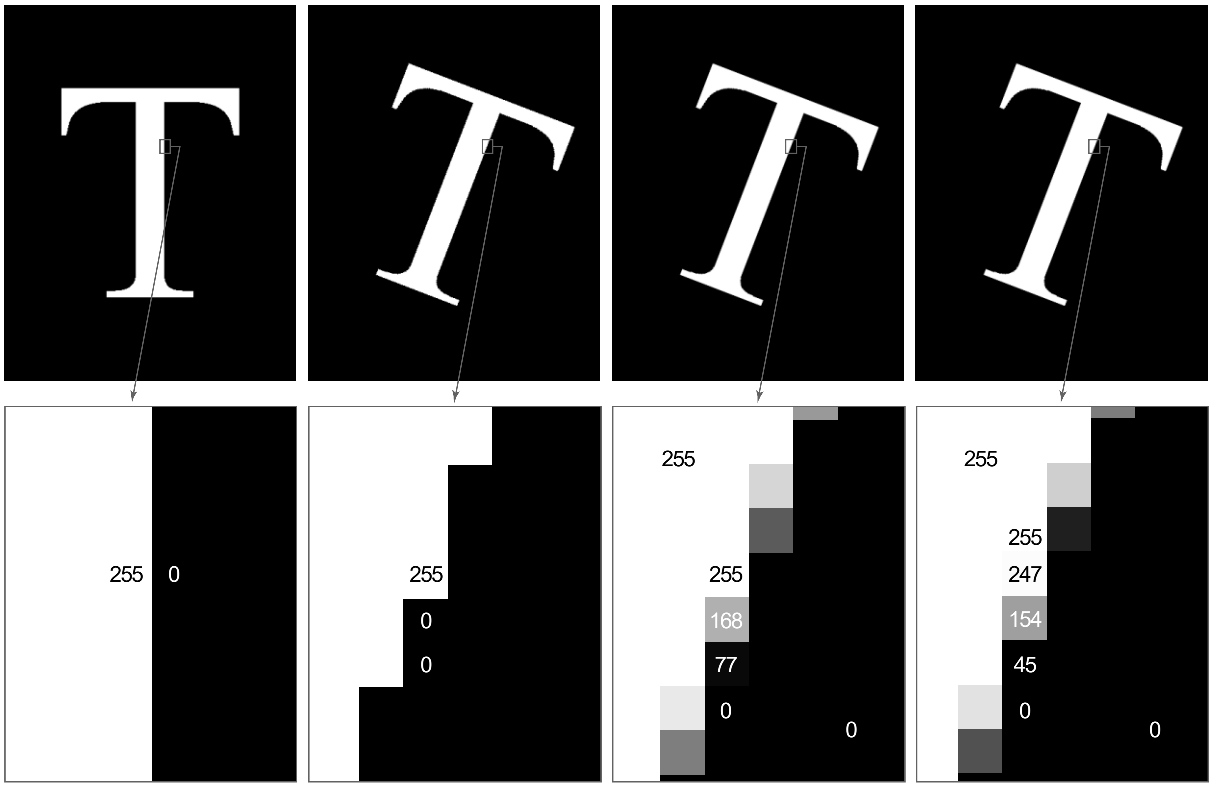


Image Registration

- Estimate the transformation parameters between two images.
- Very important application of digital image processing.
  - Single and multimodal
  - Temporal evolution and quantitative analysis (medicine, satellite images)
- A basic approach is to use control points (user defined or automatically detected) and estimate the elements of the transformation matrix by solving a linear system.

Image Registration (cont.)

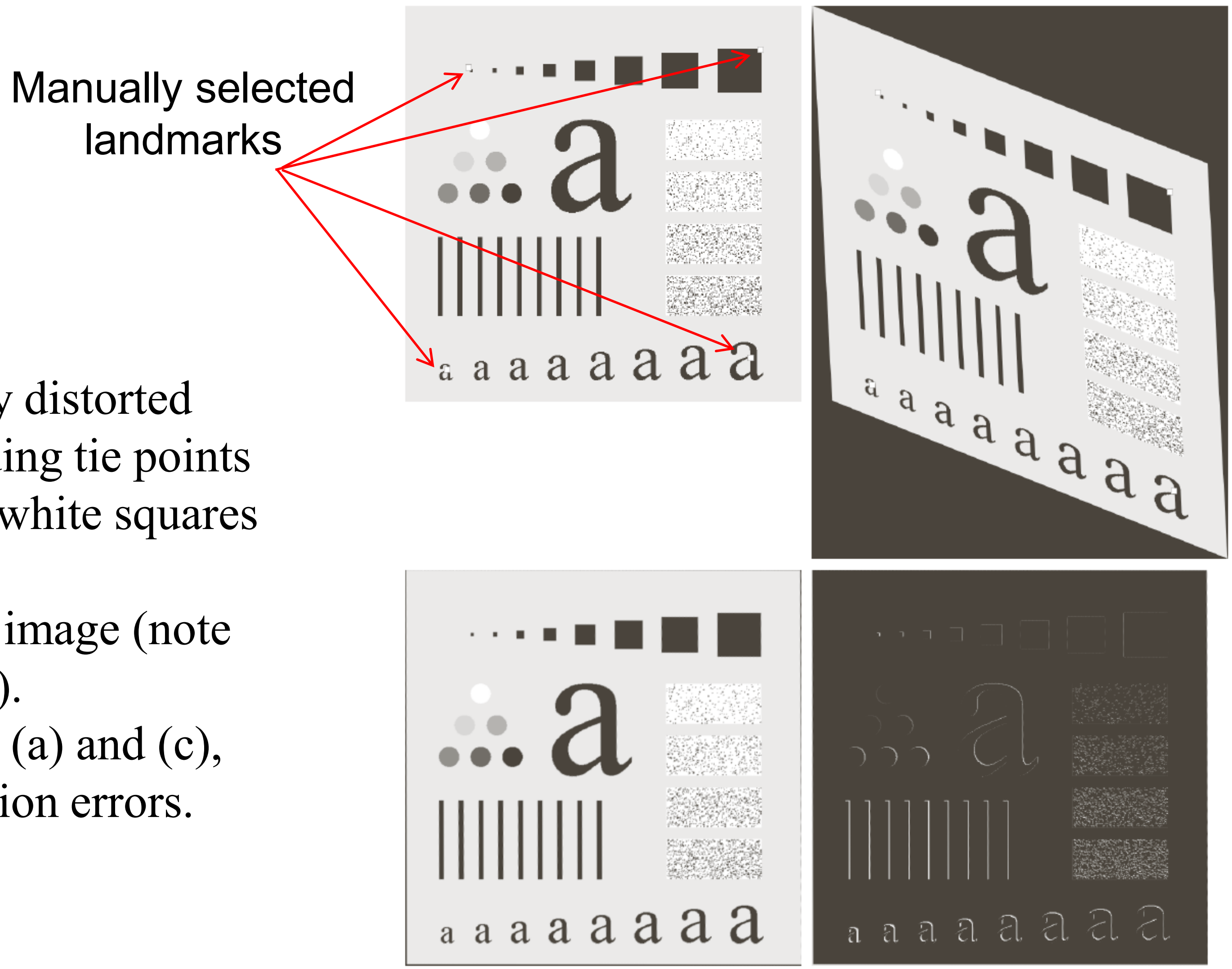


FIGURE 2.42
Image registration.
(a) Reference image.
(b) Input (geometrically distorted image). Corresponding tie points are shown as small white squares near the corners.
(c) Registered (output) image (note the errors in the border).
(d) Difference between (a) and (c), showing more registration errors.