Coding and Data Compression

Lec.8

audio Compression

Two important features of audio compression are (1) it **can be lossy** and (2) it **requires fast decoding**. Text compression must be lossless, but images and audio can lose much data without a noticeable degradation of quality. Thus, there are both lossless and lossy audio compression algorithms. It only rarely happens that a user will want to read text while it is decoded and decompressed, but this is common with audio.

Often, **audio is stored in compressed form and has to be decompressed in real-time when the user wants to listen to it**. This is why most audio compression methods are asymmetric. The encoder can be slow, but the decoder has to be fast. This is also why audio compression methods are not dictionary based. A dictionary-based compression method may have many advantages, but fast decoding is not one of them.

Sound

To most of us, sound is a very familiar phenomenon, because we hear it all the time. Nevertheless, when we try to define sound, we find that we can approach this concept from two different points of view, and we end up with two definitions, as follows:

An intuitive definition: Sound is the sensation detected by our ears and interpreted by our brain in a certain way.

A scientific definition: Sound is a physical disturbance in a medium. It propagates in the medium as a pressure wave by the movement of atoms or molecules.

We normally hear sound as it propagates through the air and hits the diaphragm in our ears. However, sound can propagate in many different media. Marine animals produce sounds underwater and respond to similar sounds. Hitting the end of a metal bar with a hammer produces sound waves that propagate through the bar and can be detected at the other end. Good sound insulators are rare, and the best insulator is vacuum, where there are no particles to vibrate and propagate the disturbance.

Sound can also be considered a wave, even though its frequency may vary all the time. It is a longitudinal wave, one where the disturbance is in the direction of the wave itself. In contrast, electromagnetic waves and ocean waves are transverse waves. Their undulations are perpendicular to the direction of the wave.

Like any other wave, **sound has three important attributes, its speed, amplitude, and period**. The frequency of a wave is not an independent attribute; it is the number of periods that occur in one time unit (one second). The unit of frequency is the hertz (Hz). The speed of sound depends mostly on the medium it passes through, and on the temperature. In air, at sea level (one atmospheric pressure), and at 20◦ Celsius (68◦ Fahrenheit), the speed of sound is 343.8 meters per second (about 1128 feet per second).

The human ear is sensitive to a wide range of sound frequencies, normally from about 20 Hz to about 22,000 Hz, depending on a person's age and health. This is the range of audible frequencies. Some animals, most notably dogs and bats, can hear higher frequencies (ultrasound). A quick calculation reveals the wavelengths associated with audible frequencies. At 22,000 Hz, each wavelength is about 1.56 cm long, whereas at 20 Hz, a wavelength is about 17.19 meters long.

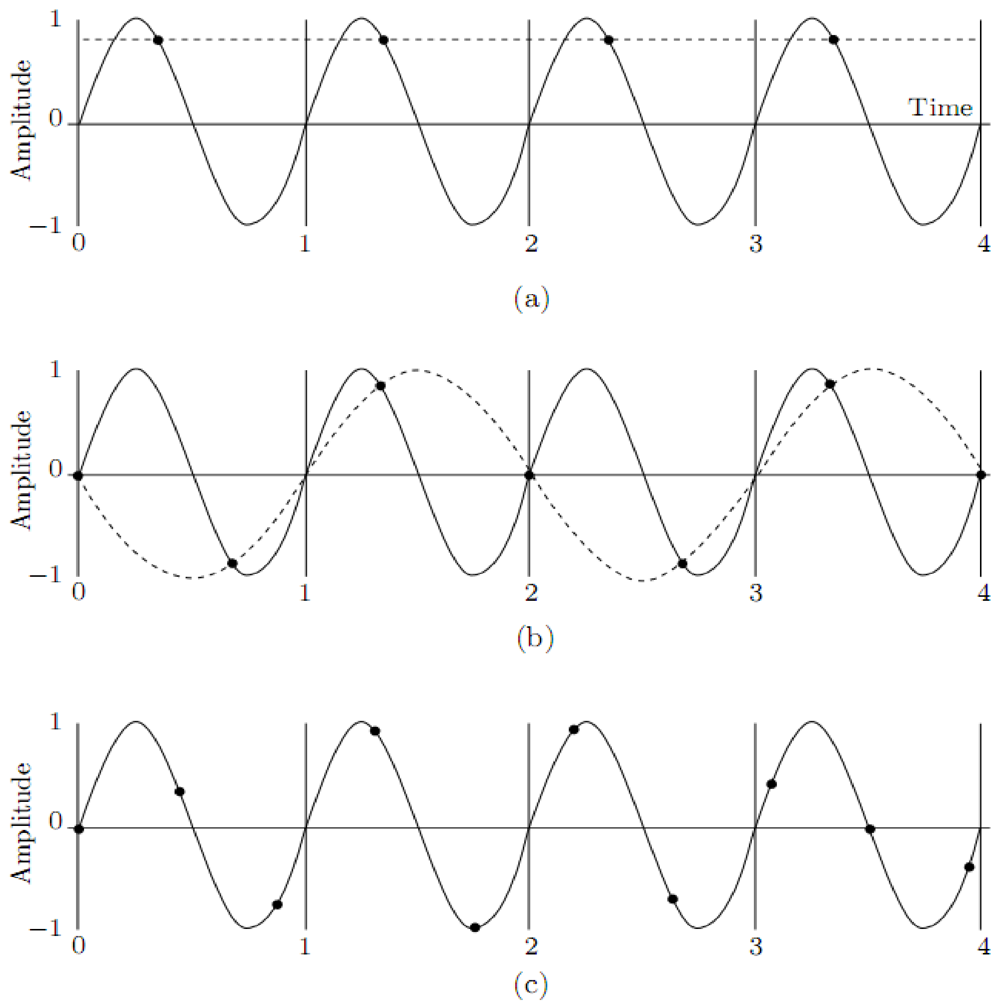
Digital Audio

Much as an **image can be digitized and broken up into pixels**, where each pixel is a number, **sound can also be digitized and broken up into numbers**. When sound is played into a microphone, it is converted into a voltage that varies continuously with time.

Next figure shows a typical example of sound that starts at zero and oscillates several times. Such voltage is the analog representation of the sound. **Digitizing sound is done by measuring the voltage at many points in time, translating each measurement into a number, and writing the numbers on a file. This process is called sampling.**

The sound wave is sampled, and the samples become the digitized sound. The device used for sampling is called an analog-to-digital converter (ADC). The difference between a sound wave and its samples can be compared to the difference between an analog clock, where the hands seem to move continuously, and a digital clock, where the display changes abruptly every second.

Since the audio samples are numbers, they are easy to edit. However, the main use of an audio file is to play it back. This is done by converting the numeric samples back into voltages that are continuously fed into a speaker. The device that does that is called a digital-to-analog converter (DAC). Intuitively, it is clear that a high sampling rate would result in better sound reproduction, but also in many more samples and therefore bigger files. Thus, the main problem in audio sampling is how often to sample a given sound.



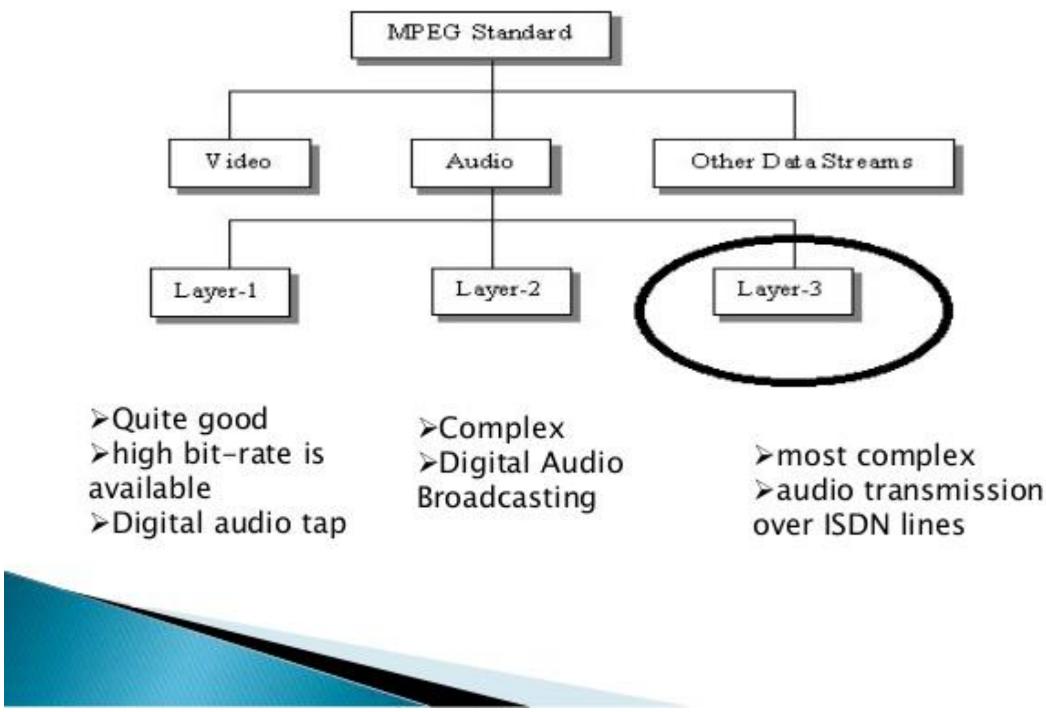
Sampling a Sound Wave.

Figure (a) shows what may happen if the sampling rate is too low. The sound wave in the figure is sampled four times, and all four samples happen to be identical. When these samples are used to play back the sound, the result is silence. Figure (b) shows seven samples, and they seem to "follow" the original wave fairly closely. Unfortunately, when they are used to reproduce the sound, they produce the curve shown in dashed. There simply are not enough samples to reconstruct the original sound wave. Figure (c) which shows 10 equally-spaced samples taken over four periods. Notice that the samples do not have to be taken from the maxima or minima of the wave; they can come from any point.

Conventional Methods

Conventional compression methods, such as RLE, statistical, and dictionary-based, can be used to losslessly compress audio data, but the results depend heavily on the specific sound. Some sounds may compress well under RLE but not under a statistical method. Other sounds may lend themselves to statistical compression but may expand when processed by a dictionary method.

It is possible to get better audio compression by developing lossy methods that take advantage of our perception of sound, and discard data to which the **human ear is not sensitive**. This is similar to lossy image compression, where visual data to which the human eye is not sensitive is discarded. In both cases we use the fact that the original information (image or sound) is analog and has already lost some quality when digitized. Losing some more data, if done carefully, may not significantly affect the played-back sound, and may therefore be indistinguishable from the original.



Categories of Audio Files

There are 3 categories in which certain Audio files belong to:-

1) Uncompressed:
   Ex) .Wav
2) Lossless:
   Ex) .WMA
3) Lossy:
   Ex) .Mp3

Standard Codecs for audio compression

For lossless compression

- LPAC (Lossless predictive audio compression):is an improved lossless audio compression algorithm developed by Tilman Liebchen, Marcus Purat and Peter Noll.
- ALAC (Apple Lossless Audio Codec):is an audio coding format, and its reference audio codec implementation, developed by Apple.
- FLAC(Free Lossless Audio Codec): can typically reduce the original size of audio file to 50–60%, and decompressed it to an identical copy of the original audio data, developed by Josh Coalson.
- WMA Lossless (Windows Media Audio Lossless): developed by Microsoft .

20

Standard Codecs for audio compression

For lossy compression

- Nero AAC Codec(Nero "advanced audio coding" codec): It was developed and distributed by Nero AG.
- FAAC(Freeware Advanced Audio Coder):is an audio compression computer program that creates AAC sound files from other formats , it is the recommended format for the company's iPod music player.

21

International organizations dealing with audio compression standardization

- ISO/IEC
  - MPEG-1 Layer III (MP3)
  - MPEG-1 Layer II
  - MPEG-1 Layer I
  - AAC
  - MPEG-4 ALS
  - MPEG-4 SLS
  - MPEG-D USAC

22

International organizations dealing with audio compression standardization

- ITU-T
  - G.711
  - G.718
  - G.719
  - G.722
  - G.723
  - G.726
  - G.728
  - G.729

23