Module 6:
Fundamentals of Digital Image

Outline of Lecture

- Digital vs. Analog Signals
- Digital Image Representation
- Fourier Transforms

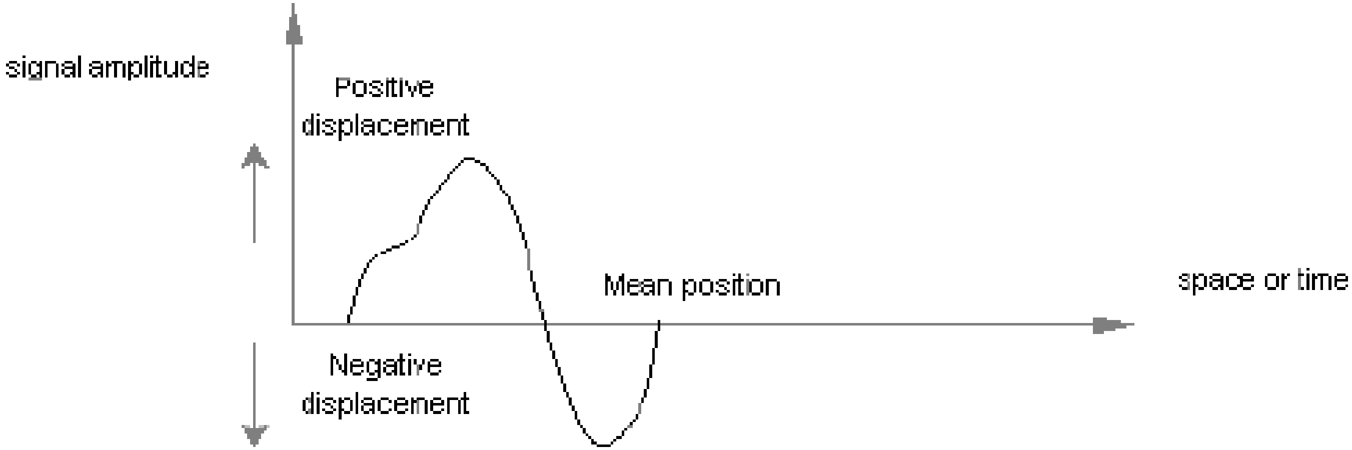
- REFERENCES:
  - R.C. Gonzalez & R.E. Woods (1992). Digital Image Processing. Addison-Wesley. (parts on Analog signals)
  - Z.N. Li and M.S. Drew (2003). Fundamentals of Multimedia. Prentice Hall (Chapter 3)

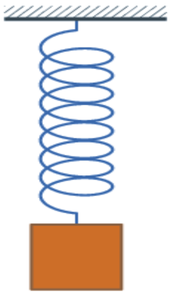
Analog Representation

- An analog quantity can vary continuously over space and/or time.
  - It is represented as $f(x,y,z,t)$
- Analog physical quantities can be transformed into electrical signals using sensors
- Signals can be represented as waves which can take up any possible real values within the instrument range (amplitude continuous)
- Value of the analog signal can be determined for any possible value of space or time variable (hence it is space or time continuous)

Waves -1

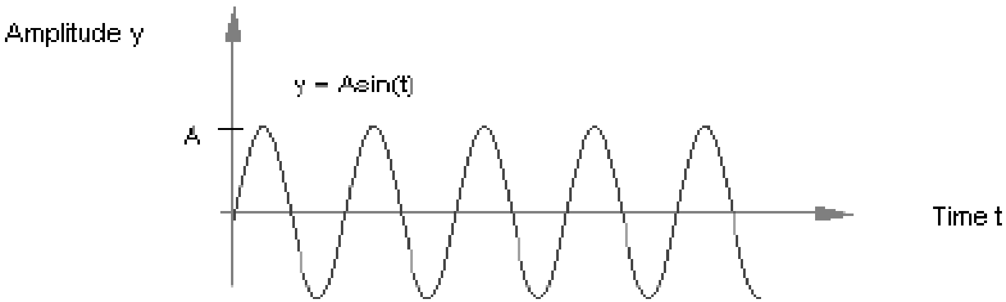
- Waves can be conceived as energy propagating from one place to another
  - A wave essentially represents a graph or plot of the motion of a set of particles in the path of a wave over space or time





Waves -2
Sinusoidal Waves

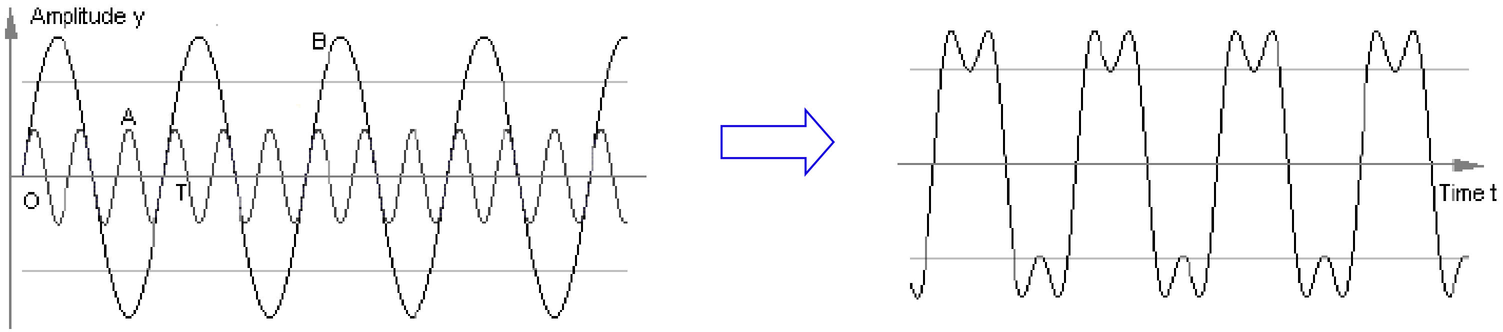
- The concepts of a wave can be explained by considering a sinusoidal function:
  - y = A sin(t) or y = B cos(t)
  - sinusoids are easier to deal with as they are periodic functions and can be represented by simple equations



- J.B. Fourier demonstrated that any irregular shaped wave can be constructed from a combination of a number of sine & cosine waves
  - Sinusoidal waves therefore can be considered as a kind of elementary wave that can be used as the building blocks to generate other waves

Waves -3
Composite Waves

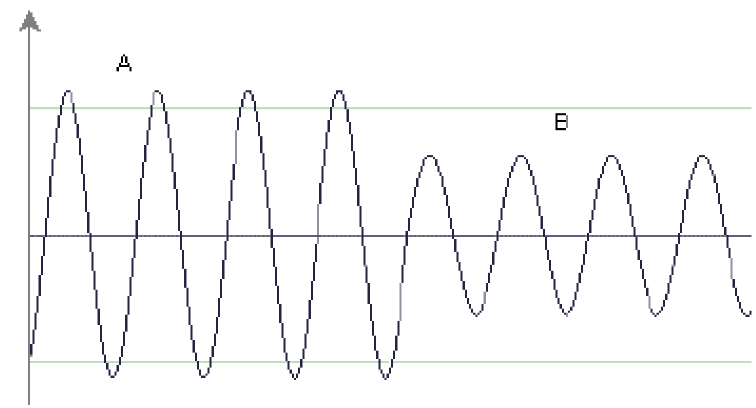
- Two or more sinusoidal waves can combined to generate a composite wave which may have a non-sinusoidal shape
    - The composite shape is obtained by adding up the amplitudes of the component waves at all points



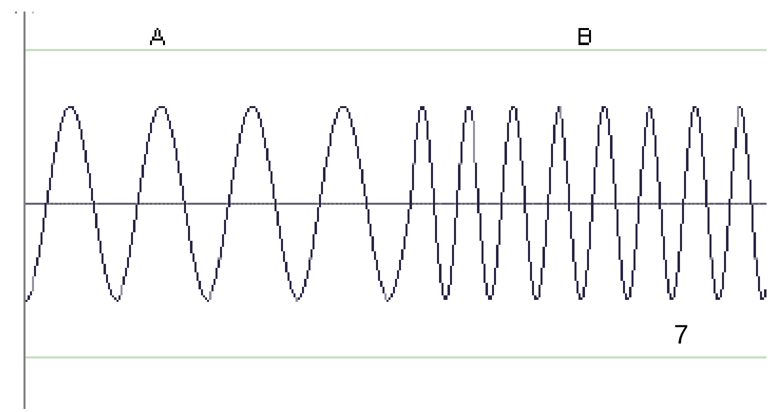
Waves -4

Amplitude and Frequency

- The peak amplitude of a wave is the max. displacement of oscillating particle from its mean position
  - It represents the intensity of the wave; e.g. the brightness of light, the loudness of sound, etc.

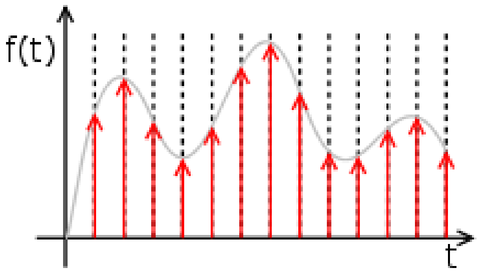


- The frequency refers to how fast the particle is oscillating:
  - Define as the # of oscillations per unit time
  - It physically represents the pitch of sound or color of light
  - A higher pitch results in a shrill sound e.g. a whistle, while a lower pitch results in a dull and flat sound e.g. sound of a padded hammer

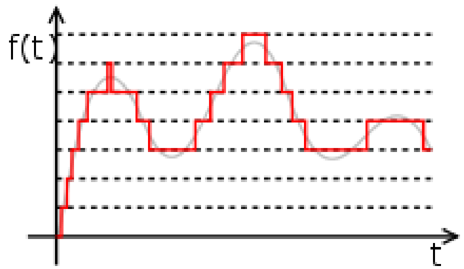


A to D Conversion -1
Sampled, Quantized, Digital

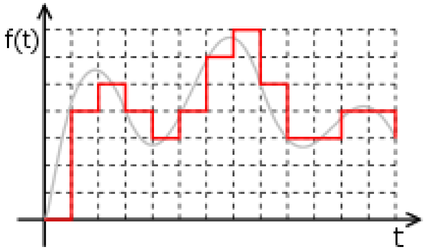
- In contrast to analog quantities,
  - **Sampled quantities** have continuous value at discrete space or time, i.e., at certain points in space or instances in time
  - **Quantized quantities** have discrete values at continuous space or time
  - **Digital quantities** have discrete values at discrete space or time



Sampled Signal:
discrete time,
continuous values



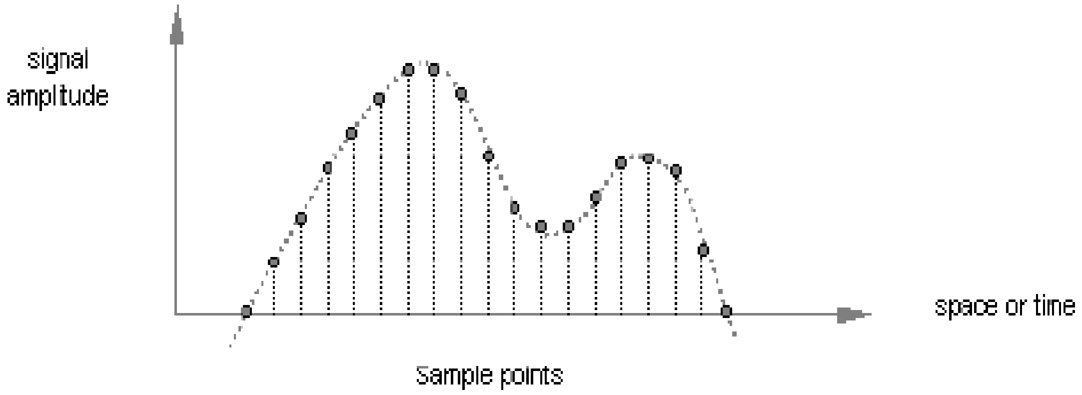
Quantized Signal:
continuous time,
discrete values



Digital Signal:
discrete time,
discrete values

A to D Conversion -2
Sampling

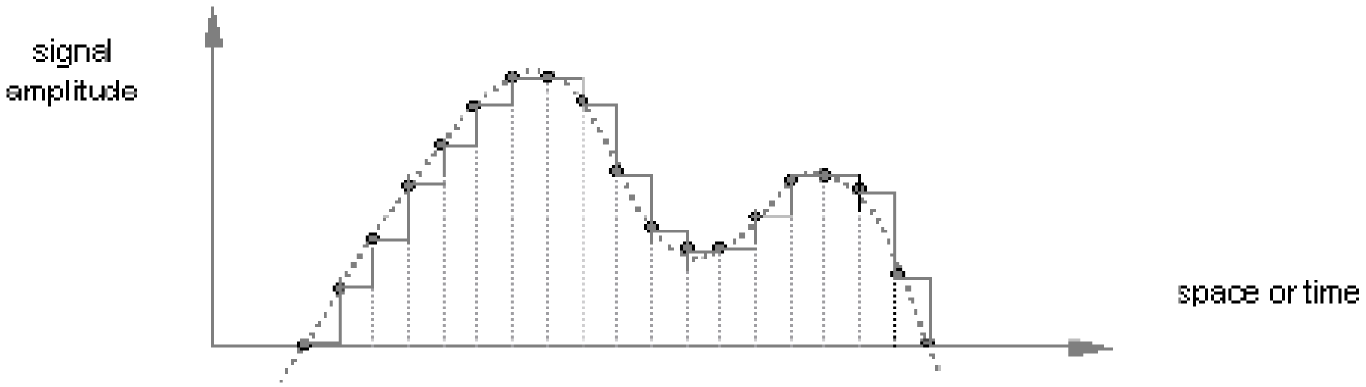
- The first step in A-D conversion is SAMPLING
  - It records values of wave at pre-determined discrete set of points and discards other values
  - For time-dependent quantities like sound, sampling is done at specific intervals of time (e.g. 10 times per second) -- time-discretization of the signal
  - For time-independent quantities like a static image, sampling is done at regular space intervals (e.g. 10 times per inch) -- space-discretization of the signal



A to D Conversion -3

Sampling Rate

- Sampling rate or sampling frequency: number of samples taken per second or per inch



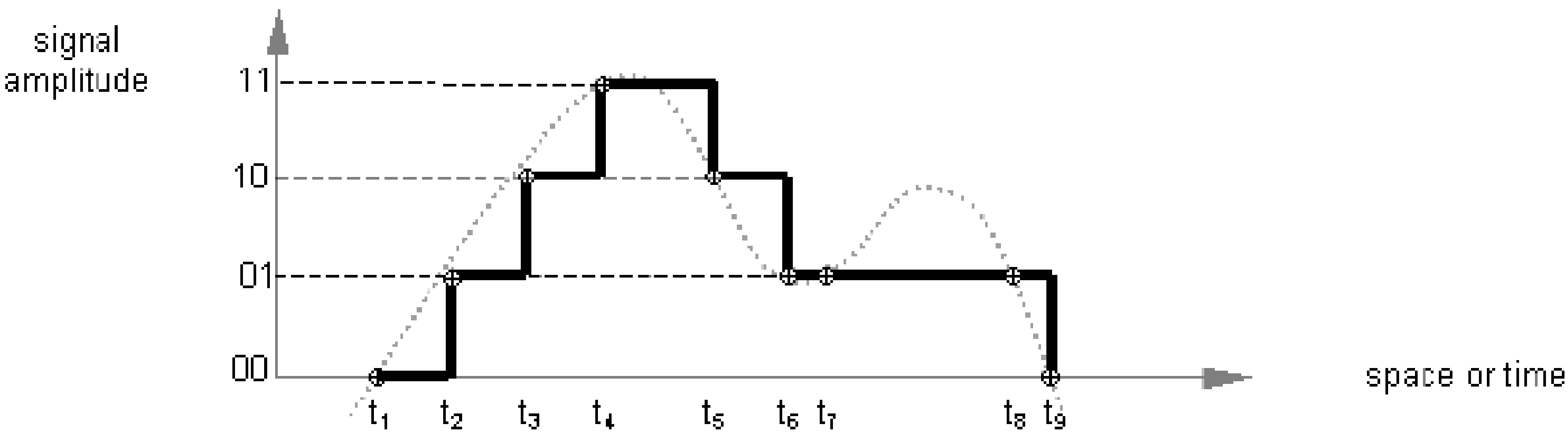
  - Balance of accuracy (higher sampling rate) and cost

- The digitized version will always be a degraded version of the original analog wave
  - Lose some data in between two sample points
  - Process irreversible

A to D Conversion -4

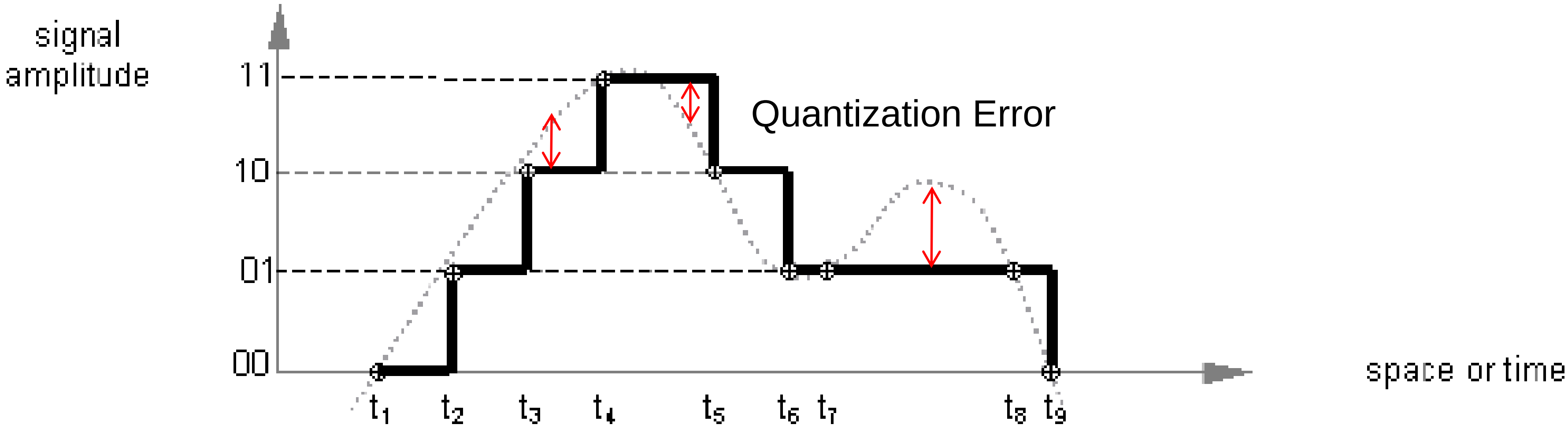
Quantization

- Quantization levels: number of different sample values used to represent a digital quantity
  - We use n-bit to represent magnitude, with values range from 0 to 2^n-1: $0,1,2, \ldots, 2^n-1$



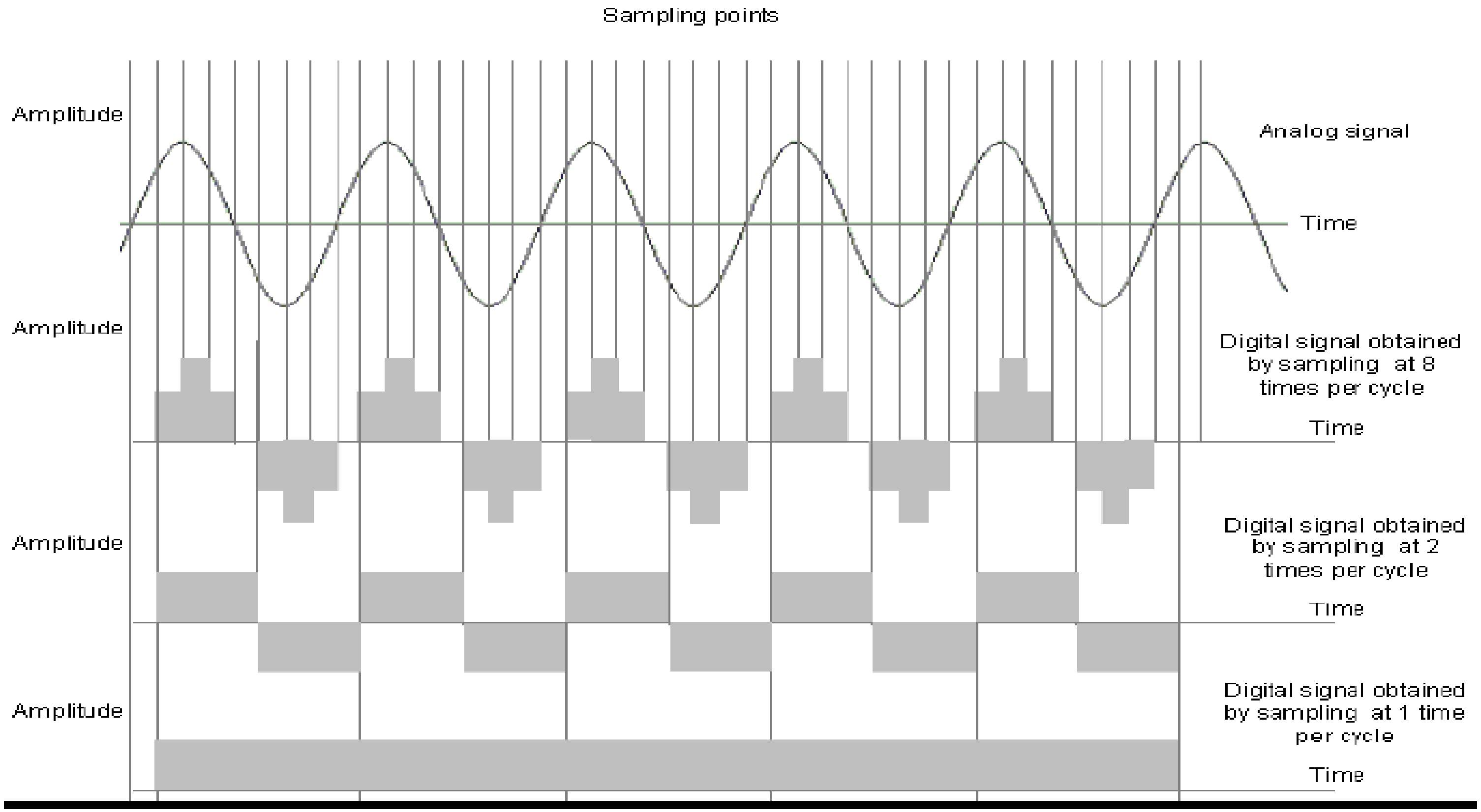
Quantization Error

- Errors in digital output
  - Because of quantization error there is always a distortion of the wave when represented digitally. This distortion effect is referred to as noise



A to D Conversion -5
Sampling Theory

- Nyquist's theorem states: T*he sampling frequency must be greater than twice the frequency of input signal in order to be able to reconstruct the original signal accurately from the sampled version.*



A to D Conversion -5

Sampling Theory

- 在进行模拟/数字信号的转换过程中，当采样频率大于信号中最高频率的2倍时，即：fs.max>=2Fmax,则采样之后的数字信号完整地保留了原始信号中的信息，就是可以不失真的恢复出原始的模拟信号。

- 一般实际应用中保证采样频率为信号最高频率的5～10倍；采样定理又称奈奎斯特抽样定理。

A to D Conversion -6

Sampling Theory

- Nyquist's theorem states: T*he sampling frequency must be greater than twice the frequency of input signal in order to be able to reconstruct the original signal accurately from the sampled version.*
  - Higher frequencies lead naturally to good approximation
  - At Nyquist frequency, the output wave is quite blocky in appearance but all the cycles are correctly reproduced
  - At < Nyquist frequency, the output fails to reproduce all the cycles of the original wave

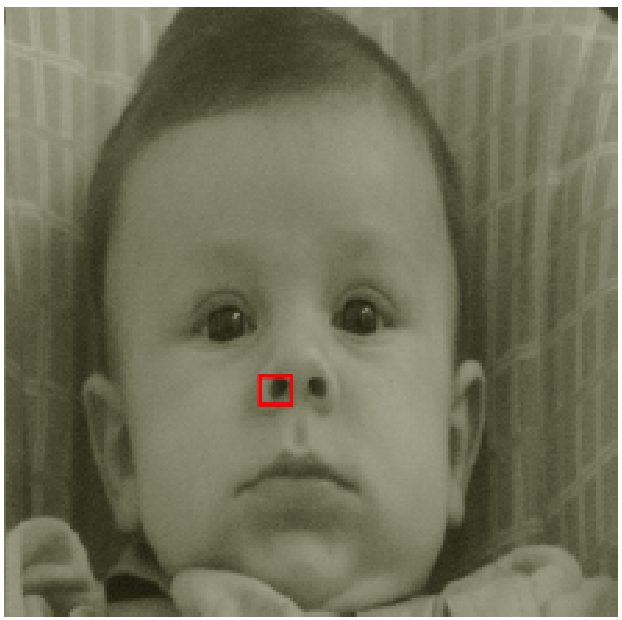
  ➔under-sampling or aliasing

Outline of Lecture

- Digital vs. Analog Signals
- Digital Image Representation
- Fourier Transforms

Digital Image Representation

- A digital image *f(x,y)* can be considered as a 2D matrix whose row and column indices identify a point in the image and the corresponding matrix element value denotes the gray level at that point

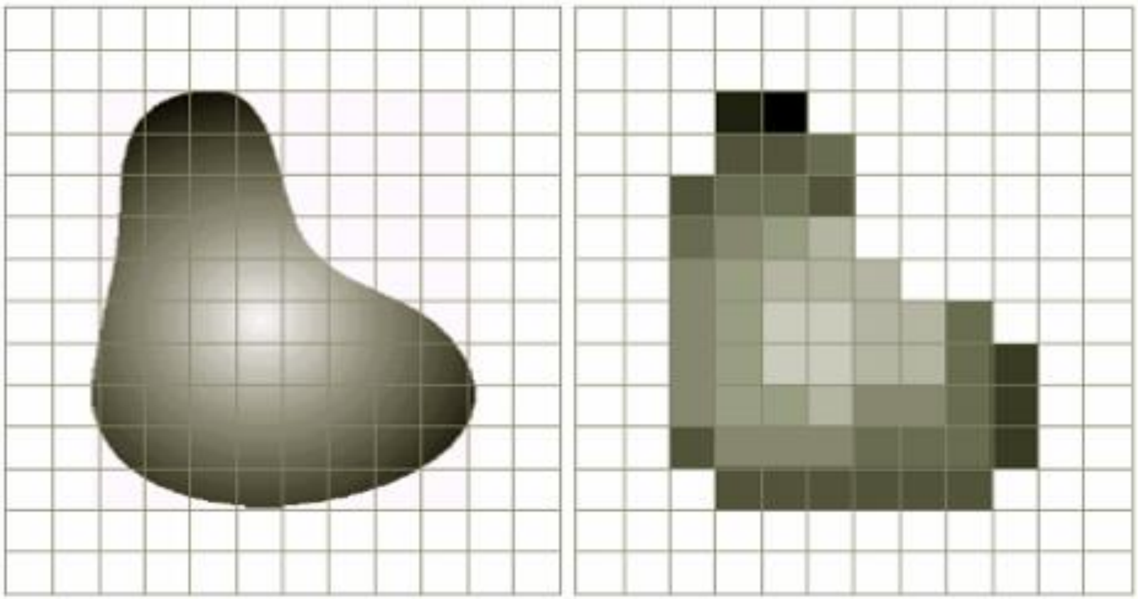


Pixel values in highlighted region

99	71	61	51	49	40	35	53	86	99
93	74	53	56	48	46	48	72	85	102
101	69	57	53	54	52	64	82	88	101
107	82	64	63	59	60	81	90	93	100
114	93	76	69	72	85	94	99	95	99
117	108	94	92	97	101	100	108	105	99
116	114	109	106	105	108	108	102	107	110
115	113	109	114	111	111	113	108	111	115
110	113	111	109	106	108	110	115	120	122
103	107	106	108	109	114	120	124	124	132

CAMERA ⇒ DIGITIZER ⇒ A set of number in 2D grid

Example of Digital Image



Continuous image projected onto a sensor array

Result of image sampling and quantization

Digital Image Resolution

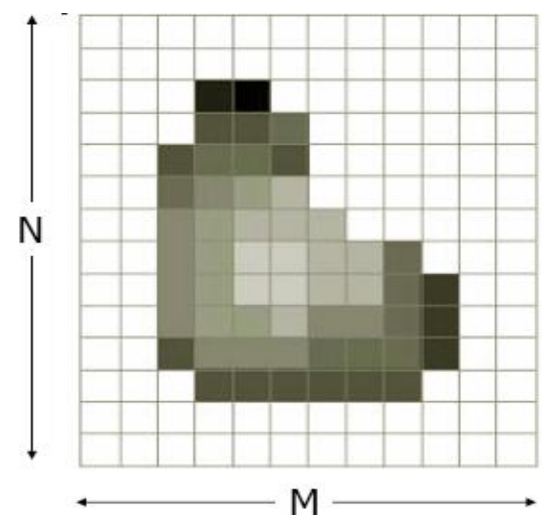
- The intensity of a monochrome image ***f*** at coordinate ***(x,y) is f(x,y)=l,*** the gray level of image at that point:
  - ***l*** is the gray level of image at ***(x,y),*** with $l = [L_{min}, L_{max}]$
  - Common practice is to shift the interval to [***0***, ***L-1***]
  - Hence ***0 = black*** , ***L = white***
  - L determines the intensity resolution of images, e.g. $L=256 = 2^8$

- Resolution depends on sampling and gray levels
  - # of gray levels:

    $L = 2^p$

  - # of bits required to store a digitized image

    $b = M \times N \times p$



Effects of Spatial Resolution:
Checkerboard Effect

a	b	c
d	e	f

(a) 1024x1024

(b) 512x512

(c) 256x256

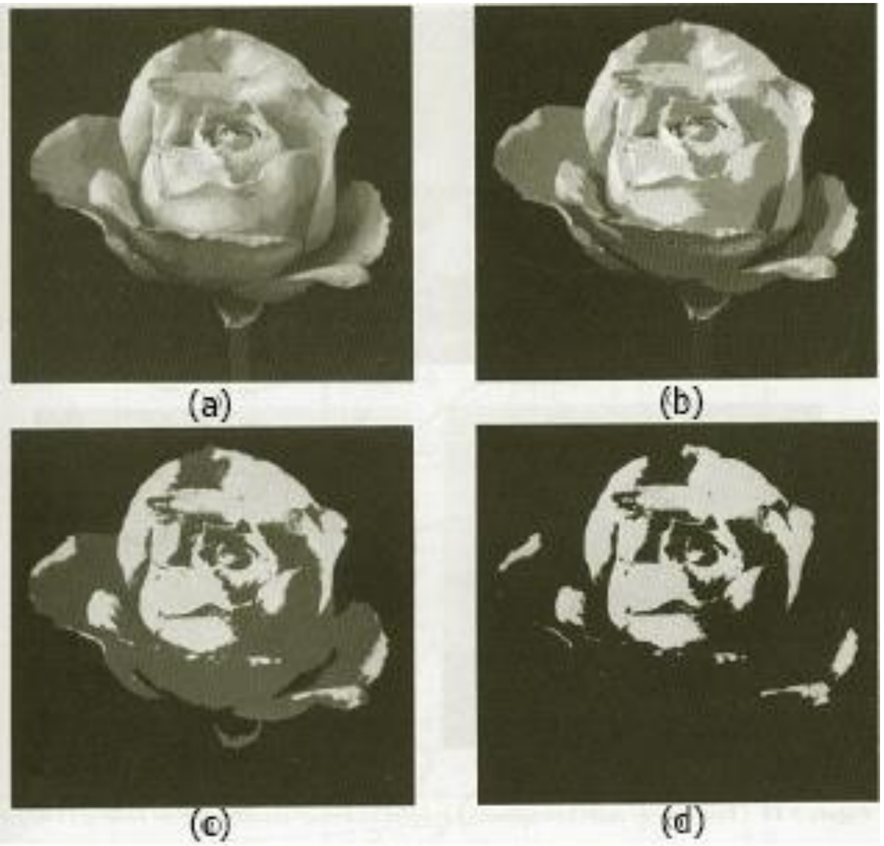
(d) 128x128

(e) 64x64

(f) 32x32

- If the resolution is decreased too much, the checkerboard effect can occur

Effects of Intensity Resolution:
False Contouring



(a) Gray level = 16 (p=4)
(b) Gray level = 8
(c) Gray level = 4
(d) Gray level = 2

- If the gray scale is not enough, false contouring can occur on the smooth area which has fine gray scales

Non-Uniform Sampling

- For a fixed value of spatial resolution, the appearance of the image can be improved by using adaptive sampling rates
  - fine sampling required in the neighborhood of sharp gray-level transitions
  - coarse sampling is necessary in relatively smooth regions

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

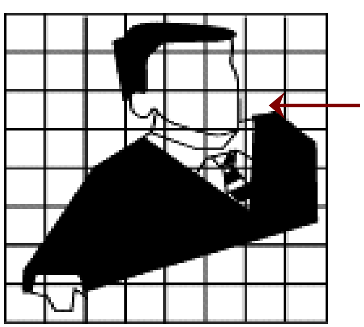
  - A simple way is to divide image into blocks and perform different sampling on each block

Gray-Scale Digital Image -1

- Digital image: represented by a matrix of pixel values
  - Square sampling grid is used, with pixels equi-spaced along the two sides of the grid

- Binary image: Pixel values are binary – (0,1) or (black, white)

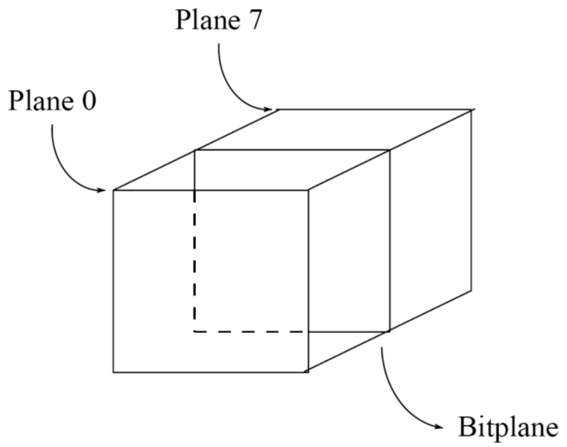
- In general, we want to represent something more complex
  - Each pixel contains a range of intensity values
  - Typical *gray scale image* contains 256 ($=2^8$) intensity values



shows an 8x8 2-D array of binary values

Gray-Scale Digital Image -2

- Intensity at each pixel is then represented by an integer
  - For p bit-planes, the value ranges from [0 - 2^P-1]
  - Binary Image: p=1
  - If p=8, we have gray levels from 0 (black) to 255 (white)



- Pixel values are represented as quantized values
  - Quantization – a way of mapping a continuous range of values (0~1) to discrete integer intervals
  - For p=3, there are 8 ranges of:

    0: 0~1./8., 1: 1./8. 1~2./8., .. 7: 7./8.~1.
  - What is the mapping function?

    V_Q = Integer (v_o * n), where v_o=original value, n=8
  - ISSUE: Is quantization reversible?

Color Digital Image

- What about Color Images?
  - Colors are modeled by {R, G, B} triplets
  - Divide the bit-planes into 3 groups, p/3 bits for R, p/3 bits for G, ..

- If three 8-bit integers are used (24-bits per pixel) , then it is a color image with 8-bits per color channel of R, G or B → 16.7 million possible colors
  - Usually an extra alpha byte (for special effect info) is stored for each pixel → 32-bit image

Outline of Lecture

- Digital vs. Analog Signals
- Digital Image Representation
- Fourier Transforms

Signal Processing

- It involves the transformation of a signal into a form which in some sense more desirable for analysis (e.g. analog to digital, or spatial to frequency)
- Analog Signal Processing Tool
  - oscilloscope: Tektronics and HP instruments
  - analog computer
- Digital Signal Processing (DSP) Tool
  - mathematical analysis and formula
  - computer software package, e.g. MATHLAB
- DSP Applications
  - DSP chips, ASIC’s, embedded systems, etc.

Fourier Transforms- Overview

- The Fourier transform produces representation of any (2D) signal as a weighted sum of sines and cosines. Because of Euler's formula:

  $e^{jq} = \cos(q) + j \sin(q)$, where $j^2 = -1$.

- Given: an image *f;* its Fourier transform *F*:
- The forward transform is: $F = T\{f\}$

  from spatial domain to (continuous) frequency domain.
- The inverse transform is: $f = T^{-1}\{F\}$

  from frequency domain back to spatial domain.

- The Fourier transform is an unique and invertible operation:

  $$f = T^{-1}\{T\{f\}\} \quad \text{and} \quad F = T\{T^{-1}\{F\}\}$$

Fourier Transforms- cont.

- ***f*** is in spatial domain

  its Fourier transform, ***F***, is in frequency domain

- The invertible property enables signals to be transformed between both domains without loss of information

- WHY we want to transform signals to frequency domain?
  - Because many operations are easier to perform in frequency domain than in spatial domain. For example:
  - Image filtering:
    - low-pass filter (to remove details): simply remove high frequency component while retain low frequency component
    - High-pass filter (to sharpen details): enhance high frequency components
  - Compression: remove high-frequency component to save storage
  - Image enhancement: scale image contents

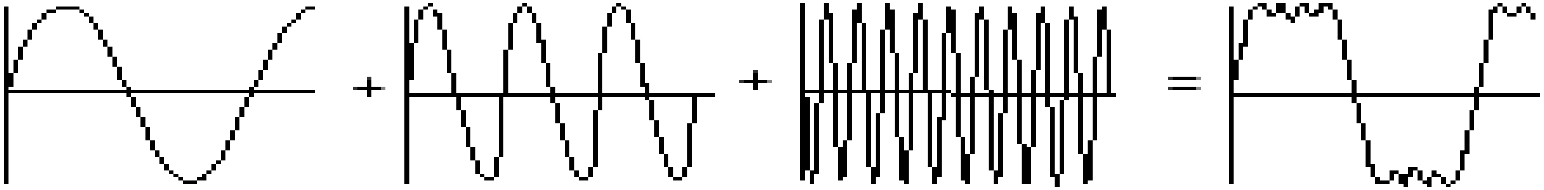
Fourier Analysis

- Trigonometric series of periodic signals
  - Basically, any periodic function (piecewise continuous with left/right derivatives existing at discontinuities) can be expressed as an infinite sum of Sines and Cosines

- Example: a periodic square wave

$$\frac{4k}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \ldots\right)$$

and the periodic function that it produces:



→ Towards a square wave

Fourier Analysis – 1D Case

- Let ***f(x)*** be a continuous function of a real variable ***x***. Its Fourier Transform ***F(u)*** is:

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx$$

where u is the frequency variable

- The inverse transform from of ***F(u)*** to ***f(x)*** is:

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{+j2\pi ux} du$$

- ***f(x)***, ***F(u)*** is called Fourier Transform pair
- They will exist if ***f(x)*** is continuous and integrable, and ***F(u)*** is integrable
- As $e^{-j2\pi ux} = \cos(2\pi ux) - j\sin(2\pi ux)$

F(u) is computed as integral sum of sine & cosine terms

Discrete Fourier Analysis- 1D

- Suppose ***f(x)*** is discretized into a series

  $f(x) = \{ f(x_0) + k\,dx \}, \;\; k = 0, 1, .. N\text{-}1$

- The forward Fourier Transform is:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N} \quad , u = 0, 1, 2, .., N-1$$

- The corresponding inverse Transform is:

$$f(x) = \sum_{u=0}^{N-1} F(u) e^{+j2\pi ux/N} \quad , x = 0, 1, 2, .., N-1$$

- Fast Fourier Transform (FFT)
  - As discussed in many textbooks and tool packages
  - has N*logN operations (efficient implementations)

Fourier Analysis – 2D Case

- Fourier transform extends easily to 2D case with ***f(x,y)*** of 2 variables. Its Fourier Transform ***F(u,v)*** is given by:

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-j2\pi(ux+vy)}\,dx\,dy$$

where u, v are the frequency variables

- The inverse transform from of ***F(u,v)*** to ***f(x,y)*** is:

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)\, e^{+j2\pi(ux+vy)}\,du\,dv$$

Discrete Fourier Analysis- 2D

- The forward Fourier Transform is:

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M + vy/N)},$$

$$u = 0,1,2,..,M-1, \;\; v = 0,1,2,..,N-1$$

- The corresponding inverse Transform is:

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) e^{+j2\pi(ux/M + vy/N)},$$

$$x = 0,1,2,..,M-1, \;\; y = 0,1,2,..,N-1$$

- If images are sampled in a square array, then M=N, and

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux+vy)/N},$$

$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) e^{+j2\pi(ux+vy)/N}$$

Properties of Fourier Transform -1

- Invertibility
- Separability

$$F(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}e^{-j2\pi ux/N}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi vy/N}\ , \quad or$$

$$or \quad F(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}F(x,v)e^{-j2\pi ux/N}$$

$$with \quad F(x,v)=N[\frac{1}{N}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi vy/N}$$

This means that F(u,v) can be computed by successive applications of 1D Fourier Transform of its inverse:

f(x,y) → F(x,v) → F(u,v)

Properties of Fourier Transform -2

- Rotation:

  If we introduce:

  $x = r\cos\theta$, $y = r \sin\theta$, then $f(x,y) \rightarrow f(r,\theta)$

  $u = w\cos\Phi$, $v = w \sin\Phi$, then $F(u,v) \rightarrow F(w,\Phi)$

  Substitute into Fourier equation yields:

  $f(r, \theta+\theta_0) \Leftrightarrow F(w, \Phi+\theta_0)$

  In other words, rotating f(x,y) by an angle θ_0 is equivalent to rotating F(u,v) by the same angle

- Scalability

  For two scalars a and b:

  $a f(x,y) \Leftrightarrow a F(u,v)$

  and $f(ax, by) \Leftrightarrow F(u/a, v/b)/|ab|$

Properties of Fourier Transform -3

- Convolution
  - It is an important operations in image processing application
  - Convolution of two functions *f(x)* and *g(x)* is:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha)\, g(x-\alpha)\, d\alpha$$

where α is a dummy variable for integration.

- Convolution theory
  - If f(x) has Fourier Transform F(u), and

    g(x) has Fourier Transform G(u), then

    f(x) * g(x) ←→ F(u) G(u)

    i.e. convolution in x domain can be obtained by taking the inverse Fourier Transform of the product F(u) G(u)
  - An analogous result is:

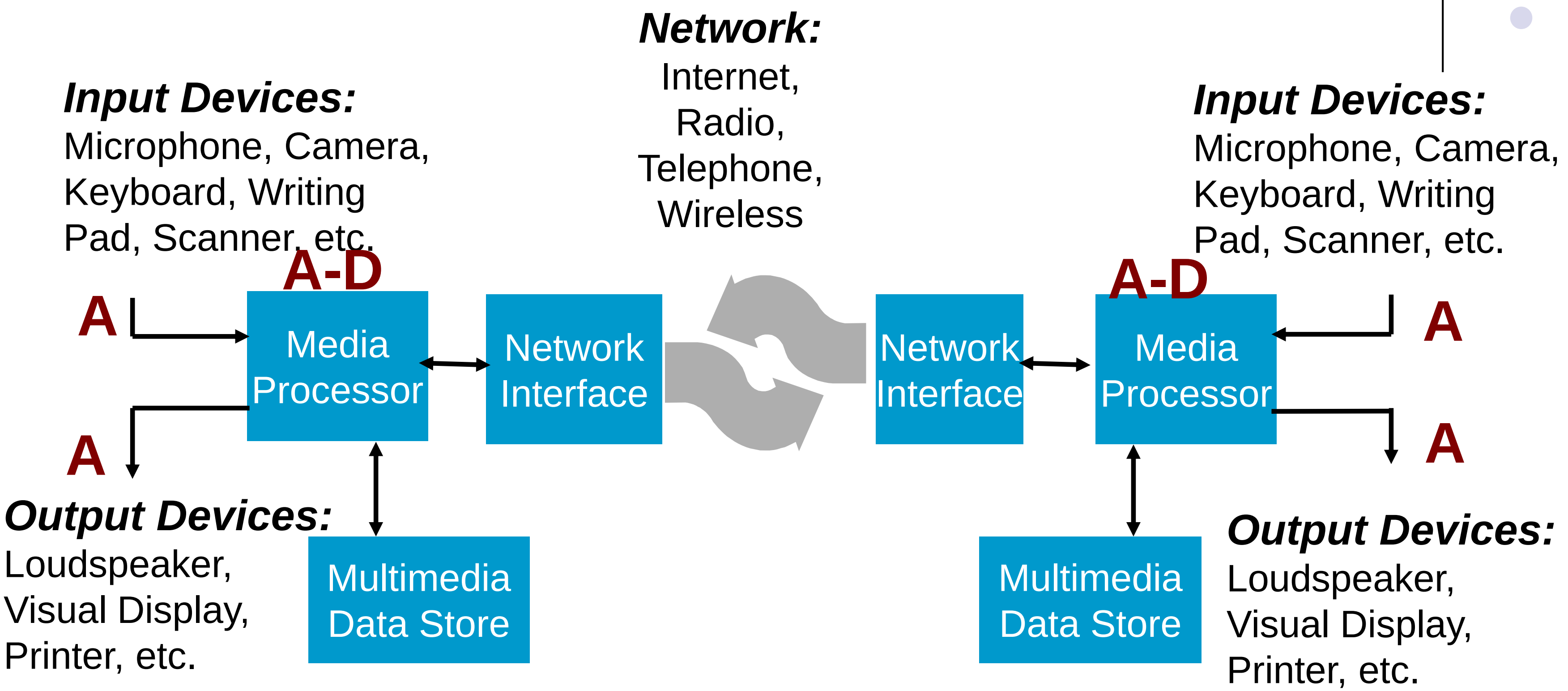
    f(x) g(x) ←→ F(u) * G(u)

  - Many applications: image enhancement, etc.

Block Transforms

- Instead of operating on whole image/audio, divide it into blocks, and process each block separately.
  - Computational complexity decreases.
  - Transform captures the local behavior better.
- DCT (Discrete Cosine Transform) has been very popular in block transform based image compression for a long time.
  - It approximates Karhunen-Loeve (KLT), the optimum transform in mean square error (MSE) sense
  - DCT is adopted for JPEG/MPEG standards
  - More about DCT transform in later part of lectures

Multimedia Signals & Systems



Multimedia Information: Text, Speech, Audio, Image, Video, Cinema (From analog to digital representations and back)

Summary

- Cover fundamentals of media processing (from analog to digital domain)
- Discuss digital signals and systems
  - Fourier transform, *Bock Transforms*, ….
- Pave way for following classes
  - Image, video and audio processing
  - many other applications later

Next Lesson

- Next Lesson

  Image Transforms and Filters

- Cover Audio part in Lesson 8

MID-TERM TEST

1.图像检索的流程及其关键技术

2. KD-Tree，Hashing基本思想