Quadrant II – Notes

Programme: BA/BCom

Subject: Computer Science

Paper Code: CSG108

Paper Title: Multimedia Technology

Unit: 05

Module Name: Sound Design, Audio Codec and File Formats

Module No: 12

Name of the Presenter: Dr Shilpa Neenad Desai

Notes:

Sound is perhaps the most important element of multimedia. Sound can be used to enrich the users experience in Multimedia. For example, a multimedia encyclopaedia of musical instruments will be vastly enriched by the addition of recordings of each instrument. If a multimedia production is destined to be used in an environment where sound hardware is not typically available, then it may be advisable to avoid the use of sound altogether. There are two types of sound that are special: music and speech. These are also the most commonly used types of sound in multimedia productions.

Sound Design

Sound comprises the spoken word, voices, music and even noise. Sound involves a complex relationship involving: a vibrating object (sound source); a transmission medium (usually air); a receiver (ear) and a preceptor (brain). When something vibrates in the air it creates waves of pressure. These waves spread like ripples from pebble tossed into a still pool, when it reaches the eardrums, the change of pressure or vibration which is understood by our brain and we experience sound. Acoustics is the branch of physics that studies sound. Sound pressure levels are measured in decibels (db).

Sound is represented in the form of a sound wave also known as waveform. Every musical instrument creates a different type of waveform. Generally, 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.

Sound is described in terms of two characteristics:

- Frequency (or pitch): 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. Optimally, people can hear from 20 Hz to 20,000 Hz (20 kHz) Sounds below 20 Hz are infrasonic and sounds above 20 kHz are ultrasonic. Individuals response to frequency responses vary greatly based on age,

wherein the upper limit decreases fairly rapidly with increasing age: few adults can hear sounds as high as 20 kHz, although children can.
- Amplitude (or loudness): 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. The amplitude relates to how loud a sound is.

The waveform of any sound is displayed by plotting its amplitude against time. Examination of waveforms can help us characterize certain types of sound.

Digital Audio

A sound is recorded by making a measurement of the amplitude of the sound at regular intervals which are defined by the "sample rate". The act of taking the measurement is often called "sampling" and each measurement is called a "sample point". Digital audio data is the representation of sound, stored in the form of samples point. Quality of digital recording depends on the sampling rate, that is, the number of samples point taken per second (Hz). The three sampling frequencies most often used in multimedia are 44.1 kHz, 22.05 kHz and 11.025 kHz. The higher the sampling rate, the more the measurements are taken (better quality). The lower the sampling rate, the lesser the measurements are taken (low quality).

Some of the quality factors for digital audio file are:
- Sampling Rate: The number of samples point taken per second (Hz)
- Sample Size (resolution): The number of bits used to record the value of a sample in a digitized signal.
- Quality of original audio source.
- Quality of capture device & supporting hardware.
- The characteristics used for capture.
- The capability of the playback environment.

Audio Codecs: A codec (compressor-decompressor) is software that compresses a stream of audio or video data for storage or transmission, then decompresses it for playback. There are many codecs that do this with special attention to the quality of music or voice after decompression. Some are “lossy” and trade quality for significantly reduced file size and transmission speed; some are “lossless,” so original data is never altered. While editing your audio files, be sure to save your files using a lossless format or codec—with repetitive saves in a lossy format, you will notice a quality degradation each time.

An audio codec is a digital electronic device or computer-based software application that aids in the compression and decompression of a digital audio data stream (stream here refers to sending sound or audio data though network). A software-based audio codec essentially consists of an implemented algorithm (set of steps) that codes and decodes an audio stream. A hardware-based audio codec is primarily for analog audio data to be encoded or decoded. An audio codec is also known as a sound codec. An audio codec is used for the compression or decompression of digital audio data from a live stream media. The purpose of using an audio codec is to effectively reduce the size of an audio file without affecting the quality of the sound so that it can be streamed. The process decreases the bandwidth requirement for transmission of an audio signal.

Audio File Formats: A sound file’s format is simply a recognized methodology for organizing and (usually) compressing the digitized sound’s data bits and bytes into a data file.
Each of the three major platforms has its own sound file format:
- AIFF (Audio Interchange File Format) for Mac OS,
- WAV (or WAVE) for Windows, and

- AU for Unix.

Although no standard audio file format has emerged, support for AIFF, WAV, and AU files is common on all platforms. AIFF is historically used for Macintosh sound files. The WAV format was introduced by Microsoft when Windows was first released. Both formats contain uncompressed sound data. But there are huge numbers of sound file formats and “multimedia containers” that store sound data (more than three hundred different file name extensions are used for sound files), and often a converter is required to read or write sound files in the format you need. The MP3 format was developed by the Moving Picture Experts Group (MPEG) and evolved during the 1990s into the most common method for storing consumer audio. MP3 stands for MPEG-1 Layer 3 audio, achieves compression ratios of around 10:1, while maintaining high quality. It incorporates a “lossy” compression algorithm to save space. An audio CD, for example, may hold an hour or so of uncompressed sound. That same CD, using MP3 compression, can store almost seven hours of the same music, but with a slight loss of quality. MP3 has become popular, therefore, as a means of compressing audio, particularly music, for downloading over the Internet — both Legally and illegally. The most common sound formats you might use are wav, aif, aac, flv, mp3, mp4, mov, swf, wma, ogg, or for mobile we use m4r, aac, midi, mmf, 3g2, 3gp, 3gp2,and 3gpp. Be sure your audio software can read and write the formats you need.

WMA (Windows Media Audio) is a proprietary Microsoft format developed to improve MP3. OGG was developed as an open-source and royalty-free “container” for sound compressed using Vorbis algorithms similar to MP3—because the Vorbis sound data resides within an Ogg container, these audio files are normally called “Ogg Vorbis.”

MIDI: Before there was a wide use of mp3 and high bandwidth network, MIDI (Musical Instrument Digital Interface) format audio was popular. MIDI was used when an audio was required to be put on a website. Musical Instrument Digital Interface or MIDI provides a standardized and efficient means of conveying musical performance information as electronic data. It is an easiest and quickest way to compose one’s own score provided one has knowledge of musical instrument and composing. MIDI is in the form of music score and not samples or recording.
To make MIDI score, we need:

- Midi keyboard / Midi keyboard software: MIDI keyboard is used to simplify the creation of music scores also called MIDI information. MIDI information is transmitted in "MIDI messages", which can be thought of as instructions which tell a music synthesizer how to play a piece of music.
- Sequencer software: A MIDI sequencer software lets us to record and edit MIDI data like a word processor and allows operations such as cut and paste; insert / delete etc.
- Sound synthesizer (built-in in to sound card): The synthesizer receiving the MIDI data must generate the actual sounds.

Characteristics of MIDI files: Since MIDI files are small, they can be easily embedded in web pages which load and play. Length of a MIDI file can be changed without affecting the pitch of the music or degrading audio quality. Working with MIDI requires knowledge of music theory. MIDI files can be generated either by recording the MIDI data from a MIDI instrument (electronic keyboard) as it is played or by using a MIDI sequencer software application.

Advantages of MIDI over digital audio:

- MIDI files are smaller that digital audio files. MIDI files embedded in web pages’ load and play more quickly as MIDI files are small in size.
- MIDI sound source of a high quality sound better than digital audio.

- One can change the length of MIDI files without changing the pitch of the music or degrading the audio quality.

Disadvantages of MIDI over digital audio:

- MIDI data does not represent the sound but musical instruments, as a result, playback will be accurate only if the MIDI playback (instrument) is identical to the device used in the production.
- The cost to produce a MIDI file is higher.
- One requires knowledge of music and skill to edit MIDI files.
- MIDI files cannot emulate voice and other effects.