www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume – 5 Issue -04 April, 2016 Page No. 16167-16171

New Ieee Standard For Advanced Audio Coding In Lossless Audio Compression : A Literature Review

Fomazou Nselapi Auristin[1], Samadhan Mali[2]

[1] Department of electronics & telecommunication engineering, Sinhgad College of engineering
fomazouauristin@gmail.com, sdmali.scoe@sinhgad.edu

Abstract: ***Due to non-stop increase of communication subscribers, the loss and distortion of data (voice, video…) is emphasized. Meanwhile, IEEE-SA (Standard Association) has developed a new standard for advanced audio coding (AAC) in August 2013, called IEEE 1857.2 for contributing to those issues. This new standard for AAC is an efficient lossless audio codec (coding decoding) technique, in improving audio quality for compression and decompression, optimizing bandwidth during transmission, saving storage space, speeding up the video streaming, audio streaming and others video or audio data. This survey consists not only of describing the different existing techniques for lossless audio compression, and also of showing the state of the art of this recent lossless standard for AAC upon its most popular predecessors.***

Keywords: codec, Advanced Audio Coding (AAC), IEEE 1857.2, audio compression, lossless coding, LPC.

1. Introduction

For a long time, the optimization of communication resource and of data recording medium is become an attractive area for many researchers. Data compression being defined as the mechanism or method of compacting 'raw' data into a smaller number of bits. Audio compression (Audio coding) is the act of converting digital audio into a format suitable for transmission or storage requiring least bandwidth consumption, whilst typically reducing the number of bits. Uncompressed digital audio coming directly from microphone typically requires a large bitrate. The data compression has undergone until today thru two categories of compression; lossy and lossless compression. If the decoded sequence differs from the original, often the original and compressed data are indiscernible this method is called lossy compression, such as MPEG-1 layer-3 (MP3), AAC, WMA lossy, Musepack etc.

Therefore, if the decoded audio sequence is identical to the original, then the coding process is lossless coding, such as FLAC, monkey's Audio, WMA, MPEG-4 SLS, MPEG-4 ALS, WavPack etc. Audio coding is essentially used for portable audio players, audio streaming, and the storage, lower bandwidth consumption, higher resolution and sampling rates, and playback of movies.

As well as for the conservation of the quality or originality of data during codec, IEEE-SA (Standard Association) is a professional and renowned association which is working in multiple domains of engineering (electrical, electronic, telecommunication, computer engineering, and so on), for contributing in such domains. Named IEEE 1857.2 or new standard for advanced audio coding (AAC) and published in August 2013, it's a new standard developed by IEEE-SA which aims to compress and decompress audio and speech signal. It enables the compression and decompression of digital audio and speech data, without any loss in quality due to a perfect reconstruction of the original signal.

2. LOSSLESS AUDIO CODING ALGORITHMS

The domain of lossless audio compression has grown rapidly during the last decade. Only a few lossless audio compressors existed in the mid-90s. Now there are already over 30 varieties downloadable from internet. This fast growth in the number of lossless compressors is due to the fact that lossless audio compression has become useful and affordable to general users as a result of high broadband penetration, powerful CPUs and low storage costs.

The lossless audio analysis and synthesis has been gone through many techniques and algorithms before reaching to the latest standard for audio coding. Those predecessors of new IEEE standard for advanced audio coding are listed in the table 1, presenting the varieties of audio lossless coding in term of their compression ratio, encoding and decoding speed and category of coding.

Table 1: List of lossless audio compressors [3].

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

2.1. MPEG4-ALS

As an addition to the MPEG-4 audio standard, Audio Lossless Coding (ALS) will define methods for lossless coding of audio signals with arbitrary sampling rates, resolutions of up to 32 bit, and up to 256 channels. In July 2003, the lossless codec from Technical University of Berlin was chosen as the first working draft. Since then, further improvements and extensions have been integrated. MPEG-4 ALS is expected to become an international standard by the end of 2005[7].

Examples for the use of lossless audio coding in general and MPEG-4 ALS in particular include both professional and consumer applications: archival systems (broadcasting, studios, record labels, libraries), studio operations (storage, collaborative working, digital transfer), high-resolution disc formats, internet distribution of audio files, online music stores (download), portable music players. The general block diagram for MPEG 4-ALS is given in [4].

2.2. MPEG4-SLS

MPEG-4 Scalable Lossless Coding () is a transform coding technology for lossless audio compression [43, 95]. It uses a hybrid lossy or lossless coding structure as shown in [3]. Input PCM audio samples are first transformed by a reversible, integer-to-integer transform called Integer Modified Discrete Cosine Transform (IntMDCT) into IntMDCT coefficients. These coefficients are passed to a two-layered structure consisting of a core layer and an enhancement layer. The core layer uses lossy audio coders such as the MPEG-4 AAC to generate a lossy-coded bit-stream. In the AAC coder, the IntMDCT coefficients are quantized at step-sizes controlled by a psychoacoustic model so that the quantization noise is best masked by the human auditory system. The quantized coefficients are subsequently entropy-coded to generate the lossy bit-stream. An error mapping process thus removes this part of information from the IntMDCT coefficients, and generates a residual signal. The residual signal is subsequently entropy-coded using bit-plane coding s to generate a scalable bit-stream that complements the lossy-coded bit-stream of the core layer. In the SLS decoder, the lossy bit-stream is decoded by an AAC decoder to generate a lossy audio signal for playback. The enhancement bit-stream is entropy-decoded back to the IntMDCT residual, which is added to the output of the error mapping process to recover the full IntMDCT coefficients. These coefficients are transformed by an Integer Inverse Modified Discrete Cosine Transform (IntMDCT) back to the original audio samples.

2.3. FLAC (Free Lossless Audio Coding)

FLAC originated in 2000, originally developed by Josh Coalson. In order to compress audio it uses a four stage method: blocking, inter-channel decorrelation, prediction, and residual coding the input to FLAC is an uncompressed audio file, either Wave or Aiff. The blocking stage divides the audio signal into blocks or portions of a specified size. This directly affects the compression ratio. If the block is too small, the total number of blocks will increase, wasting bits on encoding headers. The inter-channel decorrelation stage (or mid-side conversion) is performed by removing redundancy in the stereo signals' left and right channels. Often, the left and right channels of a stereo signal are very similar. Thus, mid-side conversion was devised to reduce the amount of bits it takes to store the left and right channels. By encoding the left and right channels into a middle channel (left and right average) and a side channel (left minus right), the amount of bits needed to store the signal can be reduced. In cases where the left and right channels are very different, it can be passed without any decorrelation. The prediction stage is essential for providing good compression. It is dependent on the efficiency of the block size chosen in the first stage. By using linear prediction and run-length encoding, FLAC predicts how each block could be most closely modelled. For example, if the current block resembles a sine wave, then FLAC would ideally choose a sine wave to model the signal in that block. Many times, the prediction takes less memory to store than the original signal. This leads to the final encoding step, residual coding

3. IEEE 1857.2 LOSSLESS AUDIO COMPRESSION

The new IEEE standard of audio coding named IEEE 1857.2 is the latest technique or algorithm for audio coding. The general block diagram of IEEE 1857.2 lossless audio compression system is shown in figure 1, in which top part denotes the encoder, and the bottom part the decoder.

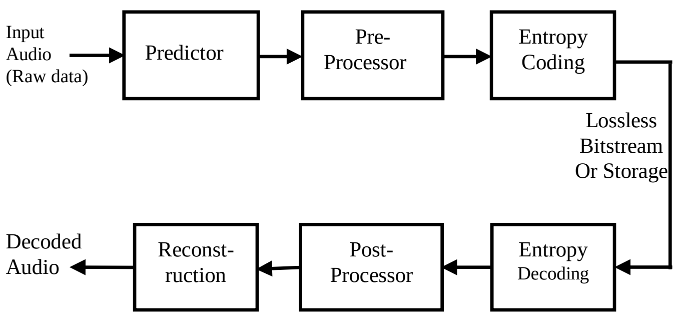

Figure 1: General lossless audio coding: encoder (top) and decoder (bottom)[1].

In encoding or compression, input audio samples are first processed by a predictor, which removes correlations in the input audio sample and, generates a prediction residue which is given by equation (1).

$$d[n]=\begin{cases} x[n] & n=0 \\ x[n]-\left\lfloor \dfrac{2^{19}+\sum_{k=1}^{n} c[n][k].x[n-k]}{2^{20}} \right\rfloor & 1\le n<lpc\_order \\ x[n]-\left\lfloor \dfrac{2^{19}+\sum_{k=1}^{lpc\_order} c[pc\_order][k].x[n-k]}{2^{20}} \right\rfloor & lpc\_order\le n<N \end{cases} \quad (1)$$

Where x is the original input for level =0 and d[n] denotes de LPC residue of the input x(n) with n= 0….., N-1.

The prediction residue is then pre-processed where the signal is flattened. The flattened prediction residue is subsequently coded by entropy coder inter lossless bitstream. In decoding part, a reserve process is performed, part in which the lossless bitstream is entropy decoded, post-processed (de-flattened), and losslessly reconstructed to a decoded signal that is an exact replication of the original input audio.

The Inter-Channel Decorrelation module is used to remove the redundancy between multi-channels. The output from this module is then fed into the Integer Lifting Wavelet Transform and denotes by figure 2, module to further decompose the signal.

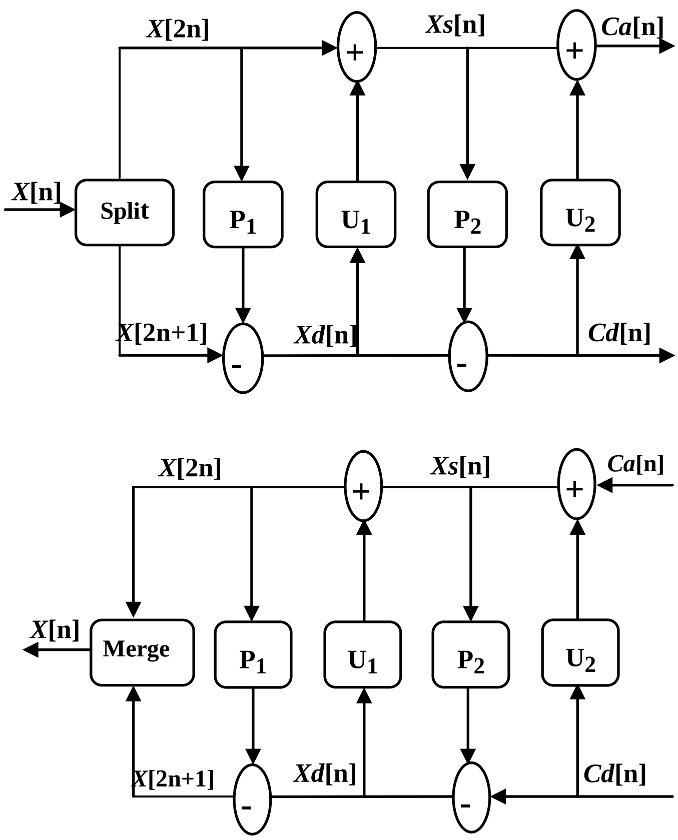

Figure 2: Diagram of integer lifting wavelet transform for analysis and synthesis, top and bottom respectively [2].

Where P denotes Prediction and U update.

Each sub-band is then LPC predicted and named PARCOR coefficient, they are computed using the Levinson-Durbin algorithm, given by the equation (3) and (4) to generate the prediction residue, for PARCOR analysis and synthesis respectively. To reduce the dynamic range of the residue signal value, residual signal is normalized by the Preprocessing module given by equation (2).

$$
\begin{aligned}
&y(0) = x(0) \\
&y(i) = \sum_{j=1}^{i} a_j^{(i)} x(i-j), \qquad 1 \le i \le p-1, \\
&y(i) = \sum_{j=1}^{p} a_j^{(p)} x(i-j), \qquad 1 \ge p,
\end{aligned}
\tag{2}
$$

With i is the time index to the samples in a frame, $x(i)$ and $y(i)$ denote input samples and their predictions in the frame, the maximal predictor order is p, and the coefficient at each predictor tap is denoted as aj, j = 1, . . . , p, the superscript in $a(i)$. j indicates that the coefficients are computed for predictor of order i.

The normalized residual signal is then entropy coded to generate the output bitstream. The output bitstream includes: entropy bitstream, LSB and signal signs from Preprocessing module, quantized LPC coefficients, and the side information from Wavelet Transform module.

$$
b_k = \begin{cases}
\left\lfloor 64\left(\ln\left(\frac{2}{3}+\frac{5}{2}\sqrt{\frac{1+p_1}{2}}\right)/\ln\left(\frac{3}{2}\right)\right)\right\rfloor & k=1 \\
\left\lfloor 64\left(\ln\left(\frac{2}{3}+\frac{5}{2}\sqrt{\frac{1-p_2}{2}}\right)/\ln\left(\frac{3}{2}\right)\right)\right\rfloor & k=2 \\
\left\lfloor 64 p_k \right\rfloor & k=3,....,lpc\_order
\end{cases}
\tag{3}
$$

Where b_k denotes the resulting quantized values and P_k is the PARCOR coefficients (P_k, $k = 1,….,lpc\_order$).

$$
par\lfloor k \rfloor = \begin{cases}
\Gamma(b_k) & k=1 \\
-\Gamma(b_k) & k=2 \\
b_k 2^{14} + 2^{13} & k=3,...,lpc\_order
\end{cases}
\tag{4}
$$

The lossless bitstream is de-multiplexed first, and the bitstream is entropy decoded to get the residue using arithmetic coding or Golomb-Rice decoding indicates by figure 3.

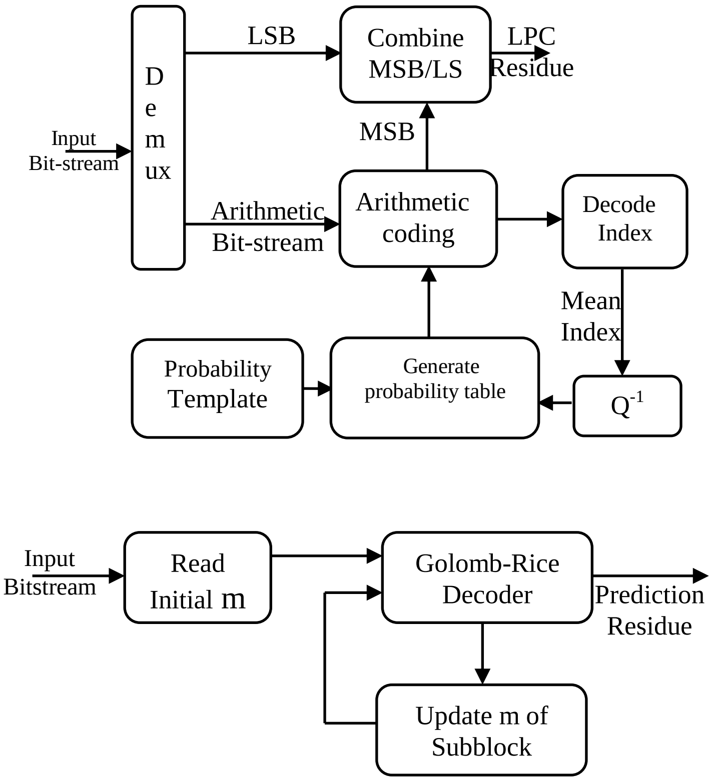

Figure 3 : Block diagram of the arithmetic decoder and Golomb-Rice ddecoder top and bottom resptively[2]

The LSB signals, the signal signs, and the LPC coefficients are used to de-normalize the residue signal in the post-processing module. The output from the Post-processing is then reconstructed by the LPC to get there constructed sub-band signals. These sub-band signals are further transformed by the wavelet transform module and inter-channel correlation module to reconstruct the original audio signal using the following equation.

$$
\begin{aligned}
h_1(z) &= (1+z^{-1}) \prod_{i=1,3,...,p-1} (1-2q_i z^{-1} - z^{-2}) \\
h_2(z) &= (1-z^{-1}) \prod_{i=2,4,...,p} (1-2q_i z^{-1} - z^{-2})
\end{aligned}
\tag{5}
$$

With $h_1(z)$ is symmetric polynomial and $h_2(z)$ is antisymmetric polynomial.

DOI: 10.18535/ijecs/v5i4.19

4. Conclusion

From years 2000, the lossless data compression has known so many algorithms for lossless compression. The most popular algorithms such as MPEG4-ALS[4,5,7], MPEG4-SLS [3], FLAC[6] are revolutionized the world of mobile communication, digital music, online music and video playing (streaming) and other areas of audio and speech are concerned. As the optimization of communication resource, storage media and other assets are become a crucial problem for researchers, the IEEE-SA (standard association) has launched in August 2013 a new standard for advanced audio coding. This algorithm for AAC is also named as 1857.2 standard, is an efficient lossless audio codec (coding decoding) technique, in improving audio quality for compression and decompression, optimizing bandwidth during transmission, saving storage space, it also speeds up the video streaming, audio streaming and others video or audio data. This standard defines a set of tools to support specific audio coding functions, including general audio coding and lossless coding.

REFERENCES

[1] H. Huang, H. Shu, and R. Yu, "LOSSLESS AUDIO COMPRESSION IN THE NEW IEEE STANDARD FOR ADVANCED AUDIO CODING," 2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP).

[2] IEEE 1857.2-2013, "IEEE Approved Draft Standard for Advanced Audio Coding," Aug. 2013.

[3] Haibin Huang, **"**Lossless Audio Coding for MPEG-4,**"** University Of Joensuu Computer Science and Statistics dissertations 17.

[4] Tilman Liebchen, Takehiro Moriya, Noboru Harada, Yutaka Kamamoto, and Yuriy A. Reznik **"**The MPEG-4 Audio Lossless Coding (ALS) Standard - Technology and Applications," Audio Engineering Society Convention Paper Presented at the 119th Convention 2005 October 7–10 New York, NY, USA.

[5] Tilman Liebchen, Yuriy A. Reznik, **"**MPEG-4 ALS : an Emerging Standard Lossless Audio Coding," Technical University of Berlin, Einsteinufer 17,D-10587 berlin, Germany RealNetwork, Inc. 2601 Elliot Avenue, Seattle Wa 98121.

[6] Microsemi,"SmartFusion cSoC: Implementation of FLAC Player Using Hardware and Software Partitioning," www.microsemi.com/.../129825-ac376.

[7] Peter Noll, Tilman Liebchen, "Lossless and Perceptual Coding of Digital Audio," Technische Universität Berlin, Fachgebiet Nachrichtenübertragung (formerly Fernmel de technik) noll@nue.tu-berlin.de, liebchen@nue.tu-berlin.de.

AUTHOR PROFILE.

FOMAZOU NSELAPI Auristin received B.E. degrees in Electrical Engineering and Telecommunication from Douala institute of technology, Cameroon in 2013 and pursuing his M.E. in Electronic and Telecommunication since 2014, at Sinhgad College of Engineering, Pune.

SAMADHAN Mali received M.E. degree in Electronics from Sinhgad College of engineering, Pune in 2010 and pursuing his Ph.D. in Electronics and Telecommunication Engineering since 2014, at Sinhgad College of engineering, Pune.