Introduction

Sound is another form of data that is stored on a computer. Similar to other types of information such as numbers, text and image, the sound is also stored in digital form. But, sound, when produced, is in analogue form, that is, continuous varying data. To store sound in the computer, it is sampled to obtain thousands of samples per second where each sample corresponds to a binary value.

Sampling

Sampling is the process that converts analogue sound into discrete digital data that can be stored in computer. An analogue sound signal is represented as shown in the figure below. This sound is sampled every two seconds to obtain discrete digital sound.

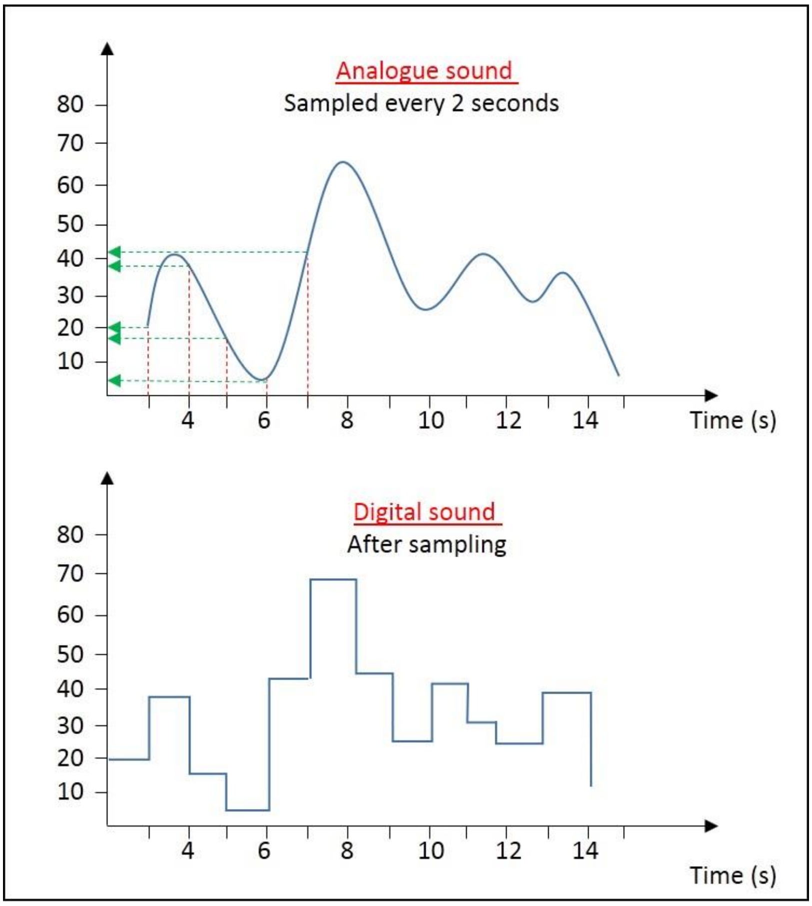

Figure 1: Sampling of an analogue audio signal

We can see that that shape of digital sound is quite similar to analogue sound but the curves are not smooth. A listener will feel that it is the digital sound is the same as the analogue sound but the quality has reduced. The audio quality of a digital sound is affected by factors such as sampling rate, bit depth and bit rate.

Factor	Definition
Sampling rate	Number of samples per second
Bit depth	Number of bits used to represent each clip
Bit rate	Number of bits used per second of audio

Sampling rate

The sampling rate is the number of samples taken per second. The higher the sampling rate, the higher the sound detail. The ups and downs of the sound wave can be recorded more clearly. The unit for sampling rate is also represented in Hertz (Hz). Each sample represents the amplitude of the wave at a certain point in time. The following image illustrates the effect of sampling rate on the quality of the audio wave. The higher the sampling rate, the more similar is the discrete form to the analogue form and better is the quality.

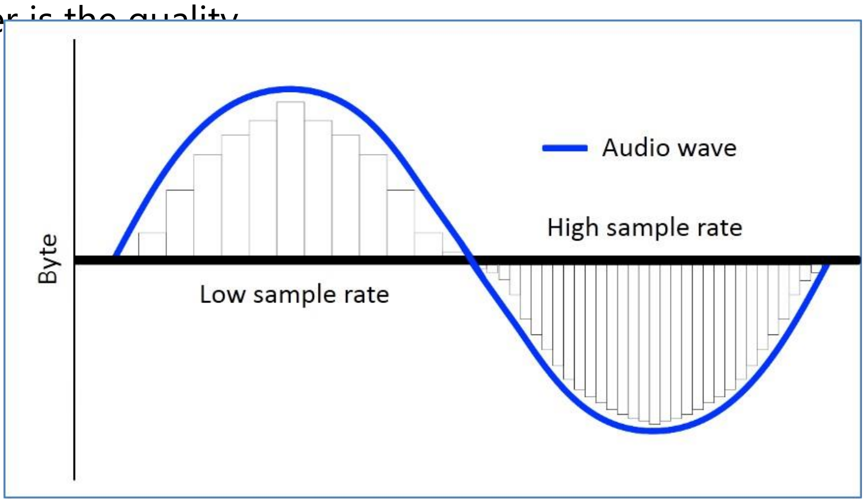

Figure 2: Sampling rate

The most common sampling rate for music is 44,100 samples per second, which is 44,100 Hz (=44.1 kHz). A voice over Internet protocol (VOIP) has a sampling rate of 8 kHz, which is enough for the human voice to be heard clearly but the quality is reduced to a certain extent.

Bit depth

Bit depth is the number of bits available for each sample. The higher the bit depth, the higher the quality the audio will be. A CD has a bit depth of 16 bits and a DVD has a bit depth of 24 bits. An n bit system can have 2^n different values. Hence, a CD can represent values from 0 to 65535(2^{16} -1).

High-quality audio files are created as pulse-code modulation (PCM). PCM is the process for digitising a sound file and creating an uncompressed file. WAV and AIFF are a few examples of uncompressed audio file formats.

Bit rate

Bit rate is the number of bits of data used to store data sampled every second. The unit for bit rates are kilobits per second (kbps).

$$\text{Bit rate} = \text{Sampling rate} \times \text{bit depth} \times \text{channels}$$

An audio file has 44,100 samples per second, bit depth of 16-bits and 2 channels (stereo audio). Bit rate of this file can be calculated as:

$\text{Bit rate} = 44100 \times 16 \times 2 = 1{,}411{,}200$ bits per second $= 1411.2$ kbps

A reasonable music audio must have a minimum bit rate of 128 kbps. The more the bit rate the better the quality. This is the reason why the audio quality of a music CD is better than downloading from the internet.

A three-minute song of audio with bit rate 1411.2 kbps, bit depth 16 bits and 2 channels, has a bit rate of 1411.2 kbps per second.

For 3 minutes, the number of bits required is, $1{,}411{,}200 \times 180 = 254{,}016{,}000$ bits.

This value is equal to $254016000 \div 8 = 31752000$ bytes $= 31.75$ megabytes (MB). This is the file size of a three-minute audio song.