IT4440
Đa phương tiện
và các ứng dụng giải trí

(MULTIMEDIA AND GAMES)

Chương VI: Audio

Phần I: Tổng quan về thông tin đa phương tiện và các kỹ thuật xử lý

Chương I: Nhập môn Multimedia

Chương II: Một số kiến thức cơ bản

Chương III: Ảnh

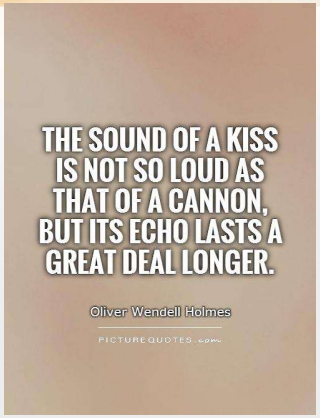
Chương IV: Màu sắc

Chương V: Video

Chương VI: Audio

- Mục tiêu của chương
- Một số khái niệm
- Âm thanh số
- Nén âm
- MIDI vs. Audio
- Tổng kết chương
- Tài liệu tham khảo

Âm thanh là gì ?



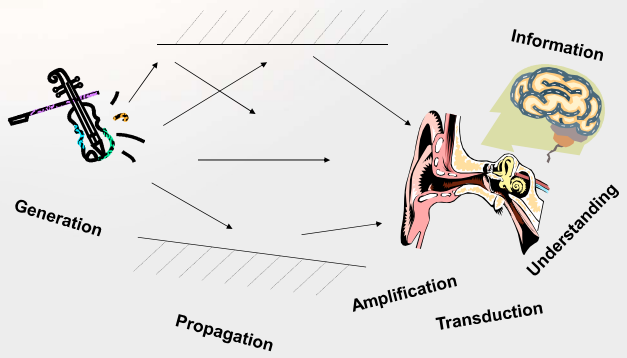
If a tree falls in the forest and nobody is there to hear it, will it make a sound?

Sound provided by
http://www.therecordist.com/downloads.html

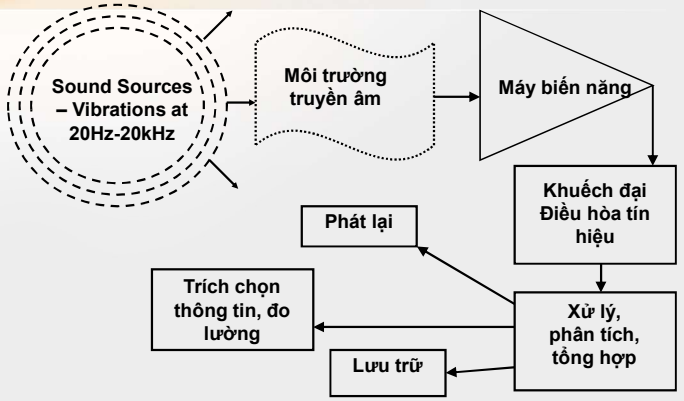
Âm thanh là gì ?

- Âm thanh có thể là
  - Tiếng nói
  - Âm nhạc
  - Tiếng ồn
- Âm thanh là một mối quan hệ phức tạp của các đối tượng sau:
  - Nguồn âm
  - Môi trường truyền (thường là không khí)
  - Bộ thu nhận âm (tai)
  - Bộ cảm nhận âm (não bộ)

Hệ thống truyền nhận âm tự nhiên



Hệ thống truyền nhận âm điện tử

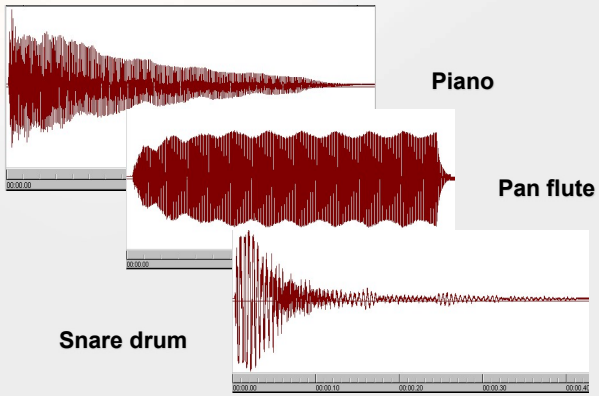


Sóng âm

Something vibrates in the air | Waves of pressure | Ear drums will translate these changes in wave Forms as sound

- Âm thanh được đo bằng decibel → **dB (decibel)**
- Âm thanh lan truyền trong môi trường: **sóng âm.**

Một số ví dụ về sóng âm



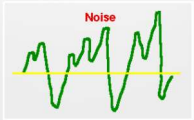
Audio vs Images

Sự khác biệt lớn nhất của âm thanh và hình ảnh là gì ?

TIME

Âm thanh và nhiễu âm

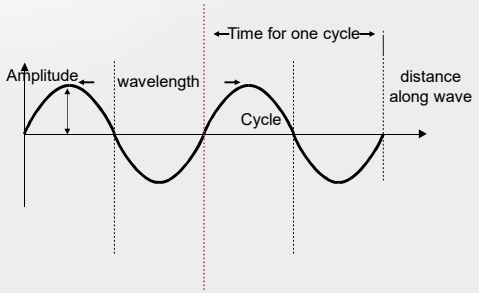
- A pleasant sound has a regular wave pattern. The pattern is repeated over and over.



- But the waves of noise are irregular. They do not have a repeated pattern.

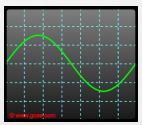
Đặc trưng của sóng âm

- Âm thanh được mô tả bởi hai đặc trưng sau:
  - **Tần số Frequency** (or pitch)
  - **Biên độ Amplitude** (or loudness)

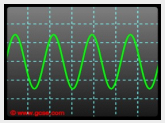


Tần số

- ***Frequency*** is a measure of how many vibrations occur in one second. This is measured in *Hertz* (abbreviation Hz) and directly corresponds to the ***pitch*** of a sound.
  - The more frequent vibration occurs the higher the pitch of the sound.



Low pitch

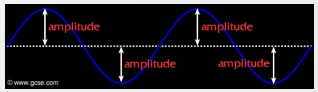


High pitch

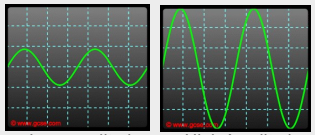
- Optimally, people can hear from **20 Hz to 20,000** Hz (20 kHz)
  - Sounds below 20 Hz are infrasonic
  - sounds above 20 kHz are ultrasonic.

Biên độ âm thanh

- ***Amplitude*** is the *maximum displacement* of a wave from an equilibrium position.
  - The louder a sound, the more energy it has. This means loud sounds have a large amplitude.



Quiet



Loud

Low amplitude High Amplitude

- The amplitude relates to how loud a sound is.

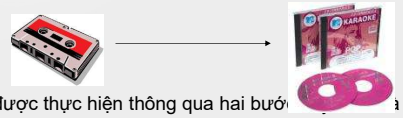
Âm thanh tương tự và âm thanh số

Âm thanh tương tự

- Tín hiệu điện mang tông tin về âm thanh như một giá trị điện thế liên tục

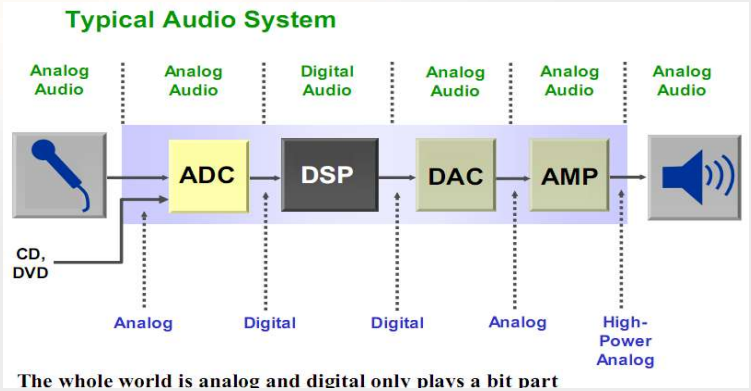
Số hóa âm thanh

- Để sử dụng trong các ứng dụng đa phương tiện, giống như ảnh, video, âm thanh phải được số hóa



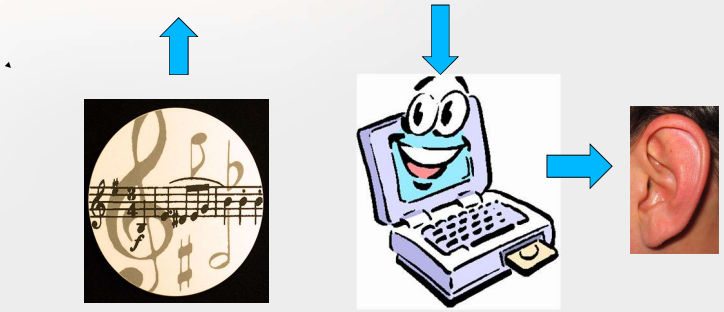
- Việc số hóa được thực hiện thông qua hai bướ lượng tử hóa

Âm thanh tương tự và âm thanh số



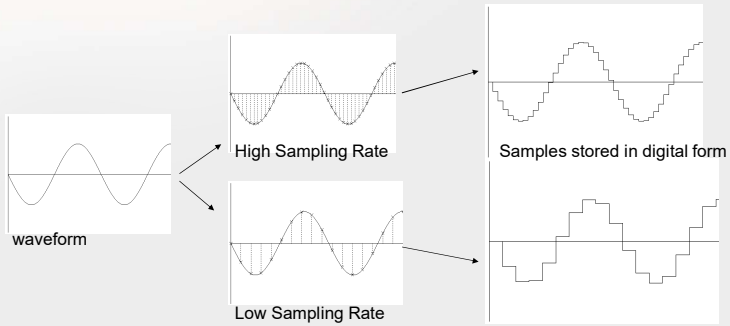
Biểu diễn số của tín hiệu âm thanh

0101010000101010101010011010001010100110100101001010001110000101010101001011111001001010…



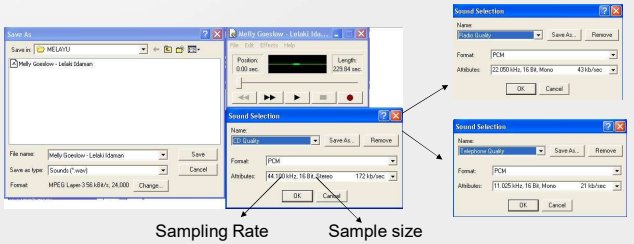
Âm thanh số

- Dữ liệu âm thanh số là biểu diễn của âm thanh, được lưu trữ dưới dạng các điểm mẫu



Âm thanh số

- **Chất lượng âm thanh số phụ thuộc vào**
  1. Tốc độ lấy mẫu (sampling rate)
  2. Kích thước mẫu (quantization step)

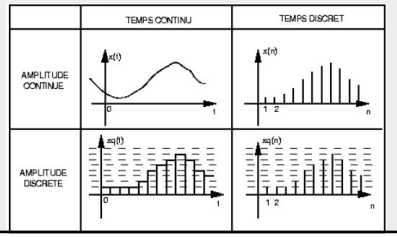


Âm thanh số

- Ngoài ra, chất lượng âm thanh số phụ thuộc vào
  - Chất lượng của nguồn âm
  - Chất lượng của thiết bị thu và các phần cứng hỗ trợ.
  - Các đặc trưng sử dụng để thu âm.
  - Khả năng phát lại của môi trường phát âm.

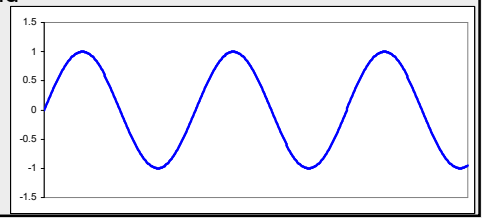
Biểu diễn số tín hiệu âm thanh

- Tín hiệu âm thanh được
  - Lấy mẫu
  - Lượng tử hóa
  - Nén



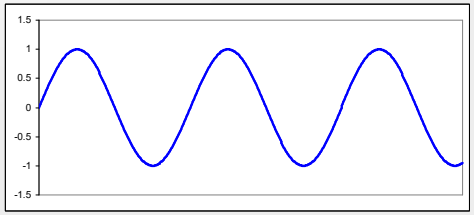
Lấy mẫu (Sampling)

- Measure amplitude at regular intervals
- How many times should we sample?
  - Suppose we have a sound wave with a frequency of 1 cycle per second



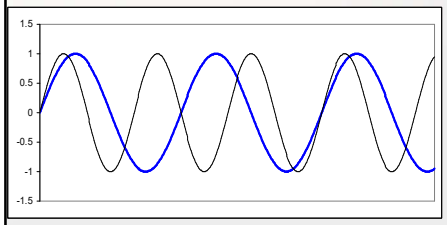
Example

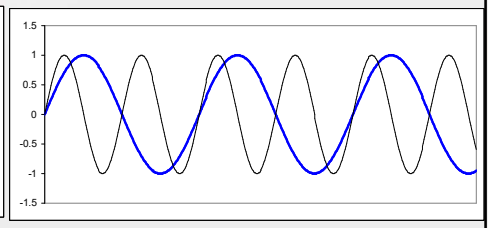
- If we sample at one cycle per second, where would the sample points fall?



Example

- If we sample at 1.5 cycles per second, where will the sample points fall?





- If we sample at two cycles per second, where do the sample points fall?

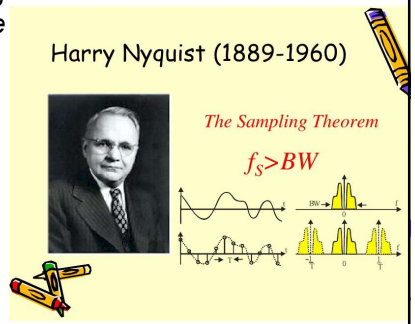
Nyquist Theorem

For lossless digitization, the sampling rate should be at least twice the maximum frequency response.

- In mathematical terms:

$$f_s > 2*f_m$$

- where f_s is sampling frequency and f_m is the maximum frequency in the signal

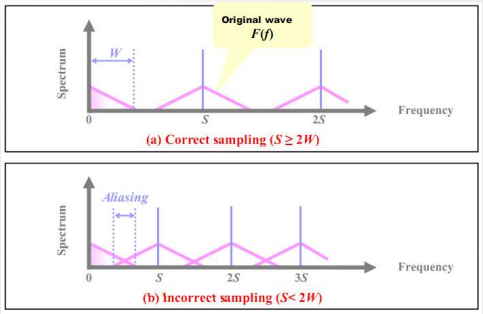


Nyquist Limit

- max data rate = 2 H $\log_2$V *bits/second,* where
  H = bandwidth (in Hz)
  V = discrete levels (bits per signal change)
- Shows the maximum number of bits that can be sent per second on a *noiseless* channel with a bandwidth of H, if V bits are sent per signal
  - Example: what is the maximum data rate for a 3kHz channel that transmits data using 2 levels (binary) ?
  - Solution (2x3,000xln2=6,000bits/second)

Limited Sampling

- But what if one cannot sample fast enough?



Sampling in the frequency domain

Các khoảng lấy mẫu

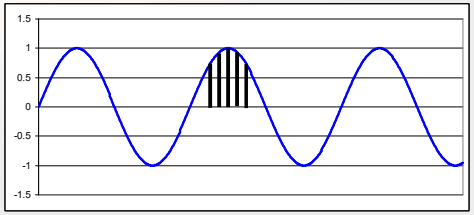
- Auditory range 20Hz to 22.05 kHz
  - must sample up to to 44.1kHz
  - common examples are 8.000 kHz, 11.025 kHz, 16.000 kHz, 22.05 kHz, and 44.1 KHz
- Speech frequency [200 Hz, 8 kHz]
  - sample up to 16 kHz
  - but typically 4 kHz to 11 kHz is used

Tốc độ lấy mẫu

The **sampling rate**, **sample rate**, or **sampling frequency** defines the number of samples per second (or per other unit) taken from a continuous signal to make a discrete signal. For time-domain signals, the unit for sampling rate is 1/s

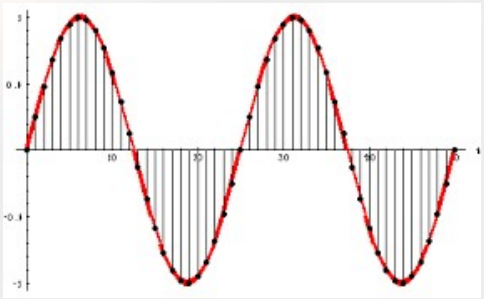
Sampling Rates	Used As...
8000	Telephony Standard, Popular in UNIX Workstations
11000	Quarter of CD rate, Popular on Macintosh
16000	G.722 Standard (Federal Standard)
18900	CD-ROM XA Rate
22000	Half CD rate, Macintosh rate
32000	Japanese HDTV, British TV audio, Long play DAT
37800	CD XA Standard
44056	Professional audio industry
44100	CD Rate
48000	DAT Rate

Lượng tử hóa

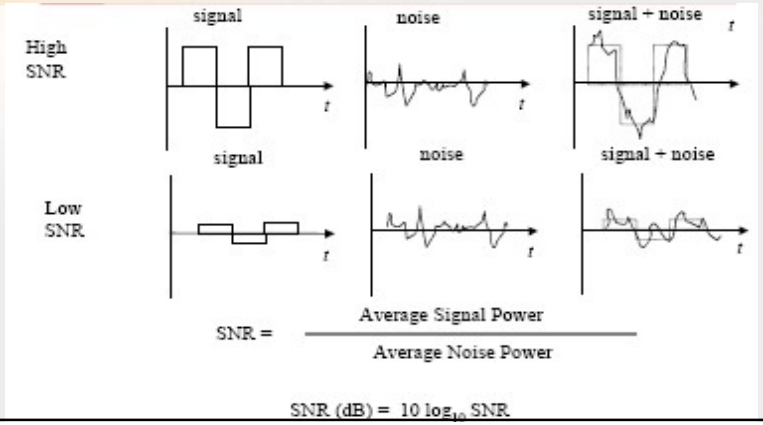


Lượng tử hóa

- Typically use
  - 8 bits = 256 levels
  - 16 bits = 65,536 levels
- How should the levels be distributed?
  - Linearly? (PCM)
  - Perceptually? (u-Law)
  - Differential? (DPCM)
  - Adaptively? (ADPCM)



Tỷ số tín hiệu trên nhiễu (SNR)

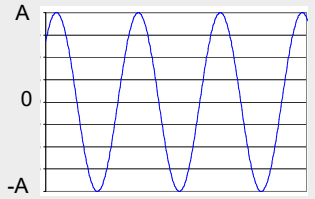


Tỷ số tín hiệu trên nhiễu (SNR)

- Measures strength of signal to noise

SNR (in DB)= $10 \log_{10}(\frac{Signal\ energy}{Noise\ energy})$

- Given sound form with amplitude in [-A, A]
- Signal energy = $\frac{A^2}{2}$



Nén âm thanh

- Tại sao phải nén âm thanh ?

So Many Bits, So Little Time (Space)

- CD audio rate: 2 * 2 * 8 * 44100 = 1,411,200 bps
- CD audio storage: 10,584,000 bytes / minute
- A CD holds only about 70 minutes of audio
- An ISDN line can only carry 128,000 bps

Security: Best compressor removes all that is recognizable about the original sound

Graphics people eat up all the space

Nén âm thanh

- Nguyên lý của nén âm ?

Take advantage of

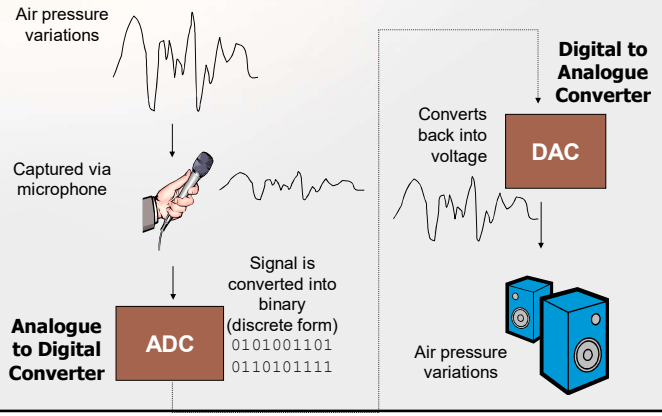
- Redundancy/Correlation
- Statistics (Local / Global)
- Assumptions / Models

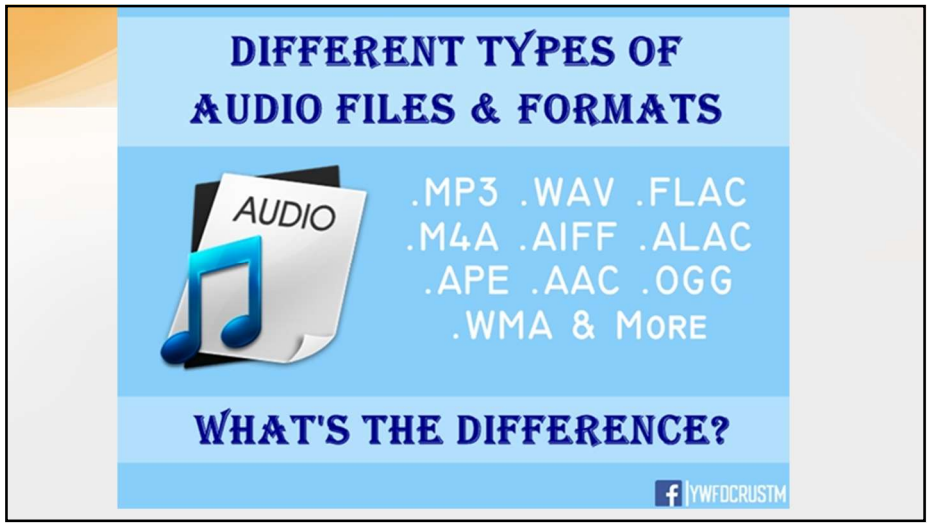
Problem: Much of this doesn't work directly on sound waveform data

Nén âm thanh

- Sound is difficult to compress using lossless methods, except for special cases.
- Some compression of audio can be obtained by run-length encoding samples that fall below a threshold that can be considered to represent silence.
- Companding uses non-linear quantization to compress speech.
- μ-law and A-law companding are used for telephony.

Thu và phát âm thanh số




DIFFERENT TYPES OF
AUDIO FILES & FORMATS
AUDIO
.MP3 .WAV .FLAC
.M4A .AIFF .ALAC
.APE .AAC .OGG
.WMA & More
WHAT'S THE DIFFERENCE?
YWFDCRUSTM