Audio Compression

Audio Processing Guide

ELECTGON
www.electgon.com
contact@electgon.com

03.06.2019

Contents

1 Introduction

sound and speech is one of the most important signals in modern systems and the growing need for audio and speech processing (transmission, storing, etc.) generates challenge for effectively performing this processing. Therefore, there is a need to have the processing devices that are able to deal with audio in different applications (telephony, multimedia, broadcasting, etc.). To evaluate audio processing systems, three main parameters are used to describe the quality of its audio: bandwidth, fidelity and spatial realism.

Bandwidth is about how much information can be stored (or transmitted) of the audio signal. The more information about the signal (its accurate level), the higher quality of the system.

Fidelity is the ability to cover the whole bandwidth of the audio signal. Compact disc (CD) technique is able to represent any audio signal within the bandwidth of 20 to 20000Hz. Traditional radio covers the bandwidth of up to 15 KHz for frequency modulation (FM) and only up to 4.5 KHz for analog modulation (AM). Telephone system has a bandwidth of merely 300-3400 Hz [4].

Spatial realism of an audio representation describes the naturalness and quality of directional information about places of particular sound sources contained in the reproduced sound. It depends on number of channels used to represent this audio. Typical configurations are 1-channel (mono), 2-channels (dual), multichannels (4-channels, 5-channels, 8-channels). There is possibility also to add a subwoofer channel to transfer low frequency range (15-150 Hz). So 5.1 channel format is 5-channels sound plus subwoofer channel [4].

2 Audio Processing Systems

For all the previously mentioned applications and specifications of the audio signals, no unique audio processing system is supposed to cover all the specified parameters but they all have general way for processing the audio signal shown in figure 1

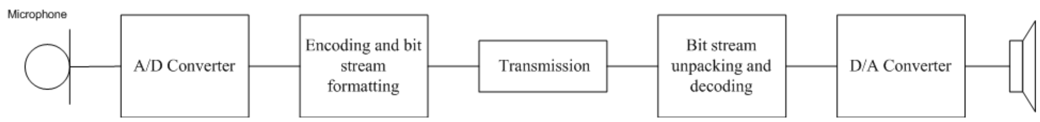

Figure 1: General Scheme for Audio Processing

What can be concluded from that figure is analog and digital processing are needed for the sound. For Analog processing, audio signal x(t) can be handled continuously in time. On the other hand, for digital processing part signal has to be represented in discrete intervals and quantized levels. Not only but also bit rate used to represent the audio signal has to be reduced to meet transmission system requirements. That is why there is different digital processing schemes to deal with the audio signal according to the requirements and the applications.

3 A/D converters

These are the interface between analog domain and digital domain, it shall receive the audio signal in time domain x(t), sample it, quantize it, generate possible code for each quantized sample. Output of A/D converters are Pulse code modulation signal (PCM). This is the original of any digital audio signal. i.e. PCM representation is a digital audio signal with maximum achievable quality. Typical resolution in bits per sample (bps) are 16, 20, 24, 32 and even 48 bps.

4 Encoding and Bit Stream Formatting

PCM representation is not an efficient method for audio storage and transmission. Therefore, compression techniques are needed. That is the main functionality of Encoding Stage. Compression techniques can be classified into lossless coding techniques or lossy techniques. In lossless techniques, all information of the signal is preserved. Lossy techniques (also called transparent coders) are corrupting or losing some information of the signal in order to drastically reduce no. of bits used to represent the digital audio signal. This corruption can be controlled in such a way that is inaudible. These lossy encoders make use of the psychoacoustic perception model of human ear that is discussed in previous section (masking, critical bands, etc) so that they can remove redundancy and irrelevant components from the audio signal. Therefore they make use of this model while allocating bits to the audio sample. In addition to this model, they use also another model to encode audio signal according to its source. For example, some encoders split input audio signal to instrumental audio and speech audio then apply different encoding algorithm to these split signals. Other encoders are just transforming the audio signal into frequency domain divided into subbands that are equivalent to critical bands of human ear then apply this psychoacoustic model to each subband. Figure 2 shows main architecture used to build transparent encoders.

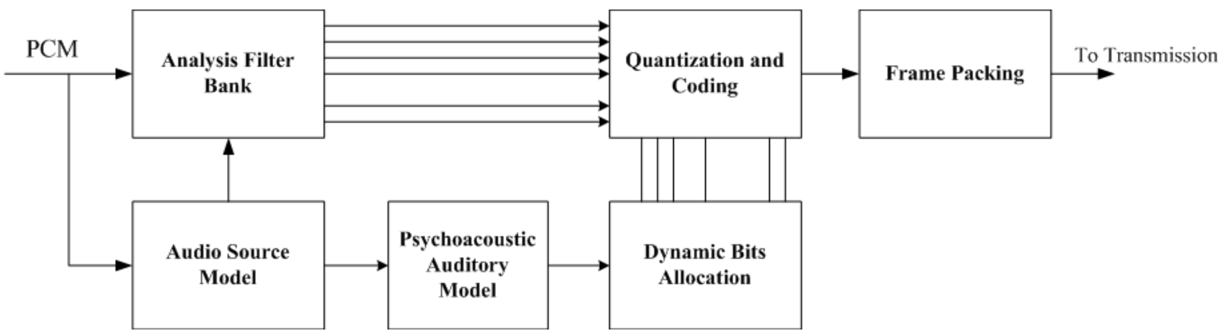

Figure 2: Main Architecture of Lossy Encoders

This figure shows that encoding process is composite of mainly three stages. First stage is the Analysis stage in which Audio Source model is applied using some analysis filters. Second stage is the coding stage in which complementary encoding process is applied according to psychoacoustic model. Last stage is packing the encoded samples into frames for transmission. A lot of encoders exist currently using this lossy compression method. Figure 3 shows different types of these encoders according to its Analysis Stage.

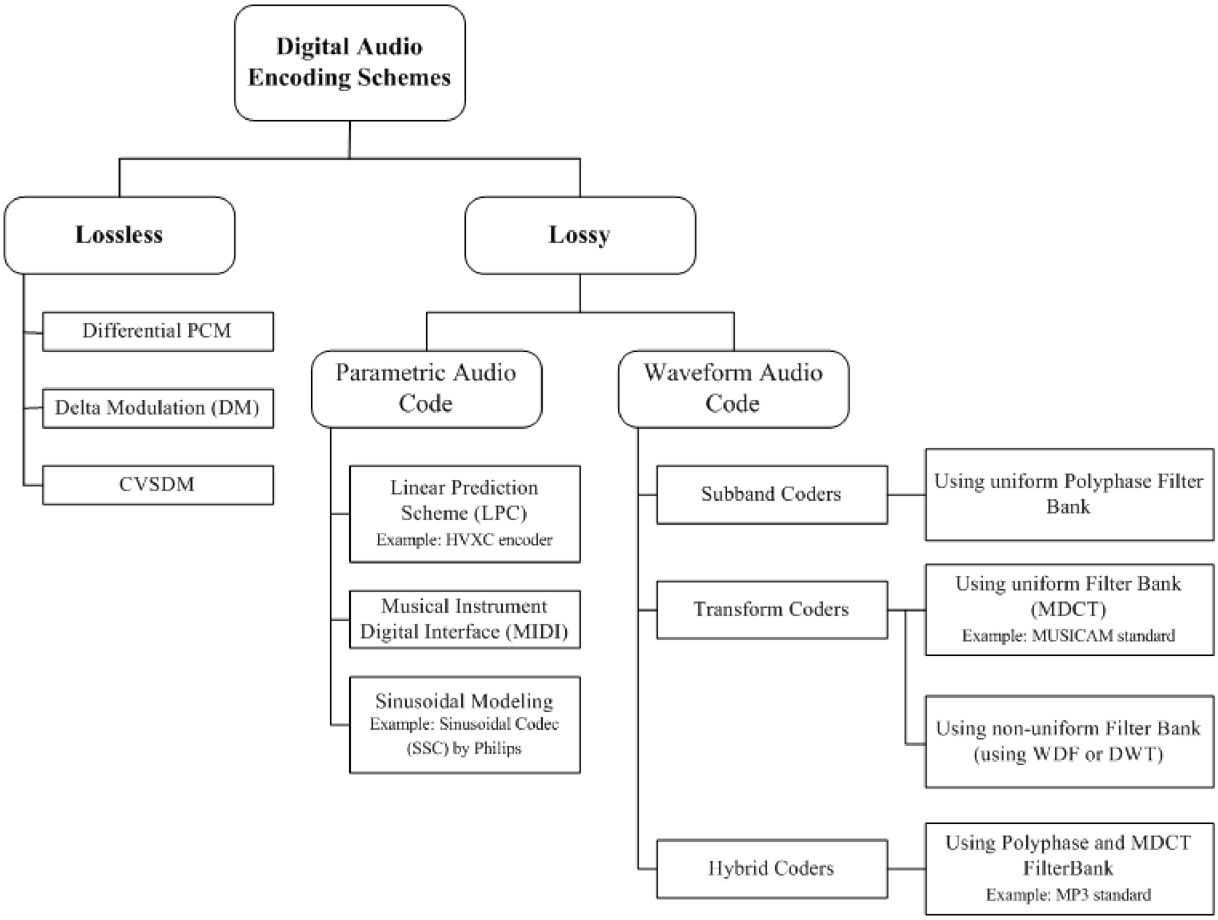

Figure 3: Different Lossy Encoding Schemes

These different encoding schemes led to different standards that are recognized currently and depicted in table 1

Lossless Encoders Standards	Transparent Encoders Standards
Apple Lossless	AAC
Audio Lossless Coding (MPEG-4ALS)	ATRAC
FLAC	Dolby Digital (AC3)
Meridian Lossless Packing	MP3
Monkey's Audio	Ogg Vorbis
Shorten	MUSICAM

Table 1: Different Standards for Audio Encoding[4]

5 Digital Audio Signal Transmission

Based on these defined standards, audio transmission is using suitable audio encoding standard according to the application. For example, for Digital Audio Broadcast (DAB) standard is using MPEG-1 layer II to broadcast its content. Digital Radio Mondiale (DRM) uses AAC for general transmission and HVXC for speech programs. Internet Transmission is using MPEG-1 layer III (MP3) for streaming the audio data.

Bibliography

[1] S. Committe, Health Risk from Exposure to Noise from Personal Music Players, sep 2008.

[2] http://www.intropsych.com, June 2015.

[3] https://www.wikipedia.org/, June 2015.

[4] V. G. Oklobdzija, Ed., Digital Systems and Applications. CRC Press, 2008.