Digital Audio Fundamentals

Dr. Ahmed GHENABZIA

Msila University

Faculty of Mathematics and Computer Science

Department of Computer Science

Email : ahmed.ghenabzia@univ-msila.dz

1.0 March 2024

Table of contents

Objectives 3

I - Content Map 4

II - Prerequisites 5

III - Prerequisites Test 6

IV - Introduction 7

V - Sound Definition 8

VI - Frequency and Pitch 9

VII - Quiz 11

VIII - Digitalization 12

1. Introduction 12
2. Sampling and Sample rate 13
3. Quantization and Bit Depth 14
4. Dynamic Range 14
5. Quiz 15

IX - Final Exam 17

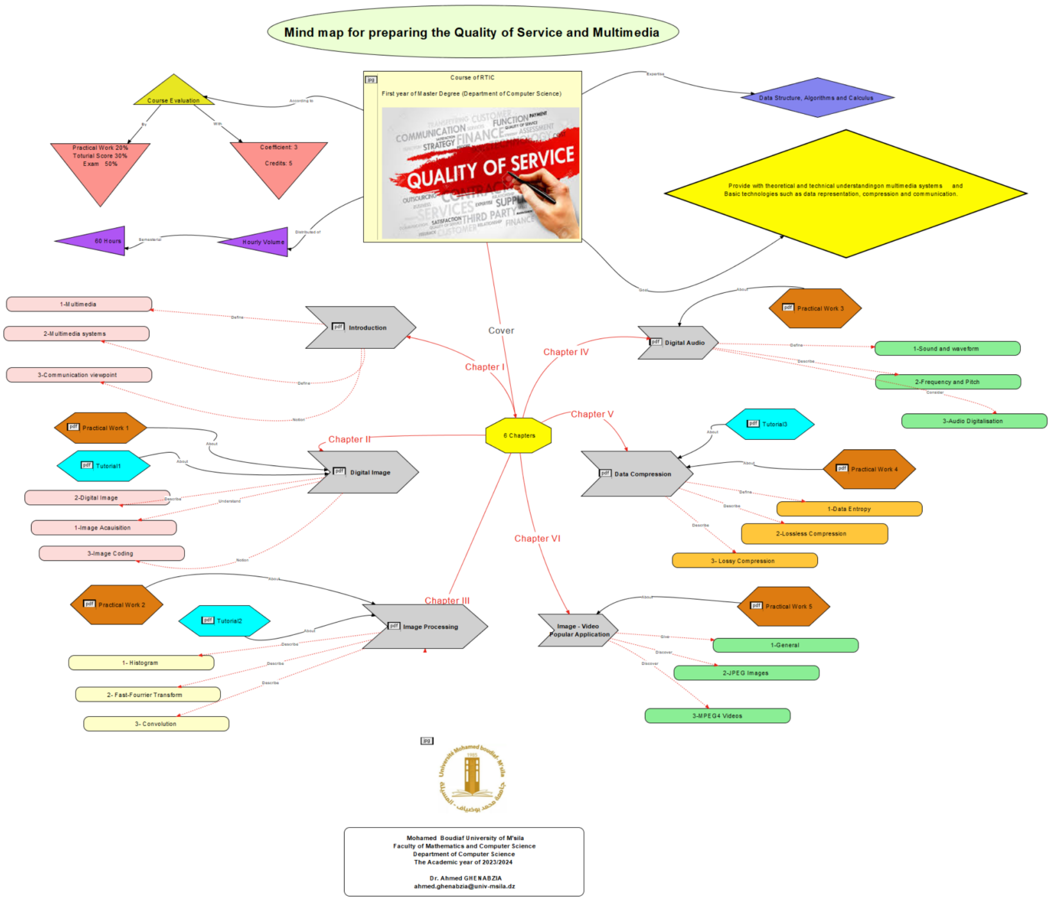
Conclusion 19

References 20

Objectives

The objective of this module seems to be providing an understanding of digital audio, covering essential concepts such as sound definition, frequency, pitch, digitalization, sampling, quantization, bit depth, and dynamic range. Through this module, students can grasp the fundamental principles of how sound is represented and manipulated in digital form, including the processes involved in converting analog sound waves into digital data, the factors influencing audio quality, and the trade-offs between accuracy, storage space, and processing time. By exploring topics like sampling rate, quantization, bit depth, and dynamic range, learners can gain insights into the technical aspects of digital audio processing, which are crucial in various multimedia applications and industries. Additionally, the module may aim to familiarize students with terminology, principles, and techniques commonly used in audio engineering, digital signal processing, and multimedia production.

Content Map



Mind map for preparing the Quality of Service and Multimedia
Course of RTIC
First year of Master Degree (Department of Computer Science)
Course Evaluation
According to
By
With
Practical Work 20%
Toturial Score 30%
Exam 50%
Coefficient: 3
Credits: 5
Hourly Volume
Distributed of
Semesterial
60 Hours
Expertise
Data Structure, Algorithms and Calculus
Goal
Provide with theoretical and technical understandingon multimedia systems and
Basic technologies such as data representation, compression and communication.
Cover
6 Chapters
Chapter I
Introduction
Define
1-Multimedia
2-Multimedia systems
Notion
3-Communication viewpoint
Chapter II
Digital Image
Practical Work 1
Tutorial1
About
Describe
2-Digital Image
Understand
1-Image Acauisition
3-Image Coding
Chapter III
Image Processing
Practical Work 2
Tutorial2
1- Histogram
2- Fast-Fourrier Transform
3- Convolution
Chapter IV
Digital Audio
Practical Work 3
1-Sound and waveform
2-Frequency and Pitch
Consider
3-Audio Digitalisation
Chapter V
Data Compression
Tutorial3
Practical Work 4
1-Data Entropy
2-Lossless Compression
3- Lossy Compression
Chapter VI
Image - Video
Popular Application
Practical Work 5
Give
Discover
1-General
2-JPEG Images
3-MPEG4 Videos
Mohamed Boudiaf University of M'sila
Faculty of Mathematics and Computer Science
Department of Computer Science
The Academic year of 2023/2024
Dr. Ahmed GHENABZIA
ahmed.ghenabzia@univ-msila.dz

Prerequisites

1. **Basic Knowledge**: Having a foundational understanding of physics, especially concepts like sound waves, frequency, amplitude, and wavelength, will be beneficial.
2. **Computer Literacy**: Familiarize yourself with basic computer operations and file management. Understanding how to navigate through folders and work with files is crucial.
3. **Mathematics**: While not mandatory, having a basic understanding of mathematics, particularly algebra and some trigonometry, can help you comprehend certain concepts in digital audio, such as signal processing and frequency modulation.

Prerequisites Test

Quiz 1

What is the relationship between frequency and wavelength in a sound wave?

- They are inversely proportional
- They are directly proportional
- They are unrelated
- They fluctuate randomly

Quiz 2

If a digital audio signal has a sampling rate of 44.1 kHz, how many samples are taken per second?

Quiz 3

Write an equation to represent the relationship between the frequency of a sound wave and its wavelength.

Introduction

Audio information plays a pivotal role in multimedia presentations, represent- ing one of the most fundamental forms of multimedia data. Despite its ap- parent simplicity, distinctions between audio and image information are significant.Unlike video frames, which can be occasionally omitted to enhance viewing speed, sound information cannot be similarly treated without losing vital perceptual cues.

In this chapter, we delve into the essential concepts of sound in multimedia,The digitization of sound inevitably involves the processes of sampling and quantization, Moving forward, we scrutinize the intricacies of audio quantization, providing foundational insights into how digital audio is managed for storage or transmission.

Sound Definition

Sound is a wave phenomenon like light, but it is macroscopic and involves molecules of air being compressed and expanded under the action of some physical device.For example, a speaker in an audio system vibrates back and forth and produces a longitudinal pressure wave that we perceive as sound. (As an example, we get a longitudinal wave by vibrating a Slinky along its length; in contrast, we get a transverse wave by waving the Slinky back and forth perpen- dicular to its length).

Without air there is no sound - for example, in space. Since sound is a pressure wave, it takes on continuous values, as opposed to digitized ones with a finite range. *MUSC 208 Winter 2014 - Digital Audio Fundamentals* p.20

Nevertheless, if we wish to use a digital version of sound waves, we must form digitized representations of audio information. Although such pressure waves are longitudinal, they still have ordinary wave properties and behaviors, such as reflection (bouncing), refraction (change of angle when entering a medium with a different density), and diffraction (bending around an obstacle).

This makes the design of “surround sound” possible. Since sound consists of measurable pressures at any 3D point, we can detect it by measuring the pressure level at a location, using a transducer to convert pressure to voltage levels.

Frequency and Pitch

A sound wave is produced by a vibrating object in a medium, say air. No matter what the vibrating object is, it is vibrating or moving back and forth at a certain frequency. This causes the surrounding air molecules to vibrate at this same frequency, sending out the sound-pressure wave. The frequency of a wave refers to the number of complete back-and forth cycles of vibrational motion of the medium particles per unit of time. The common unit for frequency is Hertz (Hz) where the *unit of time is 1 second*[*Boston University - Fundamentals of Digital Audio* p.20].

$$1Hz = \frac{1cycle}{second}$$

The period of a wave is the time for a complete back-and-forth cycle of vibrational motion of the medium particles. Shown in Figures 1 and b are two simple sine wave waveforms. If the tick mark on the horizontal axis marks the first second, then the frequency of the wave in Figure 1 has a frequency of 2 Hz, because it completes two cycles within 1 second. In Figure 1 , the wave has a frequency of 4 Hz. Sound frequency is related to the pitch of the sound. Higher frequencies correspond to higher pitches. Generally speaking, the human ear can hear sound ranging from 20 Hz to 20,000 Hz.

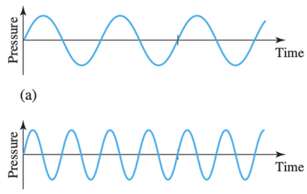


Figure 1:Sample waveforms

In general, any signal can be decomposed into a sum of sinusoids, if we are willing to use enough sinusoids. Figure 2 shows how weighted sinusoids can build up quite a complex signal. Whereas frequency is an absolute measure, pitch is a perceptual, subjective quality of sound. Generally, pitch is relative. Pitch and frequency are linked by setting the note A above middle C to exactly 440 Hz. An octave above that note corresponds to doubling the frequency and takes us to another A note. Thus, with the middle A on a piano (“A4” or “A440”) set to 440 Hz, the next A up is 880 Hz, one octave above; and so on.

Given a sound with fundamental frequency f , we define harmonics as any musical tones whose frequencies are integral multiples of the fundamental frequency, i.e., 2 f , 3 f , 4 f , ..., etc. For example, if the fundamental frequency (also known as first harmonic) is f = 100 Hz, the frequency of the second harmonic is 200 Hz. For the third harmonic it is 300 Hz, and so on. The harmonics can be linearly combined to form a new signal. Because all the harmonics are periodic at the fundamental frequency, their linear combinations are also periodic at the fundamental frequency.

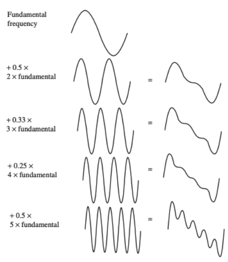


Figure 2:Building up a complex signal by superposing sinusoids

Figure 2 shows the appearance of the combination of these harmonics. Now, if we allow any (integer and noninteger) multiples higher than the fundamental frequency, we allow overtones and have a more complex and interesting resulting sound. Together, the fundamental frequency and overtones are referred to as partials. The harmonics we discussed above can also be referred to as harmonic partials.

Quiz

VII

Quiz

Which of the following properties does sound share with other wave phenomena?

- Reflection
- Refraction
- Diffraction
- All of the above

Quiz

Describe how sound is produced by a vibrating object in a medium like air.

Quiz

Calculate the frequency of a wave if the period of the wave is 0.01 seconds.

Quiz

fill the gap

The human ear can typically hear sound ranging from ______ Hz to ______ Hz.

Digitalization

VIII

1. Introduction

In Figure 3, the depiction underscores the singular dimensionality of sound. The amplitude values undergo changes over time, where pressure either increases or decreases . This singular dimensionality is denoted as a 1D signal, distinguishing it from images reliant on two variables, x and y, or videos, which further involve the variable t.

The continuous nature of the amplitude values necessitates digitization when handling such data in computer storage. Digitization involves the conversion of analog signal - representing continuous-valued voltages from micro phones - into a stream of numbers. The preference often leans towards integers

for efficiency. Despite the two-dimensional representation in Figure 3, a complete digitization of the signal necessitates sampling in each dimension—both in time and amplitude. In this context, sampling refers to the measurement of the quantity of interest at regularly spaced intervals.

The first kind of sampling—using measurements only at evenly spaced time intervals—is simply called sampling (surprisingly), and the rate at which it is performed is called the sampling frequency. Figure 4 shows this type of digitization.

For audio, typical sampling rates are from 8 kHz (8,000 samples per second) to 48 kHz. The human ear can hear from about 20 Hz (a very deep rumble) to as much as 20 kHz; above this level, we enter the range of ultrasound. The human voice can reach approximately 4 kHz and we need to bound our sampling rate from below by at least double this frequency (see the discussion of the Nyquist sampling rate, below).

Thus, we arrive at the useful range about 8 to 40 or so kHz. Sampling in the amplitude or voltage dimension is called quantization, shown in 4. While we have discussed only uniform sampling, with equally spaced sampling intervals, nonuniform sampling is possible. This is not used for sampling in time but is used for quantization (see the μ-law rule, below). Typical uniform quantization rates are 8-bit and 16-bit; 8-bit quantization divides the vertical axis into 256 levels, and 16-bit divides it into 65,536 levels.

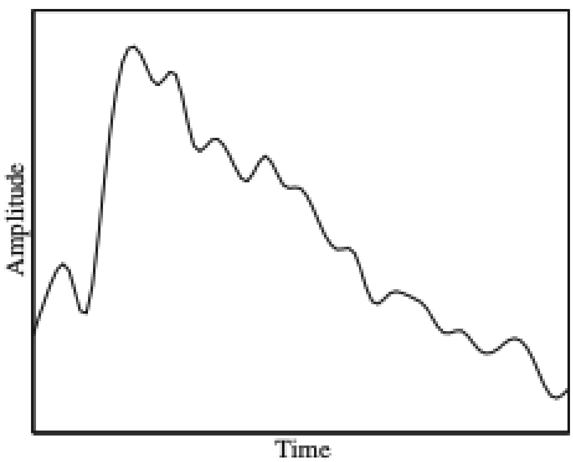


Figure 3: An analog signal: continuous measurement of pressure wave

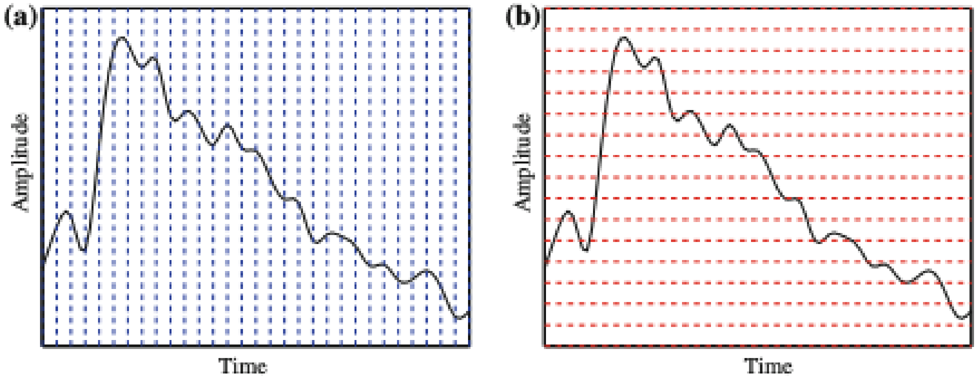


Figure 4: Sampling and quantization: a sampling the analog signal in the time dimension; b quantization is sampling the analog signal in the amplitude dimension

2. Sampling and Sample rate

In the sampling step, the sound wave is sampled at a specific rate into discrete samples of amplitude values. The higher the sampling rate, the higher the ac- curacy in capturing the sound. However, a higher sampling rate will generate more sample points, thus requiring more storage space and processing time. To give you a feel for the sampling rate for digital sound, the sampling rate for CD-quality audio is 44,100 Hz (i.e., 44,100 samples per second). To keep the illustration simple and clear, our examples will use very low sampling rates. These rates are too low to be practical for digitizing sound in real life, but they are simple enough to demonstrate the point. Figure 6 represents a continuous sound wave signal. When we digitize the sound wave, we take discrete sam- ples. Figure 6 shows a sampling rate of 10 Hz (that is, taking 10 samples of the pressure per second). Figure 6 shows a simple reconstruction of the wave by keeping the pressure value a constant between sample points. As you see, changes—crests and troughs—between any two sample points are missed.

Be careful not to confuse the sampling rate with the audio frequency. Both the sampling rate and the audio frequency are measured in Hertz (Hz), but they are not the same thing. The audio frequency relates to the pitch of the sound. The higher the frequency, the higher the pitch. The sampling rate refers to the number of samples taken per second for a sound wave. The sampling rate is a characteristic of the digitization process, but the audio frequency describes the sensory characteristics of the sound we perceive. The same sampling rate may be used for digitizing sounds of different audio frequencies. Conversely, the same sound may be digitized using different sampling rates to create different digital audio files.

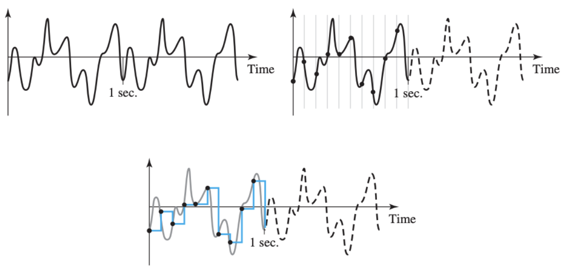


Figure 5: A theoretical continuous sound wave signal and ten sampling actions

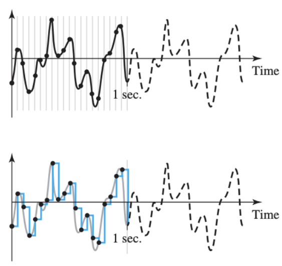


Figure 6: A sampling rate of 20 Hz and A simple reconstruction of the sound wave.

3. Quantization and Bit Depth

In the quantizing step, each of the discrete samples of amplitude values obtained from the sampling step will be mapped and rounded to the nearest value on a scale of discrete levels. Therefore, the more levels available in the scale, the higher the accuracy in reproducing the sound. For digital audio, having more levels means higher resolution. However, higher resolution will require more storage space. The number of levels in the scale is expressed in bit depth the power of 2. For example, an 8-bit audio allows 28 = 256 possible levels.

in the scale. To give you a feel of the bit depth for digital sound, CD quality audio is 16-bit (i.e., 216 = 65,536 possible levels in quantizing the samples of amplitude values). For demonstration purposes, we use 3-bit as an example, which is a scale of eight discrete levels. Again, 3-bit is not a practical bit depth used in real-life applications. The sampled data are mapped to the nearest level on the scale (Figure 7). Some samples may deviate more from their original amplitudes. With a low bit depth, data with different original amplitudes may be quantized onto the same level—for example, note the quantized sample points of the last six sample points in Figure 7.

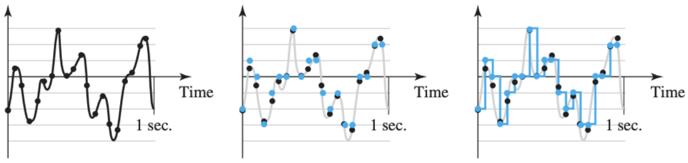


Figure 7: Quantizing with 3-bit resolution

4. Dynamic Range

In the quantization step, a scale of discrete levels of amplitude values is used to map the sample points. The range of the scale, from the lowest to highest possible quantization values in the scale, defines the dynamic range for digitizing audio. In the quantization example of the previous section, where a scale of eight levels (3-bit samples) is used, the lowest level of the scale is placed at about the lowest amplitude of the sound wave and the highest level at about the highest point of the sound wave. The remaining six levels are equally spaced in between these two levels. This scale is extended to include the highest- and lowest-amplitude values of the sound wave. That is, none of the sample points is outside of this range. The scale, shown in Figure 8, covers a full amplitude range of the sound wave.

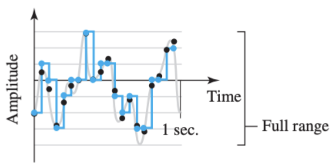


Figure 8: Quantization using a scale range equal to the full amplitude range of the sound wave

If the dynamic range is smaller than the full amplitude range of the sound wave , some data will be lost. The digitized sound wave will be “chopped” off at the limit of the range, causing clipping of the sound wave. Clipping is an undesirable effect because of loss of data. The clipped amplitude values are not recoverable. However, with a reduced dynamic range, the accuracy can be improved for the data within the range. This is an advantage, especially if most of the sample points are within a smaller middle region of the range. By sacrificing a small number of highest- and lowest-amplitude values, the accuracy of the majority of the sample points can be improved. you see that (with the same bit depth) reducing the dynamic range actually allows more subtle changes to be distinguishable (Figure 9). Instead of being quantized to the same value, some sample points can now be set apart on different levels.

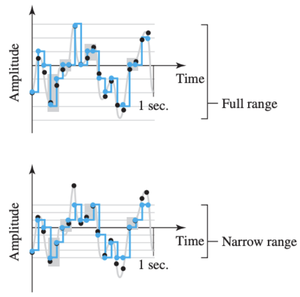


Figure 9: The quantized data (shaded in blue) in which subtle changes become more distinguishable in the reduced dynamic range than in the full range.

What if you extend the dynamic range to more than the amplitude range of the sound wave? Then you will get the opposite result of reducing the dynamic range—the accuracy will be lost. As you see in Figure 10, the amplitude of the sound wave is now within only six levels of the range. So, although this extended dynamic range has eight quantization levels available, only six levels are utilized for this sound wave.

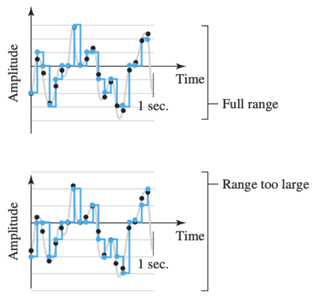


Figure 10: Comparing quantization using two ranges of scale: the full amplitude range of the sound wave, and one wider than the full amplitude range.

5. Quiz

Quiz

What is the typical range for sampling rates in audio?

- 1 Hz to 100 Hz
- 100 Hz to 1,000 Hz
- 8,000 Hz to 48,000 Hz
- 100,000 Hz to 1,000,000 Hz

Quiz

Explain why a higher sampling rate results in more accurate sound representation but requires more storage space.

Quiz

What bit depth is commonly used for CD-quality audio?

- 8-bit
- 12-bit
- 16-bit

- [ ] 24-bit

Quiz

Fill the gap

In quantization, each discrete sample of amplitude values is mapped and rounded to the nearest value on a scale of discrete levels. The more levels available in the scale, the higher the ________ in reproducing the sound.

Quiz

If a 3-bit audio quantization scale has 8 levels, how many levels does a 16-bit quantization scale have?

____

Final Exam

Quiz 1

What is the common unit for measuring frequency?

- Decibels (dB)
- Hertz (Hz)
- Amperes (A)
- Volts (V)

Quiz 2

Which of the following sampling rates is used for CD-quality audio?

- 8,000 Hz
- 16,000 Hz
- 44,100 Hz
- 48,000 Hz

Quiz 3

What is the bit depth of CD-quality audio?

- 8-bit
- 16-bit
- 24-bit
- 32-bit

Quiz 4

What happens if the dynamic range used in quantization is smaller than the full amplitude range of the sound wave?

- Improved accuracy for data within the range
- Increased storage space required
- Reduced number of quantization levels

- ○ Clipping of the sound wave

Quiz 5

Sound can travel through space.

- ○ True
- ○ False

Quiz 6

Higher sampling rates result in more accurate sound representation but require more storage space.

- ○ True
- ○ False

Quiz 7

The term 'frequency' refers to the number of cycles of a wave per second.

- ○ True
- ○ False

Quiz 8

Explain the difference between sampling rate and audio frequency.

Conclusion

In conclusion, this chapter has explored the fundamental concepts of digital audio, delving into the intricacies of sampling, quantization, and dynamic range. We have seen how these processes are essential for transforming continuous sound waves into digital representations suitable for storage, transmission, and manipulation. By understanding these core principles, we gain valuable insights into the fascinating world of digital audio and its applications in various multimedia domains.

References

Boston University - Fundamentals of Digital Audio https://www.cs.bu.edu/fac/snyder/cs583/Handouts/IntroReading.pdf

MUSC 208 Winter 2014 - Digital Audio Fundamentals https://people.carleton.edu/~jellinge/m208w14/pdf/01DigitalAudioFundamentals.pdf