Digitisation

Agenda

- Sound and images are analogue phenomena that can be represented by complex waveforms
- They must be digitised to be handled by computers: sampling and quantisation
- Sampling and quantisation rates determine the size of the digitised data

Sound and images are analogue phenomena that can be represented by complex waveforms

Analogue versus discrete phenomena

sound is a mechanical wave, which means that it results from the motion of particles through a transmission medium

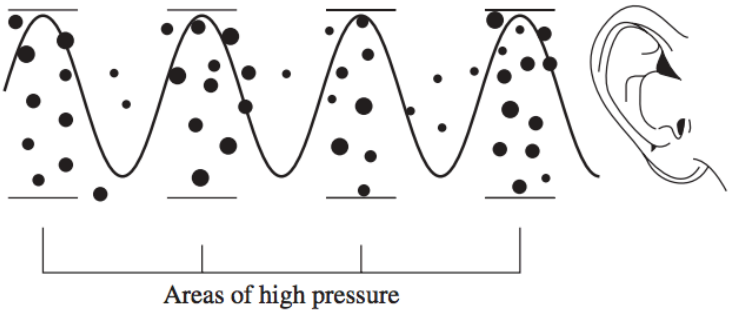


Figure 1.4 Changing air pressure caused by vibration of air molecules

Sinusoidal functions

$s(t) = A \sin(\omega * t + \Phi)$

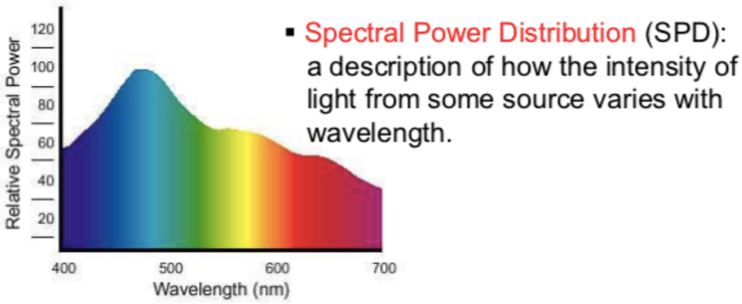
Complex Waveforms

Most waveforms are complex, i.e. their shape is the result of adding 2 or more waveforms.

Each complex waveform may be described as the sum of a number of simple sine waves, each with a particular amplitude, frequency (or wavelength) and phase.

Example: Visible Light

The wavelengths of visible light lie roughly between 400 nm and 700 nm.



Spectral Power Distribution (SPD): a description of how the intensity of light from some source varies with wavelength

Fourier analysis

Fourier analysis attempts to represent a set of data with a series of sines and cosines with different periods, amplitudes, and phases

- The complex wave can be decomposed into the sum of simple waves

the data measurements in the time domain are transformed into the period or frequency domain

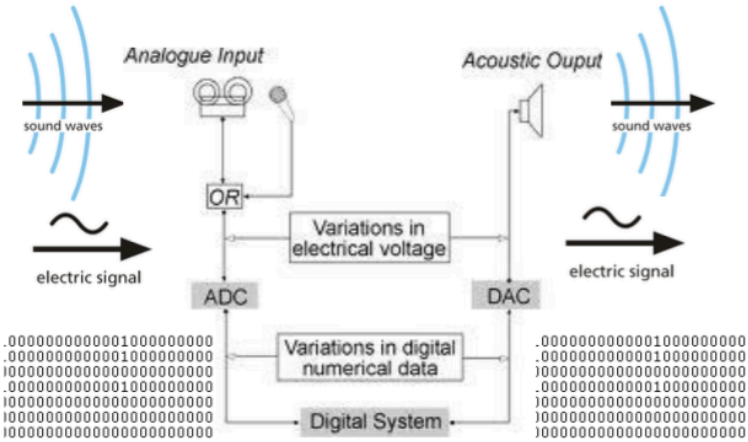
$$F(\omega) = \mathcal{F}[f(t)] = \int_{-\infty}^{\infty} f(t)e^{-iwt}\,dt$$

Information in a wave

Pure tone: A sound represented by a completely regular sine wave
Pitch: the frequency of a wave inform us about the pitch of the sound
Loud: The amplitude of a wave inform us about how loud the sound is
Color: The wavelength of a light ray

sampling and quantisation

Analogue-to-digital conversion



Analogue-to-digital conversion = converting the continuous phenomena of images, sound, and motion into a discrete representation that can be handled by a computer.

- Digitised pictures and sound can be captured in fine detail
- Digital data communication is less vulnerable to noise than is analogue
- Digital data can be communicated more compactly than analogue when compressed

Sampling

Sampling: chooses discrete points at which to measure a continuous phenomenon(a signal)

- images: evenly seperated in space
- sound: evenly seperated in time

Sampling rate: The number of samples taken per unit time or unit space

Quantisation

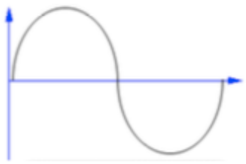
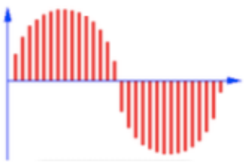
Quantisation: each sample should be represented in a fixed number of bits, called the sample size or the bit depth.(For image, color depth)

- bit depth limits the precision with which each sample can be represented

Let *n* (bit depth) be the number of bits used to quantize a digital sample. Then the maximum number of different values that can be represented, m, is

$$m = 2^n$$

sampling rate: frequency of a wave
quantization: the amplitude of a wave


Aliasing (sampling error)

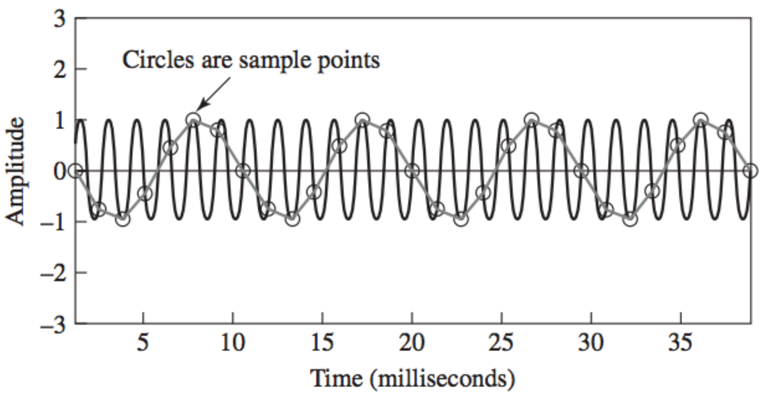


Figure 1.13 637 Hz audio wave sampled at 770 Hz

Undersampleing causes aliasing, the original data cannot be reproduced

Nyquist theorem

Let *f* be the frequency of a sine wave. Let *r* be the minimum sampling rate that can be used in the digitisation process such that the resulting digitised wave is not aliased.

r=2f (*r* is called the Nyquist rate)

Sample twice as often as the highest frequency you want to capture (Avoid aliasing)

Quantisation error

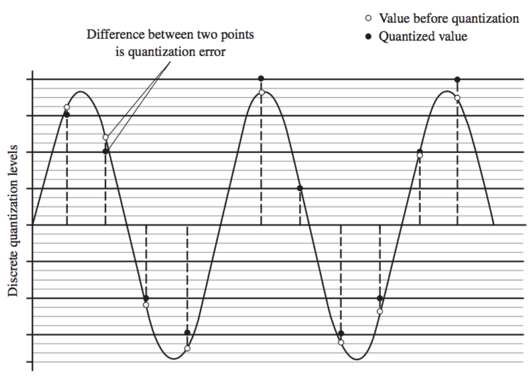


Figure 1.14 Quantization error

As bit depth increases, the precision increases. For imgae, it will lose the colors (loss details in general).

Signal to Noise Ratio (SNR)

SNR: the ratio of the meaningful content of a signal versus the associated noise

- Analogue: SNR is defined as the ratio of the average power in the signal versus the power in the noise level
- Digitised image or sound: signal-to-quantisation-noise ratio (SQNR)

$$SQNR = 20 \log_{10}\left(\frac{\max(quantization\ value)}{\max(quantization\ error)}\right)$$

SQNR

measured in decibels

related to dynamic range: the ratio of the largest sound amplitude (or color, for images) and the smallest that can be represented with a given bit depth

Let n be the bit depth of a digitised media file (e.g. digital audio). Then the signal-to-quantisation noise ratio SQNR is:

$$SQNR = 20 \log_{10}(2^n)$$

$$SQNR = 20 \log_{10}\left(\frac{\max(quantization\ value)}{\max(quantization\ error)}\right) = 20 \log_{10}\left(\frac{2^{n-1}}{1/2}\right)$$

Exercise 1:

If you are recording an audio file and you expect that the highest frequency in the file will be 10,000 Hz, what is the minimum sampling rate you should use to ensure that you will not get audio aliasing?

solution: fm = 10000Hz, fs >= 2*fm = 20000Hz

Exercise 2:

The number of possible colours in an image is determined by the quantisation rate. How many different colours can be represented with 12 bits?

solution: bit depth = 12 bits, 2^12 = 4096 colors can be represented

Exercise 3:

Prove that, if we double the number of bits used to hold a quantised value, then we square the number of quantisation levels.

Proof: assume m bits is used to hold a quantised value, now 2m bits

Original level: 2^m, Now: 2^{2m} In another form $\left(2^m\right)^2$

Sampling and quantisation rate determine the size of the digitised data

Example 1:

Data size: audio file

Sampling rate: 44.1 kHz (44,100 samples/s) Quantisation rate: 32 bits per sample (16 for each of two stereo channels)

Number of minutes: 1 minute

solution: fs = 44.1kHz, quantisation rate: 32bit/sample

R = fs * NB = 44100 $\cdot$ 32 = 1411200 bps = 1.41Mbps
File size = R * 1min = 84672000 bits = 10584000 bytes = 10.58MB

Example 2:
Sampling: 1024 pixels x 768 pixels (samples) Bits per pixel: 24
solution: File size $= \dfrac{1024 \cdot 768 \cdot 24}{8} = 2359296$ bytes = 2.36MB
Example 3:
Data size : video file
Sampling: 720 pixels x 480 pixels Bits per pixel: 24
Frame rate : 30 frames/s Length: 1 minute
Audio : 44,1 KHz, 32 bits
solution:
Image $= \dfrac{720 \cdot 480 \cdot 24 \cdot 60}{8} \cdot 30 = 1.86624$ GB
Audio $= \dfrac{44100 \cdot 32}{8} \cdot 60 = 0.010584$ GB
In total = 0.010584 + 1.86624 = 1.88 GB

Example 4: audio file
CD quality, mono, duration is 1 min
solution: CD quality = 44.1kHz | mono = 16 bits
(Mono = 16 bits, stereo = 32 bits)
File size $= \dfrac{44100 \cdot 16}{8} \cdot 60 = 5292000 = 5.292$ MB
Example 5: Compute the number of bytes needed for 1 minute of video that has 720 x 576 pixels per frame, 25 frames per second, 3 bytes per pixel, and CD-quality stereo audio.
Solution:
Audio $= \dfrac{44100 \cdot 32}{8} \cdot 60 = 10584000$ bytes
Image $= 720 \cdot 576 \cdot 3 \cdot 25 \cdot 60 = 1.86624$ GB
Total = 1.88 GB

Example in Test

Question 1

a) This question is about digitisation. Consider a sound wave W with a frequency f = 440 Hz.

[8 marks]

i) What is the sine function representing W?

(1 mark)

ii) What kind of sound is represented by a completely regular sine wave such as W?

(1 mark)

iii) What does the amplitude of W tell us about the sound it represents?

(1 mark)

iv) What is the minimum sampling rate you should use to ensure that you can digitise W without audio aliasing? Justify your answer.

(2 marks)

v) You decide to use 5 bits per sample. How many different values can W take?

(1 mark)

vi) Calculate an approximation of the Signal-to-Quantisation Noise Ratio (SQNR) of W. Explain your calculation.

(2 marks)

1) w(t) $= \sin\left(2\pi f\right) = \sin(880\pi)$
2) Pure tone
3) The loudness of the sound wave W (How loud the sound wave W is)
4) According to the Nyquist theorem, in order to avoid the audio aliasing, the sampling rate need to be greater than the twice of the largest frequency.
   $fs \geq 2 \cdot fm$. In this case, $fs \geq$ 880Hz. The minimum sampling rate is 880Hz
5) $2^5 = 32$, 32 values can W take
6) SQNR $= 20\log\left(2^5\right) = 30.1$

$$SQNR = 20\log_{10}\left(\frac{\max(quantization\ value)}{\max(quantization\, error)}\right) = 20\log_{10}\left(\frac{2^{n-1}}{1/2}\right)$$

b) This question is about digitisation.

i) How can you prevent aliasing when digitising sound?

ii) Now explain the appearance of an image that has been subsampled.

1) applying the Nyquist theorem to decide the sampling rate when digitsing sound
2) Sub-sampling =

Lose some of the details

c) This question is about image encoding.

i) Calculate the size in kilo bytes (KB) of a 200 x 300 pixels true colour image.

ii) Now calculate the size of the grayscale version of the same image.

1) color depth = 24 bits = 3 bytes = $200 \cdot 300 \cdot 3 = 180000$ B = 180 KB
2) color depth = 8 bits = 1 bytes = $200 \cdot 300 \cdot 1 = 60000$ B = 60 KB

1. What is the mathematical transformation that allows the extraction of the individual frequency components of a complex wave form?
Solution: Fourier transform
2. In the digitisation process, what is the Nyquist theorem useful for?
Solution: Avoid the aliasing issue when digitising signals
it is useful to avoid aliasing by choosing <u>an appropriate sampling rate</u>
3. Calculate the size in <u>bits</u> of a grayscale image file, which has the following characteristics: frame size is 150 pixels x 100 pixels
Solution: $150 \cdot 100 \cdot 8 = 120000$ bits
4. Calculate the size in <u>bytes</u> of a video file, which has the following characteristics: frame size is 300 pixels x 200 pixels, true colour encoding, frame rate is 30 frames/s, no audio track, duration is 2 minutes
Solution: $300 \cdot 200 \cdot 3 \cdot 30 \cdot 120 = 648$ MB
5. Calculate the size in bytes of an audio file, which has the following characteristics: CD quality, mono, duration is 1 minute
Solution: CD quality = 44.1kHz, mono = 16 bits, duration = 60s
File size = $\frac{44100 \cdot 16 \cdot 60}{8}$ = 5292000 bits = 5.292 MB