A STUDY ON AUDIO FILE FORMATS

K.Banuroopa [1], Dr. D. Shanmuga Priyaa [2]

ABSTRACT

Music which is an integrated part of the human life has to be stored so that it can be enjoyed wherever and whenever needed. The audio files can be stored in uncompressed PCM signals or stored in compressed format to storage space. The compression can be lossless or lossy. The lossless format Wav and lossy format Mp3 are most popularly used. While storing audio, it needs to be indexed, annotated and organized. It's easier to add the metadata to a compressed file rather than uncompressed file which takes lots of space. The Mp3 (MPEG-1 Layer III) formats stands out by being the most used format for streaming music through online and in storing in various cloud platforms. It seems the obvious choice to research on Music information retrieval and Content-based audio identification.

Keywords: MP3, Lossless Compression, Lossy Compression

INTRODUCTION

Audio is the sensory perception which begins even before a baby is born. A baby recognizes its mother's voice even before it can see its mother. Music plays an important role in every person's life from lullaby to funeral march. We as humans express our emotions through music. Music is in nature and has been there with mankind from time unknown. According to science a sound travels through a medium and it travels in sine waves. [1]From early 18th century there have been many attempts to record and reproduce sound. Thomas Alva Edison succeeded in his attempt by inventing gramophone in 1877, which has led to various other inventions and methods through which sound can be recorded and reproduced. According to how the sounds are recorded and reproduced they are classified as analog and digital. From 1877 till 1975 the sound was recorded as a sine wave in various devices like mechanical discs to a magnetic disc. From 1975 to till date audio is recorded in digital format.

An analog audio file resembles a sine wave form. Digital audio means the audio signal is converted into 1's and 0's using a Analog-To-Digital converter (ADC).
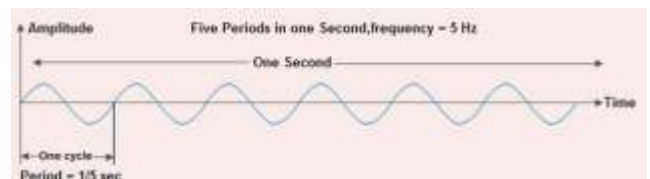

An analog file is described by its frequency, amplitude, phase and period. Frequency is the number of periods in one second or number of cycles per second where a period is time taken by the sine wave to complete one cycle. Frequency is usually measured in hertz (Hz). Amplitude is the maximum height of the signal and phase is the relative position of the signal to zero.

[1]Research Scholar, Department of CS,CA & IT, Karpagam Academy of Higher Education.

[2]Associate Professor, Department of CS,CA & IT, Karpagam Academy of Higher Education.

Pitch, Quality and Loudness are the characteristics of a sound in general. Pitch is the frequency of a sound as heard by human ear. A higher frequency signal produces a high pitch tone and vice versa. A sound file may have many frequencies. If there is only one
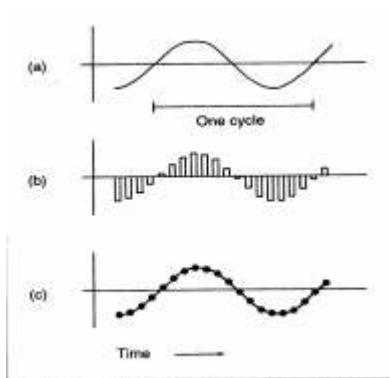
frequency then the sound file is said to be of pure tone. The note of highest amplitude is called a fundamental note. Other frequency will contribute to the quality of the sound. Loudness depends on sound pressure but also on the spectrum of the harmonics and the physical duration.

A digital audio file or a discrete file is created by sampling the analog file or continuous file and determining whether the amplitude is high or low. High amplitude produces a 1 and low amplitude produces a zero. Another way of creating a discrete file from analog file is by applying Fourier transform on it.

Fourier transform splits the continuous wave form into a sum of sine and cosine values [2]. It's really huge amount of data which is impractical to process, so, the need for compressing the audio files. The uncompressed audio formats use pulse code modulated (PCM) signal, in windows environment its WAV format and in Mac OS environment it is AIFF. WAV is a flexible file format designed to store more or less any combination of sampling rates or bitrates.

This makes it an adequate file format for storing and archiving an original recording.

Lossless compression formats

Compression of audio file is said to be lossless when no data is lost when the its uncompressed while a lossy compression does not retain all the information after compression. A lossless compressed format would require more processing for the same time recorded, but would be more efficient in terms of space used. Some of the formats with lossless compression are FLAC (Free Lossless Audio Codec), Monkey's Audio (filename extension APE), WavPack (filename extension WV), Shorten, Tom's lossless Audio Kompressor (TAK), TTA, ATRAC Advanced Lossless, Apple Lossless and lossless Windows Media Audio (WMA). The Table-1 shows the some of the audio lossless audio codecs and their general features.[4]

Codec Name	Compression Ratio	Encoding Speed	Decoding Speed	Category
Apple Lossless	low	average	fast	Proprietary
FLAC	low	fast	fast	open source
LA	high	slow	slow	proprietary
Monkey's [4]	average	average	average	open source
MPEG-4 ALS	high	average	average	standard
MPEG-4 SLS	average	average	average	standard
OptimFROG	high	slow	slow	proprietary

Real Lossless	average	average	fast	proprietary
Shorten	low	fast	fast	open source
TTA	average	fast	fast	open source
WavPack	low	fast	fast	open source
WMA Lossless	average	average	average	proprietary

Lossy Compression Audio formats

A lossy audio compression formats compresses data and then decompresses it to retrieve the data that may be different from the original, but is close enough to be useful in some way. MP3 [11] - are compressed to roughly 1/10 the size of an equivalent PCM file sacrificing quality, but achieves compression through a task known as `psychoacoustic masking which is removing the audio data which is not in human hearing range. Another format Musepack is optimized for transparent compression of stereo audio at bitrates of 160-180 kbit/s, while AAC -Windows Media Audio (WMA) is designed with Digital Rights Management (DRM) which makes it possible to have copy protection[13].

Mp3 files are to be decompressed while playing in mp3 player. Even then, Mp3 seems to give one of the best tradeoff between quality of the music and storage space. The uncompressed audio and lossless audio formats are usually used to audio analysis or editing audio data[7]. The lossy compressed audio formats are used to stream audio data through internet or storage in audio clouds[12]. Mp3 audio format is the mostly used in music analysis[8]. The few datasets available in UCI machine learning repository [5], Wikipedia's machine learning datasets [6], the datasets in International society Music information Retrieval (ISMIR) are either in Mp3 or wave format [10]. The recent dataset which was added on 24th May, 2017 consists of 106,574 tracks in Mp3 format [5].

Conclusion :

In this paper we have listed the various audio file formats available and basics of storage of an audio file. While storing audio, it needs to be indexed, annotated and organized. It's easier to add the metadata to a compressed file rather than uncompressed file which takes lots of space. The Mp3 (MPEG-1 Layer III) formats stands out by being the most used format for streaming music through online and in storing in various cloud platforms. It seems the obvious choice to research on Music information retrieval and Content-based audio identification.

References

1. https://en.wikipedia.org/wiki/History_of_sound_recording

2. Fisher, Kathleen Marie, "Towards Understanding the Compression of Sound Information" (2016). Computer Science Honors Theses. Paper 38.

3. Hacker, S. (2000). MP3: The Definitive Guide. Sebastopol, CA: O'Reilly.

4. Fomazou Nselapi Auristin, IJECS Volume 05 Issue 4 April 2016 Page No.16167-16171

5. MichaëlDefferrard, KirellBenzi, Pierre Vandergheynst, Xavier Bresson, FMA:

A Dataset For Music Analysi, https://archive.ics.uci.edu/ml/datasets.html

6. https://en.wikipedia.org/wiki/List_of_ datasets_for_machine_learning_research

7. Mark A. Bartsch, Gregory H. Wakefield, Audio Thumbnailing of Popular Music Using Chroma-Based Representations, IEEE transactions on multimedia, vol. 7, No. 1, February 2005

8. Garcia-Hernandez JJ, Feregrino-Uribe C, Cumplido R (2013) Collusion-Resistant Audio Fingerprinting System in the Modulated Complex Lapped Transform Domain. PLoS ONE 8(6): e65985. https://doi.org/10.1371/journal.pone. 0065985

9. Ruili Zhou, Yuesheng Zhu, A Robust Audio Fingerprinting Algorithm in MP3 Compressed Domain, World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering Vol:5, No:7, 2011

10. A. Wang. An industrial strength audio search algorithm. In ISMIR Proceedings, Baltimore, 2003.

11. ISO/IEC JTC1/SC29/WG11 (MPEG). Information technology - multimedia content description interface. International Standard 15938, ISO/IEC, 2001.

12. Weaver, James Roger, Best Practice - Scholars' Mine Producing Quality Audio Files" (2017). Scholars' Mine - Policies, Procedures and Guidelines.

13. Majewicz, Eric, "Analysis of digital audio compression and digital rights management techniques" (2004).